

**Development of Chemical Isotope Labeling Liquid Chromatography Mass
Spectrometry for Cellular Metabolomics**

By

Xian Luo

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

Doctor of Philosophy

Department of Chemistry

University of Alberta

© Xian Luo, 2018

Abstract

Cellular metabolomics is the global study of a whole set of metabolites and its chemical processes in cells. Different analytical platforms, such as NMR and MS, have been employed for profiling the highly complex cellular metabolome and various data processing approaches have been developed. Compared with metabolomic profiling of biofluids, more challenges remain in cellular metabolomics. For instance, additional sample handling steps, including cell harvest, metabolism quenching and cell lysis, are always required. The sample handling step should be fully optimized and compatible for downstream analysis. In some situations, the amount of cells is extremely limited, but the sensitivity of most analytical platforms cannot meet the requirement for in-depth metabolomic profiling of such small amount of samples.

Towards these challenges, the objective of my thesis is to establish efficient sample handling methods and develop high performance analytical platform for cellular metabolomics. First, a technique combining efficient cell lysis method and chemical isotope labeling (CIL) LC-MS for comprehensive metabolomic profiling of *S. cerevisiae* was developed (Chapter 2). This technique was used to investigate the metabolome change of *S. cerevisiae* under nitrogen starvation and ultrasonication stimulus (Chapter 2 and 3). Second, a rapid and efficient method was developed for harvesting and lysing adherent mammalian cells (Chapter 4). Third, high performance CIL nanoLC-MS was developed and used for high-coverage metabolomic profiling of

small numbers of mammalian cells (Chapter 5). This technique was applied to profile the metabolome of circulating exosomes from pancreatic cancer patients (Chapter 6).

Overall, this thesis describes the development of various approaches for handling different types of cells tailored to CIL LC-MS based metabolomics and the development of highly sensitive CIL nanoLC-MS and its application for analyzing small numbers cells and circulating exosomes.

Preface

A version of Chapter 2 was published as: Xian Luo, Shuang Zhao, Tao Huan, Difei Sun, R. Magnus N. Friis, Michael C. Schultz and Liang Li “High-Performance Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry for Profiling the Metabolomic Reprogramming Elicited by Ammonium Limitation in Yeast” *J. Proteome Res.* 15, 5, 1602-1612. I was responsible for data collection and analysis, and manuscript preparation. S. Zhao helped with data collection and analysis. T. Huan and D. Sun helped with metabolite identification. M. Friis and M. Schultz analyzed the related metabolic pathway. M. Schultz and L. Li proposed the original idea. L. Li supervised the project and edited the manuscript.

A version of Chapter 3 was published as: Michael C. Schultz, Jian Zhang, Xian Luo, Oleksandra Savchenko, Liang Li, Michael Deyholos and Jie Chen “Impact of Low-intensity Pulsed Ultrasound on Transcript and Metabolite Abundance in *Saccharomyces cerevisiae*” *J. Proteome Res.* 16, 8, 2975-2982. I conducted the metabolomic profiling and drafted the metabolomics data analysis discussion. M. Schultz wrote the main manuscript text and interpreted the high-throughput data, J. Zhang proposed the original idea, designed and conducted the RNA studies, O. Savchenko carried out the ultrasound experiments, L. Li supervised the metabolomic profiling, M. Deyholos supervised and helped design the RNA studies, J. Chen proposed the original ideas, wrote and finalized the manuscript text, monitored the research progress.

A version of Chapter 4 was published as: Xian Luo, Xinyun Gu and Liang Li “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*, in press. I was responsible for design of experiment, data analysis and manuscript preparation. X. Gu collected the data. L. Li supervised the project and edited the manuscript.

A version of Chapter 5 was published as: Xian Luo and Liang Li “Metabolomics of Small Numbers of Cells: Metabolomic Profiling of 100, 1000 and 10000 Human Breast Cancer Cells” *Anal. Chem.* 89, 21, 11664–11671. I was responsible for the design of experiment, data collection and analysis, and manuscript preparation. L. Li supervised the project and edited the manuscript.

A version of Chapter 6 was published as: Xian Luo, Mingrui An, Kyle C. Cuneo, David M. Lubman and Liang Li “High-Performance Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry for Exosome Metabolomics” *Anal. Chem.* in press. I was responsible for data collection, data analysis and manuscript preparation. M. An prepared exosome samples. K. Cuneo provided exosome samples. D. Lubman and L. Li supervised the project and edited the manuscript.

Acknowledgements

First of all, I would like to express my sincere appreciation to my supervisor, Professor Liang Li, for his invaluable suggestion, guidance, support and encouragement. I cannot complete this thesis work without his persistent help.

I would like to thank my supervisory committee member, Professor James Harynuk and Professor Michael Schultz, for their advice and help in the last five years. My deep gratitude also goes to all the people whom I have collaborated with, specially, Professor Michael Schultz at the Department of Biochemistry, University of Alberta, Professor Jie Chen at the Department of Electrical and Computer Engineering, Professor David M. Lubman at the Department of of Surgery, University of Michigan, Ann Arbor.

I would also like to thank all the members from Professor Liang Li group. In particular, I would like to thank Dr. Difei Sun, Dr. Chiao-Li Tseng, Dr. Yiman Wu and Professor Liyan Liu, for their training at the beginning of my research. I would like to thank Dr. Nan Wang, Jiamin Zheng, Dr. Zhendong Li, and Dr. Tao Huan for their suggestion. I would like to express my gratitude to all the other group members I worked with: Dr. Wei Han, Yunong Li, Jaspaul Tatlay, Kevin Hooton, Dorothea Mung, Shuang Zhao, Adriana Zardini Buzatto, Xiaohang Wang, Wan Chan, Xinyun Gu, Hao Li, Minglei Zhu.

Thanks to Biological Services Facility, Gareth Lambkin, for his training in cell culture. Thanks to Mass Spectrometry Facility, Dr. Randy Whittal, Bela Reiz and Jing Zheng for their training in FTICR and MALDI-TOF. Thanks to the lab coordinators

of General Chemistry and Analytical Chemistry, Dr. Norman Gee and Dr. Gregory Kiema. The skills I have learnt from being a teaching assistant are invaluable.

Finally, my special thanks to my parents, Mr. Xinlai Luo and Mrs. Hongxia Liu, for their countless efforts to support me and unconditional love over the years.

Table of Contents

Chapter 1 Introduction.....	1
1.1 General Introduction	1
1.2 Cell Harvest	3
1.3 Cell Wash.....	5
1.4 Metabolism Quench.....	6
1.5 Cell Lysis	7
1.6 Metabolite Extraction.....	10
1.7 Sample Normalization	15
1.8 Analytical Platform.....	17
1.8.1 NMR	18
1.8.2 GC-MS.....	19
1.8.3 LC-MS	21
1.8.3.1 Liquid Chromatography.....	21
1.8.3.2 Mass Spectrometry.....	23
1.8.4 CE-MS	29
1.9 Data Process.....	29
1.10 Overview of Thesis	32
Chapter 2 High-Performance Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry for Profiling the Metabolomic Reprogramming Elicited by Ammonium Limitation in Yeast	33
2.1 Introduction.....	33

2.2 Experimental Section	36
2.2.1 Workflow	36
2.2.2 Chemical and reagents	38
2.2.3 Cell growth and cell lysis.....	38
2.2.4 Dansylation labeling	39
2.2.5 DmPA bromide labeling	40
2.2.6 LC-UV quantification	40
2.2.7 LC-MS	40
2.3 Results and Discussion	42
2.3.1 Cell lysis for metabolomic analysis	42
2.3.2 Analysis of amine- and phenol-containing metabolites.....	45
2.3.3 Analysis of carboxylic acid containing metabolites	50
2.3.4 Metabolic pathway analysis	54
2.4 Conclusions.....	57
Chapter 3 Impact of Low-intensity Pulsed Ultrasound on Transcript and Metabolite Abundance in <i>Saccharomyces cerevisiae</i>	58
3.1 Introduction.....	58
3.2 Materials and Methods.....	59
3.2.1 Cell Culture.....	59
3.2.2 Inoculum preparation.....	59
3.2.3 Transcript Analysis	60
3.2.3.1 Statistical analysis.....	61

3.2.4 Metabolomic Profiling	62
3.2.4.1 Chemicals and reagents.....	62
3.2.4.2 Cell lysis and metabolite extraction ⁶	62
3.2.4.3 Dansylation labeling	62
3.2.4.4 Sample Normalization	63
3.2.4.5 LC-MS	63
3.2.4.6 Data Processing and Analysis	64
3.3 Results.....	64
3.3.1 Analysis of Annotated Transcripts.....	64
3.3.2 Metabolomic Profiling.....	67
3.4 Discussion	77
3.5 Conclusion	78
Chapter 4 Development of a Simple and Efficient Method of Harvesting and	
Lysing Adherent Mammalian Cells for Chemical Isotope Labeling LC-MS-	
Based Cellular Metabolomics	79
4.1 Introduction.....	79
4.2 Experimental	81
4.2.1. Overall workflow	81
4.2.2 Chemicals and reagents.....	83
4.2.3 Cell culture.....	83
4.2.4 Cell harvest	83
4.2.5 Cell lysis.....	84

4.2.6 Dansylation labeling	85
4.2.7 LC-UV	85
4.2.8 LC-MS	86
4.2.9 Data analysis	86
4.3 Results and Discussion	87
4.3.1 LC-UV quantification for cell harvest and lysis efficiency comparison	87
4.3.2 LC-MS results	90
4.3.3 Multivariate statistical analysis	92
4.3.4 Impact of different harvest methods on cellular metabolome	94
4.3.5 Impact of different lysis methods on cellular metabolome.....	119
4.4 Conclusions.....	122
Chapter 5 Metabolomics of Small Numbers of Cells: Metabolomic Profiling of 100, 1000 and 10000 Human Breast Cancer Cells	123
5.1 Introduction.....	123
5.2 Experimental Section	126
5.2.1 Overall Workflow	126
5.2.2 Cell Culture and Harvest.....	128
5.2.3 Cell Lysis and Metabolite Extraction	128
5.2.4 Dansylation Labeling.....	129
5.2.5 LC-UV Quantification	130
5.2.6 Microflow LC-MS	130
5.2.7 Nanoflow LC-MS	131

5.2.8 Data Processing and Metabolite Identification	132
5.3 Results and Discussion	132
5.3.1 Cell Lysis and Metabolite Extraction	132
5.3.2 Dansylation Protocol.....	136
5.3.3 Captivespray MS.....	137
5.3.4 Injection Amount	141
5.3.5 Metabolomic profiling of small numbers of cells.....	143
5.4 Conclusions.....	164
Chapter 6 High-Performance Chemical Isotope Labeling Liquid	
Chromatography Mass Spectrometry for Exosome Metabolomics	166
6.1 Introduction.....	166
6.2 Experimental Section.....	167
6.2.1 Workflow	167
6.2.2 Serum Samples.....	170
6.2.3 Serum Exosome Isolation	170
6.2.4 Metabolites Extraction and Chemical Isotope Labeling.....	170
6.2.5 LC-MS	171
6.3 Results and Discussion	172
6.3.1 Exosomes Isolation	172
6.3.2 Metabolite Extraction.....	172
6.3.3 nLC-MS	173
6.3.4 Exosome Metabolome	175

6.3.5 Comparative Metabolomics of Exosomes	185
6.4 Conclusions.....	195
Chapter 7 Conclusion and Future Work.....	196
7.1 Conclusion	196
7.2 Future Work.....	199

List of Tables

Table 1.1 Summary of sample handling methods used in cellular metabolomics.	13
Table 2.1 List of amine- or phenol-containing metabolites identified.	46
Table 2.2 List of carboxylic acid-containing metabolites identified.	51
Table 3.1 Number of reads and percent genes mapped per sample.	61
Table 3.2 Low magnitude but statistically significant transcript abundance differences between control and LIPUS-treated samples (two-tailed T test).....	65
Table 3.3 Metabolites identified by searching against DnsID.....	69
Table 3.4 Metabolites that differ significantly between CON and 100 mW/cm ² and have been definitively identified.	74
Table 4.1 Metabolites with significantly different concentrations in HeLa cell extracts prepared using different harvest methods (trypsinization/physical-scraping).....	95
Table 4.2 Metabolites with significantly different concentrations in MCF-7 cell extracts prepared using different harvest methods (trypsinization/physical-scraping)	99
Table 4.3 List of metabolites from Hela cell lysates positively identified by searching against the DnsCI-labeled standard library.....	101
Table 4.4 List of metabolites from MCF-7 cell lysates positively identified by searching against the DnsCI-labeled standard library.....	106
Table 5.1 Physical properties of chemical vapors tested in this study	139
Table 5.2 List of metabolites detected from 100 cells identified based on accurate mass and retention time matches to the Dns-labeled standard	147

Table 5.3 List of metabolites detected from 1000 cells identified identified based on accurate mass and retention time matches to the DnsCl-labeled standard library	152
Table 5.4 List of metabolites detected from 10000 cells identified based on accurate mass and retention time matches to the DnsCl-labeled standard library.	158
Table 6.1 List of peak pairs detected from the 1st set of exosomes with positive metabolite identification based on accurate mass and retention time search against the dansyl standard library.....	177
Table 6.2 List of peak pairs detected from the 2nd set of exosomes with positive metabolite identification based on accurate mass and retention time search against the dansyl standard library.....	181
Table 6.3 List of metabolites with significant change in the 1st sample set. ID Level indicates this metabolite was identified or matched to different libraries. Level 1: DnsID Library; Level 2: Zero-Reaction MCID Library; Level 3: One-Reaction MCID predicted library.....	186
Table 6.4 List of metabolites with significant change in the 2nd sample set. ID Level indicates this metabolite was identified or matched to different libraries. Level 1: DnsID Library; Level 2: Zero-Reaction MCID Library; Level 3: One-Reaction MCID predicted library.....	190

List of Figures

Figure 1.1 Applications of cellular metabolomics	2
Figure 1.2 General workflow of cellular metabolomics	3
Figure 1.3 Comparison of different analytical platforms.....	18
Figure 1.4 Schematic of a liquid chromatography system.....	21
Figure 1.5 Schematic of porous silica particle covered with C18 stationary phase	22
Figure 1.6 Separation process of liquid chromatography	23
Figure 1.7 Schematic of ESI process	24
Figure 1.8 Schematic of quadrupole mass analyser.....	25
Figure 1.9 Schematic of time-of-flight mass analyser	26
Figure 1.10 General workflow of LC-MS based metabolomics data processing.....	30
Figure 2.1 Workflow of the differential chemical isotope labeling (CIL) LC-MS method for yeast metabolomics.	37
Figure 2.2 Microscopy images of cell lysis: (A) Control group without glass beads; the yeast cells remained intact after vortexing and (B) Disruption group with glass beads in 50% MeOH after vortexing; cell debris were shown as dark dots. (C) Comparison of cell lysis efficiency of four different solvents. Cell counting with hemocytometer was used to determine the percentage of cell breakage (n=3).	44
Figure 2.3 Score plots of (A) PCA and (B) PLS-DA of dansylation labeled lysates from <i>Saccharomyces cerevisiae</i> cultured with or without ammonium limitation. The data was from experimental triplicates on biological triplicate samples. (C) Volcano plot of the changes of amine- and phenol-containing metabolites. The up-regulated or	

down-regulated metabolites by at least 1.2-fold, with p-values of smaller than 0.05, are marked in red and green, respectively. (D) Clustered heatmap showing the comparison of the relative intensity of each peak pair in the nitrogen limitation group and the control group.49

Figure 2.4 (A) PC1 versus PC2 PCA score plot, (B) PC2 versus PC3 PCA score plot, and (C) PLS-DA score plot of DmPA labeled lysates from *Saccharomyces cerevisiae* cultured with or without ammonium limitation. The data was from experimental triplicates on biological triplicate samples. (D) Volcano plot of the changes of carboxylic acid-containing metabolites. The up-regulated or down-regulated metabolites by at least 1.2-fold, with p-values of smaller than 0.05, are marked in red and green, respectively.....53

Figure 2.5 Metabolite changes in selected metabolic pathways. The graphs are for selected metabolites (blue shading) that differ significantly in abundance ($p \leq 0.05$) between cells cultured with or without ammonium limitation. Fold differences in abundance are given in the name boxes for the metabolites.....55

Figure 3.1 PCA analysis of control and LIPUS-treated samples by RNA expression values. Sample points are plotted according to PCA of the log transformed RNA-seq data that were statistically significant by ANOVA.....67

Figure 3.2 PCA and PLS-DA analysis of control and LIPUS-treated samples by relative metabolite abundance. (A) PCA and (B) PLS-DA scoring plots of control, 80 mW/cm² and 100 mW/cm² treated groups. (C) PCA and (D) PLS-DA scoring plots of control and 100 mW/cm² treated groups.72

Figure 3.3 Volcano plots of metabolite abundance in LIPUS-treated cells relative to the control. (A) 80 mW/cm² versus the control. (B) 100 mW/cm² versus the control. The up- and down-regulated metabolites (1.2-fold threshold, $p \leq 0.05$) are marked in red and green, respectively. 73

Figure 3.4 Metabolic pathway analysis. (A) Pathway relationships of dysregulated metabolites suggested by manual mapping of the CON versus 100 mW/cm² metabolite data onto the metabolic chart of yeast. (B) Overview of metabolic pathway enrichment analysis result using Pathway Analysis tools in Metaboanalyst. 76

Figure 4.1 Workflow for comparing different methods to develop a simple and efficient cell harvest and lysis method for CIL LC-MS metabolomics of adherent mammalian cells 82

Figure 4.2 Average concentrations of dansyl labeled metabolites in cell extracts (n=6) prepared using different combinations of harvest and lysis methods from (A) HeLa and (B) MCF-7 cells. T-GB = trypsinization cell harvest followed by glass-bead lysis. T-FT = trypsinization cell harvest followed by freeze-thaw-cycle lysis. S-GB = scraping cell harvest followed by glass-bead lysis. S-FT = scraping cell harvest followed by freeze-thaw-cycle lysis. 89

Figure 4.3 Average number of peak pairs (n=6) detected from (A) HeLa cell extracts and (B) MCF-7 cell extracts prepared using different combinations of harvest and lysis methods..... 91

Figure 4.4 (A) PCA and (B) PLS-DA plots of the amine/phenol submetabolomes of HeLa and MCF-7 cells from cell extracted prepared using different combinations of

harvest and lysis methods. H=HeLa cells. M=MCF-7 cells. Other abbreviations are shown in Figure 4.2 caption.....93

Figure 4.5 Volcano plots of the amine/phenol submetabolomes of (A) HeLa and (B) MCF-7 cells harvested by different methods. The p-value of each metabolite was calculated from t-test, and the fold change (FC) was calculated from the peak ratios of trypsinization group divided by the peak ratios of the scraping group (i.e., trypsinization/scraping). Using a cut-off value of $p < 0.01$ and $FC > 1.5$ or < 0.67 , the red points represent the metabolites with higher concentrations in the trypsinization group, and the green points represent the metabolites with lower concentrations in the trypsinization group. The black points represent the metabolites with no significant differences..... 112

Figure 4.6 Heat maps showing 22 selected metabolites with significant concentration differences in cell extracts prepared using different harvest methods (scraping and trypsinization) from (A) HeLa and (B) MCF-7 cells. The metabolites in (A) are 1. Argininosuccinic acid; 2. N-Methylphenylethanolamine; 3. Homovanillin; 4. Pyridoxal; 5. Leucyl-Proline; 6. 4-Hydroxybenzoic acid; 7. Allylsine; 8. 5-Hydroxy-L-tryptophan; 9. Homogentisic acid; 10. 2-Methyl-3-hydroxy-5-formylpyridine-4-carboxylate; 11. Uridine; 12. Uridine-H₂O; 13. 3-Nitrotyrosine; 14. Guanosine; 15. Inosine; 16. 4-Aminobutyraldehyde; 17. Adenosine monophosphate; 18. Uridine 5'-monophosphate; 19. Guanosine monophosphate; 20. 5-Hydroxylysine; 21. Arginine; 22. Cysteineglutathion. The metabolites in (B) are 1. Homovanillin; 2. Homogentisic acid; 3. 4-Hydroxybenzoic acid; 4. Pyridoxal; 5. 2-Methyl-3-hydroxy-5-

formylpyridine-4-carboxylate; 6. 5-Hydroxy-L-tryptophan; 7. Argininosuccinic acid; 8. Leucyl-Proline; 9. Allysine; 10. Inosine; 11. Uridine; 12. Uridine-H₂O; 13. 3-Nitrotyrosine; 14. L-Arginine; 15. 5-Hydroxylysine; 16. Guanosine monophosphate; 17. Uridine 5'-monophosphate; 18. Adenosine monophosphate; 19. Guanosine; 20. N-Methylphenylethanolamine; 21. 4-Aminobutyraldehyde; 22. Cysteineglutathion.... 114

Figure 4.7 Metabolic pathways enrichment analysis. The x-axis represents the impact of pathway, and y-axis represents the p-value. 117

Figure 4.8 Metabolite changes in selected metabolic pathways. The box plots show the relative metabolite abundances in different harvesting groups. 118

Figure 4.9 Volcano plots for comparison of the amine/phenol submetabolomes of cell extracted prepared using different lysis methods: (A) HeLa and (B) MCF-7 cells harvested by scrapping; (C) HeLa and (D) MCF-7 cells harvested by trypsinization. The p-value was from t-test, and the fold change was calculated from glass-bead/freeze-thaw-cycle. The red points represent the metabolites with higher concentrations in glass-bead lysed samples, and the green points represent the metabolites with lower concentrations in glass-bead lysed samples. The black points represent the metabolites with no significant differences in the two lysis methods.. 120

Figure 5.1 Workflow for CIL LC-MS method development. 127

Figure 5.2 (A) Comparison of peak pair numbers detected from the ultrasonic cell lysis method and the glass-bead-assisted cell lysis method. (B) Venn diagram of peak pair numbers from the two methods. (C) Comparison of peak pair numbers detected from different metabolite extraction solvents. (D) Venn diagram of peak pair numbers

detected from biological triplicate analysis using MeOH extraction. Data were from experimental triplicate analysis of three biological replicates (n=9).	134
Figure 5.3 Venn diagrams of peak pair numbers detected from biological triplicate analysis using (A) ACN extraction and (B) AMW extraction. For each biological replicate, experimental triplicate analyses were carried out.	135
Figure 5.4 (A) Peak areas of molecular ions of 9 dansyl labeled metabolites detected with the use of different dopant chemical vapors. Normalized peak areas of molecular ions of 4 dansyl labeled metabolites detected at (B) different temperatures of dry gas and (C) different capillary voltages. Data were from experimental triplicate analysis (n=3).....	140
Figure 5.5 Average peak pair numbers detected using (A) microflow LC-MS and (B) nanoLC-MS as a function of injection amount of $^{12}\text{C}/^{13}\text{C}$ -labeled cell lysates (n=3).	142
Figure 5.6 (A-C) Total ion chromatograms of labeled cell lysates, (D-F) extracted ion chromatograms of Dns-uridine, and (G-I) molecular ion regions of the $^{12}\text{C}/^{13}\text{C}$ -Dns-uridine peak pair obtained from injection of 1/10 of the labeled 10000-cell lysate (top), all labeled 1000-cell lysate (middle), and all labeled 100-cell lysate (bottom).	144
Figure 5.7 (A) Average peak pair numbers detected from $^{12}\text{C}/^{13}\text{C}$ -labeled 100-, 1000-, and 10000-cell lysates (duplicate analysis of three biological replicates or n=6). Venn diagrams of peak pair numbers detected from biological triplicate analysis of $^{12}\text{C}/^{13}\text{C}$ -labeled cell lysates of (B) 100, (C) 1000, and (D) 10000 cells.....	145
Figure 6.1 Workflow for exosome metabolomics based on CIL nLC-MS.	169

Figure 6.2 (A) Total ion chromatogram of a representative ^{12}C -/ ^{13}C -labeled exosome sample. (B) Number of peak pairs identified or matched in three libraries. (C) Venn diagram showing the commonly detected peak pairs from 1st sample set and 2nd sample set..... 174

Figure 6.3 (A,B) PCA plots, (C,D) PLS-DA and (E,F) volcano plots of exosome metabolomes from the 1st sample set (n=10) and the 2nd sample set (n=10). 193

Figure 6.4 Box plots of four identified metabolites with significant changes in exosomes of serum samples collected from cancer patients before and after chemotherapy treatment..... 194

List of Abbreviations

ACN	Acetonitrile
ANOVA	Analysis of Variance
BPC	Base Peak Chromatogram
BSA	Bovine Serum Albumin
CE	Capillary Electrophoresis
CID	Collision-Induced Dissociation
CIL	Chemical Isotope Labeling
Da	Dalton
DnsCl	Dansyl Chloride
DnsHz	Dansylhydrazine
DmPA	Dimethylaminophenacyl
EI	Electron Impact Ionization
ESI	Electrospray Ionization
FC	Fold Change
FT-ICR-MS	Fourier Transform Ion Cyclotron Resonance Mass Spectrometry
GC	Gas Chromatography
GC-MS	Gas Chromatography Mass Spectrometry
HMDB	Human Metabolome Database
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High Performance Liquid Chromatography

ICR	Ion Cyclotron Resonance
LC	Liquid Chromatography
LC-MS	Liquid Chromatography Mass Spectrometry
LC-UV	Liquid Chromatography Ultraviolet
MALDI	Matrix-assisted Laser Desorption Ionization
MeOH	Methanol
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
m/z	Mass to Charge
nm	Nano meter
NMR	Nuclear Magnetic Resonance
PCA	Principal Component Analysis
PLS-DA	Partial Least Square Discriminant Analysis
ppm	part(s) per million
QC	Quality Control
Q-TOF-MS	Quadrupole Time-Of-Flight Mass Spectrometry
ROC	Receiver Operating Characteristic
RT	Retention Time
RPLC	Reversed Phase Liquid Chromatography
RSD	Relative standard deviation

S/N	Signal to noise ratio
SPE	Solid Phase Extraction
SRM	Selected Reaction Monitoring
TOF	Time-Of-Flight
UPLC	Ultra Performance Liquid Chromatography
UV	Ultra-violet
VIP	Variable Importance on the Projection
μM	Micro Molarity

Chapter 1

Introduction

1.1 General Introduction

Cellular metabolomics aims to detect and quantify a whole set of small molecules (i.e. metabolites) in cells. Compared with animal models and human subjects, cell experiment has less variance and cellular metabolomics is easier to link with genomics and proteomics data.¹ Cellular metabolomics has broad applications in different research fields, including toxicology, pharmacology, clinical study and metabolic engineering. For instance, metabolomic profiling of human liver cancer HepG2 cell line could provide the comprehensive information to understand the molecular mechanism of drugs' liver toxicity.² Cellular metabolomics has also demonstrated its potential application in high-throughput drug discovery.³ In addition, metabolomic profiling of bacterial cell could provide the metabolic fingerprint for bacteria identification or differentiation.⁴ Figure 1.1 shows several applications of cellular metabolomics.

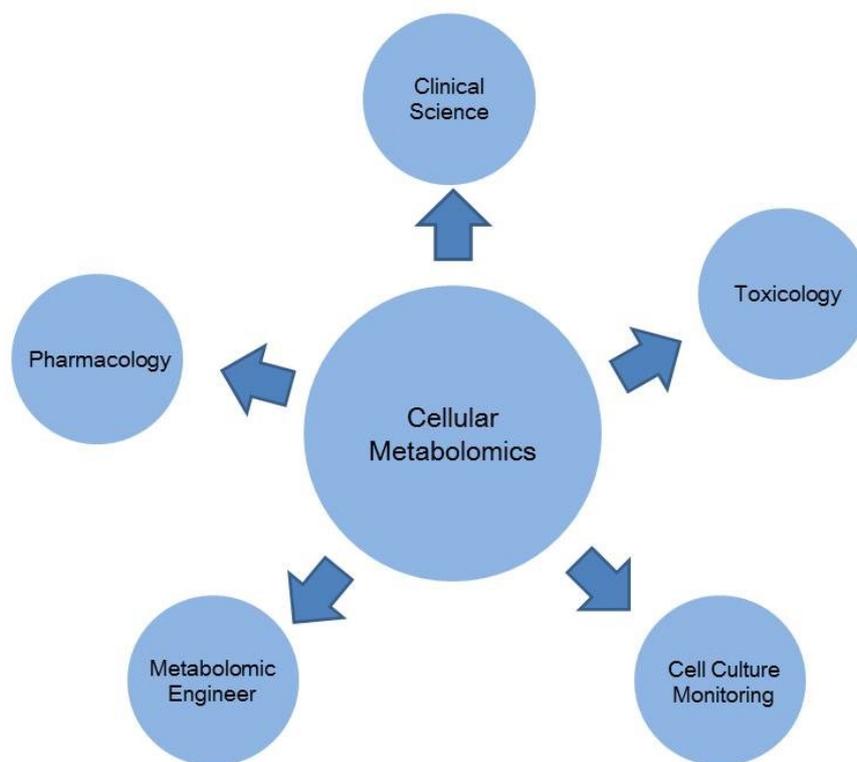


Figure 1.1 Applications of cellular metabolomics

Numberous challenges need to be addressed in cellular metabolomics. First, additional sample handling steps are usually necessary. Suitable cell harvest and metabolism quenching methods should be selected to ensure that the precise information from cellular metabolome can be acquired. Cell lysis is another key step in sample handling: on one hand, the lysis method should have high efficiency; on the other hand, lysis method should not affect the downstream sample analysis. Second, in some cases, the amount of cells is extremely low, for instance, one of the promising cancer biomarkers, circulating tumor cells (CTCs), only present 1-10 counts per mL of peripheral blood.⁵ The sensitivity of most analytical platforms is still not sufficient to achieve high-coverage metabolomic profiling of CTCs.

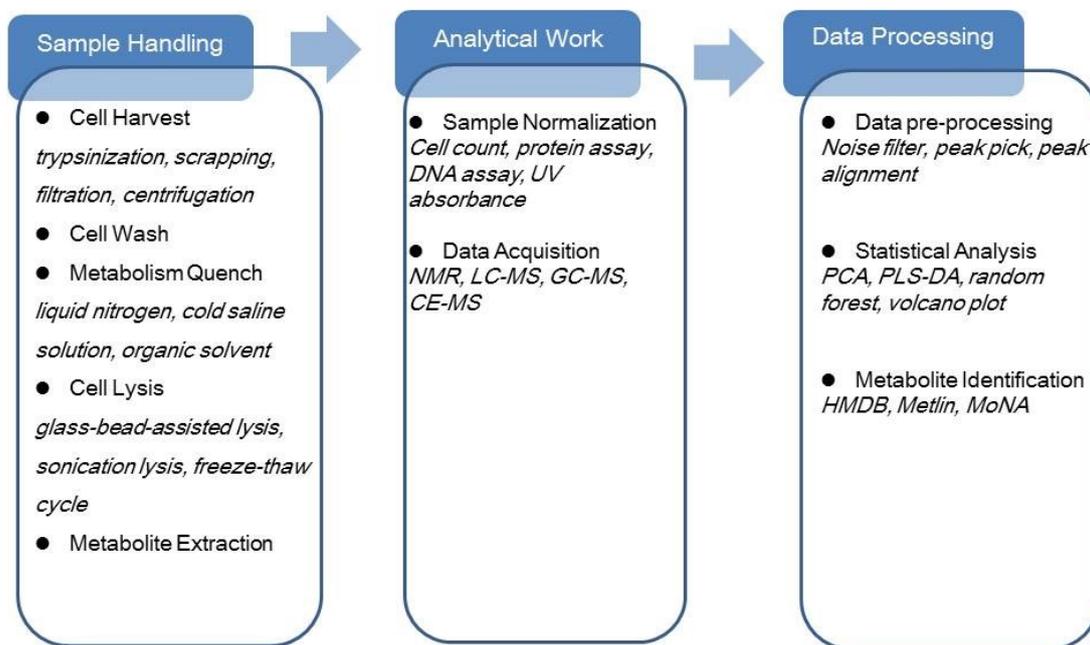


Figure 1.2 General workflow of cellular metabolomics

The general workflow of cellular metabolomics includes sample handling, analytical work and data processing, as shown in Figure 1.2. The following part of introduction will introduce the recent achievements of cellular metabolomics.

1.2 Cell Harvest

There are two basic systems in cell culture: adherent culture (anchorage-dependent culture) and suspension culture (anchorage-independent). In adherent culture system, cells need to be cultured on a specially treated substrate, whereas cells are free-floating in the culture medium in suspension culture system. Most mammalian cells are anchorage-dependent, such as the oldest and most widely used cervical cancer cell line, HeLa, and breast cancer cell line, MCF-7. Not all of the mammalian cells are anchorage-dependent, for example, lymphoblast-like cells (e.g.

Raji cell) are grown in suspension. Bacterial and yeast cells are usually cultured in suspension.

Suspension-cultured cells are usually harvested by centrifugation,⁶⁻⁸ and filtration was also used in some studies.^{9,10} Centrifugation is considered a time-consuming harvest method, as it takes a couple of minutes to pellet the cells. Filtration takes less time, but the filters may easily get blocked.¹¹ To avoid possible metabolome changes, either centrifugation or filtration should be operated under temperature control conditions (usually at 4°C).

The adherent cell harvest requires enzymatic or physical dissociation. Trypsin is the most commonly used enzyme for cell detachment and sub-culturing; however, many studies have confirmed trypsinization could cause metabolite leakage and metabolome change. For instance, K. Dettmer et al. employed GC-MS to monitor the amino acids concentration in growth medium after trypsin was added in cell culture.¹² The amino acids concentration increased as trypsinization time increased, indicating metabolite leakage during trypsinization. Using LC-MS, J. C. García-Cañaveras et al. comparatively analyzed the metabolome of Hep G2 cells harvested by trypsinization or by scraping.¹³ About 20% to 30% lower feature intensities were observed in trypsin treatment group under both RP ESI (+/-) and HILC ESI (+/-) condition. The concentrations of some energy metabolism related metabolites (e.g. ATP, AMP) and oxidative stress marker cysteine-glutathione changed during trypsinization, suggesting that the trypsinization process might induce the oxidative stress. CIL LC-MS was also applied to profile the metabolome of trypsinization or scraping treated MCF-7

breast cancer cells.¹⁴ By use of the CIL technique, the total concentration of metabolites can be directly determined by LC-UV. The quantification results demonstrated that the total concentration of scrapping treatment group is 1.8-fold higher than trypsinization group, suggesting that trypsin treatment caused severe metabolite leakage. Thus, direct scrapping should be considered as an optimal method for harvesting adherent cell culture. In addition, compared with scrapping, trypsinization is more time-consuming: it takes 1-5 min for digestion and 5-10 min for quenching and washing. However, in some special circumstances, using trypsinization for cell harvest is still required. For instance, if the exact number of cells needs to be controlled in an experiment, the cells should be detached from substrate and dissociated as single cells by trypsinization first, then counted by hemocytometer or flow cytometer.¹⁵

1.3 Cell Wash

Intercellular metabolites usually attract more attention than extracellular metabolites, thus cell wash step is designed to remove the residual metabolites from growth medium. Using an improper washing method may not remove the residual metabolites efficiently and could cause the metabolite leakage.

The most common wash solution for mammalian cell is cold phosphate buffer saline (PBS). As a buffer solution, PBS can maintain the constant pH (7.4 for most of the cultured mammalian cells) during washing processes, and the osmolality of PBS for cell washing is prepared as same as the mammalian cells (~315 mOsm/kg).¹⁶ Thus,

selecting PBS as a wash solution could avoid cell membrane damage and metabolite leakage to the utmost extent. Other wash solutions can be selected for different purposes. For example, instead of using PBS, 0.9% NaCl was selected to remove the remaining phosphate from growth medium.¹⁷ Ammonia acetate was used as a washing buffer to avoid sodium interference.¹⁸

Cell washing protocol is different for scraping and trypsinization harvest. In scraping method, the growth medium is removed by aspiration, then cell monolayer is detached, followed by washing with pre-cooled wash solution.¹⁹ In trypsinization method, the fetal bovine serum (FBS) or growth medium is added into culture to inactivate the excess trypsin after cells are detached from culture dish. To completely remove the growth medium and trypsin, the cells are centrifuged and the pellet is washed by cold wash solution for two or three times.^{15,17,20,21} Obviously, washing process is much less laborious and time-consuming in scraping harvest.

A variety of buffer solutions for bacterial cell washing were reported, such as Tris/EDTA buffer for *S. warneri*, 0.9% NaCl for *E. coli*, and potassium phosphate buffer (pH 7.0) for *E. coli*. The thick cell wall of yeast could preserve cell integrity, thus, cold LC-MS grade water can be used for yeast (*S. cerevisiae*) washing.⁶

1.4 Metabolism Quench

With enzymes' catalysis, metabolic reactions take place continuously inside of cells. Quenching step aims to stop the cellular metabolism by inhibiting or stopping

the enzymes' activity. Common cellular metabolism quenching methods include organic solvent washing, cold buffer washing and liquid nitrogen freezing.

-40 °C 60% methanol is one of the most widely used organic solvents for quenching the metabolism of *S. cerevisiae*. However, metabolite leakage was observed during the quenching.²² A. B. Canelas et al. optimized the quenching condition, and suggested that the final concentration of methanol used for quenching *S. cerevisiae* should be larger than 83% and solvent temperature should be lower than -40 °C to prevent metabolite leakage. 60% methanol was also reported to be applied in bacterial cell quenching, but metabolite leakage also existed.²³ Some studies have proved buffer additives (e.g., NaCl,¹¹ tricine,²⁴ and 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)²⁵) could be employed for reducing metabolite leakage. R. V. Kapoore et al. evaluated different types of additives and found 60% methanol with HEPES as additive could minimized metabolite leakage.²⁶

Cold salt solution (e.g. 0.9% NaCl, PBS) can also quickly slow down the cellular metabolism and avoid damaging cell membrane, it could be used in metabolism quenching.¹¹ Liquid nitrogen freezing is suggested to be used for adherent cell metabolism quenching, since it is easy to perform and could stop the metabolism reaction immediately.^{27,28} Another advantage of liquid nitrogen freezing is that it can rapidly evaporate and thus will not affect downstream sample handling.

1.5 Cell Lysis

Cell lysis is to break the cell membrane and cell wall to assist intercellular metabolites outside releasing. Chemical disruption and physical disruption are two of the major cell lysis methods.

Chemical lysis solutions are composed of high concentration of buffers (e.g. tris, acetate) or detergent (e.g., SDS, acid labile surfactant (ALS), NP-40).²⁹ Chemical disruption is gentle rapid and an efficient cell lysis method. However, these chemicals are usually difficult to remove and will affect downstream analysis in metabolomics. For instance, in GC-MS or CIL LC-MS based metabolomics, metabolites are required to be derivatized before being injected onto GC-MS or LC-MS, but the residual chemicals may affect the derivatization reactions. In addition, high concentration of salts and detergent can cause the ion suppression effect and will interfere MS analysis. More than that, salts and detergent could also damage LC column and affect separation efficiency. Desalting step is necessary if chemical disruption is used for cell lysis.

Physical disruption includes manual grinding, sonication, freeze-thaw cycles, microwave-assisted lysis, and glass-bead-assisted lysis. Manual grinding is an old and traditional lysis method that has been used for decades. Cell or tissue samples are frozen in liquid nitrogen first, and then crushed using a pestle and motor. This technique has been demonstrated to have the highest efficiency in lyzing some types of plant cells (e.g. *C. vulgaris*)³⁰.

Sonication is the technique widely used in mammalian³¹ and bacterial cell³² disruption. Ultrasonic waves create cavitation in cells, resulting in cell membrane damage. The major drawback of sonication lysis is that ultrasounds can generate a large amount of heat and could cause metabolite degradation.³³ Thus, samples should be processed in ice-water bath, and process time should not be too long. There is also a cross-contamination risk if the sonication probe is not cleaned thoroughly. Sonication cleaner was reported to be used for lysing cells in order to prevent the cross-contamination, and can also achieve high-throughput.³²

Bead-assisted lysis is to grind or vortex the samples with small beads and shear force is employed to disrupt the cells. Beads used for lysis can be made of glass, stainless steel or zirconia.³⁴ The size and vortex speed should be optimized for lysing different species of cells.³⁵ Glass-bead-assisted lysis could achieve more than 90% lysis efficiency in *S. cerevisiae* lysis.⁶

Freeze-thaw cycle lysis is to freeze the cells in liquid nitrogen or dry ice/acetone bath, and then thaw on ice-water bath. Ice crystals are formed inside of cells during the freeze-thaw process, causing cell swelling and finally resulting in cell disruption. Usually multiple rounds of freeze-thaw cycle is performed to ensure the metabolites released completely, thus freeze-thaw cycle is relatively time consuming. Adding with the proper enzyme (e.g. adding lysozyme in *E. coli*.) during freeze-thaw cycles could improve the lysis and extraction efficiency.³⁶

Microwave-assisted lysis was also reported in a few cases.^{37,38} The microwave irradiation induces the temperature shock and disrupts the cell membrane. However, some metabolites might get degraded under microwave irradiation, so it is not a commonly used lysis method in metabolomics.

1.6 Metabolite Extraction

There is no universal solvent for cellular metabolite extraction. An optimal solvent for metabolite extraction should have the following properties: 1) High efficiency and coverage. Metabolites have different chemical and physical properties, it is not possible to extract all the metabolites out by using a single solvent each time, but an optimal extraction solvent should be able to extract as many metabolites as possible. 2) No interference. Extraction solvent should not bring any interference for downstream analysis. 3) Easily to be removed. The samples usually need to be concentrated before analysis. For some derivatization methods, extraction solvent is sometimes not compatible with derivatization reaction condition, and needs to be removed. Some extraction solvents are not friendly with analytical instrument, so they must be completely removed before analysis. For instance, TFA and DMSO could significantly interfere with the MS signal, and THF could swell the PEEK tubing used on a LC system. TFA and THF could be easily removed by evaporation, however, solvents such as the high boiling point solvent like DMSO are difficult to remove, thus it should not be used for metabolite extraction.

Common extraction solvents include water, organic solvents, or water/organic solvent mixtures. In some studies, cell lysis and metabolite extraction are combined in one step. As there are so many extraction solvent systems, we will only select some of them for discussion.

ACN, methanol and chloroform are the most commonly used extraction solvents. Chloroform/water/ACN or methanol systems are used as biphasic extraction solvent for extracting both metabolites and lipids.^{19,39}

Hot ethanol or water extraction was reported for yeast intercellular metabolite extraction.^{40,41} It is suitable to extract several relatively stable metabolites like amino acids, whereas it cannot be used in untargeted metabolomics analysis as metabolite degradation occurs under high temperature.

Using acid or base solution might improve the extraction efficiency, but acid or base extraction solvent only works well for nucleotides and water soluble metabolites. For instance, J. D. Rabinowitz et al. reported using acidic acetonitrile as extraction solvent could minimize the triphosphates degradation.⁴² Maharjan et al. compared different extraction solvents and found that KOH solution has the highest efficiency for UDP-Glucose extraction.⁴³ However, metabolite destruction, such as NAD, indole compounds, was always taken place during strong acid/base extraction process. Neutralization step is required in acid or base solution extraction and the precipitation may happen in this step, and some metabolites may be absorbed onto precipitates

causing sample loss. Table 1 summarizes sample handling methods used in cellular metabolomics.

Table 1.1 Summary of sample handling methods used in cellular metabolomics.

Cell Type	Harvest Method	Rinse	Quenching	Lysis	Extraction	Reference
Human Cervical Cancer Cell (HeLa)	Scrape	Ice cold PBS, 1×	Liquid nitrogen	Freeze-thaw cycle	biphasic methanol/chloroform/water	Gielisch, I. et al. ⁴⁴
Human Cervical Cancer Cell Line (Hela)	Tyrpsinization	Ice cold PBS, 2×	N/A	N/A	N/A	Feng, J. et al. ⁴⁵
Human Breast Cancer Cell Line (MCF-7)	Scrape	Ice cold PBS, 2×	Methanol	NA	biphasic methanol/chloroform/water	Teng, Q. et al. ¹⁹
Human Breast Cancer Cell Line (MCF-7)	Tyrpsinization	Ice cold PBS, 3×	Liquid Nitrogen	Glass-bead lysis	1:1 MeOH: water	Luo, X. et al. ¹⁵
Human liver cancer cell line (Hep G2)	Scrape	Ice cold PBS, 2×	3:1 methanol:water	Sonication	3:1 methanol: water	Meissen, J. K. et al. ⁴⁶
Prostate Epithelial Cells (WPE1-NB11)	Scrape	Cold PBS	N/A	N/A	biphasic methanol/chloroform/water	Teahan, O. et al. ⁴⁷
Colon Adenocarcinoma Cell Line (SW480)	Scrape or trypsinization	PBS	Liquid nitrogen	Freeze-thaw cycle	Different solvent systems	Dettmer, K. et al. ¹²
Human Liver cell line (THLE-2) Proximal tubular cell line (HK-2)	Scrape or trypsinization	Warm PBS	88% methanol	Glass-bead lysis	88% methanol	Muschet, C. et al. ⁴⁸
Human Pancreatic Adenocarcinoma Cells	Trypsinization	Cold PBS	N/A	Sonication	biphasic methanol/chloroform/water	Watanabe, M. et al. ⁴⁹

(Panc-1, Miapaca-2)						
Yeast (<i>S. cerevisiae</i>)	Centrifugation	Cold water	Liquid nitrogen	Glass-bead lysis	50% methanol	Luo, et al. ⁶
Yeast (<i>S. cerevisiae</i>)	Filtration	N/A	-20 °C extraction solvent	N/A	40:40:20 acetonitrile/methanol/w ater	Xu, Y. et al. ⁹
Bacteria (<i>E. coli</i>)	Centrifugation	Cold 0.9% NaCl	Cold 0.9% NaCl	Sonication	50% methanol	Wu Y. et al. ⁷
Bacteria (<i>E. coli</i>)	Centrifugation	Cold PBS	Incubation on ice	Freeze-thaw cycle	K ₂ HPO ₄ /NaH ₂ PO ₄ in acetonitrile	Ye Y. et al. ⁵⁰
Bacteria (<i>E. coli</i>)	Filtration	N/A	N/A	N/A	Hot 80% ethanol	Zampieri, M. et al. ¹⁰
Bacteria (<i>S. warneri</i>)	Centrifugation	100 mM Tris/5 mM EDTA	N/A	Cell disruptor Or enzymic lysis	100 mM Tris/5 mM EDTA buffer	Fu F., et al. ⁸

1.7 Sample Normalization

Quantitative metabolomics aims to compare the concentration of a variety of metabolites from different samples. The weight, size and volume of a given sample could affect the metabolite quantification result. To achieve accurate quantification result, sample normalization is necessary.

Generally, there are two types of sample normalization method: pre-acquisition normalization or post-acquisition normalization. In pre-acquisition normalization, the total concentration of metabolites of each sample is determined first, and by adjusting the injection or loading volume, equal amount of samples is then analyzed by analytical platforms. In post-acquisition normalization, samples are loaded on analytical platforms without controlling sample amount. On NMR and MS, in most of situations, the total signal intensity is proportional to the total concentration of metabolite. Thus, the concentration of each individual metabolite is normalized based on the total intensity of the signal.

The advantage of pre-normalization method is that the sample concentration is known, so the optimal loading amount of sample could be determined. Since the same amount is loaded onto an analytical platform in pre-acquisition method, the instrument response is similar for all of the samples. In practice, as the total concentration of metabolites is difficult to quantify, people only select one or several metabolites as references to normalize samples. For instance, creatinine is a well-known reference to reflect the urine concentration. However, there is no good reference in cellular metabolomics, so several metabolites are selected as references and used for sample normalization.

Compared with pre-acquisition normalization, post-acquisition normalization doesn't consume extra samples and is less laborious. However, injected sample amount can't be controlled if post-acquisition normalization is used, the total signal intensity may not reflect the real total concentration of metabolites when signal saturation and ion suppression exist.

There are some specific normalization methods for cellular metabolomics. The most straightforward normalization method in cellular metabolomics is to count the number of cells. Hemocytometer and flowcytometer are commonly used for cell counting. In hemocytometer method, one part of cells is sampled and counted under microscope, and the total number of cells could be determined by multiplying the dilution factor. The adherent cell cultures need to be detached from culture dish and suspended evenly to get the accurate counting. Using hemocytometer to count the cells is a time-consuming and labor-intensive process, whereas imaging systems are equipped on some of state-of-art microscopes (e.g. BioTek Lionheart FX Automated Microscope), the live cell can be imaged and counted within a second. Flowcytometer can count millions of cells in a short time and are also used for cell sorting. Cells need to be stained by fluorescent dye or monoclonal antibodies with fluorochromes before counting.⁵¹ The incubation process for cell staining usually takes more than half an hour, and cellular metabolome may alter during this process.

For yeast and bacterial cells, optical density at 600 nm (OD600) is often used for monitoring the culture density.^{52,53} The calibration curve of the number of cells and OD600 value could be established, then the number of cells could be determined rapidly by measuring OD 600. However, the OD 600 measurement only works for suspension cell cultures, not for adherent cell culture. The correlation between OD600 and cell count can be affected by environmental factors: such as, pH, temperature and water activity.⁵⁴

The protein amount could also be used for sample normalization. BCA and Bradford assays are classic protein quantification assays which have been used for years, and they have already been applied in cellular metabolomics for sample normalization.^{47,55} However, the extraction and quenching solvent used in cellular metabolomics may not be compatible with protein extraction and sample loss is always observed, so the total protein amount sometimes failed to produce expected correlations with total concentration of metabolites.⁵⁶

Some studies demonstrated that DNA concentration has a good correlation with the total concentration of metabolites.^{48,56} In these studies, DNA was extracted out and its concentration was determined by absorbance- or fluorescence-based methods. The major drawback of DNA normalization methods is that additional extraction steps and expensive DNA extraction kit are required. Thus, it is still not commonly used in metabolomics normalization.

Our group reported a method that could determine total metabolite concentration by LC-UV directly.⁵⁷ In this method, the metabolites were derivatized with DnsCl. Dansyl tag is a chromophore which has a unique absorption at 338 nm, and a LC-UV method was established and used for total metabolite quantification. Our group further optimized this method to achieve the high-throughput quantification. Dansyl labeled metabolites were extracted out by ethyl acetate and quantified by using of a plate reader.⁵⁸ Including cell lines, dansyl derivatization technique has been successfully used for quantifying various samples.^{15,59,60}

1.8 Analytical Platform

A variety of analytical platforms were used in metabolomics. Figure 1.3 shows a comparison of different analytical platforms.

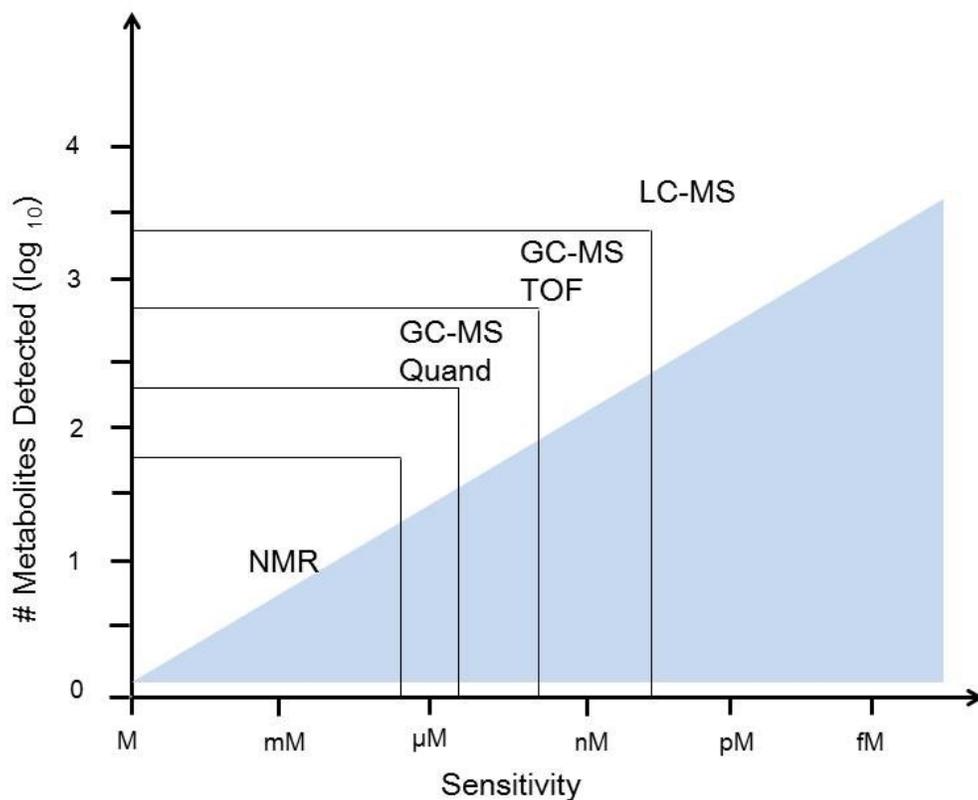


Figure 1.3 Comparison of different analytical platforms. Adapted from reference 61

1.8.1 NMR

MS and NMR are two leading analytical platforms in metabolomics. The sensitivity of MS is much higher than NMR, but fortunately, the concentration of some important metabolites, such as TCA cycle intermediates, amino acids and sugars, are high enough for NMR analysis.⁶² Compared with MS, NMR is a non-destructive analytical approach, and it is more powerful in unknown metabolite structure identification. NMR can also be used for detecting or quantifying those metabolites which are not easily ionized by MS. ¹H, ¹³C, ¹⁵N and ³¹P are the most widely used NMR spectroscopy in metabolomics.⁶³ Among them, ¹H NMR spectroscopy has high sensitivity and it is the most widely used NMR spectroscopy in cellular metabolomics. For

instance, Feng et al. studied the metabolome change of HeLa cells after exposure to silica nanoparticles by using of ^1H NMR.⁴⁵ ^{31}P NMR is useful in studying some energy metabolism related metabolites in cells (e.g. ADP, ATP), however, the overlap of other phosphate metabolites' signal limits its application. $^{63}\text{ }^{13}\text{C}$ NMR can provide more metabolite structure information, but the sensitivity of ^{13}C NMR is quite low due to the low abundance of ^{13}C (1.1%). Because of the high complexity of metabolomics samples, 2D NMR (e.g., ^1H - ^1H COSY, ^1H - ^1H TOCSY, and ^1H - ^{13}C HSQC) has been widely applied for improving metabolite identification.⁶⁴

66

Sample preparation in NMR based cellular metabolomics is not laborious. Cell lysates need to be dried to remove the extraction solvent completely, and dried materials are then re-dissolved in buffer solutions. Phosphate in deuterated water (D_2O) is a commonly used buffer solution in NMR metabolomics. The control of buffer pH is necessary to avoid any shifting of NMR resonances.⁶⁷ Different strategies have been proposed for metabolite identification and quantitation in NMR metabolomics.⁶⁸

Over all, NMR provides complementary information for metabolites which are not ionized well on MS. If the concentrations of target metabolites are high enough, NMR is the most straightforward and a robust analytical technique.

1.8.2 GC-MS

GC-MS is one of the most widely used analytical platforms in metabolomics, especially for volatile metabolite analysis. The advantage of GC-MS is that reproducible fragment pattern can be achieved when compounds are ionized with an electron impact (EI) ion source. In addition, the retention time of GC-MS is more stable than LC-MS, thus, the metabolite standards retention

time information could be collected and used for metabolite identification. The U.S. National Institute of Standards and Technology (NIST) collected GC-MS mass spectra of 242,477 compounds and one third of them have retention index information. In comparison, NIST14 LC-MS/MS spectra library only contains 8,171 unique compounds and lacks retention time information.⁶⁹

However, in GC-MS metabolomics, non-volatile metabolites need to be derivatized and converted to volatile compounds before analysis, thus an extra derivatization step is always required and needs to be optimized. Silylation is the most commonly used derivatization reaction in GC-MS based metabolomics. In silylation, a silyl group replaces active hydrogen (e.g., amine, phenol, alcohol, carboxylic acid) of metabolites. The common silylation reaction reagents include N-tert-butyldimethylsilyl- N-methyltrifluoroacetamide (MTBSTFA)⁷⁰ and N-Methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA)⁷¹. MTBSTFA derivatization can give a better sensitivity for amine-containing metabolites, but MTBSTFA cannot react with carbohydrate completely, and MSTFA should be used as instead.⁶⁹ Ketone and aldehydes can also be derivatized by silylation, as the active hydrogen present during the equilibrium.⁷² However, silylation derivatized ketone or aldehydes are not thermally stable, and methoxymation is suggested to be carried out before GC-MS analysis.

2D GC-MS has been applied in cellular metabolomics to further extend the metabolome coverage. Compared with 1D GC-MS, 2D GC-MS has better sensitivity, resolution and peak capacity.⁷³ Z. Yu et al. demonstrated that the sensitivity of their 2D GC-MS system was 10-20 fold higher than 1D GC-MS, and more than 600 putative metabolites were detected from U2OS cell line by their 2D GC-MS platform.⁷⁴

Sample preparation is critical in GC-MS based metabolomics. For example, colon cancer cell metabolome profile was highly depended on extraction solvent selection.⁷⁵ The selection of derivatization reagent is also important, as performance of derivatization reagent could also affect the cellular metabolomic profiling.⁷⁶

1.8.3 LC-MS

1.8.3.1 Liquid Chromatography

Liquid chromatography is one of the chromatographic techniques which is used for separating each component from a mixture. A liquid chromatography system usually contains solvent reservoir, pump systems, injection valve, column and detector. Figure 1.4 shows each component of liquid chromatography.

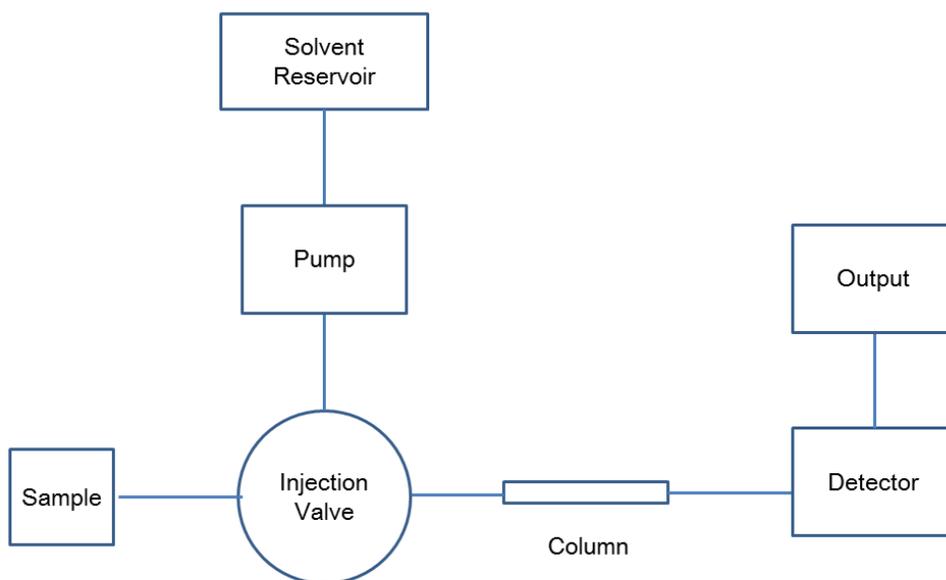


Figure 1.4 Schematic of a liquid chromatography system

The most commonly used liquid chromatography for metabolomic profiling is reversed phase chromatography (RPLC). In RPLC, the stationary phase is made of non-polar materials, and the mobile phase is usually water combined with organic solvents. Methanol and acetonitril are two of most widely used mobile phase. The reversed phase column is packed with porous silica particles with with alkyl group modification. As an example, Figure 1.5 shows the s porous silica particle covered with C18 stationary phase

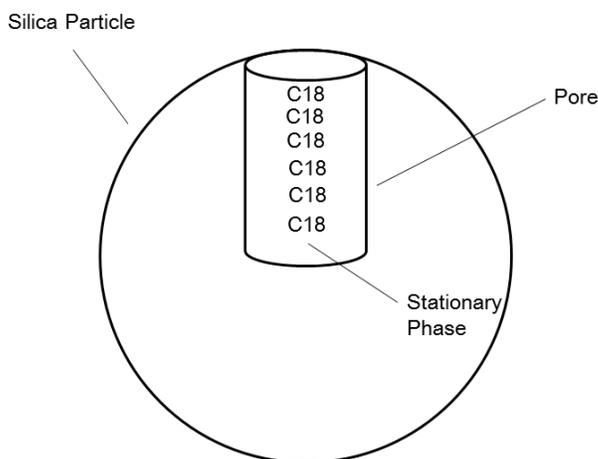


Figure 1.5 Schematic of porous silica particle covered with C18 stationary phase

Figure 1.6 illustrates the separation process of liquid chromatography. A, B and C represent three components in the mixture, respectively. The mixture is injected onto column and carried by mobile phase. Due to the difference of chemical and physical properites of Component A, B and C, the migration speed of three comopnents is different on column. In this case, Component A has the highest migration speed, while the migration speed of Component B is slower than Component A, and Component C has the slowest migration speed. The result is that, Component A is the first component elutes out from LC column, Component B and Component C elute out later in order. The different migration speed of different analytes is the basis of chromatographic separation.

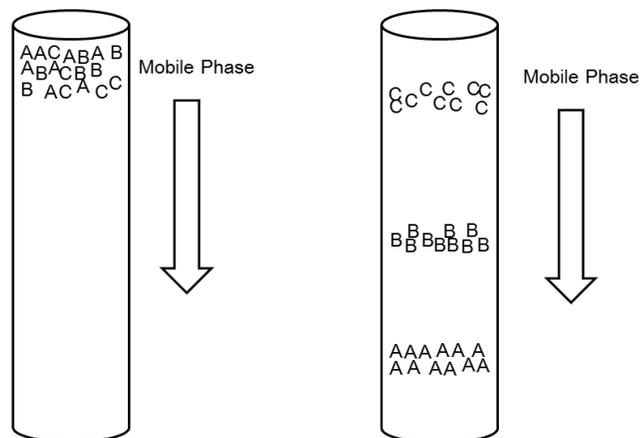


Figure 1.6 Separation process of liquid chromatography

1.8.3.2 Mass Spectrometry

A mass spectrometer includes ion source, mass analyzer, detector, vacuum system and data process system. Electrospray ionization (ESI) interface is commonly equipped on LC-MS system as an ion source. There are four major processes in ESI: (1) In LC-ESI-MS, eluent from LC flows through a capillary with a high voltage applied, and then breaks to highly charged droplets. (2) With heated dry gas blowing, the solvent of droplets is evaporated and the size of droplet is shrunk. (3) The droplet disintegrations are repeated, and small “offspring” droplets are formed. (4) The electrostatic force of droplet becomes high, and the solute ions “escape” from surface of droplet to the gas phase. The schematic of ESI process is in Figure 1.7.

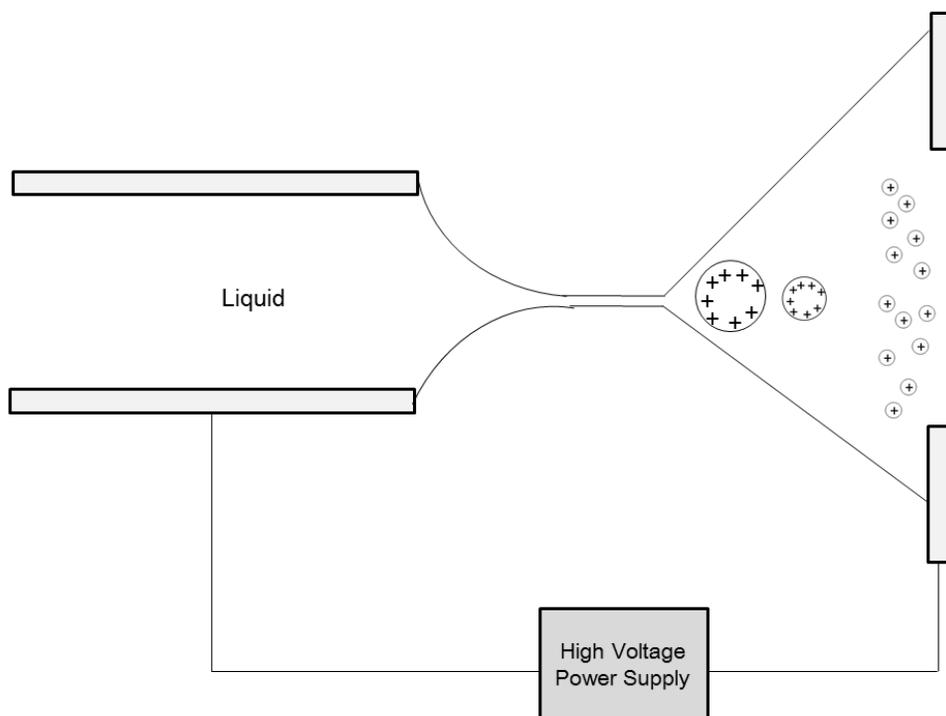


Figure 1.7 Schematic of ESI process

Mass analyzer of quadrupole time-of-flight (Q-TOF) MS is consisted of quadrupole, collision cell and time-of-flight instrument. Quadrupole is consisted of four cylindrical rods where creates a hyperbolic field. Both DC and RF voltage are applied on four rods. In particular electric field defined by DC and RF, only ions with specific m/z could transmit the quadrupole, other ions collides with rod and get quenched. Thus, quadrupole mass analyzer is also called mass filter. By increasing DC and RF voltage, the ions with increasing m/z value could transmit quadrupole. Collision cell is a quadrupole or hexapole. In MS mode, only RF voltage is applied on quadrupole and collision cell, and all the ions with different m/z value could transmit quadrupole and collision cell. In MS/MS mode, only the selected precursor ions could pass the quadrupole, and then the precursor ions collide with neutral gas in collision cell and induce fragmentation. The schematic of quadrupole is showed in Figure 1.8.

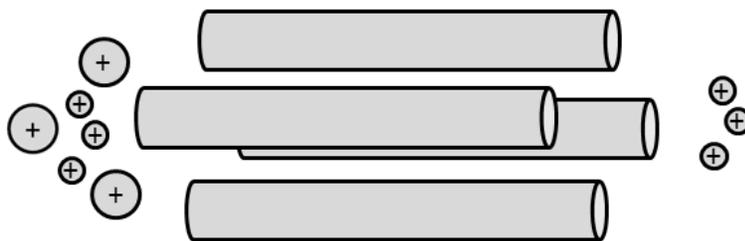


Figure 1.8 Schematic of quadrupole mass analyser

In time-of-flight mass analyzer, a pulsed voltage is applied to push the ions into flight tube. The ions gain initial kinetics from pulsed voltage. The flight tube is under high vacuum and without electronic field, where ions fly freely inside. Theoretically, all the ions gain same kinetic energy at the beginning, thus the ions with high m/z value will have low flying speed, while the ions with low m/z value will fly faster in flight tube. After calibrated by standard, time-of-flight instrument could measure the m/z value of ions. With installation of reflectron, the resolution of time-of-flight instrument could reach more than 20,000. The detector installed on TOF instrument is usually multi-channel plate (MCP) detector. The schematic of quadrupole is showed in Figure 1.9.

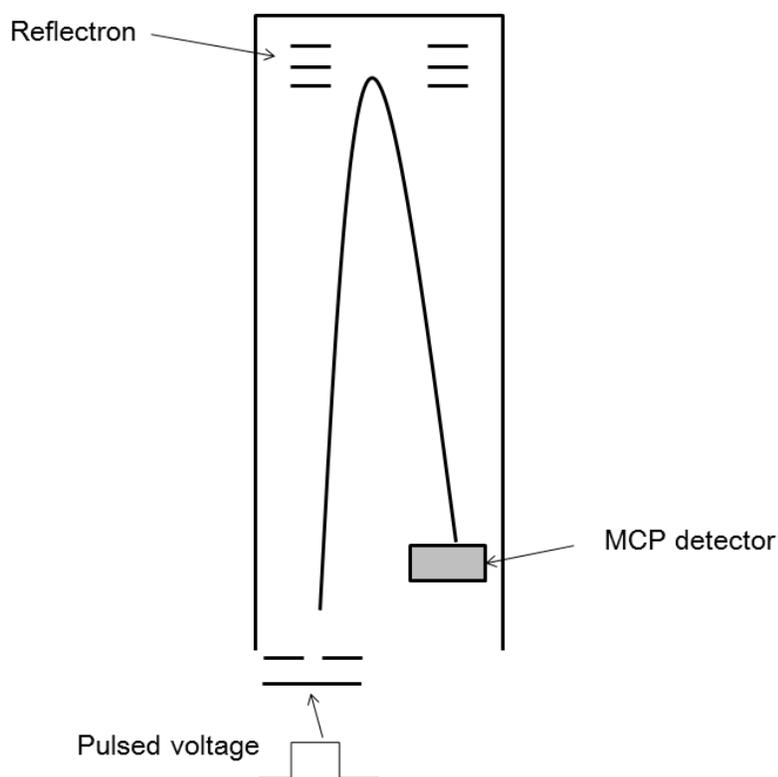


Figure 1.9 Schematic of time-of-flight mass analyser

LC-MS is a very powerful analytical tool in metabolomics. Reversed phase LC column is the most widely used LC column for non-polar or semi-polar metabolites separation (e.g. glycerides, steroids, flavonoids), and hydrophilic interaction LC (HILIC) is used for polar metabolite separation (e.g. sugar, amino acid, nucleotides). Several different types of ion sources are available for LC-MS metabolomics, including ESI, APCI and APPI. ESI is the most commonly used ion source in LC-MS based metabolomics, as both the semi-polar and polar metabolites can be ionized efficiently. As a complementary ionization technique, APCI and APPI were sometimes employed for non-polar metabolites analysis, such as lipids.⁷⁷ Various mass analyzers have been used in LC-MS based cellular metabolomics, including low resolution mass analyzers, such as triple quadrupole (QqQ) and triple quadrupole linear ion traps (Q-TRAP),

and high resolution mass analyzers, such as TOF, orbitrap, and FT-ICR. Among them, QqQ and Q-TRAP are commonly used for trace-level metabolites absolute quantification when operating under selected reaction monitoring (SRM) or multiple reaction monitoring (MRM) mode.⁷⁸ TOF and orbitrap are widely used for untargeted metabolomic profiling.

Due to the matrix effect and ion suppression, MS quantification always requires internal standards as references. Stable isotope analogues are usually used as internal references, but sometimes it is difficult to obtain or very expensive to synthesize. Instead of synthesizing stable isotope analogues, the differential isotope labeling (DIL) or chemical isotope labeling (CIL) method introduces an isotope tag with mass difference into metabolites. The isotope tag structures can be tuned for different purposes. For instance, the isotope tags which contain an aromatic ring could enhance the hydrophobicity of metabolites and thus achieve better separation by RPLC. The isotope tags which includes easily charged structures could help increase ionization efficiency on ESI and thus enhance the detection sensitivity. A divide-and-conquer strategy is often applied in CIL LC-MS metabolomics, which means different isotope labeling reagents are developed for labeling metabolites with different functional groups. Our group developed a series of ¹²C-/¹³C- isotope labeling reagents for different submetabolome profiling (e.g. DnsCl for amine-/phenol-, DmPA for carboxylic acid, base activated DnsCl for hydroxyl, DnsHz for carbonyl) and they have been applied in different types of biological sample analysis including cell lines.⁷⁹⁻⁸² Y. Feng group reported on a deuterium based isotope labeling technique (e.g. 2-Dimethylaminoethylamine for carboxylic acid, 2-(2-hydrazinyl-2-oxoethyl) isoquinolin-2-ium bromide for carbonyl, 4-(N,Ndimethylamino)phenyl isothiocyanate for amine, ω-bromoacetylquinolinium bromide for thiols) and of its application in metabolomic profiling of a variety of samples.⁸³

Further sensitivity enhancement in LC-MS based metabolomics is still highly desirable, especially for the analysis of small numbers of cells or a single cell. Nanoflow LC-MS has been employed for metabolomic analysis of limited amount of sample.⁸⁴ The sensitivity enhancement of nanoLC-MS can be attributed to the reduction of sample dilution and improvement of ionization efficiency. The inner diameter (ID) of nanoLC column is usually from 75 μm -200 μm , compared with a 2.1 mm ID microflow column, the sample dilution is hundreds of folds less. The droplets formed on nano emitter are 100-1000 fold smaller than those from a conventional ESI source. More efficient desolvation of smaller droplets results in more ions being formed and entering into a mass spectrometer.^{85,86} Recent advances of nanoLC-MS based metabolomics has been reviewed by Chetwynd, A. J. et al.⁸⁷

Two dimensional LC (2D-LC) has also been applied in cellular metabolomics to improve the metabolite separation and metabolome coverage. In 2D-LC, samples are injected onto two independent separation systems under different chromatographic conditions. For instance, M. Navarro-Reig et al. coupled HILIC and RPLC columns to develop a comprehensive 2D LC-MS platform for metabolomic profiling work.⁸⁸ Kennedy group has developed off-line 2D LC-MS for *E. coli* metabolome profiling. The samples were fractionated on a strong anion exchange column before being injected onto a reversed phase LC, and 391 metabolites were detected from *E. coli* with optimal conditions.⁸⁹

LC-MS is one of the most powerful analytical platforms in cellular metabolomics. However, there is still a room for LC-MS sensitivity improvement to achieve high-coverage metabolomic profiling of small numbers of cells or even from single cells. Metabolite identification is currently still a bottleneck in LC-MS metabolomics.

1.8.4 CE-MS

Although poor reproducibility has hampered application of CE-MS in metabolomics, CE-MS is still a complementary and important technique for highly polar and charged metabolites analysis.⁹⁰ A lot of important metabolites are anionic (e.g. nucleotides, carboxylic acids) or cationic compounds (e.g. tyramine, hypoxanthine), can be separated, detected and quantified by CE-MS.^{91,92} CE-MS has been applied in cellular metabolome profiling. C. Ibáñez et al. applied sheath-liquid CE-MS for detecting cationic metabolites from HT-29 colon cancer cells.²¹ T. Soga et al. detected 1692 metabolites from *B. subtilis* cells by using of CE-MS, and among them 150 could be identified.⁹³ The small sampling volume of CE-MS makes it a promising tool for single cell analysis. Sweedler group has developed CE-MS methods for analyzing metabolites from a single neuron cell, and more than hundred of metabolites were detected from a single metacerebral cell.^{94,95}

1.9 Data Process

As there are many different methods and strategies for processing the data generated from different analytical platforms, I will only discuss general data processing procedures of LC-MS based metabolomics. In LC-MS metabolomics, the data processing procedure usually includes filtering and background subtraction, peak picking, multiple LC-MS data alignment, statistical analysis and metabolite identification. Figure 1.4 shows the general workflow of LC-MS based metabolomics data processing.

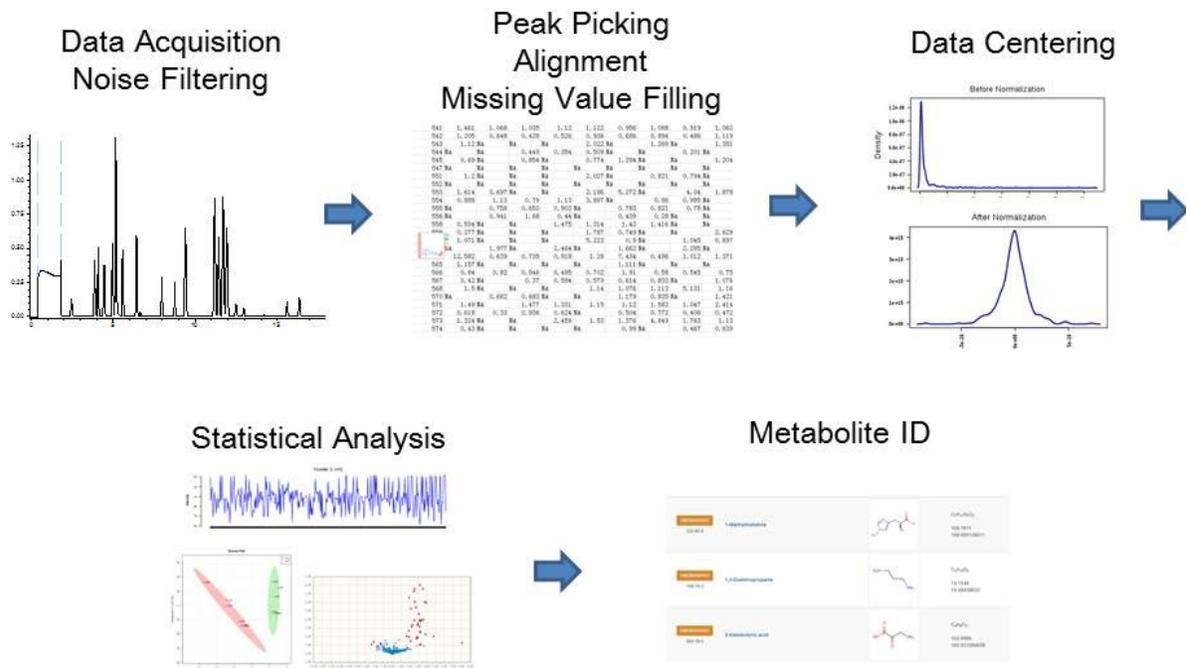


Figure 1.10 General workflow of LC-MS based metabolomics data processing

Filtering and background subtraction aims to remove contaminants and noise from LC-MS spectra to get rid of any interferences for downstream data analysis. After filtering and background subtraction, “clean” LC-MS spectra are achieved and ready for peak picking. In this step, centroids of MS peaks are calculated and then ion chromatographic peaks are constructed, thus each peak represents one putative metabolite.⁹⁶ In metabolomics, multiple LC-MS data sets are usually acquired, and those LC-MS runs are aligned together based on retention time and accurate mass to generate a data matrix. Missing values may exist in the data matrix due to misalignment or missing peaks. Missing values could cause problems in statistical analysis, so various approaches have been developed for missing value imputation in metabolomic dataset.⁹⁷⁻

⁹⁹ Data centering and scaling are usually performed before statistical analysis, and different methods, include mean centering, auto scaling and pareto scaling, can be used to centering or scaling metabolomics data.

A variety of statistical tools are performed for biomarker discovery. Univariate statistical tools include fold change analysis, t-test, volcano plot and one-way Analysis of Variance (ANOVA). Fold change analysis, t-test and volcano plot are commonly used statistical tools for binary comparison (e.g. healthy control and disease group), and ANOVA is used for three or more groups comparison. Multivariate analysis, including PCA and PLS-DA, and OPLS-DA, is also very common in metabolomic data analysis. One of the advantages of multivariate analysis is that, the metabolomics dataset could be visualized, so it is more straightforward for comparison. However, overfitting may exist in supervised analysis (e.g. PLS-DA, OPLSDA), and permutation test is required to validate the model.¹⁰⁰ Other statistical analysis, such as heatmap, k-means and random forest are also widely used in metabolomics.^{101,102} After potential biomarkers are discovered, receiver operating characteristic (ROC) curves are performed for its performance evaluation. Lots of software has been developed for metabolomics data processing and statistical analysis, for instance, Bruker MetaboScope, Waters TransOmic, Gary Siuzdak group's XCMS,^{103,104} David Wishart group's Metaboanalyst,¹⁰⁵ Oliver Fiehn group's Metabox,¹⁰⁶ Lloyd Sumner group's MET-IDEA,¹⁰⁷ and our group's MCID.¹⁰⁸

Metabolite identification and database searching is the number one grand challenge in LC-MS based metabolomics. In LC-MS based metabolomics, accurate mass, MS/MS and retention time are the most important information for metabolite identification. Retention time can vary greatly when different columns and chromatographic conditions are used. Thus, retention time information is not commonly used for metabolite identification, accurate mass and MS/MS are more often used as a searching parameter for metabolite identification. Some well-known libraries, such as HMDB,¹⁰⁹ Metlin,¹¹⁰ Mass Bank of North America (MoNA), mzCloud, and NIST14 MS/MS libraries, contain hundreds of thousands of standard MS and MS/MS spectra.

Our group constructed an evidence-based metabolome library for metabolite identification.¹¹¹ Well characterized retention time can still be a piece of important information for metabolite identification, research groups have developed different approaches for normalizing experimental retention time and normalized RT can be used for library search.^{112,113} Predicted retention time libraries were also constructed and achieved varying degree of success.^{114,115}

1.10 Overview of Thesis

The main objective of this thesis is to develop CIL LC-MS for cellular metabolome profiling. In Chapter 2, a highly efficient lysis method was developed and used for yeast (*S. cerevisiae*) metabolite extraction. CIL LC-MS was performed to comparatively metabolomic profile the yeast cultured with or without nitrogen limitation. A potential link of pantothenate accumulation and nitrogen limitation was discovered and it might have potential application in fermentation engineer. In Chapter 3, we applied the technique developed in Chapter 2 to investigate the impact of ultrasonication on *S. cerevisiae*. Our results indicated that ultrasonication might trigger the reprogramming of carbon metabolism of *S. cerevisiae*. Chapter 4 describes the development of a simple and rapid method for adherent cell harvesting and lysis. In this chapter, the efficiency of different cell harvesting and lysis methods were evaluated by LC-UV. Based on the findings, physical scarping and frozen-thaw cycles were suggested to be used for cell harvesting and lysis in CIL LC-MS metabolomics. In Chapter 5, CIL nanoLC-MS was developed and performed for metabolomic profiling of small number of breast cancer cells. More than a thousand metabolites could be detected from a hundred breast cancer cells by the use of this platform. In Chapter 6, by using the technique developed in Chapter 5, the metabolome of exosomes from pancreatic cancer patients before and after chemotherapy were analyzed and significant changed metabolites were discovered.

Chapter 2

High-Performance Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry for Profiling the Metabolomic Reprogramming Elicited by Ammonium Limitation in Yeast

2.1 Introduction

The yeast species *Saccharomyces cerevisiae* is widely used as a model system for studying biological processes¹¹⁶ as well as in baking and the production of alcoholic beverages.¹¹⁷ Improvement of yeast strains used in these traditional applications has shifted from conventional methods for isolation of beneficial variants¹¹⁸ to directed engineering of desirable metabolic traits.^{117,119} The latter metabolic engineering approach is also having a dramatic effect on the development of yeast for more recent industrial applications ranging from biofuel to drug production.^{120,121} The workflow in contemporary metabolic engineering projects now often includes broad coverage metabolic profiling implemented with the goal of determining if and to what extent yeast metabolism is affected by a particular engineering strategy. The information obtained by this approach can guide further engineering steps that improve system performance.^{122,123} Therefore, high coverage quantitative analysis of the yeast metabolome is highly desirable. However, analytical challenges remain in achieving this goal, because yeast produces a large range of metabolites of different classes at different concentrations.

There are various analytical platforms used for yeast's metabolome analysis based on NMR¹²⁴⁻¹²⁶ or mass spectrometry (MS) combined with gas chromatography (GC),^{127,128} liquid chromatography (LC)¹²⁹⁻¹³¹ and capillary electrophoresis (CE).¹³² Among them, LC-MS provides relatively high sensitivity and detectability. However, some challenges still remain for LC-MS.

First, ion suppression, matrix effect and instrument drift reduce quantification accuracy and precision. Second, hydrophilic and highly polar metabolites are poorly retained on a reversed phase chromatography (RPLC) column. Hydrophilic interaction liquid chromatography (HILIC) is required as a complementary technique to separate these metabolites.¹³³ And both positive and negative ion scans are needed to detect various classes of metabolites. All these will increase the workload of metabolome analysis. Third, metabolites that have low concentration or low ionization efficiency during electrospray ionization (ESI) will not be detected or detected with low signals, preventing their accurate quantification.

To address these challenges, we have developed a high-performance chemical isotope labeling (CIL) LC-MS platform for quantitative analysis of various sub-metabolomes with high coverage. For example, differential $^{12}\text{C}/^{13}\text{C}$ -dansylation (Dns) labeling LC-MS is used for profiling the amine/phenol submetabolome⁷⁹ and $^{12}\text{C}/^{13}\text{C}$ -p-dimethylaminophenacyl (DmPA) bromide labeling is used for analyzing the carboxylic acid submetabolome.⁸⁰ In differential CIL, a ^{12}C -labeling reagent is used to label an individual sample, while a pooled sample produced from mixing equal amounts of aliquots of all individual samples is labeled with the ^{13}C -labeling reagent. The ^{13}C -labeled pool is spiked into an individual ^{12}C -labeled sample and the mixture is then subjected to LC-MS analysis. The differentially labeled metabolites are detected as peak pairs in mass spectra and their peak intensity ratio provides the basis for relative quantification. Since the same ^{13}C -labeled pool is used as a reference for all the ^{12}C -labeled individual samples, the peak intensity ratios of a given metabolite reflect the relative concentration difference of the metabolite in the comparative samples. These peak ratio values can be used for metabolomic comparison using statistical tools. In high-performance CIL LC-MS, the labeling reagents such as Dns and DmPA are rationally designed to alter the metabolite chemical and physical

properties to such an extent that simultaneous improvement in separation, detection and quantification can be achieved. As a result, only RPLC-MS using positive ion detection is needed. This provides a sensitivity improvement of 10 to 1000-fold, allowing quantification of both high and low abundance metabolites.

In this work, we report the development of CIL LC-MS for yeast metabolome analysis with high submetabolome coverage and the application of this method for investigating metabolic reprogramming provoked by removing a nitrogen source from log phase cells in batch culture. Specifically, cells were switched from high ammonium sulfate medium to medium without ammonium sulfate. We chose to study the effect of ammonium sulfate withdrawal on metabolism for several reasons. First, ammonium sulfate is a high quality nitrogen source for yeast that on its own can provide all the nitrogen needed for growth.^{134,135} Second, physiological homeostasis of yeast and other eukaryotes depends critically on the regulation of nitrogen metabolism. Third, in our experiments ammonium sulfate withdrawal is expected to affect nitrogen but not sulfate metabolism. That is because the store of sulfate in log phase cells is likely sufficient for several divisions,¹³⁶ and in our experiments the culture time in 0% ammonium sulfate (6 hours) allows for only 2 or 3 divisions.¹³⁵ Finally, the amount of ammonium sulfate provided to cells is an important determinant of product yield in yeast that have been metabolically engineered to synthesize biofuels from fatty acids.^{137,138} Therefore, information about the metabolic response of cells to ammonium sulfate withdrawal is potentially relevant to continued development of metabolic engineering strategies in yeast. Because the base medium used in our experiments contains 13 L-amino acids that can support growth,¹³⁹ the withdrawal of ammonium sulfate is not expected to trigger a strong physiological response. We chose this modest starvation in order to determine if CIL LC-MS can provide the necessary

sensitivity to confidently identify intercellular metabolites which displayed small changes in abundance in response to a subtle culture manipulation, thereby illustrating the potential general applicability of the developed workflow for yeast metabolomics studies.

2.2 Experimental Section

2.2.1 Workflow

Figure 2.1 shows the overall workflow of CIL LC-MS for yeast metabolomics. In general, cells are washed with water and harvested by centrifugation. Washed cell pellets are resuspended in extraction/suspension solvent and lysed by vortexing with glass beads. The lysate is separated from the glass beads and cellular debris, transferred to a new vial and dried down. The lysates are re-dissolved in water and subjected to ^{12}C -labeling. The labeled samples are injected into LC-UV to determine the total concentration of labeled metabolites in individual samples for sample amount normalization. Based on the total concentration, equal amounts of individual samples are taken and mixed to generate a pooled sample. This pooled sample is labeled using ^{13}C -labeling. Each ^{12}C -labeled individual sample is mixed with an equal amount of the ^{13}C -labeled pool and the mixture is injected onto LC-MS for metabolite detection and relative quantification. The resultant quantitative metabolomic data are analyzed by using statistical tools to determine the significant metabolites differentiating different groups of cells. These metabolites can be mapped to metabolic pathways for biological studies.

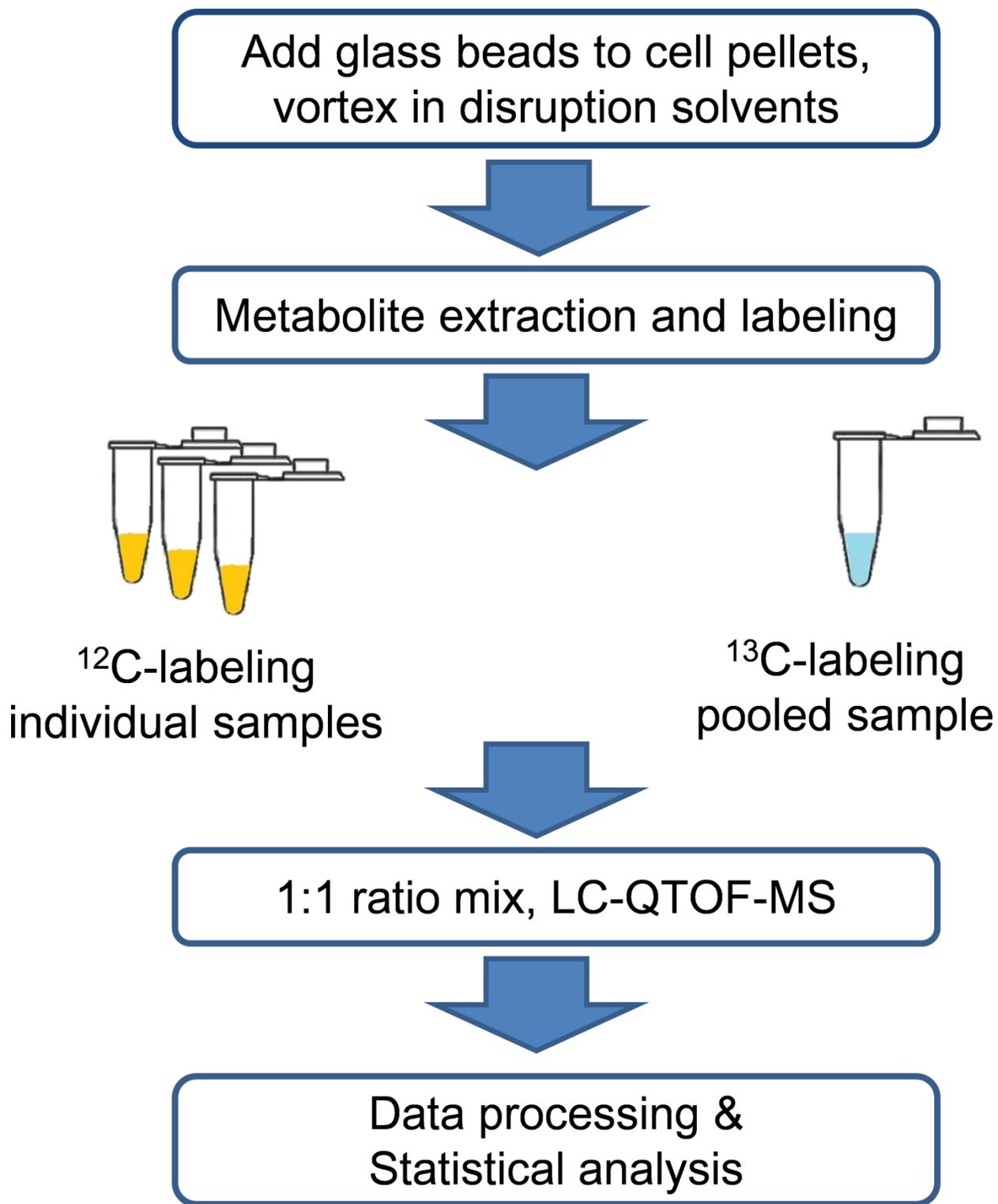


Figure 2.1 Workflow of the differential chemical isotope labeling (CIL) LC-MS method for yeast metabolomics.

2.2.2 Chemical and reagents

All the chemicals and reagents, unless otherwise stated, were purchased from Sigma-Aldrich Canada (Markham, ON, Canada). Glass beads (0.5 mm diameter) were purchased from Biospec Products. ^{13}C -labeling reagents were synthesized in our lab using the procedures published previously^{79,80} and are available from MCID.chem.ualberta.ca. LC-MS grade water, formic acid, acetonitrile (ACN) and methanol (MeOH) were purchased from Thermo Fisher Scientific (Edmonton, AB, Canada).

2.2.3 Cell growth and cell lysis

Strain BY4741¹⁴⁰ was selected for this work. BY4741 and its S288c relatives are widely used in metabolic engineering,^{137,141-145} and like most laboratory strains, can use ammonium and all the L-amino acids for growth.^{134,135} The synthetic culture medium employed here, complete minimal with 2% w/v glucose,¹³⁹ contains 0.5% w/v ammonium sulfate as recommended for BY4741;¹⁴⁶ we refer to this medium as CMD+AS. While CMD+AS contains a high amount of ammonium sulfate, it includes only a subset of amino acids. This amino acid formulation is preferred for BY4741 because its proliferation is inhibited in synthetic media containing all amino acids.¹⁴⁷ Cells were cultured to mid-log phase CMD+AS at 30 °C and 225 rpm in a shaking incubator for 24 h. The medium was replaced with fresh CMD+AS (control group) or CM without ammonium sulfate + 2% glucose (CMD-AS; nitrogen starvation group) and cultures were returned to the 30 °C shaking incubator. The cells were harvested by centrifugation (4640 *g* for 10 min at 4 °C) 6 hr after resuspension in fresh medium. The pellets were resuspended in 1 mL cold water (LC-MS grade), and spun in an Eppendorf 5415C microcentrifuge at 16000 *g* for 1 min at 4 °C. After removing the water, this washing process was repeated two more times to

remove excess growth medium. After removal of the water the final cell pellets were snap-frozen in liquid nitrogen and then stored at $-80\text{ }^{\circ}\text{C}$ until further use.

For cell lysis, 0.5 cc (mL) of glass beads and 100 μL of extraction/suspension solvent (MeOH or ACN in water; see Results) were added to a 1.5 mL microcentrifuge tube which contained the frozen cell pellets. Cell lysis was achieved via five 1 min periods of bead-beating at 3200 rpm using a VORTEX-GENIE 2 Mixer (Fisher Scientific) alternated with five 1 min incubations in an ice-water bath. After cell lysis, another 800 μL of extraction solvent was added for metabolite extraction. To determine the percent cell breakage, cells were counted with hemocytometer under LEICA DM 1L inverted contrasting microscope (Wetzlar, Germany). Unbroken cells and cell debris were removed by centrifugation at 16000 g at $4\text{ }^{\circ}\text{C}$ for 10 min, and the supernatant was transferred to a new vial and dried down in a Savant SC110A Speed Vac at room temperature. Drying metabolite extracts at room temperature was used by others.¹⁴⁸ After metabolite extraction, the metabolite levels were not expected to be affected by the sample drying process, as most of the enzymes should be removed. The dried metabolites were re-dissolved in LC-MS grade water, and stored at $-80\text{ }^{\circ}\text{C}$.

2.2.4 Dansylation labeling

25 μL of the metabolite extract was mixed with 12.5 μL of sodium carbonate/sodium bicarbonate buffer and 12.5 μL of ACN. The solution was spun down and mixed with 25 μL of freshly prepared ^{12}C -dansyl chloride solution (18 mg/mL in ACN) (for light labeling) or ^{13}C -dansyl chloride solution (18 mg/mL in ACN) (for heavy labeling). The reaction was allowed to proceed for 1 hr at $40\text{ }^{\circ}\text{C}$. After 1 hr, 5 μL of 250 mM NaOH was added to the reaction mixture to quench the excess dansyl chloride. The solution was then incubated at $40\text{ }^{\circ}\text{C}$ for another 10

min. Finally, 25 μL of formic acid in 1:1 ACN/ H_2O was added to consume excess NaOH and to acidify the solution.

2.2.5 DmPA bromide labeling

Prior to isotopic labeling, a liquid-liquid extraction step was carried out to remove amine-containing compounds in order to increase the specificity of the reaction for carboxylic acids¹⁴⁹. Cellular metabolites were extracted with 150 μL of ethyl acetate. The organic layer was dried and dissolved in 30 μL of 20 mg/mL triethylamine, and then mixed with 30 μL of freshly prepared ^{12}C -DmPA bromide solution (20 mg/mL in ACN) (light labeling) or ^{13}C -DmPA bromide solution (20 mg/mL in ACN) (heavy labeling). The reaction was incubated at 85 $^\circ\text{C}$ for 60 min.

2.2.6 LC-UV quantification

A LC-UV quantification step was carried out prior to mass analysis in order to control the amount of sample used for metabolome comparison (i.e., sample normalization).¹⁵⁰ 2 μL of the labeled solution was injected onto a Phenomenex Kinetex C18 column (2.1 mm \times 5 cm, 1.7 μm particle size, 100 \AA pore size) linked to a Waters ACQUITY UPLC system (Waters, Milford, MA) for step-gradient LC-UV. The UV detector was operated at 338 nm. Solvent A was 0.1% (v/v) formic acid in 5% (v/v) ACN, and solvent B was 0.1% (v/v) formic acid in acetonitrile. The gradient started with 100% A for 1 min and was increased to 95% B within 0.1 min and hold at 95% B for 1.5 min. The gradient was restored to 100% A in 0.5 min and hold at this condition for 3 min to re-equilibrate the column. The flow rate used was 0.45 mL/min.

2.2.7 LC-MS

The ^{13}C -labeled pool was mixed with an equal amount of the ^{12}C -labeled individual sample and the mixture was analyzed using a Bruker Impact HD Quadrupole Time-of-flight (Q-TOF) mass spectrometry (Bruker, Billerica, MA) linked to an Agilent 1100 series binary HPLC system (Agilent, Palo Alto, CA). The samples were injected onto an Agilent reversed phase Eclipse Plus C18 column (2.1 mm \times 10 cm, 1.8 μm particle size, 95 \AA pore size) for separation. Solvent A was 0.1% (v/v) formic acid in 5% (v/v) acetonitrile, and solvent B was 0.1% (v/v) formic acid in acetonitrile. The chromatographic conditions for dansyl labeling were: t = 0 min, 20% B; t = 3.5 min, 35% B; t = 18 min, 65% B; t = 21 min, 95% B; t = 26 min, 95% B. The gradient for DmPA labeling was t = 0 min, 20% B; t = 9 min, 50% B; t = 22 min, 65% B; t = 26 min, 80% B; t = 29 min, 98% B; t = 40 min, 98% B. The flow rate was 180 $\mu\text{L}/\text{min}$. All MS spectra were obtained in the positive ion mode. The MS conditions used for Q-TOF were as follows: nebulizer, 1.0 bar; dry temperature, 230 $^{\circ}\text{C}$; dry gas, 8 L/min; capillary voltage, 4500 V; end plate offset, 500V; spectra rate, 1.0 Hz.

2.2.8 Data Processing and Analysis

A software tool, IsoMS,¹⁰⁸ was used to process the raw data generated from multiple LC-MS runs by peak picking, peak pairing, peak-pair filtering, and peak-pair intensity ratio calculation. After alignment of multiple data files by retention time and accurate mass, missing values in the aligned files were filled by the Zerofill program.⁹⁷ The data file was uploaded to the MetaboAnalyst website (www.metaboanalyst.ca) and multivariate statistical analysis was performed. Principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA) and heatmap were used to analyze the data. Volcano plots and box plots were generated using Origin 9.0. Metabolite identification was performed based on accurate mass and retention time match to a dansyl or DmPA standard library.¹¹² Putative identification was done

based on accurate mass match to the metabolites in the yeast metabolome database (YMDB) (www.ymdb.ca), the human metabolome database (HMDB) (www.hmdb.ca) and the predicted human metabolite library in MyCompoundID (MCID) (www.mycompoundid.org).¹⁵¹

2.3 Results and Discussion

2.3.1 Cell lysis for metabolomic analysis

Efficient and complete cell lysis is critical to generate the metabolome profile that reflects the whole cell population. If only a fraction of cells are lysed, only a partial population is investigated, leaving one to wonder if the un-lysed cells might have different metabolomic compositions than the lysed ones. Another critical consideration is related to the use of chemicals for cell lysis. It is important not to use any chemicals that may cause interference in downstream sample processing such as chemical labeling and LC-MS analysis. In our work, we chose to employ a glass-bead-assisted cell lysis method to obtain intracellular metabolites from yeast. Using this method, LC-MS analysis can be conducted using metabolites extracted from the cells harvested from <1 mL of culture. We have optimized this bead method, including the cell extraction/suspension solution and lysis condition, in order to achieve high-yield cell lysis.

Figure 2.2 A,B shows the microscopy images of cells. The control group was vortexed in 50% MeOH without glass beads for five rounds, and the disruption group was vortexed with glass beads in 50% MeOH. The intact cells can refract light and appear as bright dots, while the cell debris appears as dark dots. Almost all the cells were cracked after 5 rounds of vortexing with glass beads in MeOH. Since one cell may break and generate several pieces of debris, directly calculating the ratio of intact cells (light dots) to black dots after lysis, would generate an artificially high estimate of cell lysis efficiency. In our work, the cell breakage percentage was determined by the number of intact cells before and after lysis.

Since methanol (MeOH) and acetonitrile (ACN) are widely used as solvents in cell lysis and metabolites extraction, we compared four solvents: 50% MeOH, 80% MeOH, 50% ACN, and 80% ACN to see which gives the best cell lysis efficiency using the bead method. The results are plotted in Figure 2.2 C. Among them, the 50% MeOH group showed the highest cell lysis efficiency with 96.1% of cells lysed. The lowest efficiency group was 80% ACN with only 15.9% lysis efficiency. The efficiency of using 80% MeOH (76.6%) was higher than that of using 50% ACN (58.6%). The solvent dependence of lysis efficiency may be attributed to the different levels of cell suspension and clumping in different solvents. In the bead beating method, agitation by vortexing causes beads to collide with cells and crack their walls and thus a solvent offering the best dispersal of cells in suspension likely provides the most efficient cell lysis. In all the subsequent experiments, the glass-bead method using 50% MeOH for cell suspension was used for cell lysis, which is fully compatible with downstream labeling and LC-MS analysis.

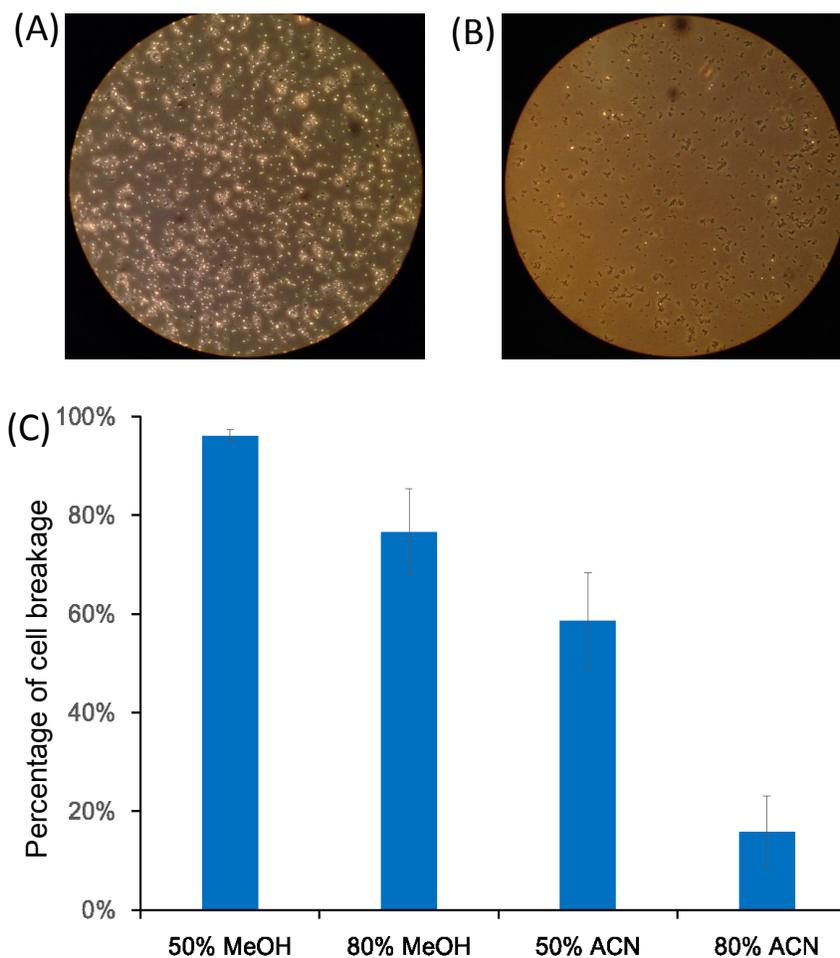


Figure 2.2 Microscopy images of cell lysis: (A) Control group without glass beads; the yeast cells remained intact after vortexing and (B) Disruption group with glass beads in 50% MeOH after vortexing; cell debris were shown as dark dots. (C) Comparison of cell lysis efficiency of four different solvents. Cell counting with hemocytometer was used to determine the percentage of cell breakage (n=3).

2.3.2 Analysis of amine- and phenol-containing metabolites

The amine-containing metabolites play an important role in metabolic pathways. For example, the 275 amines and phenols in the current Dns-metabolite standard library cover 43 different metabolic pathways. Most of the metabolites in amino acid metabolic pathways and nucleotide metabolic pathways are amine-containing metabolites. Phenol-containing metabolites are common in tyrosine, phenylalanine and tryptophan metabolic pathways. Thus, dansylation LC-MS offers a way to figure out how the nitrogen starvation affects these metabolic pathways. On average, 3763 ± 271 peak pairs were detected in triplicate experiments on biological triplicate samples of the ammonium limitation group and control group lysates. In total, there were 5719 peak pairs or metabolites detected from all the samples combined. The IsoMS program filtered out all the adducts, dimers, multimers, etc., in mass spectra to retain only the $[M+H]^+$ peak pairs and thus in most cases each pair represents a unique metabolite; in cases of a metabolite containing more than one amine or phenol group, multiple products may be generated. We performed metabolite identification based on mass and retention time match to a dansyl library containing 275 metabolite standards.¹¹² A total of 120 metabolites were positively identified and they are listed in Table 2.1. Detailed information on the matches is shown in supplemental information which is available from Dr. Liang Li. We also searched the accurate masses of the detected metabolites against the HMDB database and found 926 matches (In supplemental information which is available from Dr. Liang Li). In addition, using the MCID library with one reaction, we further matched 1857 metabolites (In supplemental information which is available from Dr. Liang Li). Thus, among the 5719 peak pairs found, we could match 3040 metabolites (i.e., 53%). It should be noted that we also searched the yeast metabolome database (YMDB) which is much smaller than HMDB. We matched 672 metabolites in YMDB using mass search.

Among them, only 8 metabolites were uniquely found in YMDB, while 664 metabolites overlapped with the human metabolites (In supplemental information which is available from Dr. Liang Li). And these 8 unique metabolites could be matched to the MCID library. Thus, the use of YMDB did not yield any additional matches. Based on the large number of peak pairs detected and metabolite matches found, it is clear that the yeast metabolome is very complex. Unfortunately there are a limited number of metabolite standards currently available to confirm the identities of the putatively matched metabolites. Future work of confirming these identities may lead to the discovery of new intermediates in known pathways as well as revealing totally unknown metabolic pathways.

Table 2.1 List of amine- or phenol-containing metabolites identified.

No.	HMDB ID	Name	No.	HMDB ID	Name
1	HMDB29306	4-Ethylphenol	61	HMDB00721	Glycylproline
2	HMDB29118	Tyrosyl-Valine	62	HMDB00719	Homoserine
3	HMDB29109	Tyrosyl-Leucine	63	HMDB00719	Homoserine - H ₂ O
4	HMDB29105	Tyrosyl-Glycine	64	HMDB00706	Aspartyl-phenylalanine
5	HMDB29098	Tyrosyl-Alanine	65	HMDB00696	Methionine
6	HMDB29095	Tryptophyl-Tyrosine	66	HMDB00687	Leucine
7	HMDB29065	Threoninyl-Leucine	67	HMDB00669	Ortho-Hydroxyphenylacetic acid
8	HMDB29007	Phenylalanyl-Tyrosine	68	HMDB00650	Alpha-Aminobutyric acid
9	HMDB28988	Phenylalanyl-Alanine	69	HMDB00641	Glutamine
10	HMDB28941	Leucyl-Tyrosine	70	HMDB00557	Alloisoleucine
11	HMDB28937	Leucyl-Proline	71	HMDB00517	Arginine
12	HMDB28878	Histidinyl-Alanine	72	HMDB00500	4-Hydroxybenzoic acid
13	HMDB28854	Glycyl-Valine	73	HMDB00469	5-Hydroxymethyluracil
14	HMDB28848	Glycyl-Phenylalanine	74	HMDB00455	Allocystathionine
15	HMDB28844	Glycyl-Isoleucine	75	HMDB00455	Allocystathionine - Isomer

16	HMDB28694	Alanyl-Phenylalanine	76	HMDB00452	Alpha-aminobutyric acid
17	HMDB28691	Alanyl-Leucine	77	HMDB00450	5-Hydroxylysine
18	HMDB28689	Alanyl-Histidine	78	HMDB00446	N-Alpha-acetyllysine
19	HMDB13243	Leucyl-phenylalanine	79	HMDB00440	3-Hydroxyphenylacetic acid
20	HMDB11737	Gamma Glutamylglutamic acid	80	HMDB00301	Urocanic acid
21	HMDB11177	Phenylalanyl-proline	81	HMDB00300	Uracil
22	HMDB06050	o-Tyrosine	82	HMDB00296	Uridine
23	HMDB04987	Alpha-Aspartyl-lysine	83	HMDB00296	Uridine - H ₂ O
24	HMDB03911	3-Aminoisobutanoic acid	84	HMDB00292	Xanthine
25	HMDB03464	4-Guanidinobutanoic acid - H ₂ O	85	HMDB00279	Saccharopine
26	HMDB03423	Glutamine	86	HMDB00279	Saccharopine - H ₂ O
27	HMDB03337	Oxidized glutathione	87	HMDB00228	Phenol
28	HMDB03320	Indole-3-carboxylic acid	88	HMDB00214	Ornithine
29	HMDB03012	Aniline	89	HMDB00210	Pantothenic acid
30	HMDB02658	6-Hydroxynicotinic acid	90	HMDB00206	N6-Acetyl-Lysine
31	HMDB02393	N-methyl-aspartic acid	91	HMDB00191	Aspartic Acid
32	HMDB02390	3-Cresotinic acid	92	HMDB00182	Lysine
33	HMDB02362	2,4-Diaminobutyric acid	93	HMDB00177	Histidine
34	HMDB02322	Cadaverine	94	HMDB60003	Isovanillic acid
35	HMDB02141	N-Methyl- α -aminoisobutyric acid	95	HMDB00172	Isoleucine
36	HMDB02064	N-Acetylputrescine	96	HMDB00168	Asparagine
37	HMDB02005	Methionine Sulfoxide	97	HMDB00168	Asparagine - H ₂ O
38	HMDB02005	Methionine Sulfoxide - Isomer	98	HMDB00167	Threonine
39	HMDB01964	Caffeic acid	99	HMDB00162	Proline
40	HMDB01906	2-Aminoisobutyric acid	100	HMDB00161	Alanine
41	HMDB01891	m-Aminobenzoic acid	101	HMDB00159	Phenylalanine
42	HMDB01842	Guanidine	102	HMDB00158	Tyrosine
43	HMDB01545	Pyridoxal	103	HMDB00157	Hypoxanthine - multi-tags
44	HMDB01431	Pyridoxamine	104	HMDB00149	Ethanolamine
45	HMDB01414	1,4-diaminobutane	105	HMDB00148	Glutamic Acid - H ₂ O
46	HMDB01392	p-Aminobenzoic acid	106	HMDB00148	Glutamic Acid
47	HMDB01257	Spermidine	107	HMDB00133	Guanosine
48	HMDB01232	4-Nitrophenol	108	HMDB00123	Glycine
49	HMDB01173	5'-Methylthioadenosine	109	HMDB00112	Gamma-Aminobutyric acid

50	HMDB01149	5-Aminolevulinic acid	110	HMDB00112	Gamma-Aminobutyric acid - H ₂ O
51	HMDB01065	2-Hydroxyphenethylamine - Isomer	111	HMDB00101	Deoxyadenosine
52	HMDB01049	Gamma-Glutamylcysteine	112	HMDB00099	Cystathionine - Isomer
53	HMDB00957	pyrocatechol	113	HMDB00099	Cystathionine
54	HMDB00929	Tryptophan	114	HMDB00070	Pipecolic acid
55	HMDB00883	Valine	115	HMDB00056	Beta-Alanine
56	HMDB00819	Normetanephrine	116	HMDB00050	Adenosine
57	HMDB00763	5-Hydroxyindoleacetic acid	117	HMDB00045	Adenosine monophosphate
58	HMDB00759	Glycyl-Leucine	118	HMDB00020	p-Hydroxyphenylacetic acid
59	HMDB00755	Hydroxyphenyllactici acid	119	316000000*	Phenyl-Leucine
60	HMDB00750	3-Hydroxymandelic acid	120	HMDB00750	3-Hydroxymandelic acid - COOH

*This metabolite does not have an HMDB number, but is present in the dansyl standards library.

The number shown is a designated number in the dansyl standards library.

Using the data generated by dansylation LC-MS, PCA and PLS-DA were first applied to evaluate whether ammonium starvation affects the metabolomic profiles of *Saccharomyces cerevisiae*; the relevant score plots are shown in Figure 2.3A,B. The two groups with and without the ammonium starvation can be well separated. In the PCA plot, 43.7% and 17.3% variances were captured by the first principal component (1st PC) and the 2nd PC, respectively. Most variances were captured by the first two components reflected a good quality of the model. This was further validated in the cross-validation test for the PLS-DA model as $R^2=0.999$ and $Q^2=0.981$.

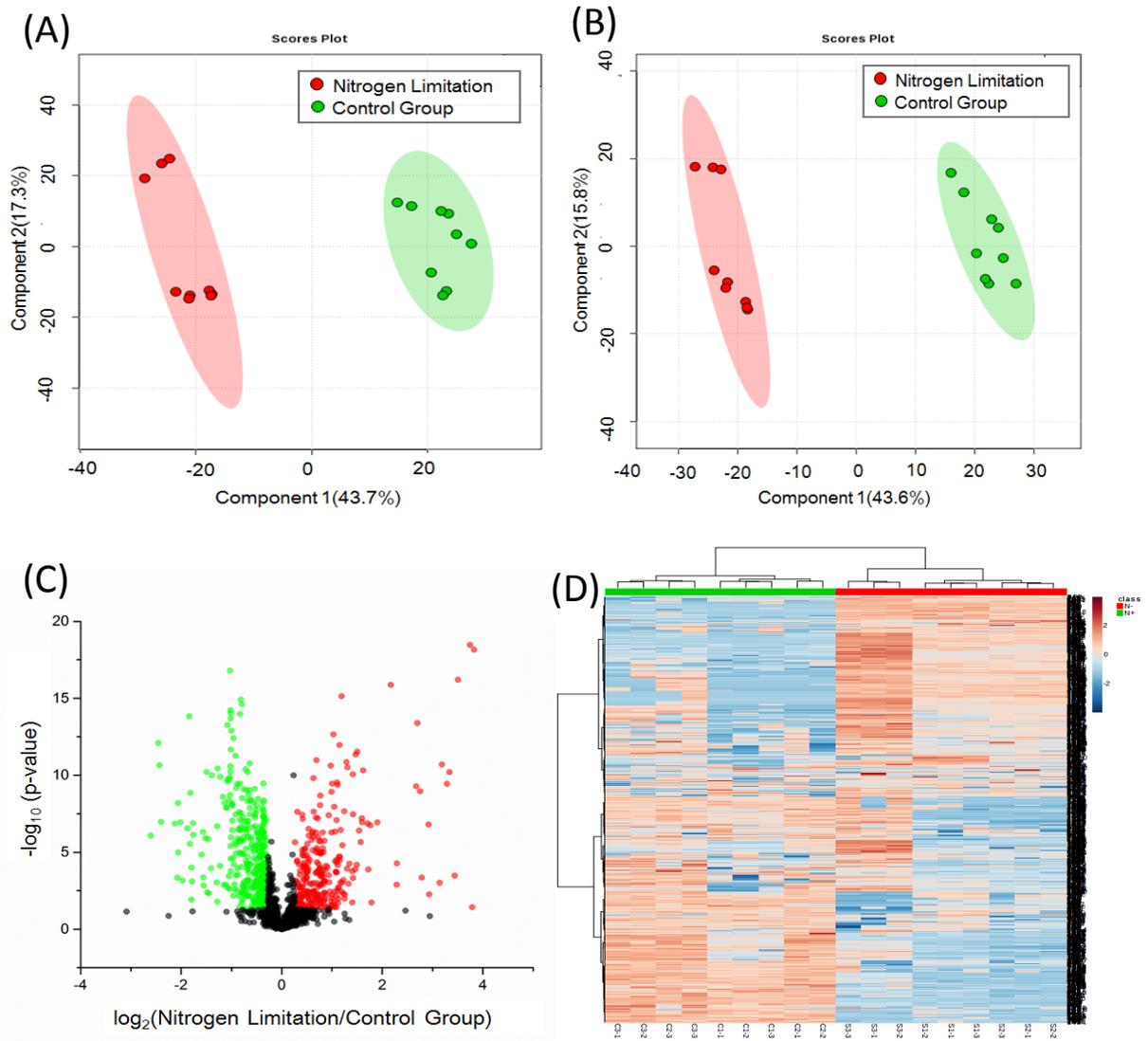


Figure 2.3 Score plots of (A) PCA and (B) PLS-DA of dansylation labeled lysates from *Saccharomyces cerevisiae* cultured with or without ammonium limitation. The data was from experimental triplicates on biological triplicate samples. (C) Volcano plot of the changes of amine- and phenol-containing metabolites. The up-regulated or down-regulated metabolites by at least 1.2-fold, with p-values of smaller than 0.05, are marked in red and green, respectively. (D)

Clustered heatmap showing the comparison of the relative intensity of each peak pair in the nitrogen limitation group and the control group.

We also analyzed the data using a Volcano plot which combines the fold-change and the p-value from a t-test to determine the significant metabolites that separate the two groups. The Volcano plot is shown in Figure 2.3C using thresholds of ≥ 1.2 -fold change and ≤ 0.05 p-value.

There were 603 peak pairs up-regulated and 706 peak pairs down-regulated in response to nitrogen limitation (In supplemental information which is available from Dr. Liang Li). To compare the intensity change of each metabolite, a heatmap was generated (see Figure 2.3D) with deeper red representing higher intensity and deeper blue representing lower intensity. The color distribution is similar within each group, but very different between the groups, indicating that there are characteristic and reproducible differences between the metabolic programs of yeast cultured in the presence or absence of ammonium sulphate. All these statistical analyses show that the amine/phenol submetabolome of *Saccharomyces cerevisiae* changes significantly in response to nitrogen starvation and the CIL LC-MS method can reveal the change of a large number of metabolites.

2.3.3 Analysis of carboxylic acid containing metabolites

Carboxylic acid containing metabolites are another important submetabolome in metabolic pathways. Among the cellular metabolites of central carbon metabolism, there are multiple carboxylic acids including di-, tri-, and keto- carboxylic compounds.¹⁵² Some of these are intermediates in the tricarboxylic acid (TCA) cycle, which is a core pathway for the metabolism of sugars, amino acids and lipids.¹⁵³ To analyze this submetabolome, we used DmPA bromide labeling method to increase the detection sensitivity of these metabolites.

An average of 1961±71 peak pairs or metabolites were detected from triplicate experiments on biological triplicate samples of the nitrogen limitation group and the control group. In total, 2286 peak pairs were detected. Identification of carboxylic acid containing metabolites was performed based on mass and retention time match to a DmPA standard library containing 188 acid standards. 33 metabolites were positively identified and they are shown in Table 2.2. Supplemental Information lists more information on the matches. By searching accurate masses of the remaining peak pairs, we found 248 matches to the metabolites in HMDB and 1347 matches to the MCID library (In supplemental information which is available from Dr. Liang Li) for a total of 1595 matches representing 70% of the 2286 peak pairs detected. Using the yeast database, we matched 133 metabolites and among them only 11 metabolites were not found in HMDB (In supplemental information which is available from Dr. Liang Li). However, these 11 unique metabolites could be found in the MCID library. As in the case of analyzing the amine/phenol submetabolome, searching YMDB did not generate any additional match. The large number of metabolites detected in this acid submetabolome profiling work suggests again that the composition of the yeast metabolome is complex. Many of the acid-containing metabolites remain to be positively identified.

Table 2.2 List of carboxylic acid-containing metabolites identified.

No.	HMDB ID	Name	No.	HMDB ID	Name
1	HMDB00634	Citraconic acid	18	HMDB00892	Valeric acid
2	HMDB00202	Methylmalonic acid	19	HMDB00718	Isovaleric acid
3	HMDB01987	2-Hydroxy-2-methylbutyric acid	20	HMDB00134	Fumaric acid
4	HMDB00440	3-Hydroxyphenylacetic acid	21	HMDB00695	Methyloxovaleric acid
5	HMDB02466	3-Hydroxybenzoic acid	22	HMDB00408	2-Methyl-3-ketovaleric acid
6	HMDB00576	Ethyl Malonate	23	HMDB00689	Isocaproic acid

7	HMDB01975	2-Ethyl-2-Hydroxybutyric acid	24	HMDB02097	4-Ethylbenzoic acid
8	HMDB00858	Monomethyl glutaric acid	25	HMDB00666	Heptanoic acid
9	HMDB02428	Terephthalic acid	26	HMDB00784	Azelaic acid
10	HMDB00669	Hydroxyphenylacetic acid	27	HMDB00392	2-Octenoic acid
11	HMDB00254	Succinic acid	28	HMDB01877	Valproic acid
12	HMDB00955	3-Hydroxy-4-methoxycinnamic acid	29	HMDB00847	Nonanoic acid
13	HMDB00779	Phenyllactic acid	30	HMDB00623	Dodecanedioic acid
14	HMDB00209	Phenylacetic acid	31	HMDB00511	Capric acid
15	HMDB00208	Oxoglutaric acid	32	HMDB00529	5-Dodecenoic acid
16	HMDB01870	Benzoic acid	33	HMDB00638	Dodecanoic acid
17	HMDB02176	Ethylmethylacetic acid			

This acid submetabolome data set was examined using PCA and PLS-DA. The PCA plot (see Figure 2.4A) indicates that the separation of the starvation and control groups based on the 1st and 2nd PCs is not as large as that in the amine/phenol submetabolome. However, the two groups are well separated based on the 2nd and 3rd PCs (Figure 2.4B). There were 31.8%, 11.6% and 10.9% of variances captured by the 1st, 2nd and 3rd PC, respectively. We also applied PLS-DA to evaluate the starvation and control groups. As shown in Figure 2.4C, there is a clear distinction between the ammonium limitation group and the control group ($R^2=0.999$ and $Q^2=0.873$). However, lower percentage (11.6%) of variances was captured by the 1st PC, compared to the amine/phenol submetabolome, which suggests that the extent of the acid submetabolome change is less than that of the amine/phenol submetabolome. This conclusion is supported by the Volcano plot shown in Figure 2.4D. Using thresholds of ≥ 1.2 -fold change and ≤ 0.05 p-value, there are only 83 metabolites up-regulated and 189 metabolites down-regulated (In supplemental information which is available from Dr. Liang Li). These numbers are much less than the number of changed amine/phenol metabolites. Nevertheless, these changes still

reflect a large difference in acid submetabolome between the two groups. The sensitive CIL LC-MS method with high coverage submetabolomic profiling allows the detection of these changes.

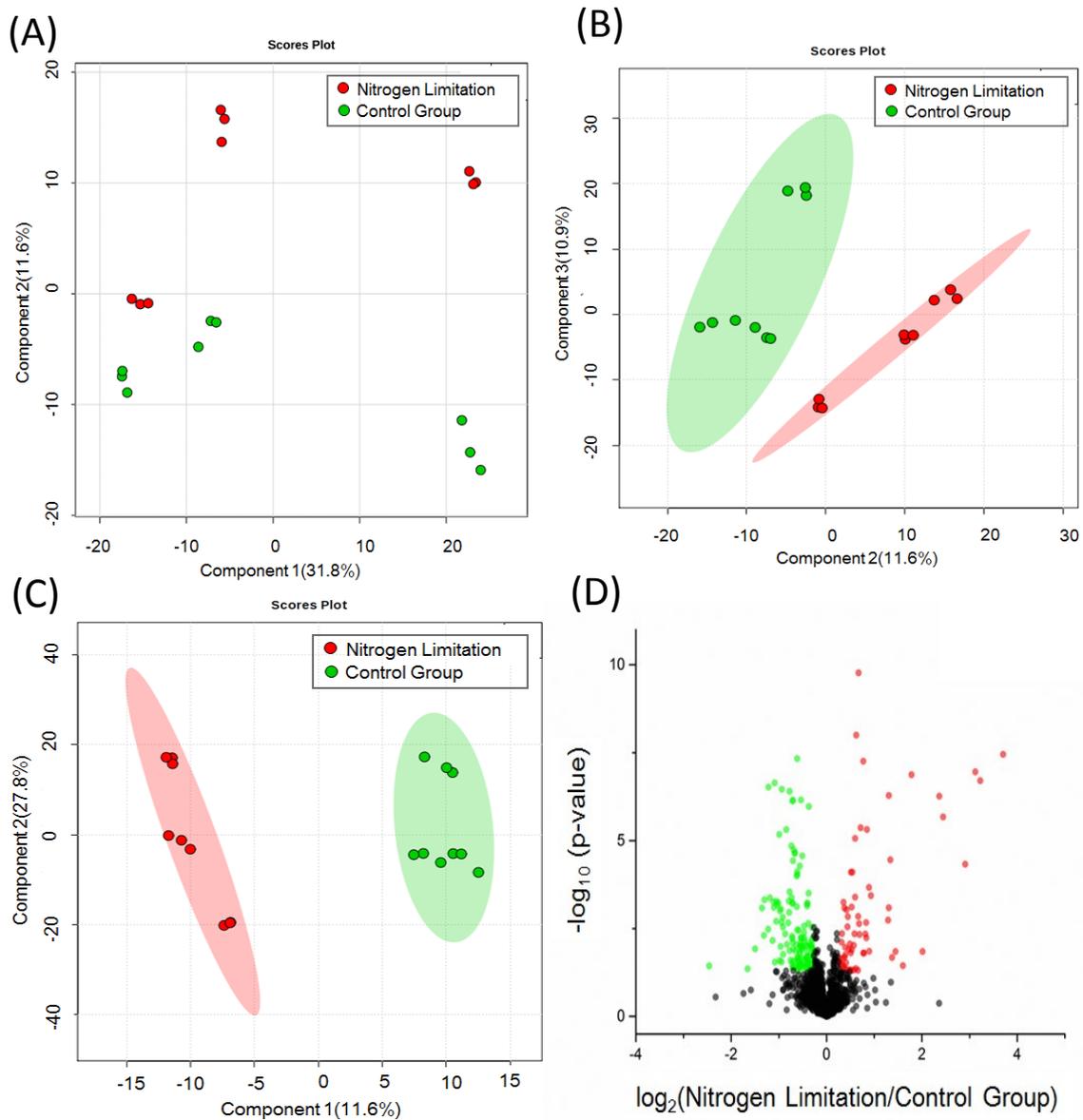


Figure 2.4 (A) PC1 versus PC2 PCA score plot, (B) PC2 versus PC3 PCA score plot, and (C) PLS-DA score plot of DmPA labeled lysates from *Saccharomyces cerevisiae* cultured with or without ammonium limitation. The data was from experimental triplicates on biological triplicate samples. (D) Volcano plot of the changes of carboxylic acid-containing metabolites. The up-

regulated or down-regulated metabolites by at least 1.2-fold, with p-values of smaller than 0.05, are marked in red and green, respectively.

2.3.4 Metabolic pathway analysis

The large coverage of amine/phenol and acid submetabolomes enables us to examine some of the metabolic pathways significantly altered by ammonium starvation. Budding yeast is well equipped to fine-tune its metabolism to fluctuations in nitrogen source availability. The way that metabolism is reprogrammed by nitrogen source fluctuation depends on whether its quality and/or amount changes. In our experiment the nitrogen source manipulation is not dramatic: removal of ammonium, a preferred nitrogen source, but not amino acids that can be used in its stead. The CIL LC-MS data support the prediction that metabolic reprogramming under this condition is modest. For example, while complete nitrogen withdrawal is associated with α -ketoglutarate induction signaling ammonia deprivation,¹⁵⁴ withdrawal of just ammonium sulfate has little if any effect on ammonium (1.17-fold increase at 6 hr in CMD-AS, $P = 4.32 \times 10^{-2}$) and causes depletion of α -ketoglutarate (oxoglutarate, see Figure 2.5).

The most informative changes in metabolite abundance associated with ammonium sulfate limitation are shown in Figure 2.5 (fold changes are the CMD-AS/CMD+AS ratios). These include induction of aspartate and depletion of arginine and its immediate precursor L-argininosuccinate. Changes in polyamine metabolism, which is fed by aspartate and arginine, are also associated with ammonium starvation (see Figure 2.5). Most striking is the buildup of N-acetylputrescine, an intermediate in polyamine breakdown.¹⁵⁵ In yeast, increased synthesis of N-acetylputrescine likely causes depletion of the polyamines spermine and spermidine.¹⁵⁵ Similarly, overexpression of an enzyme in the pathway of polyamine catabolism in mammalian cells causes accumulation of a polyamine catabolic intermediate, and drawing down of spermine

and spermidine.¹⁵⁶ It is surprising then that induction of N-acetyl-putrescine in ammonium-starved yeast cells is not associated with depletion of spermidine or spermine; spermidine in fact accumulates (Figure 2.5). Collectively these data suggest that the reprogramming of polyamine metabolism in ammonium-starved yeast cells is not simply the result of a change in flux through the polyamine catabolism pathway compared to the flux through the pathway of synthesis.

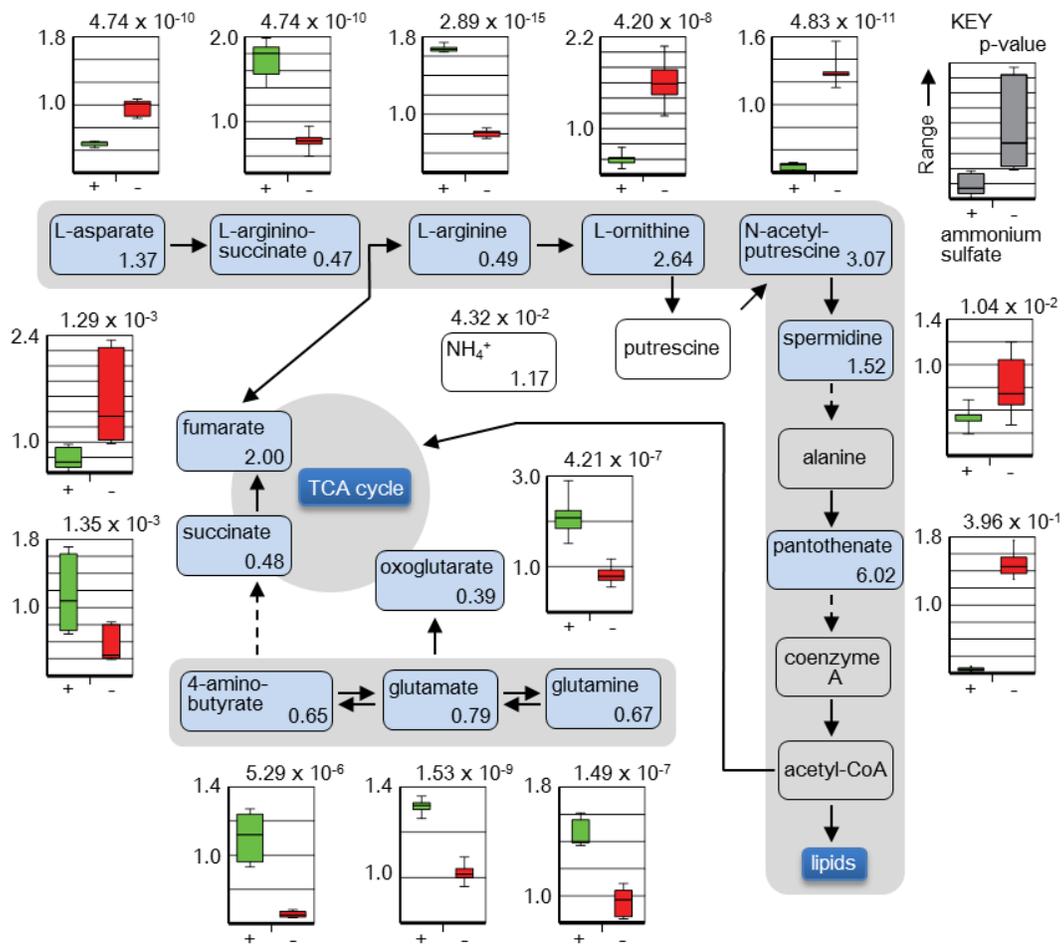


Figure 2.5 Metabolite changes in selected metabolic pathways. The graphs are for selected metabolites (blue shading) that differ significantly in abundance ($p \leq 0.05$) between cells cultured

with or without ammonium limitation. Fold differences in abundance are given in the name boxes for the metabolites.

The strongest effect of ammonium starvation on metabolite abundance is the 6-fold induction of pantothenate (Figure 2.5). This induction is potentially linked to the metabolism of aspartate and arginine, since pantothenate biosynthesis can involve conversion of these amino acids to ornithine, followed by spermine production in a sub-pathway that generates pantothenate (Figure 2.5). Pantothenate is in the biosynthetic pathway for acetyl-CoA, a central intermediate in overall metabolism.^{157,158} The fact that pantothenate accumulates in ammonium-starved cells raises the possibility that altered nitrogen metabolism affects the flux through pathways that require acetyl-CoA. Because a large number of compounds targeted for production in engineered yeast require acetyl-CoA for their synthesis, strategies for achieving high acetyl-CoA yield from glucose have been and continue to be developed.¹⁵⁹⁻¹⁶² Our evidence that ammonium limitation has a strong effect on pantothenate metabolism raises the possibility that systems engineering for optimal production of acetyl-CoA could benefit from interventions that target this axis of metabolism.

All the nitrogen in the macromolecular constituents of yeast cells is derived from glutamate and glutamine, which themselves are formed by reactions in which ammonium is the nitrogen donor.^{135,163} CMD contains glutamate, but not glutamine. Therefore in CMD+AS and CMD-AS, cells must use ammonium and glutamate to produce the glutamine required for synthesis of asparagine, tryptophan, purines and pyrimidines.¹³⁵ It follows that ammonium starvation likely affects the abundance of molecules in this module of metabolism. Indeed, there are modest changes in the abundance of glutamate and three molecules it can be used to make: glutamine, 2-

oxoglutarate (α -ketoglutarate) and 4-aminobutyrate (GABA) (see Figure 2.5). Since 2-oxoglutarate is a key intermediate in the TCA cycle, it is not surprising that other TCA intermediates (succinate, fumarate) also change in abundance in response to ammonium starvation (Figure 2.5). Indeed the buildup of fumarate may be causally related to depletion of two of its precursors, succinate in the TCA cycle and L-arginino-succinate in the pathway of arginine synthesis (Figure 2.5). The steady state metabolite abundance data we have obtained will provide a starting point for metabolic flux analysis aimed at developing a more comprehensive understanding of the cellular response to ammonium limitation.

2.4 Conclusions

We have developed and applied a high-performance metabolomic profiling workflow for quantitative yeast metabolomics with high coverage. Dansyl and DmPA labeling have been used to quantify the changes of the amine/phenol submetabolome and the acid submetabolome in yeast cells, respectively. The same workflow should be applicable to profile the submetabolomes of other chemical groups such as ketones/aldehydes and hydroxyl using new reagents that are currently under development in our laboratory. We demonstrated the utility of this workflow for comparative yeast metabolomics in cells cultured with and without nitrogen limitation. It was found that the abundances of a large number of metabolites were significantly altered in cells cultured with nitrogen limitation, proving the importance of performing metabolomic profiling with high sensitivity and accuracy. Among the metabolic pathway changes examined, our results suggest that systems engineering of acetyl-CoA production in yeast might benefit from manipulation of the links between ammonium assimilation and accumulation of pantothenate, which is an essential precursor of acetyl-CoA.

Chapter 3

Impact of Low-intensity Pulsed Ultrasound on Transcript and Metabolite Abundance in

Saccharomyces cerevisiae

3.1 Introduction

Ultrasound has a frequency greater than 20 kHz (above the normal hearing range).^{164,165} By inducing the formation and collapse of microscale gas bubbles, ultrasound can generate an environment in which macromolecules are subjected to high hydrodynamic shear stress and temperature.^{166,167} The effects of these physical forces on biological materials and cells are highly variable and dependent on the type of cell, the irradiation protocol and environmental context. At the tissue level, the ultrasound methods in medical imaging are known to cause little overt damage,¹⁶⁸ and low-intensity pulsed ultrasound (LIPUS) can in fact promote tissue repair by stimulating the proliferation of hematopoietic stem cells as well as fibroblasts and osteoblasts.¹⁶⁸⁻¹⁷² Tissue metabolism may also be altered by ultrasound as a result of microbubble collapse; the latter can induce microstreaming and microjets which may change the flow of nutrients to and from cells.¹⁷³ At the molecular level, ultrasound can affect cells by stimulating mass transfer which in turn increases the efficiency of cellular enzyme reactions including those involved in expression of proteins and other molecules (e.g., lipids in algae).^{167,174,175}

Because the biological applications of ultrasound extend from medicine to biotechnological production in microorganisms, there is intense interest in deeper understanding of the cellular responses to ultrasound. Such responses have been assessed at the population level (for example cell proliferation), and at the molecular level by application of such methods as mRNA expression profiling^{176,177} and analysis of protein post-translational modification.¹⁷⁸ One powerful

tool for studying cell physiology, global analysis of metabolite abundance, has been little used in work on cellular effects of ultrasound. Here, we address this experimental shortfall by characterizing the metabolome of a model eukaryote, the budding yeast *Saccharomyces cerevisiae*, as it converts glucose to the biofuel ethanol. The ultrasound treatment during this fermentation was a proprietary LIPUS treatment (frequency 1.5 MHz, duty cycle 20%) that has been previously reported to improve ethanol production by two microorganisms during fermentation.^{179,180} The metabolome of control and LIPUS-treated cells was characterized by chemical isotopic labeling liquid chromatography mass spectrometry (CIL-LC-MS).⁷⁹ Metabolomic profiling was complemented by parallel global transcript profiling using RNA-seq. Our work shows that metabolic reprogramming is clearly revealed by CIL-LC-MS though cells do not exhibit a strong transcriptomic signature having undergone LIPUS treatment.

3.2 Materials and Methods

3.2.1 Cell Culture

In our previous study and the current work, we used *Saccharomyces cerevisiae* strain SSL3 (spent sulfite liquor fermentation strain 3), one of the most stress tolerant strains for glucose fermentation.¹⁸¹ This strain was purchased from the American Type Culture Collection (ATCC 96581). Single colonies of SSL3 for growth in liquid medium were obtained from plates incubated at 30°C for 48 hours on mYPD agar (0.3% yeast extract; 0.5% proteose peptone; 0.1% glucose; 2% agar). The designation ‘mYPD’ is given to highlight the fact that this YPD formulation is nutrient-poor compared to the standard in yeast molecular biology (YPD, which is 1% yeast extract; 1% bactopectone; 2% glucose).

3.2.2 Inoculum preparation

One or two colonies were inoculated into 100 mL of mYPD and incubated at 180 rpm in a rotary shaker for 24 hours at 30°C. Cells were harvested by centrifugation for 5 minutes at 3000 rpm, washed with 0.9% NaCl, and then resuspended in 10 mL of medium with low yeast extract and high glucose (mYD; 0.1% yeast extract, 15% glucose, 0.5g/L (NH₄)₂HPO₄, 1.38 g/L NaH₂PO₄, 0.03 g/L MgSO₄·7H₂O - adjusted to pH 5.5 with 2 M NaOH before sterilization). One mL of this cell concentrate was inoculated into 100 mL mYD for the fermentation culture.

The 100 mL fermentations were carried out in 250 mL Erlenmeyer culture flasks at 30°C for 5 days in a rotary shaker as above. LIPUS (1.5 MHz, 20% duty cycle) was applied to the culture by placing each culture flask in its own water bath chamber equipped with an ultrasound transducer. Ultrasound treatments of 5 minutes each were given 12 times/day (in short, 5 minute treatments every 2 hours for the 5-day fermentation). This protocol was selected based on the previous results for *S. cerevisiae*.¹⁸⁰ Ultrasound was applied at 80 or 100 mW/cm²; these conditions are referred to as 80 mW/cm² and 100 mW/cm² respectively. The control fermentations (CON) did not receive LIPUS treatment. The control and LIPUS-treated flasks were sampled for transcriptomic and metabolomic analysis at a single time point, specifically the 5-day end-point of the experiment.

This endpoint was chosen because cell number increased during the first 5 days of fermentation and then started to drop off (data not shown). A possible reason for declining activity after 5 days is that the exposure to ethanol in the closed bioreactor began to exceed the level that is readily tolerated by the strain used.¹⁸²

3.2.3 Transcript Analysis

RNA was from triplicate yeast cultures under CON, 80 mW/cm² and 100 mW/cm² conditions using a Qiagen RNA plant mini kit (RNeasy Plant Mini Kit, 74904). The highly

purified total RNA was made into a TrueSeq Paired End 100 bp library and sequenced on an Illumina HighSeq 2000 system. The sequencing data was mapped to the Ensemble *S. cerevisiae* genome (http://uswest.ensembl.org/Saccharomyces_cerevisiae/Info/Index) annotated with genes and transcripts using CLC Genomics Workbench 7.0.3. The number of reads per sample averaged to just under 10 million, which resulted in a mapping average of 85% of the 7126 annotated gene transcripts in the *S. cerevisiae* genome (Table 3.1).

Table 3.1 Number of reads and percent genes mapped per sample.

Sample	Number of reads	Percent genes mapped
100L-1	9265779	84.79
100L-2	9447157	83.93
100L-3	8893095	84.35
80L-1	14003689	86.44
80L-2	7888644	84.73
80L-3	8115397	83.5
CON-1	9214649	85.41
CON-2	9786535	84.42
CON-3	10502568	85.74

3.2.3.1 Statistical analysis

ANOVA was used to identify expression differences within the entire dataset at $p < 0.05$. This data subset was further analyzed as follows. Principle component and cluster analyses were performed following the procedures in the tutorial of RNA-Seq analysis (http://www.clcbio.com/wp-content/uploads/2012/08/RNA-Seq_analysis_part_I.pdf), as well as a series of t-tests comparing all three groups of samples. Fold-differences in the original normalized comparisons were used in ReVigo (<http://revigo.irb.hr/>) to generate gene ontology

(GO) annotation enrichment profiles. These profiles were not informative because of the low overall effect of LIPUS on end-point RNA expression (see Results).

3.2.4 Metabolomic Profiling

3.2.4.1 Chemicals and reagents

LC-MS grade solvent (acetonitrile, methanol and water) was purchased from Thermo Fisher Scientific Canada. Glass beads (0.5 mm diameter) were purchased from Biospec Products. ^{13}C -dansyl chloride was synthesized in our lab using the protocol published previously.⁷⁹ All the other reagents and chemicals were purchased from Sigma-Aldrich Canada.

3.2.4.2 Cell lysis and metabolite extraction⁶

100 μL of 50% MeOH and 0.5 mL of glass beads were added to yeast pellets in a 1.5 mL microcentrifuge tube. Five-rounds of bead beating (one minute/round) were performed for cell lysis. After lysis, 800 μL of 50% MeOH was added for metabolite extraction. Cell debris was removed by centrifugation at 16000 x g at 4°C for 10 min, and the supernatant was transferred to another microcentrifuge tube and dried down in a Speed Vac (Savant SC110A). The dried extract was re-dissolved in LC-MS grade water and stored at -80°C.

3.2.4.3 Dansylation labeling

25 μL of the metabolite extract was mixed with 12.5 μL of acetonitrile (ACN) and 12.5 μL of sodium carbonate-sodium bicarbonate buffer. The solution was then mixed with 25 μL of 18 mg/mL ^{12}C -dansyl chloride in ACN for light labeling, or 18 mg/mL ^{13}C -dansyl chloride solution in ACN for heavy labeling. The reaction was carried out at 40°C for 1 hr. After 1 hr, the reaction was cooled in an ice-water bath and 5 μL of 250 mM NaOH was added to quench the excess dansyl chloride. The solution was then incubated at 40°C for another 10 min. Finally, 25 μL of

425 mM formic acid in 1:1 ACN:H₂O (v/v) was added to consume excess NaOH and to acidify the solution.

3.2.4.4 Sample Normalization

A sample normalization step was performed before LC-MS analysis.¹⁵⁰The total concentration of labeled metabolites was quantified by LC-UV in order to use the same amount of each sample for metabolome comparison. 2 μ L of the labeled solution was injected onto a Phenomenex Kinetex C18 column (2.1 mm \times 5 cm, 1.7 μ m particle size, 100 \AA pore size) linked to a Waters ACQUITY UPLC system (Waters, Milford, MA) for step-gradient LC-UV. Mobile phase A was 5% (v/v) ACN in water with 0.1% (v/v) formic acid added, and mobile phase B was acetonitrile with 0.1% (v/v) formic acid added. The step-gradient used for LC-UV was: t=0 min, 100% A; t=1 min, 100% A; t=1.1 min, 5% A; t=2.5 min, 5% A; t=3min, 100% A; t=6 min, 100% A. The flow rate was 450 μ L/min. The UV detector was operated at 338 nm.

3.2.4.5 LC-MS

The ¹²C-/¹³C-mixtures were injected onto a Dionex Ultimate 3000 UHPLC system (Thermo Fisher Scientific) linked to a Bruker Maxis Impact Quadrupole Time-of-flight (Q-TOF) mass spectrometer (Bruker, Billerica, MA). Separations were performed on an Agilent reversed phase Eclipse Plus C18 column (2.1 mm \times 10 cm, 1.8 μ m particle size, 95 \AA pore size). Mobile phase A was 5% (v/v) ACN in water with 0.1% (v/v) formic acid, and mobile phase B was acetonitrile with 0.1% (v/v) formic acid. The chromatographic conditions were: t = 0 min, 20% B; t = 3.5 min, 35% B; t = 18 min, 65% B; t = 21 min, 95% B; t = 26 min, 95%; t=34 min, 95% B. The flow rate was 180 μ L/min. The mass spectrometer conditions were as follows: capillary voltage, 4500 V; end plate offset, 500V; dry temperature, 230°C; spectra rate, 1.0 Hz; nebulizer, 1.0 bar; dry gas, 8 L/min. All MS spectra were obtained in the positive ion mode.

3.2.4.6 Data Processing and Analysis

The raw data was exported as a .csv file, which included m/z, peak intensity, peak width and retention time. Peak pairs extraction, peak filter and peak ratio calculations were conducted using the software tool IsoMS.¹⁰⁸ The files were aligned by each feature's retention time and accurate mass. The missing values in features were filled in using a Zerofill script.⁹⁷ Principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and pathway enrichment analysis were all performed using the website-based statistical tool MetaboAnalyst (www.metaboanalyst.ca).¹⁸³ Volcano plots were generated by Origin 2015. Metabolite positive identification was done based on retention time and accurate mass match to a DnsID library (www.mycompoundid.org).¹¹² Metabolite putative identification was performed based on accurate mass match to the metabolites in the human metabolome database (HMDB) (www.hmdb.ca)¹⁸⁴ and the evidence-based metabolome library in MyCompoundID (MCID) (www.mycompoundid.org)¹⁵¹ with one reaction. The MCID library is composed of 8,021 known human endogenous metabolites and 375,809 predicted metabolites from one metabolic reaction.

3.3 Results

This initial exploration of global molecular effects of LIPUS on *S. cerevisiae* is an end-point study of steady state RNA and metabolite abundance.

3.3.1 Analysis of Annotated Transcripts

Data obtained by RNA-seq of total RNA were used to compare the expression level of annotated transcripts in control and LIPUS-treated cells. ANOVA without a fold-change threshold filter revealed that 354 annotated transcripts showed a significant difference in expression (reads) between the three groups of samples (in Supporting Information which is available from Dr. Liang Li). Consistent with an effect of LIPUS on transcript abundance, PCA

of the log transformed expression data separated the 100 mW/cm² sample set from CON and 80 mW/cm² (Figure 3.1). CON and 80 mW/cm², on the other hand, were not widely separated. While global statistical analysis did not separate both 80 mW/cm² and 100 mW/cm² from CON, the data for the treatment groups does hint at an effect of LIPUS on RNA expression levels. Specifically, considering differences that satisfy the $p \leq 0.05$ threshold, the 80 mW/cm² and 100 mW/cm² groups share 23 genomic features that correspond to transcripts that differ in abundance from the CON samples (Table 3.2).

Table 3.2 Low magnitude but statistically significant transcript abundance differences between control and LIPUS-treated samples (two-tailed t-test).

Systematic name	Gene	Brief description	CON vs 80 mW/cm ²		CON vs 100 mW/cm ²	
			Fold change	p value	Fold change	p value
YDL126C	<i>CDC48</i>	AAA ATPase; subunit of polyUb-selective segregase complex	-1.13	0.02	-1.19	1.53E-03
YDR099W	<i>BMH2</i>	14-3-3 protein, minor isoform	1.17	0.05	1.31	0.01
YER074W	<i>RPS24A</i>	Ribosomal 40S subunit protein S24A	-1.09	0.04	-1.1	5.15E-03
YGL067W	<i>NPY1</i>	Peroxisomal NADH diphosphatase (pyrophosphatase)	1.21	0.05	1.26	0.02
YIL009C-A	<i>EST3</i>	Component of the telomerase holoenzyme	-1.3	0.03	-1.4	0.03
YIR035C	<i>NRE1</i>	Putative cytoplasmic short-chain dehydrogenase/reductase	1.37	0.04	1.21	0.03
YJL075C		ORF, Dubious	-1.22	0.05	-1.33	0.05
YKL053W		ORF, Dubious	-1.48	5.14E-03	-1.32	0.02
YKL187C	<i>FAT3</i>	Required for fatty acid uptake; mitochondrion-associated	1.52	0.05	1.75	0.05
YLL037W		ORF, Dubious	-1.36	0.04	-1.56	0.03
YLR174W	<i>IDP2</i>	Cytosolic NADP-specific isocitrate dehydrogenase	-1.23	0.02	-1.23	0.03

YML079W		ORF , Uncharacterized	-1.19	0.02	-1.21	0.04
YMR260C	<i>TIF11</i>	Translation initiation factor eIF1A	-1.14	0.01	-1.11	0.01
YNL113W	<i>RPC19</i>	RNA polymerase subunit AC19	1.32	0.02	1.43	9.71E-03
YNL298W	<i>CLA4</i>	Cdc42p-activated protein kinase	-1.18	0.02	-1.29	8.79E-03
YNR026C	<i>SEC12</i>	Guanine nucleotide exchange factor involved in ER to Golgi transport	1.26	0.05	1.25	0.01
YNR035C	<i>ARC35</i>	Subunit of the ARP2/3 complex	-1.19	0.04	-1.2	0.04
YOL054W	<i>PSH1</i>	E3 ubiquitin ligase	1.1	0.05	1.18	0.01
YOL096C	<i>COQ3</i>	Mitochondrial O-methyltransferase	-1.14	0.05	-1.24	0.01
YOR040W	<i>GLO4</i>	Mitochondrial glyoxalase II	1.43	5.59E-03	1.59	0.04
YPL179W	<i>PPQ1</i>	PP1 family protein phosphatase	1.34	0.05	1.36	0.05
YPL274W	<i>SAM3</i>	High-affinity S-adenosylmethionine permease	-1.12	0.03	-1.25	0.05
YPR126C		ORF , Dubious	1.49	0.03	1.49	0.03

The analyses in Figure 3.1 were refined by filtering to include only those mRNA expression differences that likely affect protein synthesis in the cell, that is, expression differences higher than 2-fold.^{185,186} When this filter is applied to the data, it is evident that neither 80 mW/cm² nor 100 mW/cm² differs substantially from CON (In supplemental information which is available from Dr. Liang Li). That is, only YML039W (retrotransposon TYA Gag and TYB Pol genes) differs between 80 mW/cm² and CON (-2.6 fold), and only *YPT6*, *RIP1* and uncharacterized ORF YDL071C differ between 100 mW/cm² and CON (respectively 2.1, 2 and 3.6 fold). Furthermore, no mRNA that differs in abundance by 2-fold or more between 80 mW/cm² and CON also differs in expression between 100 mW/cm² and CON, and the gene that differs in expression between CON and 80 mW/cm² is not in the same pathway or functional category as any gene that differs between CON and 100 mW/cm². Since 80 mW/cm² and 100 mW/cm² have the same effect on ethanol production,¹⁸⁰ it follows that the molecular mechanism by which these

treatments improve biofuel yield is not reflected in end-point mRNA concentration differences that could underlie population-wide differences in protein expression level.

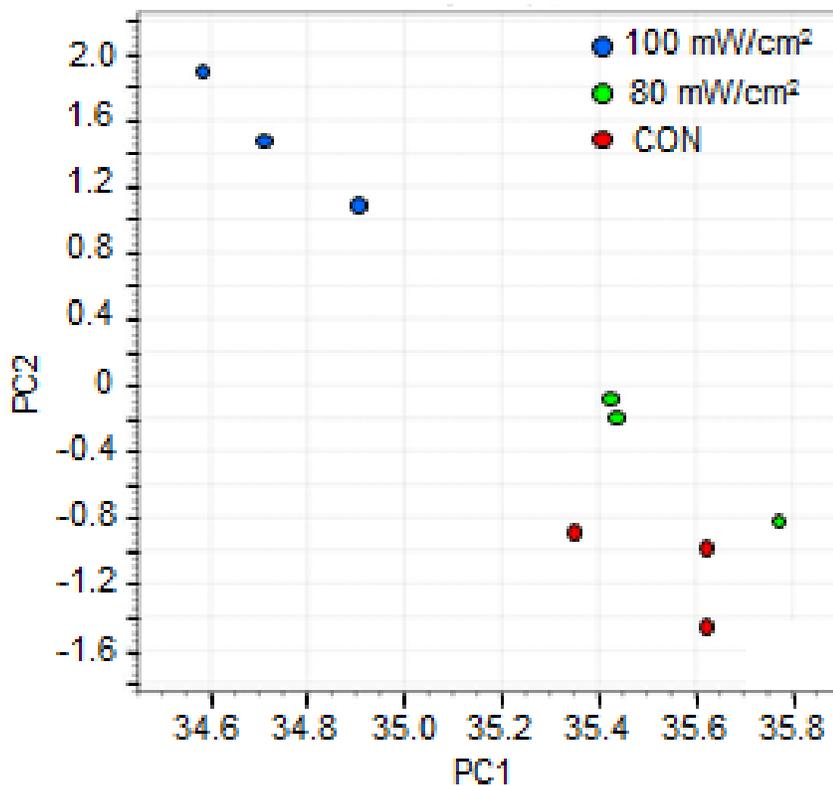


Figure 3.1 PCA analysis of control and LIPUS-treated samples by RNA expression values. Sample points are plotted according to PCA of the log transformed RNA-seq data that were statistically significant by ANOVA.

3.3.2 Metabolomic Profiling

RNA-seq analysis revealed only a modest effect of LIPUS on transcript abundance. Since LIPUS causes increased cellular production of ethanol,¹⁸⁰ we turned our attention to the possibility that steady state levels of other intracellular metabolites differ between the control and the LIPUS-treated cells. We used metabolomics to explore this possibility. The analytical

method employed was chemical isotopic labeling liquid chromatography mass spectrometry (CIL-LC-MS). Isotope labeling was achieved by dansylation, which is advantageous because it improves the sensitivity of metabolite detection, and metabolite separation by reversed phase chromatography.⁷⁹ First, each individual sample was subject to ¹²C-dansylation labeling. The total concentration of dansylation-labeled metabolites in each individual sample was then determined by LC-UV. Based on these measurements, equal amounts of each individual sample were mixed to generate a pooled sample, which was labeled using ¹³C-dansyl chloride. After that, each ¹²C-dansylation-labeled individual sample was mixed with an equal amount of the ¹³C-dansylation-labeled pooled sample, and the ¹²C- /¹³C-mixture injected onto LC-MS for analysis. Metabolite identification was based on retention time and accurate mass match to appropriate libraries. The data were analyzed by statistical tools to identify the dysregulated metabolites.

In total, 4035 peak pairs or putative metabolites were detected from aligned files, which combined the CON, 80 mW/cm² and 100 mW/cm² groups. We first searched against the DnsID library, which contains 275 standards with accurate mass and retention time, for positive metabolite identification. Ninety-three metabolites have a match in the DnsID library (Table 3.3). Putative metabolite identification was also performed by searching accurate mass against HMDB, and 640 peak pairs were matched to metabolites (In supplemental information which is available from Dr. Liang Li). We further searched against the predicted metabolome library MCID and identified an additional 1694 metabolite matches (In supplemental information which is available from Dr. Liang Li). Therefore, we identified 2334 metabolites out of 4035 peak pairs for a 58% matching rate.

Multivariate PCA was performed to visualize all the metabolite information obtained in the profiling experiment (Figure 3.2A). In the plot, the 80 mW/cm² and CON samples overlap, while

the 100 mW/cm² samples are separated well on principal component 2 (PC 2) from other two. The PCA is an unsupervised data reduction technique, and thus the correlation between predictive variables and target variables is not considered. As a complement to PCA, PLS-DA was also performed to examine the metabolome dataset. Figure 3.2B shows that there is only a slight separation between the CON and 80 mW/cm² samples, but a clear separation on component 1 between the 100 mW/cm² and CON. This indicates that LIPUS can affect the steady state abundance of intracellular metabolites in yeast. Ethanol production does not differ between 80 mW/cm² and 100 mW/cm²,¹⁸⁰ but 100 mW/cm² caused metabolite alterations that were not elicited by the 80 mW/cm² treatment. From these observations, we consider it unlikely that the metabolite abundance differences between CON and 100 mW/cm² are, on their own, reflective of metabolic events that lead to higher biofuel production.

Table 3.3 Metabolites identified by searching against DnsID.

HMDB.No.	Name	HMDB.No.	Name
HMDB29306	4-Ethylphenol	HMDB00719	Homoserine
HMDB29118	Tyrosyl-Valine	HMDB00716	Pipecolic acid
HMDB29109	Tyrosyl-Leucine	HMDB00696	Methionine
HMDB29098	Tyrosyl-Alanine	HMDB00687	Leucine
HMDB29008	Phenylalanyl-Valine	HMDB00670	Homo-arginine
HMDB29007	Phenylalanyl-Tyrosine	HMDB00669	Ortho-Hydroxyphenylacetic acid
HMDB28941	Leucyl-Tyrosine	HMDB00641	Glutamine
HMDB28941	Leucyl-Tyrosine	HMDB00557	Alloisoleucine
HMDB28937	Leucyl-Proline	HMDB00517	Arginine
HMDB28878	Histidinyl-Alanine	HMDB00500	4-Hydroxybenzoic acid
HMDB28854	Glycyl-Valine	HMDB00484	Vanillic acid
HMDB28853	Glycyl-Tyrosine	HMDB00469	5-Hydroxymethyluracil
HMDB28848	Glycyl-Phenylalanine	HMDB00455_2	Allocystathionine - Isomer

HMDB28844	Glycyl-Isoleucine	HMDB00455	Allocystathionine
HMDB28694	Alanyl-Phenylalanine	HMDB00440	3-Hydroxyphenylacetic acid
HMDB28691	Alanyl-Leucine	HMDB00306	Tyramine
HMDB28691	Alanyl-Leucine	HMDB00301	Urocanic acid
HMDB28689	Alanyl-Histidine	HMDB00300	Uracil
HMDB13243	Leucyl-phenylalanine	HMDB00296_2	Uridine - H2O
HMDB06050	o-Tyrosine	HMDB00296	Uridine
HMDB04987	Alpha-Aspartyl-lysine	HMDB00279_2	Saccharopine - H2O
HMDB04811	2,4-Dichlorophenol	HMDB00279	Saccharopine
HMDB03464_2	4-Guanidinobutanoic acid - H2O	HMDB00214	Ornithine
HMDB03423	D-Glutamine	HMDB00210	Pantothenic acid
HMDB03337	Oxidized glutathione	HMDB00206	N6-Acetyl-Lysine
HMDB03320	Indole-3-carboxylic acid	HMDB00192	Cystine
HMDB03012	Aniline	HMDB00191	Aspartic Acid
HMDB02706_2	Canavanine - Isomer	HMDB00187	Serine
HMDB02706	Canavanine	HMDB00182	Lysine
HMDB02390	3-Cresotinic acid	HMDB00177	Histidine
HMDB02199	Desaminotyrosine	HMDB00172	Isoleucine
HMDB02064	N-Acetylputrescine	HMDB00168	Asparagine
HMDB02005_2	Methionine Sulfoxide - Isomer	HMDB00167	Threonine
HMDB02005	Methionine Sulfoxide	HMDB00162	Proline
HMDB01392	p-Aminobenzoic acid	HMDB00161	Alanine
HMDB01232	4-Nitrophenol	HMDB00159	Phenylalanine
HMDB01169	4-Aminophenol	HMDB00158	Tyrosine
HMDB01123	2-Aminobenzoic acid	HMDB00149	Ethanolamine
HMDB00957	pyrocatechol	HMDB00148_2	Glutamic Acid - H2O
HMDB00939	S-Adenosylhomocysteine	HMDB00148	Glutamic Acid
HMDB00929	Tryptophan	HMDB00123	Glycine
HMDB00883	Valine	HMDB00112	Gamma-Aminobutyric acid
HMDB00759	Glycyl-L-Leucine	HMDB00099_2	Cystathionine - Isomer
HMDB00750	3-Hydroxymandelic acid	HMDB00099	Cystathionine
HMDB00721	Glycylproline	HMDB00070	Pipecolic acid
HMDB00719_2	Homoserine - H2O	HMDB00020	p-Hydroxyphenylacetic acid
HMDB00056	Beta-Alanine		

We plotted the PCA and PLS-DA of only the CON and 100 mW/cm² groups to further evaluate the influence of LIPUS on metabolite abundance in yeast cells. These two groups clearly separate from each other on PC1 in the PCA score plot (Figure 3.2C). The two groups also separate on component 1 in the PLS-DA score plot (Figure 3.2D). The model that the CON and 100 mW/cm² groups are distinct was subjected to a cross-validation test. R^2 , which estimates how well the model fits the data, is 0.99; Q^2 , which describes predictive ability of the model, is 0.93.^{187,188} Together these analyses confirm the robustness of the PLS-DA model.

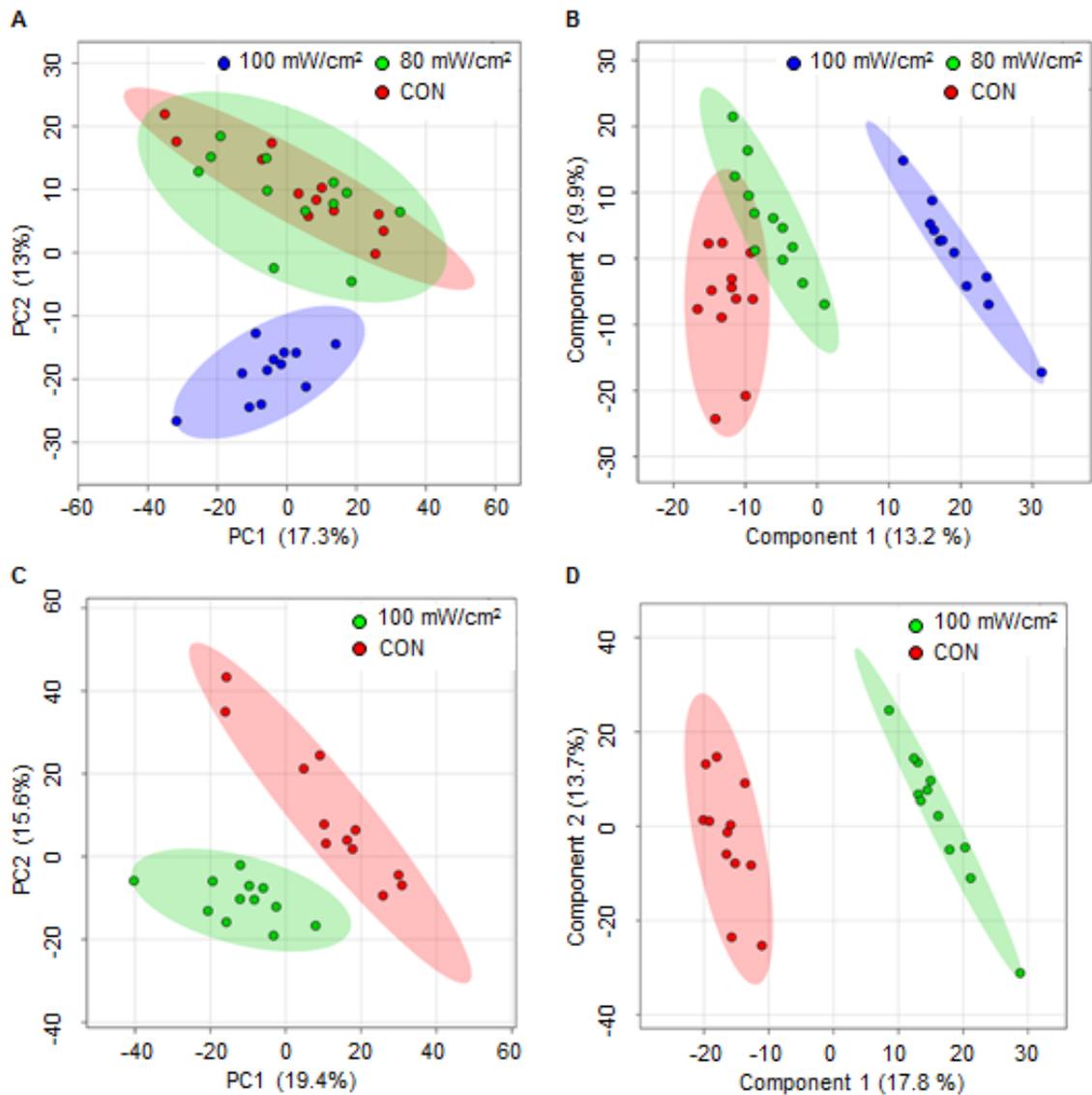


Figure 3.2 PCA and PLS-DA analysis of control and LIPUS-treated samples by relative metabolite abundance. (A) PCA and (B) PLS-DA scoring plots of control, 80 mW/cm² and 100 mW/cm² treated groups. (C) PCA and (D) PLS-DA scoring plots of control and 100 mW/cm² treated groups.

Volcano plots of CON versus 80 mW/cm² (Figure 3.3A) and CON versus 100 mW/cm² (Figure 3.3B) were also used to visualize the metabolomics data. At the threshold of $p=0.05$, the

fold change cut-off value for up-regulated and down-regulated metabolites was 1.20 and 0.83, respectively. Comparing CON and 100 mW/cm² there are 434 upregulated metabolites and 229 down regulated metabolites (Figure 3.3B). The number of dysregulated metabolites is lower when comparing CON and 80 mW/cm²: there are only 87 upregulated metabolites and 36 downregulated metabolites (Figure 3.3A). These results further confirm that the 100 mW/cm² treatment has a more significant effect on yeast metabolism than the lower dose 80 mW/cm² treatment.

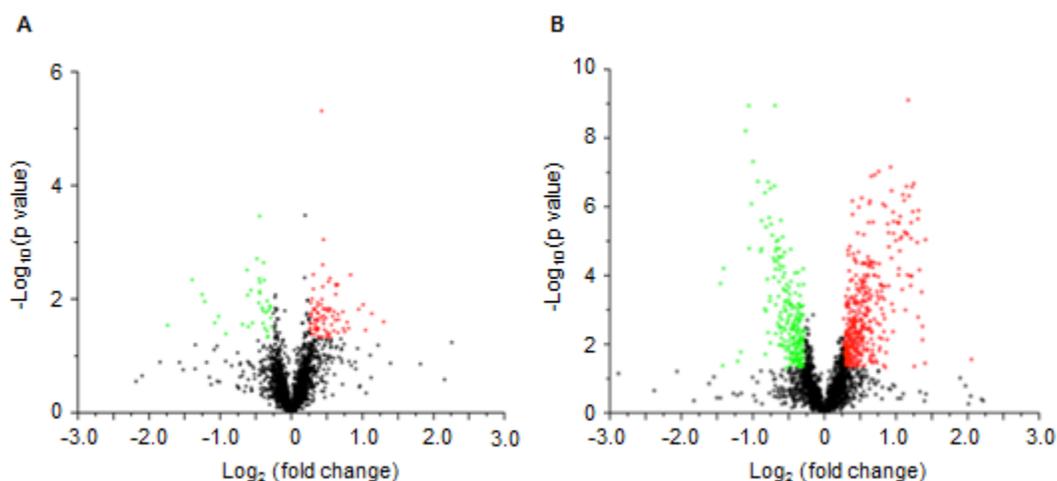


Figure 3.3 Volcano plots of metabolite abundance in LIPUS-treated cells relative to the control. (A) 80 mW/cm² versus the control. (B) 100 mW/cm² versus the control. The up- and down-regulated metabolites (1.2-fold threshold, $p \leq 0.05$) are marked in red and green, respectively.

Table 3.4 Metabolites that differ significantly between CON and 100 mW/cm² and have been definitively identified.

HMDB No.	Name	Fold change	p value
HMDB00300	Uracil	0.75	0.001898
HMDB00214	Ornithine	0.73	0.007194
HMDB00279	Saccharopine	1.22	0.049021
HMDB00670	Homo-L-arginine	1.87	2.15E-06
HMDB00641	L-Glutamine	1.29	1.03E-05
HMDB00149	Ethanolamine	1.24	0.000549
HMDB02064	N-Acetylputrescine	1.30	0.007016
HMDB00056	Beta-Alanine	1.96	2.01E-06
HMDB00296	Uridine	1.35	5.35E-05
HMDB03337	Oxidized glutathione	1.20	0.031509
HMDB00210	Pantothenic acid	1.48	2.93E-05
HMDB00296_2	Uridine - H2O	1.43	0.000121
HMDB00939	S-Adenosylhomocysteine	1.33	0.007566
HMDB00759	Glycyl-L-Leucine	1.35	0.003507
HMDB28691	Alanyl-Leucine	1.34	0.004264
HMDB28848	Glycyl-Phenylalanine	1.48	0.002568
HMDB28694	Alanyl-Phenylalanine	1.28	0.024969
HMDB04987	Alpha-Aspartyl-lysine	1.44	0.002311
HMDB00440	3-Hydroxyphenylacetic acid	1.29	0.008094
HMDB28853	Glycyl-Tyrosine	1.41	0.000138

Table 3.4 lists the metabolites that differ significantly in abundance between CON and 100 mW/cm², and have been definitively identified. Two metabolites with increased steady state abundance in the 100 mW/cm² treatment group, namely β -alanine and pantothenic acid, have a direct precursor-product relationship.¹⁸⁹⁻¹⁹¹ Interestingly four other dysregulated metabolites – uridine, uracil, N-acetyl-putrescine and ornithine – all potentially contribute to the synthesis of β -

alanine. Considering this data in the context of the organization of yeast metabolic pathways,¹⁹² it is plausible that LIPUS affects the metabolism of pyrimidine and five amino acids (proline, alanine, aspartate, glutamate and arginine) (Figure 3.4A). In part this conclusion is consistent with the result of an *in silico* analysis in which metabolites that differ in abundance between the three sample groups were input into the Pathway Analysis tool in Metaboanalyst. This tool combines enrichment analysis based on metabolite concentrations with pathway topology analysis, which takes into account the structure of pathways. In the diagram shown in Figure 3.4B, the y-axis represents the p-value calculated from pathway enrichment analysis, while x-axis represents the pathway impact values from topological analysis. The deeper red color represents a larger p-value, and a larger node radius represents a larger impact value. This pathway impact visualization reveals some of the same relationships as manual analysis. Overall the metabolomics data are consistent with an effect of LIPUS on mechanisms that tie into the pathways of pyrimidine, proline, alanine, aspartate, glutamate and arginine metabolism.

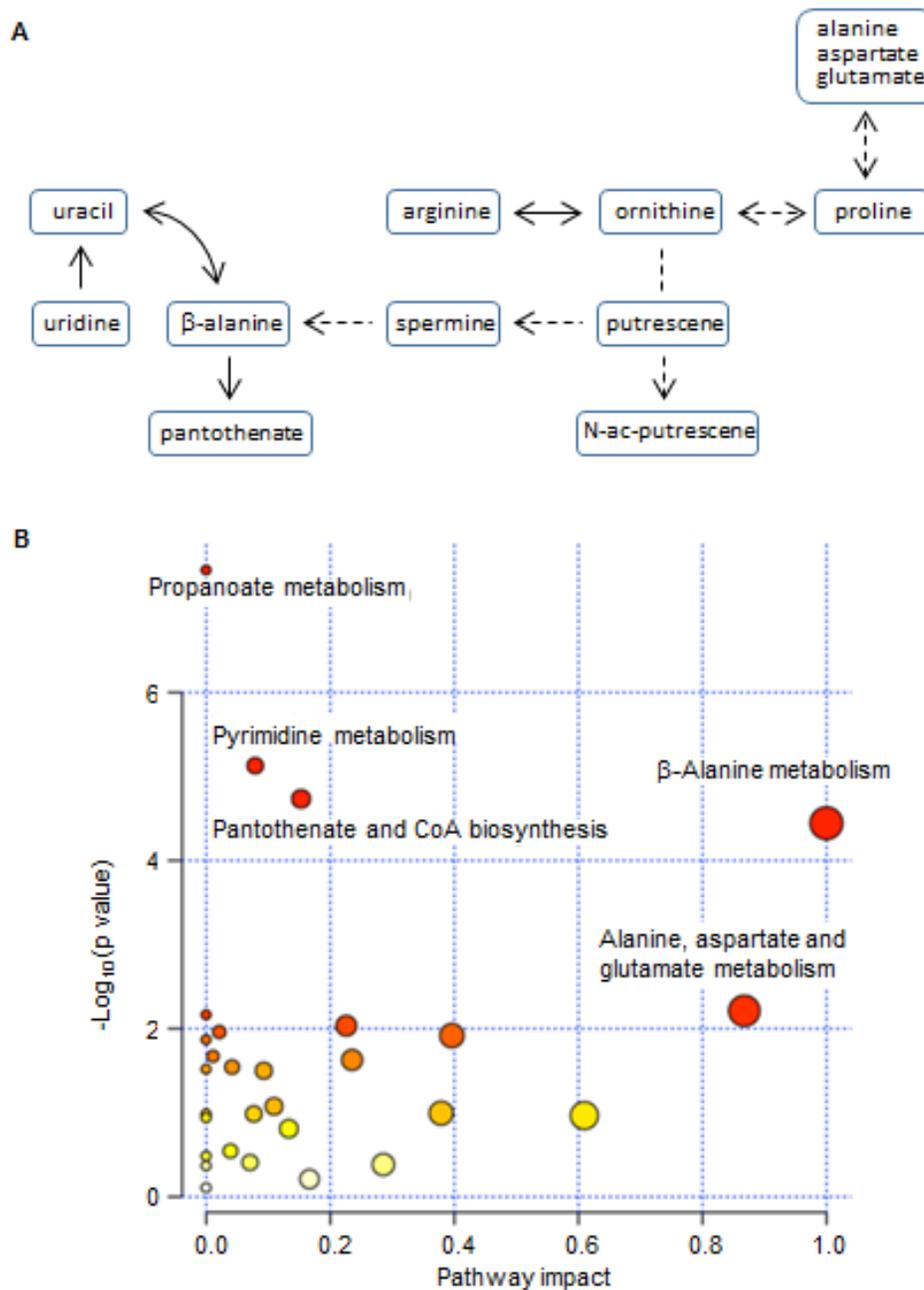


Figure 3.4 Metabolic pathway analysis. (A) Pathway relationships of dysregulated metabolites suggested by manual mapping of the CON versus 100 mW/cm² metabolite data onto the metabolic chart of yeast. (B) Overview of metabolic pathway enrichment analysis result using Pathway Analysis tools in Metaboanalyst.

3.4 Discussion

The yield of ethanol from high-glucose cultures of *S. cerevisiae* strain SSL3 is significantly improved by LIPUS.¹⁸⁰ Here we explored the possibility that transcriptomic and metabolomic profiling will reveal biological mechanisms that contribute to this effect.

The current endpoint analysis of transcript abundance suggests that modulation of mRNA availability for translation is not an important mechanism of physiological reprogramming by LIPUS. This is somewhat surprising because transcriptional and translational regulation make a major contribution to the normal control of glucose and ethanol metabolism in yeast. It is possible that some effects of LIPUS on transcript abundance in cells were not detected because they occurred early during the LIPUS treatment and were transient. This scenario is not unreasonable considering that transcription can be reprogrammed during glucose depletion from the medium even when its total level remains sufficient for robust fermentative metabolism and exponential growth.¹⁹³ Future detailed time course studies in which transcript abundance is monitored in parallel with the proliferation state of cells and glucose and ethanol levels in the culture will be required to test this hypothesis.

For cells harvested at the completion of the growth phase under fermentation conditions, metabolomics was a much more effective tool than transcriptomics for revealing physiological responses to LIPUS treatment. That is, while the small number of transcript changes revealed by RNA-seq could not be mapped onto a plausible cell response pathway, the pattern of metabolite changes under LIPUS suggested that flux through interdependent pathways of metabolism is affected by this treatment (Figure 3.4A). Therefore, metabolomics analysis is a promising avenue for further characterization of how cells react to LIPUS. The time course studies proposed for transcriptomics analysis may serve as an important guide for future metabolomics experiments.

In particular, if the temporal studies reveal a transient effect of LIPUS on the transcriptome, then it would likely be profitable to perform a detailed study of metabolomic reprogramming during this time interval.

3.5 Conclusion

We studied the impact of LIPUS on RNA expression and metabolism of *S. cerevisiae*. Metabolomic profiling by CIL-LC-MS indicated that LIPUS has an impact on the pathways of alanine, arginine, aspartate, glutamate, proline, and pyrimidine metabolism, though the transcript expression signature of LIPUS treated *S. cerevisiae* did not differ significantly compared to the untreated cells after five days. LIPUS activates metabolic effects beyond reprogramming of the core pathways of carbon metabolism.

Chapter 4

Development of a Simple and Efficient Method of Harvesting and Lysing Adherent Mammalian Cells for Chemical Isotope Labeling LC-MS-Based Cellular Metabolomics

4.1 Introduction

Cellular metabolomics uses analytical techniques to detect and quantify a large set of metabolites or the metabolome in cells. Even with the analysis of a subset of the whole metabolome using techniques currently available, cellular metabolomics has become an important tool in biological research.¹⁹⁴ Comparing to analyzing the metabolomes of biofluids such as urine and blood, cellular metabolomics requires additional sample handling steps, i.e., cell harvest and lysis. In order to profile the metabolome properly, a robust and reproducible method for cell harvest and lysis is required. Several studies have shown that improper handling of the cell harvest and lysis process could artificially alter the metabolite concentrations due to sample loss as well as residual metabolic enzyme activities.^{26,27,195-205}

Because current analytical techniques can only cover a fraction of the entire metabolome and metabolite detectability varies from one technique to another, the extent of any effects on the metabolome data caused by cell harvest and lysis methods is technique-dependent. The reported studies mainly based on the use of NMR, GC-MS and LC-MS methods for metabolome analysis.^{26,27,195-205} Among these methods, LC-MS provides higher sensitivity for metabolite detection. However, conventional LC-MS techniques, even with the use of multiple methods (e.g., various combinations of reversed phase LC separation, hydrophilic interaction LC separation, positive ion detection, and negative ion detection), do not offer high-coverage metabolome analysis with high quantification accuracy. One alternative approach of metabolome analysis is to use

chemical isotope labeling (CIL) to alter the chemical and physical properties of the metabolites for improving detection sensitivity and quantification accuracy.⁷⁹ There are a number of labeling reagents that have been reported for targeted and untargeted metabolite analysis with varying degrees of success.²⁰⁶⁻²¹⁶

Our laboratory has been involved in developing a “divide-and-conquer” approach based on CIL LC-MS for comprehensive and quantitative metabolomics. We have reported four rationally designed isotope labeling reagents for analyzing the amine/phenol⁷⁹, carboxyl⁸⁰, hydroxyl⁸² and carbonyl⁸¹ submetabolomes separately. The combined results of the four submetabolomes offer a high-coverage analysis of the whole metabolome. In addition, the labeled metabolites can be efficiently separated using reversed phase (RP) LC and effectively ionized as mainly protonated ions, rendering the possibility of using a single setup, RPLC-MS with positive ion mode detection, for metabolite analysis.

In this study, we report a simple and efficient method of harvesting and lysing adherent mammalian cells tailored to CIL LC-MS-based cellular metabolomics. Being able to detect thousands of cellular metabolites with high quantification accuracy, ¹²C-/¹³C-dansylation LC-MS was employed to examine the amine/phenol submetabolomes of cell extracted prepared using different cell harvest and lysis methods. Amines/phenols are major groups of metabolites in the cellular metabolome and involved in most of the metabolic pathways in metabolisms. Thus, the impact of cell harvesting and lysis processes on the amine/phenol submetabolome should reflect their impact on the whole metabolome. Using MCF-7 cells and HeLa cells as representatives of cultured adherent cells widely used in biological studies, we examined and compared the performance of trypsinization method vs. physical scraping method for cell harvest, and glass-bead-assisted method vs. freeze-thaw-cycle method for cell lysis.

4.2 Experimental

4.2.1. Overall workflow

Figure 4.1 shows the overall workflow of this study. MCF-7 cells and HeLa cells were cultured in 6-well plates in replicates with the same cell number. Cells were harvested by two different methods: trypsinization or physical scraping. The cell pellets were then treated by two lysis methods: freeze-thaw-cycle lysis or glass-bead-assisted lysis. The cell lysates were extracted and subjected to chemical labeling using ^{12}C -dansyl chloride (DnsCl). A pooled sample from aliquots of individual samples was prepared and labeled by ^{13}C -dansyl chloride. The total concentration of the labeled metabolites in each sample was measured by LC-UV. The ^{12}C -labeled sample and ^{13}C -labeled pool were mixed by equal mole amount. The mixture was injected into LC-MS for analysis. The peak pairs detected in MS were extracted by IsoMS, and individual peak-pairs from different LC-MS runs were aligned together based on accurate mass and retention time to produce a metabolite peak ratio table. Multivariate data analysis was performed by MetaboAnalyst (<http://www.metaboanalyst.ca>). The metabolites were identified by searching against MyCompoundID library (<http://www.mycompoundid.org>).

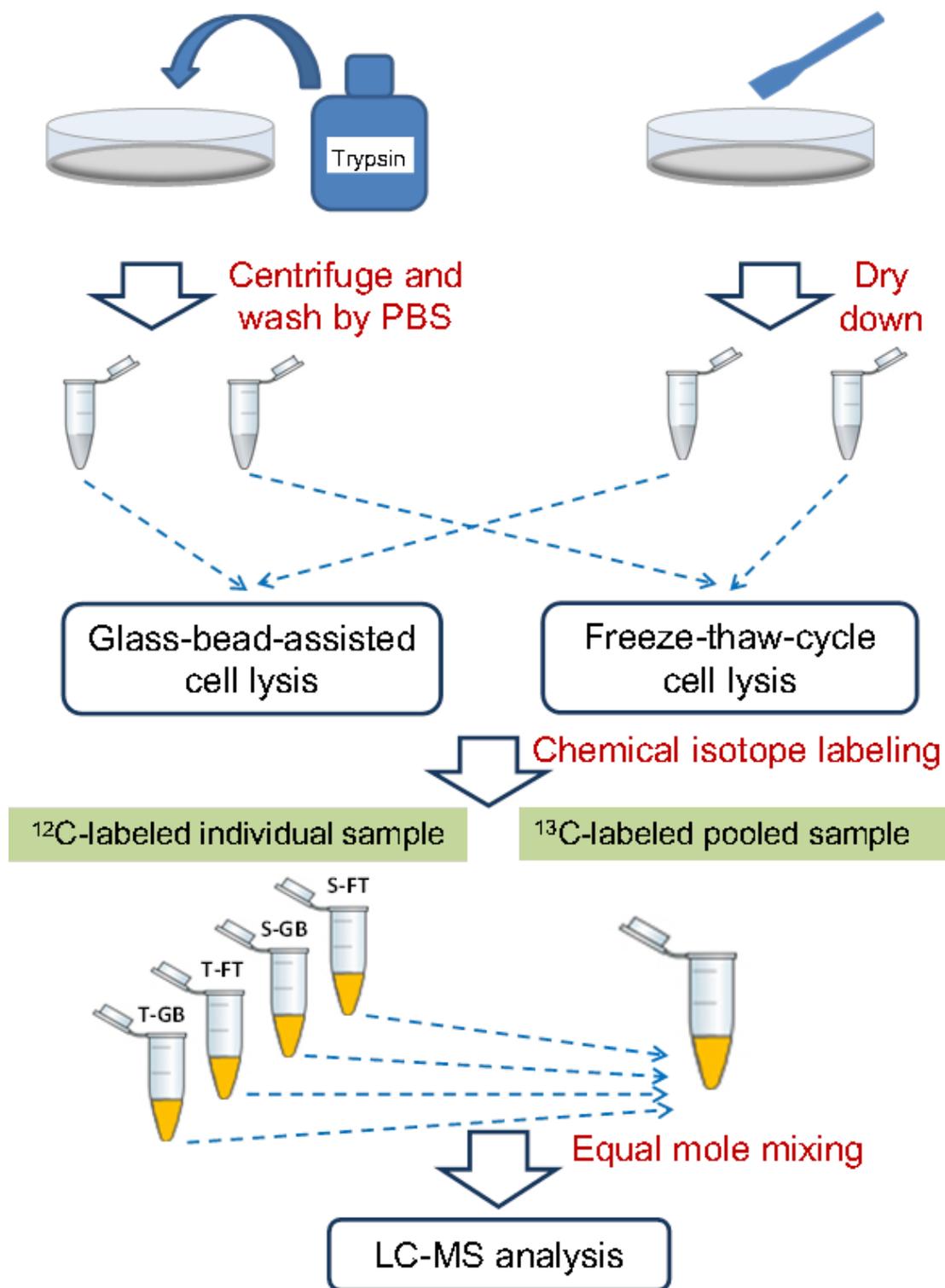


Figure 4.1 Workflow for comparing different methods to develop a simple and efficient cell harvest and lysis method for CIL LC-MS metabolomics of adherent mammalian cells.

4.2.2 Chemicals and reagents

The LC-MS grade reagent, including water, acetonitrile, methanol and formic acid, were purchased from Fisher Scientific (Ottawa, ON). 0.5-mm-diameter glass beads were purchased from Biospec Products (Bartlesville, OK). ^{13}C -dansyl chloride was available from the University of Alberta (<http://mcid.chem.ualberta.ca>).

4.2.3 Cell culture

Two types of cell lines, HeLa (ATCC CCL-2) and MCF-7 (ATCC HTB-22), were selected in this study. The growth medium for HeLa cell was Hyclone Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10% fetal bovine serum (FBS). For MCF-7 cell culture, additional 0.01 mg/mL human recombinant insulin was supplemented as suggested by American Type Culture Collection (ATCC). The same number of cells was cultured in Falcon 6-well plates. The cultures were incubated at 37 °C in a humidified atmosphere with 5% CO₂. The growth medium was renewed every two days.

4.2.4 Cell harvest

Cells were harvested by either trypsinization or physical scraping. For trypsinization, the cells were washed by cold phosphate buffer saline (PBS). 0.5 mL of 0.25% trypsin/EDTA (Hyclone, Logan, Utah) was added and incubated with cell cultures at 37 °C. The trypsinization process was monitored under inverted microscope, and quenched by growth medium when cells appeared rounded. The cultures were then transferred into the 15-mL centrifuge tubes. Trypsin and growth medium were removed by 7-min 125 g centrifugation at 4 °C. The cell pellets were suspended in 5 mL of PBS and centrifuged at 125 g for 7 min at 4 °C. The cells were washed three times. The cells were then snap-frozen in liquid nitrogen and stored in -80 °C freezer. Note that cell washing to remove the growth medium is important to reduce contamination of

metabolites from the medium in the cellular metabolome. We examined the cell washing efficiency using an approach described previously ⁷. Basically, the washing elute from each round of PBS washing of cells was labeled and injected into LC-MS for analysis. Any metabolites detected in a washing elute would indicate that the washing was not complete. We found that, in general, after three times washing cycles, metabolites from growth medium could be removed.

For physical scraping, the growth medium was removed, and cell cultures were washed by cold PBS for three times. 1 mL of cold methanol was added for metabolism quenching. The cells were then detached by scraping, and transferred into 1.5-mL vials. After methanol was removed using Savant SC110A Speed Vac, the sample vials were stored in -80 °C freezer for further use.

4.2.5 Cell lysis

Cell lysis by using the glass-bead-assisted lysis method followed a previously published protocol ⁶. In brief, the cell pellets were suspended in 100 µL of 50% MeOH and 50% water, and 0.5 mL of glass beads were added. Cells were lysed via five 1-min periods of bead-beating at 3200 rpm alternated with five 1-min incubations in an ice-water bath. After cell lysis, 800 µL of 50% MeOH and 50% water was added for metabolite extraction. Glass beads, cell debris and unbroken cells were removed by centrifugation at 16000 g for 10 min at 4 °C. The supernatant was then transferred into another vial and dried down in Speed Vac.

For freeze-thaw-cycle lysis, 300 µL of 50% MeOH and 50% water was added into cell pellets. The vial was placed in liquid nitrogen for 2 min, and thawed in water for 2 min with vortex. The freeze-thaw cycle was repeated for four more times. Then the vial was centrifuged at 16000 g for 10 min, and the supernatant was transferred to another vial and dried down. The dried metabolites were re-dissolved in water and stored in -80 °C freezer.

It should be noted that we did not use ultrasonication for cell lysis in this work. While ultrasonication is a widely used method for lysing mammalian cells, we have recently shown that cell lysis efficiency was similar for the ultrasonication method and the glass-bead-assisted method²¹⁷. However, a lot of energy is absorbed during the ultrasonication process, which may cause metabolite degradation²¹⁷. In addition, comparing to freeze-thaw-cycle or bead-assisted method, ultrasonication has a relatively low throughput: only one sample can be lysed using a sonication tip. Moreover, there is also a risk of cross-contamination if the tip is not washed thoroughly.

4.2.6 Dansylation labeling

The labeling protocol was the same as that previously reported⁷⁹. In brief, 25 μL of cell extract was mixed with 12.5 μL of ACN, and 12.5 μL of sodium carbonate/sodium bicarbonate buffer and 25 μL of ^{12}C -dansyl chloride (18 mg/mL in ACN) or ^{13}C -dansyl chloride (18 mg/mL in ACN). The reaction vial was incubated at 40 °C for one hour. 5 μL of 250 mM NaOH was added and incubated for another 10 min to quench the excess DnsCl. Finally, 25 μL of 425 mM formic acid in 1:1 ACN/H₂O was added to the reaction mixture to acidify the solution.

4.2.7 LC-UV

The total concentration of dansyl labeled metabolites was measured by a step-gradient LC-UV method¹⁵⁰. 5 μL of labeled sample was injected into a Phenomenex Kinetex C18 column (2.1 mm \times 5 cm, 1.7 μm particle size, 100 Å pore size) connected to a Waters ACQUITY UPLC system (Waters, Milford, MA). Mobile phase A was 0.1% (v/v) formic acid in 5% (v/v) ACN, and solvent B was 0.1% (v/v) formic acid in acetonitrile. The LC gradient was as follows: t = 0, 0% B; t = 1 min, 0% B; t = 1.1 min, 95% B; t = 2.6 min, 95% B; t = 3.1 min 0% B; t = 6.5 min, 0% B. The flow rate was 0.45 mL/min. The PDA detector was operated at 338 nm.

4.2.8 LC-MS

Each ^{12}C -labeled individual sample was mixed with ^{13}C -labeled pool sample by equal mole amount. LC-MS was done using a Thermo Scientific Dionex Ultimate 3000 UHPLC System (Sunnyvale, CA) linked to a Bruker Maxis II quadrupole time-of-flight (Q-TOF) mass spectrometer (Bruker, Billerica, MA). The LC column was an Agilent reversed phase Eclipse Plus C18 column (2.1 mm \times 10 cm, 1.8 μm particle size, 95 \AA pore size). The mobile phases were the same as those used for LC-UV. The LC gradient was: t = 0 min, 20% B; t = 3.5 min, 35% B; t = 18 min, 65% B; t = 21 min, 99% B; t = 34 min, 99% B. The flow rate was 0.18 mL/min. The MS conditions were as follows: polarity, positive; dry temperature, 230 $^{\circ}\text{C}$; dry gas, 8 L/min; capillary voltage, 4500V; nebulizer, 1.0 bar; end plate offset, 500V; spectra rate, 1.0 Hz.

4.2.9 Data analysis

All the spectra were first converted to .csv files by Bruker Daltonics Data Analysis 4.3 software. The peak pairs were extracted from .csv files by IsoMS ¹⁰⁸. Data generated from multiple runs were aligned together based on peak's accurate mass and retention time. The missing values in aligned file were filled by Zerofill software ⁹⁷. The principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were performed by MetaboAnalyst ²¹⁸ (www.metaboanalyst.ca). The metabolites were positively identified by searching against DnsID Library, which contains retention time, MS and MS/MS information of 275 unique amine/phenol-containing metabolite standards ¹¹² (www.mycompoundid.org). Putative identification or match was performed by searching accurate mass against MyCompoundID library, which contains 8,021 known human metabolites and 375,809 predicted metabolites ¹⁵¹ (www.mycompoundid.org).

4.3 Results and Discussion

4.3.1 LC-UV quantification for cell harvest and lysis efficiency comparison

Cellular metabolomics involves the comparison of the metabolomes of different groups of cells. To compare the concentration differences of individual metabolites in different cell samples, it is critical to normalize the sample amounts before performing LC-MS analysis of samples. We have reported a step-gradient LC-UV method to measure the total concentration of dansyl labeled metabolites in a sample and use the total concentration for sample amount normalization¹⁵⁰. In this work, we applied this approach to gauge the relative performance of cell harvest and lysis methods. Briefly, we started with the seeding of the same number of cells in individual wells of a 6-well plate for replicate culturing, which ensured that the same number of cultured cells was used as the starting material from each well for cell harvest, cell lysis and cell-extract labeling. We performed the LC-UV analysis of labeled metabolites from the processed samples and then compared their LC-UV quantification results which should reflect the differences in efficiencies of cell harvest and cell lysis done by different methods.

In our study, cells were harvested by trypsinization (abbreviated as T) or physical scraping (abbreviated as S) and lysed by glass-bead-assisted lysis (GB) or freeze-thaw cycle lysis (FT). In total, there were four combinations for comparison: T-GB, T-FT, S-GB and S-FT (see Figure 4.1). Two commonly used cell lines in biological studies, HeLa cells and MCF-7 cells, were selected for our study to represent adherent mammalian cells. Figures 4.2A and 2B show plots of the average concentration of labeled metabolites determined in each of the four combination methods for HeLa and MCF-7 cells, respectively. For the HeLa cells, the total concentration in T-GB, T-FT, S-GB and S-FT was found to be 0.52 ± 0.13 , 0.52 ± 0.19 , 0.90 ± 0.16 and 1.15 ± 0.23 mM, respectively. The standard deviation for each concentration measurement was the result of

combined variations in biological replicates (n=6). The total metabolite concentration in the physical scraping group is about 1.8-fold higher than that of the trypsinization group, with either GB lysis or FT lysis. Thus, for the HeLa cells, the scraping harvest method was more efficient than trypsinization. The same finding was obtained for the MCF-7 cells. In this case, the total metabolite concentration of T-GB, T-FT, S-GB and S-FT was 0.71 ± 0.25 , 1.1 ± 0.24 , 1.30 ± 0.09 and 1.79 ± 0.20 mM, respectively. The concentration of the scraping group was also about 1.8-fold higher than that of the trypsinization group. The reduced concentration might be caused by metabolite loss during trypsinization process through cell membrane damage and metabolites leakage.

The concentration plots shown in Figure 4.2 can also be used to gauge the differences in cell lysis efficiencies. For both HeLa and MCF-7 cells, the total metabolite concentrations of the FT groups were higher than those of the GB groups except the case of T-GB and T-FT groups of HeLa cells where the total concentrations of the two groups had no significant difference. In the GB lysis method, to ensure that we could recover most of the metabolites, a relatively larger volume (800 μ L) of extract solvent was used to rinse the beads, followed by drying. During the drying process, some relatively volatile metabolites might be lost, while other metabolites might adsorb onto the container walls and could not be re-dissolved, resulting in sample loss.

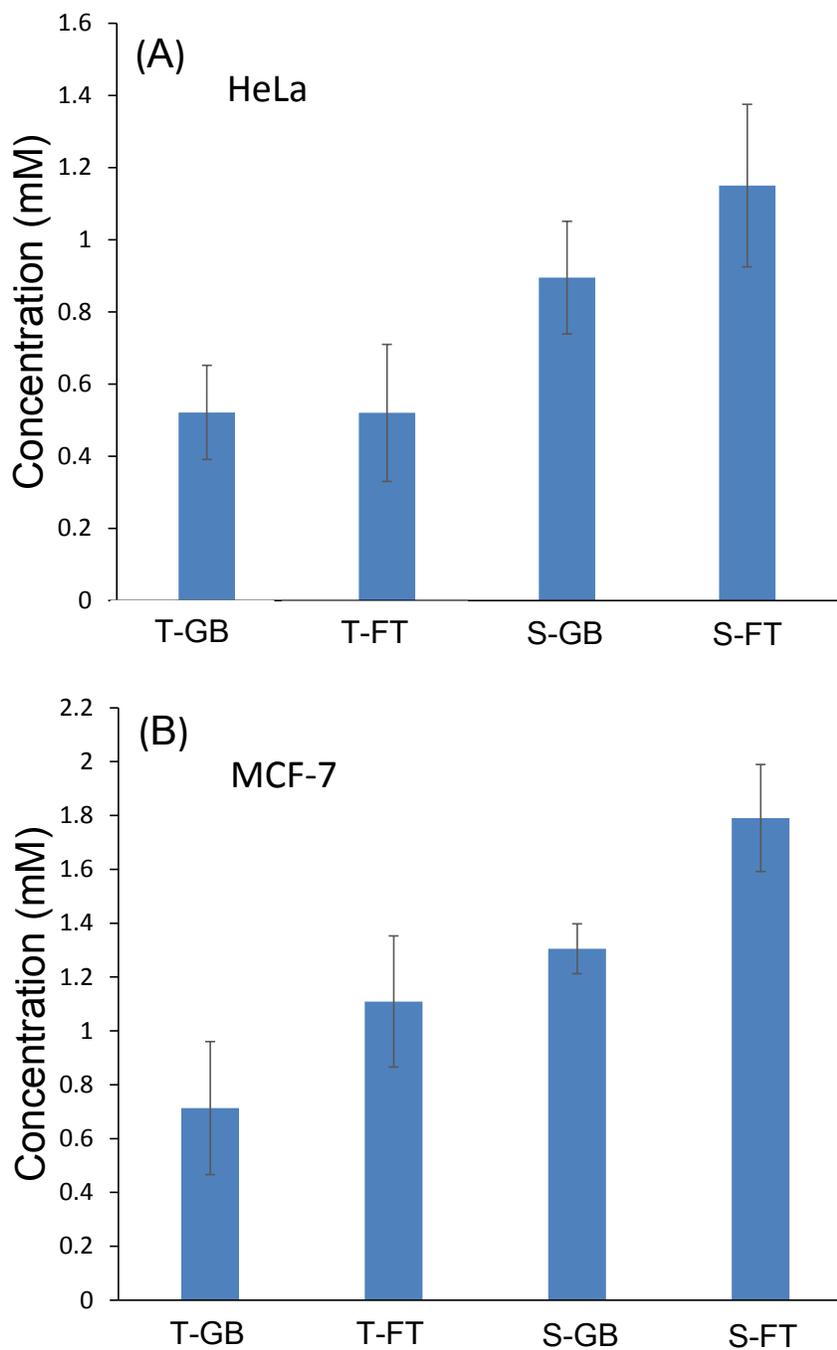


Figure 4.2 Average concentrations of dansyl labeled metabolites in cell extracts (n=6) prepared using different combinations of harvest and lysis methods from (A) HeLa and (B) MCF-7 cells.

T-GB = trypsinization cell harvest followed by glass-bead lysis. T-FT = trypsinization cell

harvest followed by freeze-thaw-cycle lysis. S-GB = scraping cell harvest followed by glass-bead lysis. S-FT = scraping cell harvest followed by freeze-thaw-cycle lysis.

The above results obtained from LC-UV measurement of labeled metabolites indicate that the combination of physical scraping for cell harvest and freeze-thaw-cycle for cell lysis gave the highest efficiencies. However, the LC-UV data only gauges the total metabolite amount difference, not the metabolite composition difference. Moreover, the experimental conditions used in harvest and cell lysis may affect the downstream process and analysis. Thus, from the metabolomic profiling point of view, we need to determine which combination method generates the optimal metabolomic result. We proceeded to use LC-MS and statistical analysis to examine the differences of the metabolome profiles generated from different combination methods.

4.3.2 LC-MS results

In our LC-MS analysis, the ^{13}C -labeled pool served as a global internal standard and was mixed with the ^{12}C -dansyl labeled individual sample by equal mole amount. The same amount of mixtures prepared from all individual samples was injected into LC-MS. On average, for the HeLa cells, 3079 ± 50 , 3033 ± 71 , 3045 ± 68 and 3016 ± 73 peak pairs were detected from T-GB, T-FT, S-GB and S-FT, respectively (see Figure 4.3 A). There was no significant difference among the four groups prepared by different harvest and lysis combinations. These results show that, if the same injection amount was used in LC-MS, the harvest and lysis methods would not affect the number of peak pairs detected. These results were confirmed in the analysis of MCF-7 cells. An average of 2768 ± 127 , 2773 ± 49 , 2641 ± 16 , and 2604 ± 72 peak pairs were detected from T-GB, T-FT, S-GB and S-FT, respectively (see Figure 4.3B). It is interesting to note that we detected about 400 peak pairs less from the MCF-7 cell lysates, compared to the HeLa cells. Judging from

the number of peak pairs or metabolites detected alone, these results indicate that the amine/phenol submetabolome profiles of MCF-7 and HeLa cells are different. Nevertheless, in both cases, thousands of peak pairs or metabolites were detected, illustrating the high metabolomic coverage achievable by the dansylation LC-MS method.

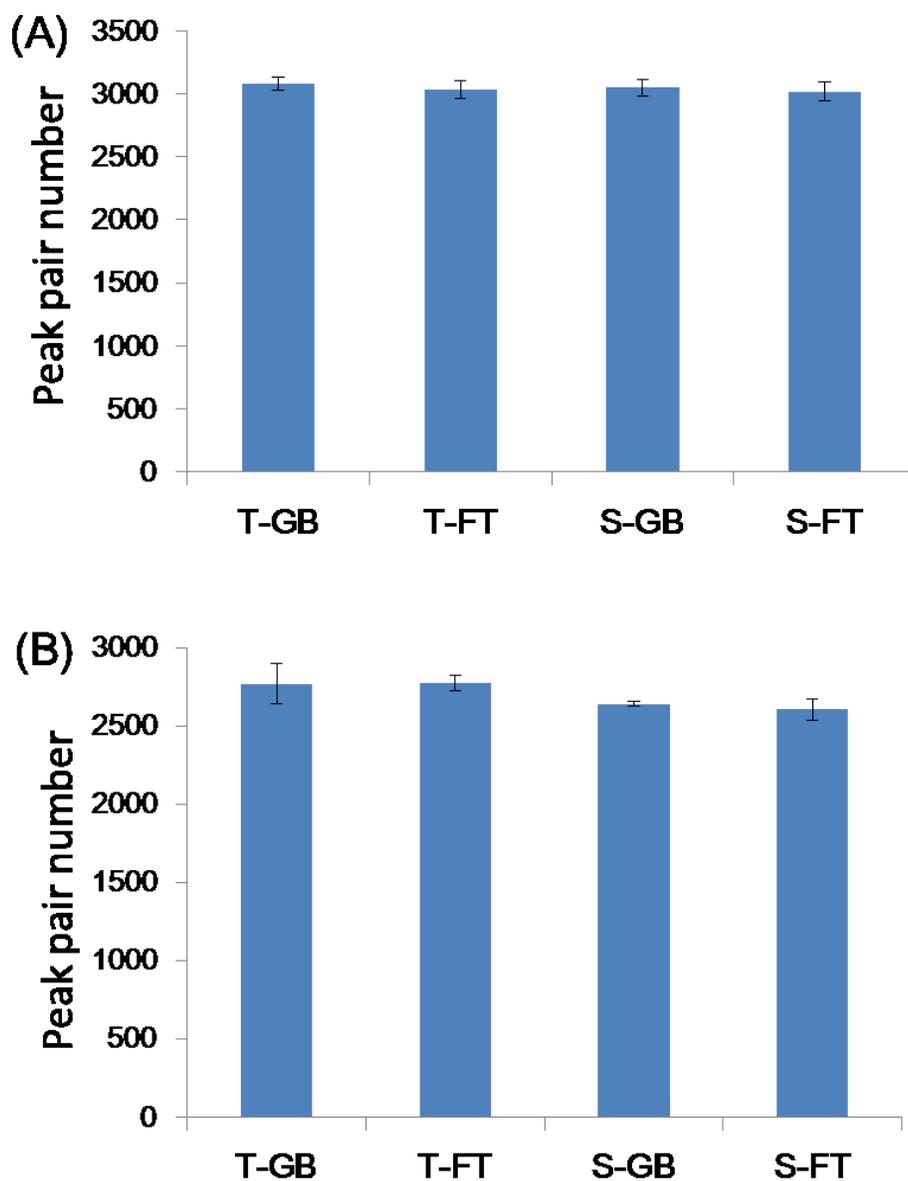


Figure 4.3 Average number of peak pairs (n=6) detected from (A) HeLa cell extracts and (B) MCF-7 cell extracts prepared using different combinations of harvest and lysis methods.

4.3.3 Multivariate statistical analysis

We applied multivariate statistical analysis to the metabolome data set obtained from the samples prepared using different harvest and lysis methods in order to examine the overall metabolome profile differences and similarities (i.e., the number and type of metabolites detected as well as their relative concentration differences in different samples). The score plot from the unsupervised PCA analysis is shown in Figure 4.4A. In this plot, 38.8% of data was captured by the first principal component (PC), and 10.6% of data was captured by the second PC. Overall, ~50% of data could be captured by 1st and 2nd PCs, indicating an excellent model. As Figure 4.4A shows, in both HeLa and MCF-7 cells, the samples of the trypsinization group (T) were separated from those of the scraping group (S), while the glass-bead-assisted lysis group (GB) and the freeze-thaw-cycle lysis group (FT) were overlapped. These results suggest that the trypsinization process might not only cause metabolite leakage or sample loss, but also cause concentration changes for some of the detectable metabolites. The lysis method (GB or FT) had minor impact on the cellular metabolome, although the GB group gave a lower total metabolite concentration as discussed earlier. In the PCA plot, no matter which harvest and lysis methods were used, HeLa cells (H) and MCF-7 cells (M) are clearly separated. Thus, the metabolomes of the two different cell lines are significantly different.

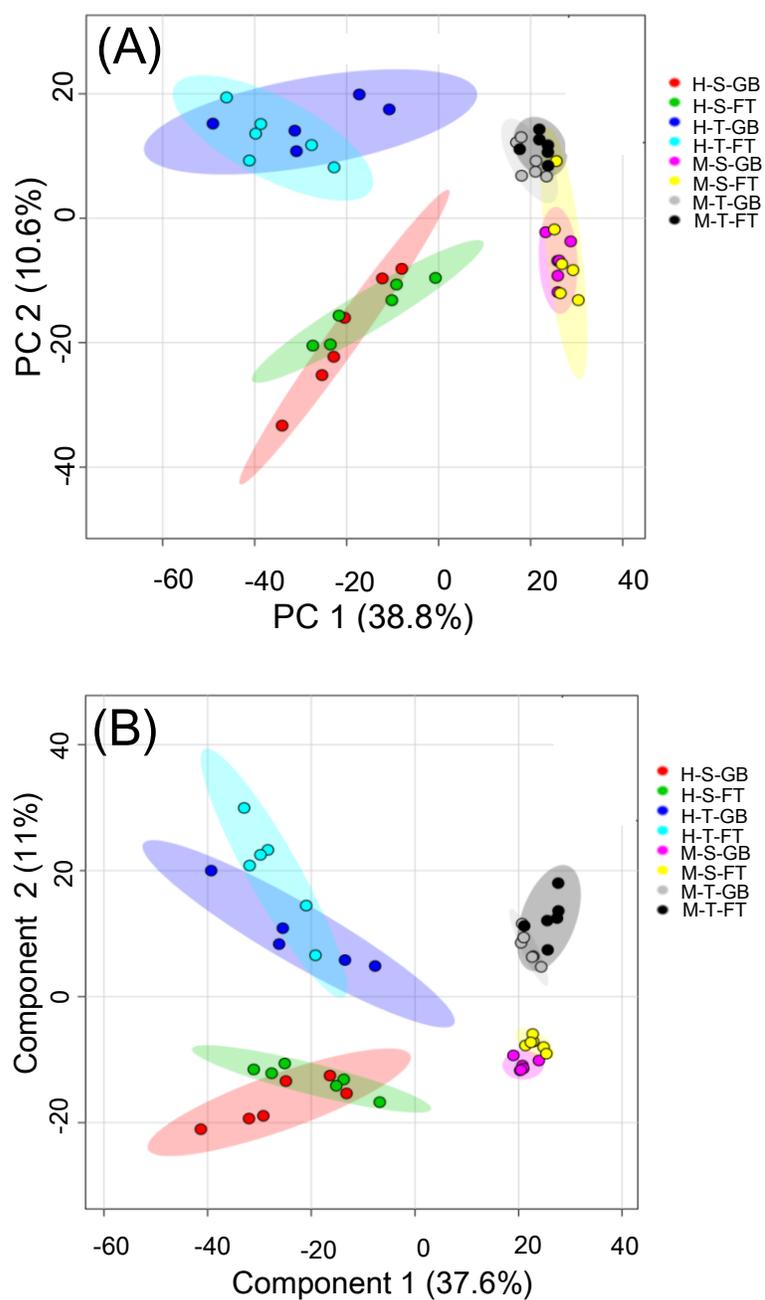


Figure 4.4 (A) PCA and (B) PLS-DA plots of the amine/phenol submetabolomes of HeLa and MCF-7 cells from cell extracted prepared using different combinations of harvest and lysis methods. H=HeLa cells. M=MCF-7 cells. Other abbreviations are shown in Figure 4.2 caption.

Supervised PLS-DA analysis was also applied to the metabolome data set and the score plot is shown in Figure 4.4B. The two cell lines are separated on component 1 and the two cell harvest groups are separated on component 2 with $R^2=0.968$ and $Q^2=0.932$ from cross-validation test. The high scores of R^2 and Q^2 confirm the robustness of the model. These PLS-DA analysis results confirmed the findings of the PCA analysis.

4.3.4 Impact of different harvest methods on cellular metabolome

To further analyze the impact of different harvest methods on cellular metabolome, univariate analysis using volcano plots was performed on the metabolome data set (Figure 4.5). In the volcano plot, the x-axis is the fold change (FC) of trypsinization/scraping groups, and y-axis is the p-value from t-test of the two groups. For HeLa cells (Figure 4.5A), there were 429 metabolites with significantly higher fold changes and 305 metabolites with significantly lower fold changes, using the criteria of $p\text{-value}<0.01$ and $FC>1.5$. We used a relatively large fold-change (>1.5) as a criterion to examine the metabolites showing relative large abundance differences in comparative samples. For MCF-7 cells (Figure 4.5B), there were 131 metabolites with higher fold changes and 88 metabolites with lower fold changes. These results show that there were a large number of metabolites having significantly different concentrations in samples prepared using the two different harvest methods. The metabolites with significant changes are listed in Tables 4.1 and 4.2 along with the metabolite identification results shown in Tables 4.3 and 4.4.

Table 4.1 Metabolites with significantly different concentrations in HeLa cell extracts prepared using different harvest methods (trypsinization/physical-scraping).

No.	HMDB	Compound Name	fold change	p-value	Level
59	HMDB01397	Guanosine monophosphate	2.93	1.47E-03	1
68	HMDB00001	1-Methylhistidine	0.46	2.67E-04	1
101	HMDB00045	Adenosine monophosphate	9.56	1.18E-06	1
105	HMDB00133	Guanosine	4.96	2.25E-03	1
107	HMDB12114	(3S)-3,6-Diaminohexanoate	1.92	1.69E-03	2
144	HMDB00517	L-Arginine	1.77	1.27E-03	1
146	HMDB00052	Argininosuccinic acid	0.55	7.64E-03	2
196	HMDB01410	2-Amino-4-oxo-6-(1',2'-dioxopropyl)-7,8-dihydroxypteridin	2.17	5.16E-03	2
202	HMDB01325	N6,N6,N6-Trimethyl-L-lysine	1.69	7.14E-05	2
206	HMDB00045	Adenosine monophosphate	9.90	7.91E-07	1
215	HMDB03331	1-Methyladenosine	2.34	1.15E-04	2
247	HMDB03276	Hydrogen sulfide	2.83	1.72E-05	2
255	HMDB00299	Xanthosine	2.55	1.25E-03	2
256	HMDB00095_2	Cytidine monophosphate - Isomer	4.08	1.85E-06	1
264	HMDB00195	Inosine	62.38	1.12E-03	2
275	HMDB03334	Symmetric dimethylarginine	2.64	3.74E-04	1
364	HMDB00641	L-Glutamine	0.25	1.04E-05	1
364	HMDB03423	D-Glutamine	0.25	1.04E-05	1
371	HMDB00904	Citrulline	0.56	6.85E-03	1
380	HMDB00856	N-a-Acetylcitrulline	1.57	7.65E-03	2
381	HMDB11737	Gamma Glutamylglutamic acid	0.60	2.57E-03	1
429	HMDB00187	L-Serine	0.66	7.30E-04	1
503	HMDB00125	Glutathione	1.83	4.02E-03	2
515	HMDB02005	Methionine Sulfoxide	0.59	8.66E-03	1
529	HMDB00187	L-Serine	0.51	1.14E-04	1
542	HMDB00191	L-Aspartic Acid	2.95	3.24E-06	1
560	HMDB12326	L-Gulose	6.74	9.64E-04	2
577	HMDB00288	Uridine 5'-monophosphate	4.81	7.24E-07	2
581	HMDB00191	L-Aspartic Acid	2.79	6.65E-04	1
622	HMDB04437	Diethanolamine	0.55	6.81E-03	1
707	HMDB02335	Aspartyl-L-proline	0.43	1.47E-05	2

763	HMDB06555	dIMP	1.51	2.35E-04	2
811	HMDB01263	Allysine	6.25	3.31E-05	2
850	HMDB00174	L-Fucose	0.59	2.94E-03	2
854	HMDB00721	Glycylproline	0.35	8.35E-06	1
907	HMDB00079	Dihydrothymine	0.29	3.06E-05	2
978	HMDB00296	Uridine	10.74	1.22E-03	1
984	HMDB00056	Beta-Alanine	0.64	8.47E-03	1
986	HMDB00585	Glucosylgalactosyl hydroxylysine	1.82	8.47E-05	2
1001	HMDB00721	Glycylproline	0.50	3.69E-04	1
1015	HMDB00323	3-Amino-2-piperidone	1.92	1.66E-04	2
1138	HMDB02284	N-Acetylcadaverine	1.87	3.17E-03	2
1155	HMDB00576	Monoethyl malonic acid	4.39	1.62E-04	2
1202	HMDB00296_2	Uridine - H2O	11.63	7.67E-05	1
1208	HMDB12136	1-Amino-propan-2-ol	3.18	4.40E-04	2
1251	HMDB11166	L-beta-aspartyl-L-leucine	2.29	3.11E-04	2
1268	HMDB00292	Xanthine	2.70	5.17E-03	1
1310	HMDB11170	L-gamma-glutamyl-L-isoleucine	1.93	9.44E-03	2
1354	HMDB03911	3-Aminoisobutanoic acid	0.44	7.35E-05	1
1407	HMDB01080	4-Aminobutyraldehyde	2.85	3.65E-05	2
1430	HMDB00600	Galactosylhydroxylysine	1.79	3.53E-05	2
1431	HMDB03609	2-Aminoacrylic acid	0.57	9.40E-04	2
1435	HMDB01257	Spermidine	2.09	2.43E-03	1
1492	HMDB00594	Glutamylphenylalanine	2.23	9.90E-03	2
1549	HMDB02201	N-Carboxyethyl-g-aminobutyric acid	0.15	2.30E-06	2
1575	HMDB28691	Alanyl-Leucine	4.19	3.02E-05	1
1661	HMDB00883	L-Valine	0.66	1.94E-03	1
1677	HMDB00759	Glycyl-L-Leucine	1.99	2.14E-03	1
1677	HMDB28844	Glycyl-Isoleucine	1.99	2.14E-03	1
1695	HMDB00300	Uracil	48.47	8.47E-04	1
1736	HMDB28691	Alanyl-Leucine	7.62	2.64E-05	1
1799	HMDB28848	Glycyl-Phenylalanine	2.02	2.96E-03	1
1846	HMDB01545	Pyridoxal	0.49	4.22E-05	1
1893	HMDB03581	Dethiobiotin	0.53	3.78E-04	2
1929	HMDB00159	L-Phenylalanine	0.59	1.40E-03	1
1959	HMDB00243	Pyruvic acid	2.20	1.45E-03	2

2034	HMDB02248	Gamma glutamyl ornithine	1.52	4.74E-03	2
2080	HMDB28937	Leucyl-Proline	0.17	1.97E-05	1
2084	HMDB00687	L-leucine	0.46	5.70E-04	1
2091	HMDB00159	L-Phenylalanine	0.57	2.52E-04	1
2133	HMDB03869	Epsilon-(gamma-Glutamyl)-lysine	0.47	2.98E-04	2
2151	HMDB01491	Pyridoxal 5'-phosphate	1.93	2.27E-03	2
2216	HMDB00450	5-Hydroxylysine	2.43	1.68E-04	1
2451	HMDB11162	L-beta-aspartyl-L-alanine	0.58	6.72E-04	2
2463	HMDB00339	2-Methylbutyrylglycine	0.26	8.22E-06	2
2523	HMDB01889	Theophylline	0.24	3.29E-05	1
2569	HMDB12230	Gamma-glutamyl-L-putrescine	12.15	1.90E-03	2
2580	HMDB06045	Dityrosine	0.33	5.02E-05	2
2609	HMDB00656	Cysteineglutathione disulfide	14.81	6.97E-04	2
2647	HMDB01256	Spermine	2.15	1.02E-05	2
2667	HMDB00214	Ornithine	3.42	6.98E-07	1
2733	HMDB03454	4-Pyridoxolactone	0.39	2.28E-06	2
2760	HMDB00955	Isoferulic acid	0.52	3.46E-03	1
2819	HMDB02135	S-(3-oxo-3-carboxy-n-propyl)cysteine	1.51	6.19E-03	2
2827	HMDB02107	Phthalic acid	0.53	1.26E-03	2
2888	HMDB12134	1,2-Dihydroxy-3-keto-5-methylthiopentene	1.83	6.19E-05	2
2956	HMDB00500	4-Hydroxybenzoic acid	0.32	7.04E-06	1
2960	HMDB00701	Hexanoylglycine	0.19	2.87E-07	2
2993	HMDB00512	N-Acetyl-L-phenylalanine	0.21	2.22E-06	2
3038	HMDB03227	Methanethiol	1.72	6.52E-03	2
3041	HMDB01276	N1-Acetylspermidine	1.66	2.58E-03	2
3097	HMDB11686	p-Cresol glucuronide	4.79	4.07E-03	2
3106	HMDB00132	Guanine	4.04	2.70E-03	2
3143	HMDB00177	L-Histidine	0.64	1.28E-03	1
3266	HMDB06524	3-Indoleacetonitrile	0.55	7.32E-03	2
3282	HMDB01526	S-Acetyldihydroliipoamide	7.81	5.62E-05	2
3316	HMDB03320	Indole-3-carboxylic acid	0.50	8.81E-04	1
3362	HMDB00209	Phenylacetic acid	0.62	3.76E-03	2
3528	HMDB29105	Tyrosyl-Glycine	0.50	4.53E-03	1
3622	HMDB02044	8-Hydroxyguanosine	1.65	1.80E-03	2

3643	HMDB12286	S-Prenyl-L-cysteine	0.45	6.02E-03	2
3751	HMDB29098	Tyrosyl-Alanine	2.01	8.69E-03	1
3793	HMDB01414	1,4-diaminobutane	1.80	1.80E-03	1
3953	HMDB02043	5-Phenylvaleric acid	2.13	2.04E-03	2
4027	HMDB02322	Cadaverine	1.91	4.46E-03	1
4095	HMDB11176	L-phenylalanyl-L-hydroxyproline	0.51	4.39E-03	2
4097	HMDB00107	Galactitol	0.40	1.91E-03	2
4101	HMDB00158	L-Tyrosine	0.37	8.20E-04	1
4127	HMDB00375	3-(3-Hydroxyphenyl)propanoic acid	0.06	8.65E-04	2
4299	HMDB05809	Eugenol	2.00	6.63E-03	2
4342	HMDB04586	Perillic acid	1.84	1.94E-03	2
4558	HMDB04072	4-Hydroxystyrene	0.62	2.09E-03	2
4612	HMDB00866	N-Acetyl-L-tyrosine	0.10	1.46E-06	2
4884	HMDB04058	5,6-Dihydroxyindole	0.59	9.77E-04	2
4926	HMDB01387	N-Methylphenylethanolamine	0.28	2.44E-04	2
5010	HMDB03905	Imidazole-4-acetaldehyd	0.61	1.00E-04	2
5024	HMDB11687	Phenylbutyrylglutamine	0.38	2.43E-03	2
5054	HMDB00226	Orotic acid	0.54	9.36E-03	2
5104	HMDB00030	Biotin	0.15	6.12E-06	2
5272	HMDB11718	4-Hydroxybenzaldehyde	0.46	3.83E-03	2
5341	HMDB12219	Dopamine quinone	0.23	4.07E-05	2
5503	HMDB06954	2-Methyl-3-hydroxy-5-formylpyridine-4-carboxylate	0.18	6.35E-04	2
5521	HMDB01430	L-Dopachrome	0.60	5.62E-03	2
5667	HMDB12182	8-Hydroxypurine	0.38	1.35E-04	2
5791	HMDB00472	5-Hydroxy-L-tryptophan	0.12	3.62E-04	2
6264	HMDB06779	Indole-5,6-quinone	0.48	5.05E-03	2
6916	HMDB00175	Inosinic acid	0.49	5.19E-03	2

Table 4.2 Metabolites with significantly different concentrations in MCF-7 cell extracts prepared using different harvest methods (trypsinization/physical-scraping).

No.	HMDB No.	Compound Name	Fold change	p-value	Level
59	HMDB01397	Guanosine monophosphate	2.75	2.32E-04	1
83	HMDB01410	2-Amino-4-oxo-6-(1',2'-dioxopropyl)-7,8-dihydroxypteridine	1.70	6.02E-03	2
84	HMDB00229	Nicotinamide ribotide	2.00	2.59E-04	2
101	HMDB00045	Adenosine monophosphate	49.20	4.15E-07	1
105	HMDB00133	Guanosine	9.73	6.16E-05	1
114	HMDB01397	Guanosine monophosphate	5.35	9.85E-07	2
127	HMDB02022	Glycineamideribotide	1.81	2.26E-03	2
144	HMDB00517	L-Arginine	2.44	9.35E-05	1
146	HMDB00052	Argininosuccinic acid	0.42	6.48E-03	2
206	HMDB00045	Adenosine monophosphate	45.43	3.70E-07	1
264	HMDB00195	Inosine	104.89	4.15E-03	2
288	HMDB11168	L-beta-aspartyl-L-serine	1.68	7.63E-05	2
365	HMDB00912	Succinyladenosine	0.66	9.05E-03	2
380	HMDB00856	N-a-Acetylcitrulline	0.33	3.79E-03	2
383	HMDB02335	Aspartyl-L-proline	0.48	7.16E-03	2
392	HMDB00802	Pterin	0.48	7.82E-03	2
532	HMDB02278	2-(acetylamino)-1,5-anhydro-2-deoxy-3-O-b-D-galactopyranosyl-D-arabino-Hex-1-enitol	3.83	2.01E-03	2
563	HMDB00167	L-Threonine	0.59	3.93E-03	1
563	HMDB00719	L-Homoserine	0.59	3.93E-03	1
568	HMDB05765	Ophthalmic acid	0.57	3.55E-03	2
577	HMDB00288	Uridine 5'-monophosphate	5.45	1.22E-05	2
619	HMDB00149	Ethanolamine	10.37	1.70E-06	1
850	HMDB00174	L-Fucose	0.57	3.31E-03	2
892	HMDB00854	Formiminoglutamic acid	1.65	6.24E-03	2
970	HMDB01263	Allysine	3.73	4.27E-06	2
978	HMDB00296	Uridine	10.71	3.07E-05	1
983	HMDB00056	Beta-Alanine	0.64	2.79E-03	1
983	HMDB00161	L-Alanine	0.64	2.79E-03	1

1034	HMDB03338	Hydroxylamine	1.71	5.83E-03	2
1046	HMDB12201	Cis-zeatin-7-N-glucoside	1.78	4.91E-03	2
1202	HMDB00296_2	Uridine - H2O	14.54	4.01E-06	1
1407	HMDB01080	4-Aminobutyraldehyde	1.93	1.59E-03	2
1575	HMDB28691	Alanyl-Leucine	0.37	2.08E-03	1
1761	HMDB11105	5-Acetylamino-6-formylamino-3-methyluracil	0.50	1.74E-03	2
1846	HMDB01545	Pyridoxal	0.26	1.93E-03	1
2080	HMDB28937	Leucyl-Proline	0.22	1.00E-03	1
2100	HMDB01263	Allysine	0.28	3.10E-03	2
2216	HMDB00450	5-Hydroxylysine	2.77	2.85E-03	1
2314	HMDB00130	Homogentisic acid	0.41	7.87E-03	2
2956	HMDB00500	4-Hydroxybenzoic acid	0.39	6.28E-03	1
3159	HMDB00195	Inosine	125.60	8.41E-04	2
3165	HMDB01904	3-Nitrotyrosine	12.29	2.01E-06	2
3190	HMDB05199	(R)-Salsolinol	0.08	1.95E-03	2
3192	HMDB01488	Nicotinic acid	3.20	1.78E-04	2
3928	HMDB12176	5-Aminopentanamide	1.76	9.89E-03	2
4127	HMDB00375	3-(3-Hydroxyphenyl)propanoic acid	0.16	7.91E-04	2
4187	HMDB00656	Cysteineglutathione disulfide	1.64	9.15E-03	2
4303	HMDB01257	Spermidine	2.22	9.17E-04	2
4382	HMDB03747	Resveratrol	1.91	8.31E-05	2
4409	HMDB00206	N6-Acetyl-L-lysine	0.31	4.53E-04	2
4807	HMDB11150	Deoxyhypusine	1.75	4.60E-03	2
4898	HMDB01084	D-1-Piperidine-2-carboxylic acid	3.92	2.25E-05	2
4915	HMDB00555	3-Methyladipic acid	0.46	8.04E-03	2
4926	HMDB03633	N-Methyltyramine	0.57	3.07E-03	2
5225	HMDB02338	Biochanin A	0.16	5.86E-05	2
5243	HMDB00132	Guanine	6.47	9.37E-03	2
5320	HMDB00299	Xanthosine	0.20	4.82E-03	2
5374	HMDB05199	(R)-Salsolinol	5.24	1.05E-03	2
5503	HMDB06954	2-Methyl-3-hydroxy-5-formylpyridine-4-carboxylate	0.30	2.16E-04	2
5719	HMDB04089	Formylanthranilic acid	0.44	6.59E-03	2
5733	HMDB01314	Cyclic GMP	0.54	4.53E-03	2
5791	HMDB00472	5-Hydroxy-L-tryptophan	0.34	3.63E-03	2

5911	HMDB02393	N-Methyl-D-aspartic acid	0.33	7.17E-04	2
6442	HMDB00252	Sphingosine	0.35	5.95E-03	2
6571	HMDB00269	Sphinganine	0.30	1.16E-04	2

Table 4.3 List of metabolites from Hela cell lysates positively identified by searching against the DnsCl-labeled standard library.

Peak Pair Information						Identification Result	
Peak Pair #	T _R (min)	Corrected T _R (min)	mz_light	mz_heavy	monoisotopic mass (Da)	HMDB.No.	Name
68	2.16	2.12	403.1438	405.1505	169.0854	HMDB00001	1-Methylhistidine
68	2.16	2.12	403.1438	405.1505	169.0854	HMDB00479	3-methyl-histidine
101	2.29	2.24	581.1216	583.1276	347.0632	HMDB00045	Adenosine monophosphate
105	2.30	2.26	517.1504	519.1570	283.0921	HMDB00133	Guanosine
130	2.40	2.35	375.0777	377.0843	141.0193	HMDB00224	O-Phosphoethanolamine
144	2.42	2.37	408.1702	410.1766	174.1118	HMDB00517	L-Arginine
153	2.47	2.42	510.1910	512.1973	276.1327	HMDB00279	Saccharopine
154	2.47	2.43	388.1077	390.1137	154.0494	HMDB00157	Hypoxanthine + H ₂ O
213	2.75	2.68	422.1862	424.1925	188.1279	HMDB00670	Homo-L-arginine
256	2.97	2.89	557.1126	559.1211	323.0543	HMDB00095_2	Cytidine monophosphate - Isomer
275	3.09	3.01	436.2016	438.2082	202.1433	HMDB03334	Symmetric dimethylarginine
313	3.28	3.19	366.1118	368.1184	132.0535	HMDB00168	L-Asparagine
364	3.62	3.51	380.1276	382.1342	146.0693	HMDB00641	L-Glutamine
364	3.62	3.51	380.1276	382.1342	146.0693	HMDB03423	D-Glutamine
371	3.79	3.67	409.1545	411.1611	175.0961	HMDB00904	Citrulline
381	3.82	3.70	510.1555	512.1616	276.0971	HMDB11737	Gamma Glutamylglutamic

							acid
421	3.94	3.82	501.1551	503.1611	267.0968	HMDB00050	Adenosine
443	4.07	3.93	399.1050	401.1111	165.0466	HMDB02005	Methionine Sulfoxide
443	4.07	3.93	399.1050	401.1111	165.0466	HMDB02005_2	Methionine Sulfoxide - Isomer
494	4.25	4.11	353.1167	355.1234	119.0584	HMDB00719	L-Homoserine
499	4.29	4.15	339.1008	341.1078	105.0425	HMDB00187	L-Serine
574	4.92	5.06	381.1114	383.1184	147.0531	HMDB00148	L-Glutamic Acid
579	5.01	5.11	365.1160	367.1222	131.0577	HMDB00725	Trans-4-Hydroxyl-L-Proline
581	5.11	5.18	367.0959	369.1027	133.0376	HMDB00191	L-Aspartic Acid
591	5.32	5.51	422.1744	424.1816	188.1161	HMDB00206	N6-Acetyl-L-Lysine
611	5.40	5.64	492.1806	494.1870	258.1223	HMDB00279_2	Saccharopine - H2O
616	5.49	5.78	353.1167	355.1235	119.0584	HMDB00167	L-Threonine
618	5.52	5.82	395.1274	397.1339	161.0691	HMDB00510	Aminoadipic acid
619	5.63	5.94	295.1111	297.1177	61.0528	HMDB00149	Ethanolamine
622	5.65	5.97	339.1375	341.1439	105.0791	HMDB04437	Diethanolamine
740	6.19	6.58	309.0912	311.0976	75.0329	HMDB00123	Glycine
812	6.49	6.96	364.1693	366.1758	130.1109	HMDB02064	N-Acetylputrescine
854	6.60	7.09	406.1435	408.1501	172.0852	HMDB00721	Glycylproline
862	6.63	7.13	323.1060	325.1125	89.0476	HMDB00161	L-Alanine
978	6.99	7.57	478.1282	480.1348	244.0699	HMDB00296	Uridine
984	7.04	7.64	323.1067	325.1127	89.0484	HMDB00056	Beta-Alanine
987	7.08	7.67	337.1219	339.1285	103.0636	HMDB00112	Gamma-Aminobutyric acid
1143	7.52	8.16	453.1690	455.1754	219.1107	HMDB00210	Pantothenic acid
1178	7.64	8.29	492.1444	494.1498	258.0861	HMDB00884_2	Ribothymidine - Isomer
1202	7.73	8.38	460.1178	462.1243	226.0595	HMDB00296_2	Uridine - H2O
1218	7.79	8.45	370.0973	372.1040	136.0390	HMDB00157_2	Hypoxanthine - multi-tags
1268	7.99	8.67	386.0922	388.0989	152.0339	HMDB00292	Xanthine
1272	8.01	8.69	337.1220	339.1286	103.0637	HMDB00452	L-Alpha-aminobutyric acid

1272	8.01	8.69	337.1220	339.1286	103.0637	HMDB03911	3-Aminoisobutanoic acid
1273	8.02	8.71	351.1375	353.1440	117.0791	HMDB03355	5-Aminopentanoic acid
1329	8.26	8.97	408.1587	410.1656	174.1004	HMDB28854	Glycyl-Valine
1340	8.32	9.03	376.0964	378.1030	142.0380	HMDB00469	5-Hydroxymethyluracil
1354	8.39	9.10	337.1217	339.1286	103.0634	HMDB00650	D-Alpha-aminobutyric acid
1354	8.39	9.10	337.1217	339.1286	103.0634	HMDB01906	2-Aminoisobutyric acid
1401	8.56	9.29	335.1061	337.1154	101.0478	HMDB00719_2	L-Homoserine - H2O
1426	8.68	9.42	370.0969	372.1035	136.0386	HMDB00157_3	Hypoxanthine - Isomer
1429	8.77	9.52	363.1011	365.1076	129.0427	HMDB00148_2	L-Glutamic Acid - H2O
1537	9.38	10.19	349.1218	351.1287	115.0635	HMDB00162	L-Proline
1607	9.75	10.56	470.1753	472.1806	236.1170	HMDB28988	Phenylalanyl-Alanine
1660	9.93	10.74	426.1197	428.1263	384.1227	HMDB00939	S-Adenosylhomocysteine
1661	10.03	10.84	351.1374	353.1442	117.0791	HMDB00883	L-Valine
1662	10.08	10.90	383.1096	385.1162	149.0512	HMDB00696	L-Methionine
1665	10.12	10.94	365.1532	367.1601	131.0949	HMDB03640	Beta-Leucine
1677	10.22	11.04	422.1747	424.1814	188.1164	HMDB00759	Glycyl-L-Leucine
1677	10.22	11.04	422.1747	424.1814	188.1164	HMDB28844	Glycyl-Isoleucine
1695	10.30	11.12	346.0859	348.0926	112.0276	HMDB00300	Uracil
1715	10.38	11.20	438.1494	440.1554	204.0910	HMDB00929	L-Tryptophan
1736	10.46	11.29	436.1903	438.1970	202.1320	HMDB28691	Alanyl-Leucine
1799	10.85	11.69	456.1582	458.1665	222.0999	HMDB28848	Glycyl-Phenylalanine
1846	11.01	11.85	401.1172	403.1239	167.0589	HMDB01545	Pyridoxal
1879	11.09	11.94	373.0854	375.0923	139.0270	HMDB02658	6-Hydroxynicotinic acid
2026	11.57	12.43	399.1377	401.1445	165.0794	HMDB00159	L-Phenylalanine
2080	11.74	12.61	462.2064	464.2129	228.1481	HMDB28937	Leucyl-Proline

2084	11.76	12.62	365.1533	367.1604	131.0950	HMDB00172	L-Isoleucine
2094	12.10	13.06	365.1532	367.1601	131.0948	HMDB00687	L-leucine
2102	12.36	13.33	365.1533	367.1602	131.0950	HMDB00557	L-Alloisoleucine
2216	13.05	14.05	315.1094	317.1160	162.1022	HMDB00450	5-Hydroxylysine
2282	13.29	14.31	416.1168	418.1234	182.0585	HMDB00755	Hydroxyphenyllactici acid
2401	14.15	15.22	425.1205	427.1254	191.0622	HMDB00763	5- Hydroxyindoleacetic acid
2523	14.63	15.72	414.1245	416.1315	180.0662	HMDB01889	Theophylline
2608	14.97	16.07	319.1113	321.1180	85.0530	HMDB03911_ 2	3-Aminoisobutanoic acid - H2O
2667	15.19	16.30	300.1039	302.1104	132.0911	HMDB00214	Ornithine
2787	15.54	16.67	386.1062	388.1128	152.0479	HMDB00020	p- Hydroxyphenylacetic acid
2787	15.54	16.67	386.1062	388.1128	152.0479	HMDB00440	3- Hydroxyphenylacetic acid
2787	15.54	16.67	386.1062	388.1128	152.0479	HMDB00669	Ortho- Hydroxyphenylacetic acid
2787	15.54	16.67	386.1062	388.1128	152.0479	HMDB02390	3-Cresotinic acid
2904	16.01	17.16	327.1163	329.1230	93.0580	HMDB03012	Aniline
2913	16.04	17.20	402.1004	404.1064	168.0421	HMDB00484	Vanillic acid
2956	16.24	17.40	372.0904	374.0969	138.0320	HMDB00500	4-Hydroxybenzoic acid
2959	16.31	17.48	307.1112	309.1179	146.1057	HMDB00182	L-Lysine
3143	17.18	18.34	389.1281	391.1347	155.0698	HMDB00177	L-Histidine
3316	17.81	18.98	395.1067	397.1131	161.0484	HMDB03320	Indole-3-carboxylic acid
3528	18.73	19.89	353.1068	355.1135	238.0970	HMDB29105	Tyrosyl-Glycine
3751	19.56	20.73	360.1140	362.1211	252.1113	HMDB29098	Tyrosyl-Alanine
3793	19.71	20.88	278.1085	280.1152	88.1004	HMDB01414	1,4-diaminobutane
4027	21.00	22.17	285.1164	287.1231	102.1163	HMDB02322	Cadaverine
4101	21.48	22.64	324.5955	326.6023	181.0744	HMDB00158	L-Tyrosine

4101	21.48	22.64	324.5955	326.6023	181.0744	HMDB06050	o-Tyrosine
4112	21.65	22.81	374.1302	376.1368	280.1437	HMDB29118	Tyrosyl-Valine
4170	21.92	23.08	328.1011	330.1073	94.0428	HMDB00228	Phenol
4221	22.17	23.33	373.0859	375.0925	139.0276	HMDB01232	4-Nitrophenol
4331	22.64	23.80	381.1380	383.1440	294.1593	HMDB29109	Tyrosyl-Leucine
4525	23.30	24.46	342.1164	344.1230	108.0581	HMDB01858	p-Cresol
4525	23.30	24.46	342.1164	344.1230	108.0581	HMDB02048	m-Cresol
4525	23.30	24.46	342.1164	344.1230	108.0581	HMDB02055	o-Cresol
4601	23.56	24.72	322.1044	324.1115	176.0922	HMDB00259	Serotonin
4846	24.27	25.43	356.1316	358.1383	122.0733	HMDB29306	4-Ethylphenol
5026	24.61	25.77	302.6005	304.6074	137.0844	HMDB00306	Tyramine
5287	25.55	26.71	289.0790	291.0837	110.0414	HMDB00957	pyrocatechol

Table 4.4 List of metabolites from MCF-7 cell lysates positively identified by searching against the DnsCl-labeled standard library

Peak Pair Information						Identification Result	
Peak Pair #	T _R (min)	Corrected T _R (min)	mz _{light}	mz _{heavy}	monoisotopic mass (Da)	HMDB.No.	Name
68	2.16	2.12	403.1438	405.1505	169.0854	HMDB00001	1-Methylhistidine
						HMDB00479	3-methyl-histidine
101	2.29	2.24	581.1216	583.1276	347.0632	HMDB00045	Adenosine monophosphate
105	2.30	2.26	517.1504	519.1570	283.0921	HMDB00133	Guanosine
130	2.40	2.35	375.0777	377.0843	141.0193	HMDB00224	O-Phosphoethanolamine
144	2.42	2.37	408.1702	410.1766	174.1118	HMDB00517	L-Arginine
153	2.47	2.42	510.1910	512.1973	276.1327	HMDB00279	Saccharopine
154	2.47	2.43	388.1077	390.1137	154.0494	HMDB00157	Hypoxanthine + H ₂ O
213	2.75	2.68	422.1862	424.1925	188.1279	HMDB00670	Homo-L-arginine
256	2.97	2.89	557.1126	559.1211	323.0543	HMDB00095_2	Cytidine monophosphate - Isomer
275	3.09	3.01	436.2016	438.2082	202.1433	HMDB03334	Symmetric dimethylarginine
313	3.28	3.19	366.1118	368.1184	132.0535	HMDB00168	L-Asparagine
364	3.62	3.51	380.1276	382.1342	146.0693	HMDB00641	L-Glutamine
						HMDB03423	D-Glutamine
371	3.79	3.67	409.1545	411.1611	175.0961	HMDB00904	Citrulline
381	3.82	3.70	510.1555	512.1616	276.0971	HMDB11737	Gamma Glutamylglutamic acid

421	3.94	3.82	501.155 1	503.1611	267.0968	HMDB00050	Adenosine
443	4.07	3.93	399.105 0	401.1111	165.0466	HMDB02005	Methionine Sulfoxide
443	4.07	3.93	399.105 0	401.1111	165.0466	HMDB02005_ 2	Methionine Sulfoxide - Isomer
494	4.25	4.11	353.116 7	355.1234	119.0584	HMDB00719	L-Homoserine
499	4.29	4.15	339.100 8	341.1078	105.0425	HMDB00187	L-Serine
574	4.92	5.06	381.111 4	383.1184	147.0531	HMDB00148	L-Glutamic Acid
579	5.01	5.11	365.116 0	367.1222	131.0577	HMDB00725	Trans-4-Hydroxyl-L- Proline
581	5.11	5.18	367.095 9	369.1027	133.0376	HMDB00191	L-Aspartic Acid
591	5.32	5.51	422.174 4	424.1816	188.1161	HMDB00206	N6-Acetyl-L-Lysine
611	5.40	5.64	492.180 6	494.1870	258.1223	HMDB00279_ 2	Saccharopine - H2O
616	5.49	5.78	353.116 7	355.1235	119.0584	HMDB00167	L-Threonine
618	5.52	5.82	395.127 4	397.1339	161.0691	HMDB00510	Amino adipic acid
619	5.63	5.94	295.111 1	297.1177	61.0528	HMDB00149	Ethanolamine
622	5.65	5.97	339.137 5	341.1439	105.0791	HMDB04437	Diethanolamine
742	6.21	6.60	309.090 6	311.0970	75.0322	HMDB00123	Glycine
812	6.49	6.96	364.169 3	366.1758	130.1109	HMDB02064	N-Acetylputrescine
854	6.60	7.09	406.143 5	408.1501	172.0852	HMDB00721	Glycylproline
862	6.63	7.13	323.106 0	325.1125	89.0476	HMDB00161	L-Alanine

978	6.99	7.57	478.128 2	480.1348	244.0699	HMDB00296	Uridine
984	7.04	7.64	323.106 7	325.1127	89.0484	HMDB00056	Beta-Alanine
987	7.08	7.67	337.121 9	339.1285	103.0636	HMDB00112	Gamma-Aminobutyric acid
114 3	7.52	8.16	453.169 0	455.1754	219.1107	HMDB00210	Pantothenic acid
117 8	7.64	8.29	492.144 4	494.1498	258.0861	HMDB00884_ 2	Ribothymidine - Isomer
120 2	7.73	8.38	460.117 8	462.1243	226.0595	HMDB00296_ 2	Uridine - H2O
121 8	7.79	8.45	370.097 3	372.1040	136.0390	HMDB00157_ 2	Hypoxanthine - multi- tags
127 2	8.01	8.69	337.122 0	339.1286	103.0637	HMDB00452	L-Alpha-aminobutyric acid
127 2	8.01	8.69	337.122 0	339.1286	103.0637	HMDB03911	3-Aminoisobutanoic acid
127 3	8.02	8.71	351.137 5	353.1440	117.0791	HMDB03355	5-Aminopentanoic acid
132 9	8.26	8.97	408.158 7	410.1656	174.1004	HMDB28854	Glycyl-Valine
134 0	8.32	9.03	376.096 4	378.1030	142.0380	HMDB00469	5- Hydroxymethyluracil
135 4	8.39	9.10	337.121 7	339.1286	103.0634	HMDB00650	D-Alpha-aminobutyric acid
135 4	8.39	9.10	337.121 7	339.1286	103.0634	HMDB01906	2-Aminoisobutyric acid
142 6	8.68	9.42	370.096 9	372.1035	136.0386	HMDB00157_ 3	Hypoxanthine - Isomer
147 4	9.08	9.86	349.121 9	351.1294	115.0636	HMDB00162	L-Proline
160 7	9.75	10.56	470.175 3	472.1806	236.1170	HMDB28988	Phenylalanyl-Alanine
165 5	9.91	10.72	365.152 9	367.1600	131.0945	HMDB03640	Beta-Leucine

1660	9.93	10.74	426.1197	428.1263	384.1227	HMDB00939	S-Adenosylhomocysteine
1661	10.03	10.84	351.1374	353.1442	117.0791	HMDB00883	L-Valine
1662	10.08	10.90	383.1096	385.1162	149.0512	HMDB00696	L-Methionine
1677	10.22	11.04	422.1747	424.1814	188.1164	HMDB00759	Glycyl-L-Leucine
1677	10.22	11.04	422.1747	424.1814	188.1164	HMDB28844	Glycyl-Isoleucine
1695	10.30	11.12	346.0859	348.0926	112.0276	HMDB00300	Uracil
1736	10.46	11.29	436.1903	438.1970	202.1320	HMDB28691	Alanyl-Leucine
1765	10.61	11.44	438.1486	440.1552	204.0903	HMDB00929	L-Tryptophan
1799	10.85	11.69	456.1582	458.1665	222.0999	HMDB28848	Glycyl-Phenylalanine
1846	11.01	11.85	401.1172	403.1239	167.0589	HMDB01545	Pyridoxal
1879	11.09	11.94	373.0854	375.0923	139.0270	HMDB02658	6-Hydroxynicotinic acid
2026	11.57	12.43	399.1377	401.1445	165.0794	HMDB00159	L-Phenylalanine
2080	11.74	12.61	462.2064	464.2129	228.1481	HMDB28937	Leucyl-Proline
2084	11.76	12.62	365.1533	367.1604	131.0950	HMDB00172	L-Isoleucine
2094	12.10	13.06	365.1532	367.1601	131.0948	HMDB00557	L-Alloisoleucine
2094	12.10	13.06	365.1532	367.1601	131.0948	HMDB00687	L-leucine
2216	13.05	14.05	315.1094	317.1160	162.1022	HMDB00450	5-Hydroxylysine
2282	13.29	14.31	416.1168	418.1234	182.0585	HMDB00755	Hydroxyphenyllactic acid

247 5	14.47	15.55	425.118 8	427.1238	191.0605	HMDB00763	5-Hydroxyindoleacetic acid
252 3	14.63	15.72	414.124 5	416.1315	180.0662	HMDB01889	Theophylline
260 8	14.97	16.07	319.111 3	321.1180	85.0530	HMDB03911_ 2	3-Aminoisobutanoic acid - H2O
266 7	15.19	16.30	300.103 9	302.1104	132.0911	HMDB00214	Ornithine
278 7	15.54	16.67	386.106 2	388.1128	152.0479	HMDB00020	p-Hydroxyphenylacetic acid
278 7	15.54	16.67	386.106 2	388.1128	152.0479	HMDB00440	3-Hydroxyphenylacetic acid
278 7	15.54	16.67	386.106 2	388.1128	152.0479	HMDB00669	Ortho-Hydroxyphenylacetic acid
278 7	15.54	16.67	386.106 2	388.1128	152.0479	HMDB02390	3-Cresotinic acid
290 4	16.01	17.16	327.116 3	329.1230	93.0580	HMDB03012	Aniline
291 3	16.04	17.20	402.100 4	404.1064	168.0421	HMDB00484	Vanillic acid
295 6	16.24	17.40	372.090 4	374.0969	138.0320	HMDB00500	4-Hydroxybenzoic acid
295 9	16.31	17.48	307.111 2	309.1179	146.1057	HMDB00182	L-Lysine
314 3	17.18	18.34	389.128 1	391.1347	155.0698	HMDB00177	L-Histidine
331 6	17.81	18.98	395.106 7	397.1131	161.0484	HMDB03320	Indole-3-carboxylic acid
352 8	18.73	19.89	353.106 8	355.1135	238.0970	HMDB29105	Tyrosyl-Glycine
375 1	19.56	20.73	360.114 0	362.1211	252.1113	HMDB29098	Tyrosyl-Alanine
379 3	19.71	20.88	278.108 5	280.1152	88.1004	HMDB01414	1,4-diaminobutane

4020	20.97	22.14	324.5954	326.6024	181.0742	HMDB06050	o-Tyrosine
4027	21.00	22.17	285.1164	287.1231	102.1163	HMDB02322	Cadaverine
4101	21.48	22.64	324.5955	326.6023	181.0744	HMDB00158	L-Tyrosine
4112	21.65	22.81	374.1302	376.1368	280.1437	HMDB29118	Tyrosyl-Valine
4170	21.92	23.08	328.1011	330.1073	94.0428	HMDB00228	Phenol
4221	22.17	23.33	373.0859	375.0925	139.0276	HMDB01232	4-Nitrophenol
4331	22.64	23.80	381.1380	383.1440	294.1593	HMDB29109	Tyrosyl-Leucine
4525	23.30	24.46	342.1164	344.1230	108.0581	HMDB01858	p-Cresol
4525	23.30	24.46	342.1164	344.1230	108.0581	HMDB02048	m-Cresol
4525	23.30	24.46	342.1164	344.1230	108.0581	HMDB02055	o-Cresol
5026	24.61	25.77	302.6005	304.6074	137.0844	HMDB00306	Tyramine
5287	25.55	26.71	289.0790	291.0837	110.0414	HMDB00957	pyrocatechol

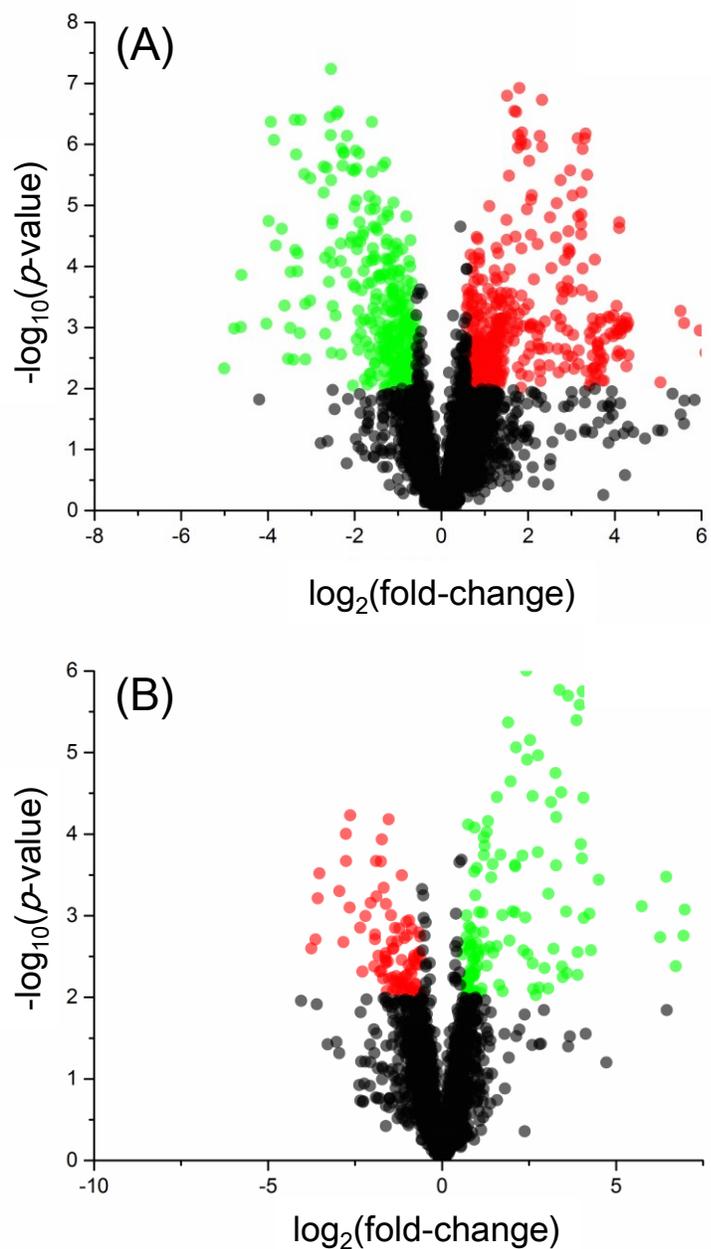


Figure 4.5 Volcano plots of the amine/phenol submetabolomes of (A) HeLa and (B) MCF-7 cells harvested by different methods. The p-value of each metabolite was calculated from t-test, and the fold change (FC) was calculated from the peak ratios of trypsinization group divided by

the peak ratios of the scraping group (i.e., trypsinization/scraping). Using a cut-off value of $p < 0.01$ and $FC > 1.5$ or < 0.67 , the red points represent the metabolites with higher concentrations in the trypsinization group, and the green points represent the metabolites with lower concentrations in the trypsinization group. The black points represent the metabolites with no significant differences.

To illustrate the concentration differences of individual metabolites, we selected some of the metabolites commonly detected in both HeLa and MCF-7 cells that were also positively identified with very large concentration differences between the trypsinization group and the scraping group to produce heap maps (Figure 4.6). As Figure 4.6 shows, there are a variety of metabolites with different structures having significant differences in concentration in the two groups of samples. The observed concentration differences could be attributed to the residual enzyme activity or metabolism in cells harvested by trypsinization. In the scraping harvest method, the cellular metabolism should stop immediately after MeOH was added in. In contrast, cell metabolism could still take place during trypsinization till three-time cell washing procedure was finished (about 30 min). During the trypsinization and washing process, cells would have quick and multiple changes in enzyme levels and metabolic activities, resulting in changes in concentration of some metabolites ²¹⁹. Our results are consistent with those of a previous study which concluded that trypsinization is a more suitable technique for sub-culturing the cells, but not for metabolomics study, as they observed some metabolites related with oxidative stress changed significantly by the trypsinization process ²⁷.

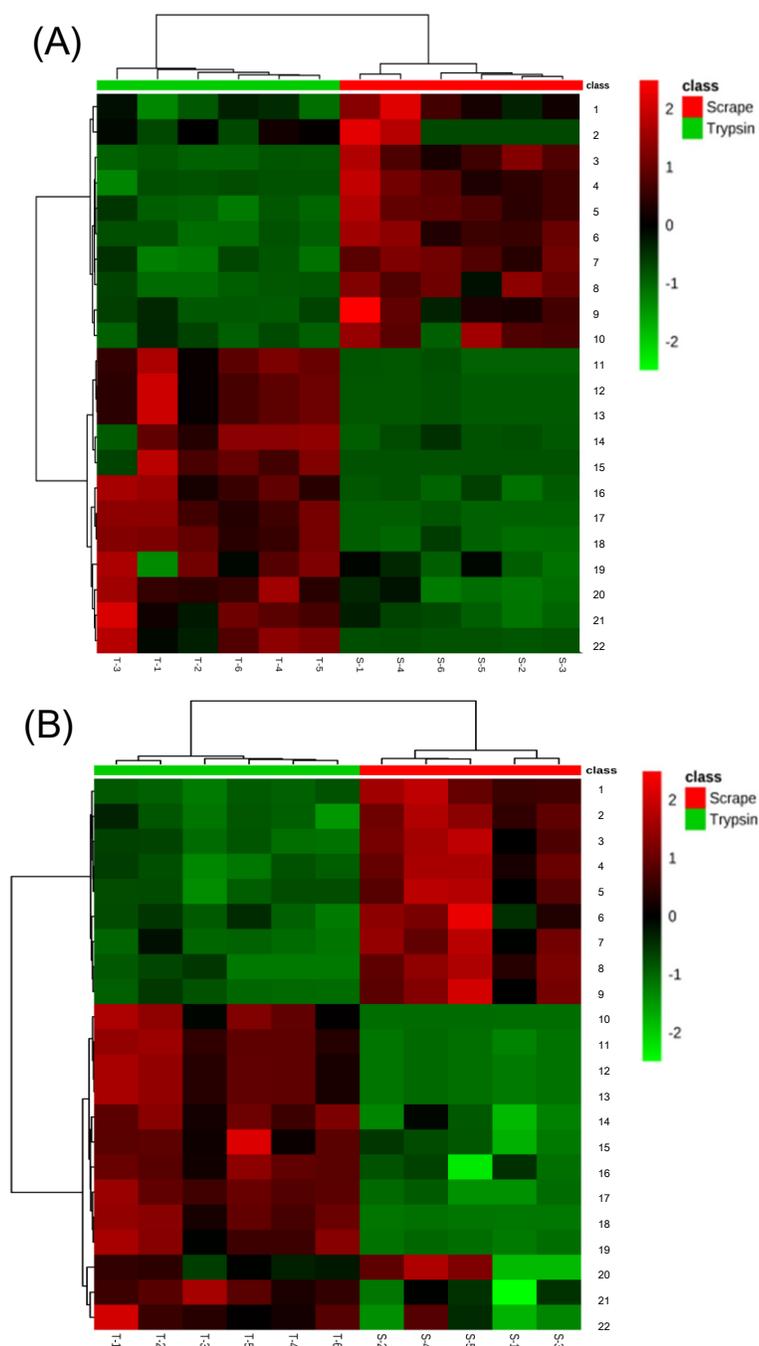


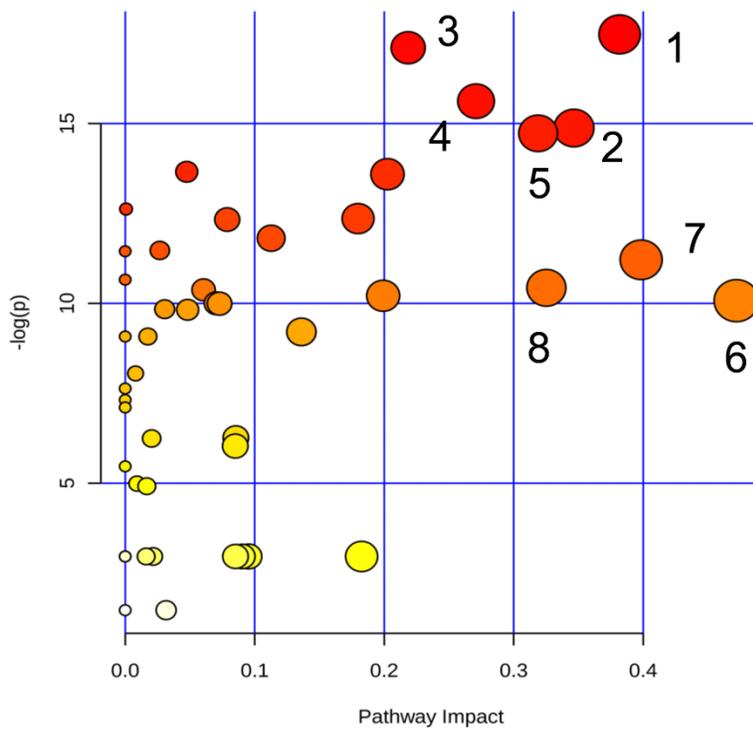
Figure 4.6 Heat maps showing 22 selected metabolites with significant concentration differences in cell extracts prepared using different harvest methods (scraping and trypsinization) from (A) HeLa and (B) MCF-7 cells. The metabolites in (A) are 1. Argininosuccinic acid; 2. N-

Methylphenylethanolamine; 3. Homovanillin; 4. Pyridoxal; 5. Leucyl-Proline; 6. 4-Hydroxybenzoic acid; 7. Allysine; 8. 5-Hydroxy-L-tryptophan; 9. Homogentisic acid; 10. 2-Methyl-3-hydroxy-5-formylpyridine-4-carboxylate; 11. Uridine; 12. Uridine-H₂O; 13. 3-Nitrotyrosine; 14. Guanosine; 15. Inosine; 16. 4-Aminobutyraldehyde; 17. Adenosine monophosphate; 18. Uridine 5'-monophosphate; 19. Guanosine monophosphate; 20. 5-Hydroxylysine; 21. Arginine; 22. Cysteineglutathion. The metabolites in (B) are 1. Homovanillin; 2. Homogentisic acid; 3. 4-Hydroxybenzoic acid; 4. Pyridoxal; 5. 2-Methyl-3-hydroxy-5-formylpyridine-4-carboxylate; 6. 5-Hydroxy-L-tryptophan; 7. Argininosuccinic acid; 8. Leucyl-Proline; 9. Allysine; 10. Inosine; 11. Uridine; 12. Uridine-H₂O; 13. 3-Nitrotyrosine; 14. L-Arginine; 15. 5-Hydroxylysine; 16. Guanosine monophosphate; 17. Uridine 5'-monophosphate; 18. Adenosine monophosphate; 19. Guanosine; 20. N-Methylphenylethanolamine; 21. 4-Aminobutyraldehyde; 22. Cysteineglutathion.

To examine the impact of trypsinization on metabolic pathways in our study, we uploaded the identified/matched metabolites onto the Pathway Enrichment Analysis tool in Metaboanalyst. Figure 4.7 shows the enrichment analysis result. The x-axis represents the pathway impact, and the y-axis represents the negative logarithm of p-value. Figure 4.7 shows several amino-acid and purine related metabolic pathways were significantly affected by trypsinization. One interesting pathway affected was the glutathione pathway. This finding is not surprising, as glutathione is an important antioxidant in cells. The arginine and proline pathway have the least p-values and the most impact in enrichment analysis. As an example, we mapped the detected metabolites into this pathway, and the results are shown in Figure 4.7. In this pathway, the levels of some upstream metabolites such as glutamine, citrulline and argininosuccinate were decreased, while

the downstream metabolite such as arginine was increased after trypsinization. All the polyamines, including spermine, spermidine and putrescine, were increased.

Taken together, the above results show that there were metabolite level differences observed in cell samples prepared using trypsinization and scraping methods. The differences were likely caused by the trypsinization process where cell metabolism was not immediately stopped.



Metabolic Pathway

1	Arginine and proline metabolism
2	beta-Alanine metabolism
3	Sulfur metabolism
4	Purine metabolism
5	Vitamin B6 metabolism
6	Alanine, aspartate and glutamate metabolism
7	Biotin metabolism
8	Glutathione metabolism

Figure 4.7 Metabolic pathways enrichment analysis. The x-axis represents the impact of pathway, and y-axis represents the p-value.

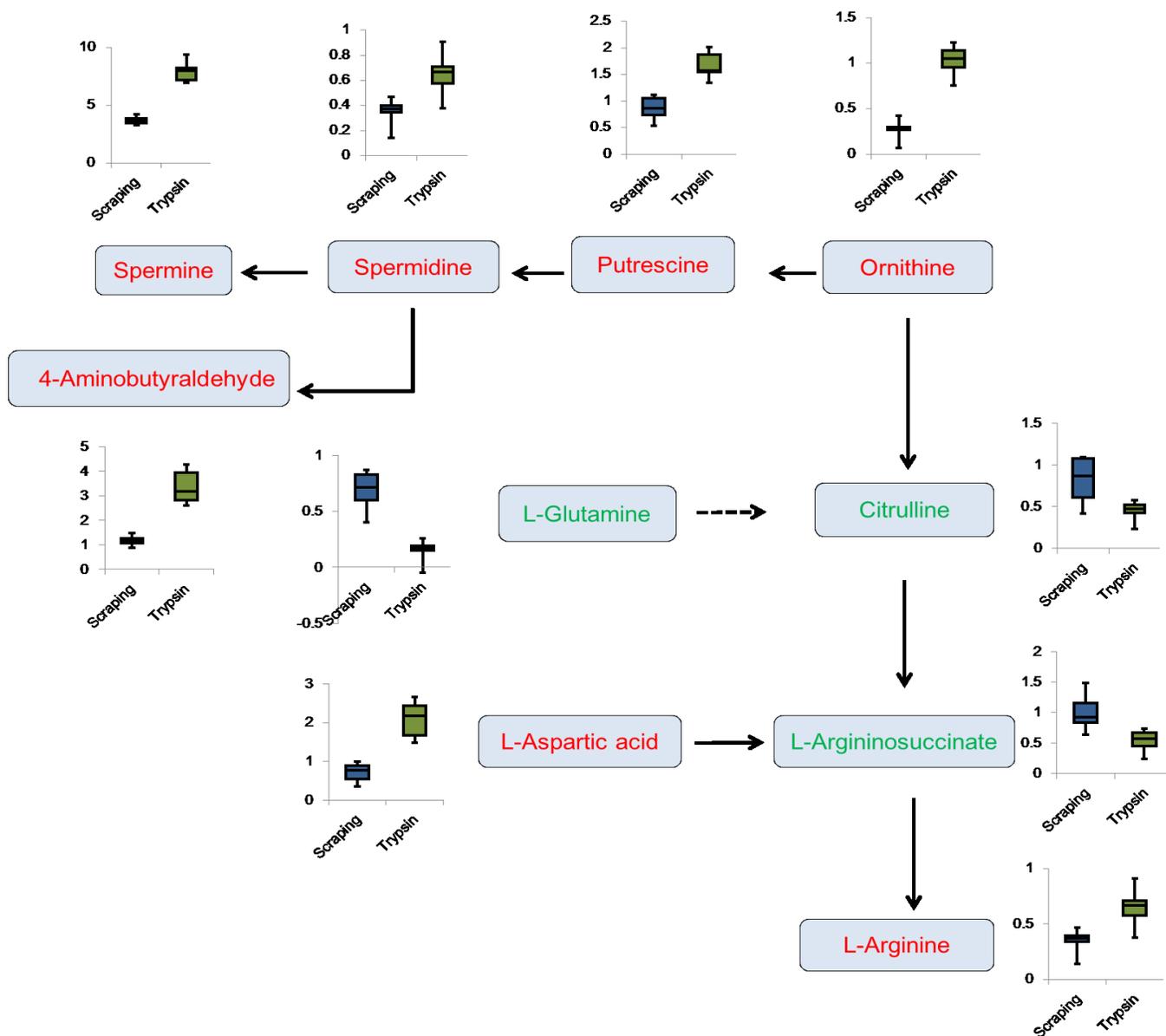


Figure 4.8 Metabolite changes in selected metabolic pathways. The box plots show the relative metabolite abundances in different harvesting groups.

4.3.5 Impact of different lysis methods on cellular metabolome

The PCA and PLS-DA analyses shown in Figure 4.4 indicate that the cellular metabolomes of samples prepared by FT and GB lysis methods do not differ as significantly as those from the two harvest methods. We used the volcano plots to further examine the impact of lysis methods on the cellular metabolomes of HeLa and MCF-7 cells (Figure 4.9). For the HeLa cells (Figure 4.9A), there are only 70 metabolites with significantly higher fold changes and 77 metabolites with significantly lower fold changes found in the two lysis methods with scraping for cell harvest, compared to 429 metabolites with higher fold changes and 305 metabolites with lower fold changes found in the two harvest methods. Similarly, for the MCF-7 cells (Figure 4.9B), only 85 metabolites with higher fold changes and 37 metabolites with lower fold changes were detected with scraping harvest. This observation of smaller impact by the lysis method is not surprising, considering that the cellular metabolism had been already quenched in the cell harvest step and the cellular metabolite levels should not change without active enzymes. Some differences in metabolite levels were observed in the samples prepared by the GB and FT methods. These differences could be attributed to the variations in lysis efficiencies and the extent of metabolite loss in these two methods. These results suggest that using the same method for cell lysis is important for comparing the metabolomes of different groups of cells.

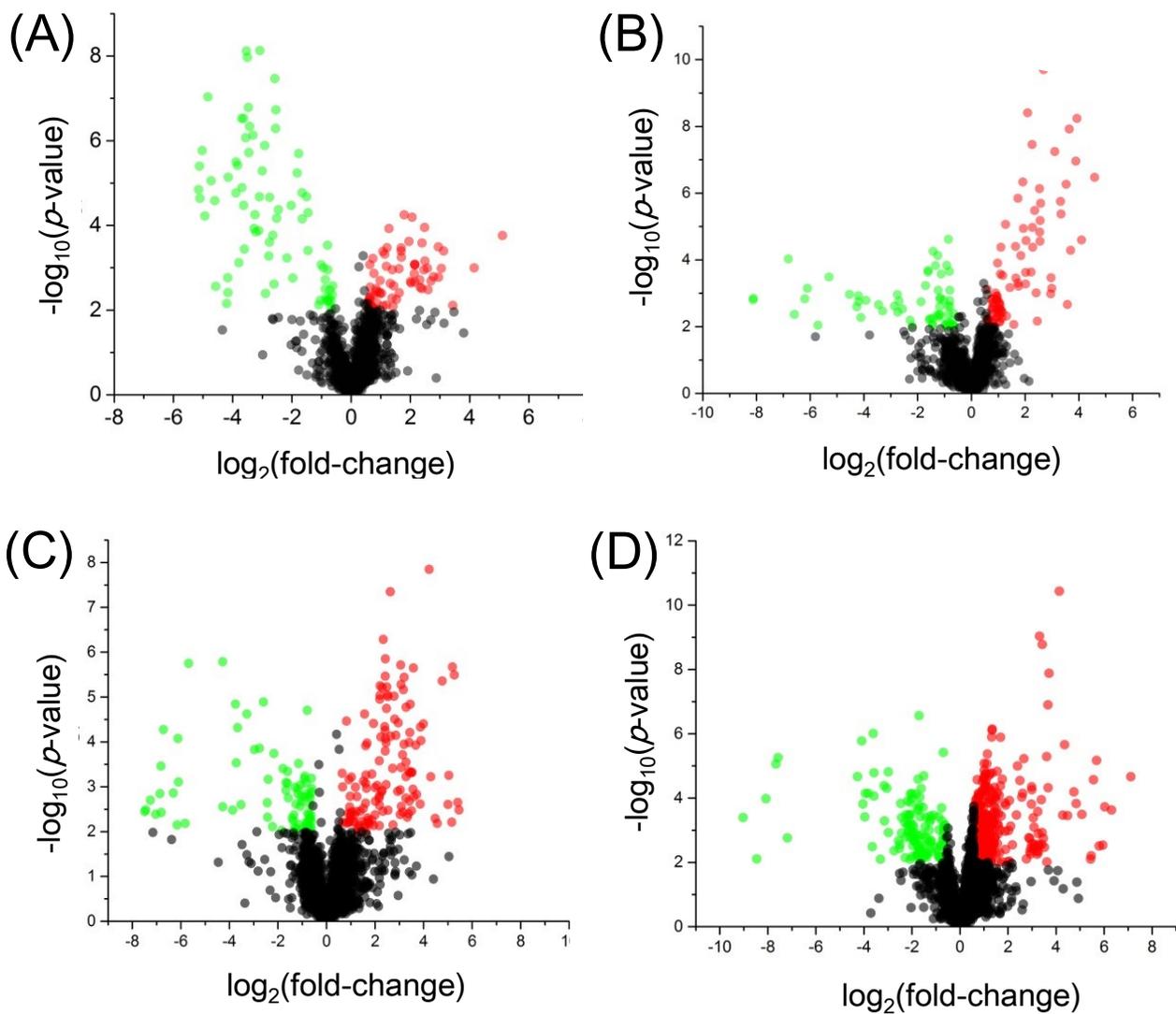


Figure 4.9 Volcano plots for comparison of the amine/phenol submetabolomes of cell extracted prepared using different lysis methods: (A) HeLa and (B) MCF-7 cells harvested by scrapping; (C) HeLa and (D) MCF-7 cells harvested by trypsinization. The p-value was from t-test, and the fold change was calculated from glass-bead/freeze-thaw-cycle. The red points represent the metabolites with higher concentrations in glass-bead lysed samples, and the green points

represent the metabolites with lower concentrations in glass-bead lysed samples. The black points represent the metabolites with no significant differences in the two lysis methods.

We detected more significantly changed metabolites between the FT and GB lysis methods from the cells with trypsinization harvest. For the HeLa cells (Figure 4.9C), 185 metabolites with higher concentrations and 81 metabolites with lower concentrations metabolites were found in the two lysis methods. For the MCF-7 cells (Figure 4.9D), 341 metabolites with higher concentrations and 134 metabolites with lower concentrations were detected. This larger difference may be caused by the cell membrane damage or metabolite leak during the trypsinization process. However, even with a larger number of significantly changes metabolites, the cells harvested by trypsinization treated with different lysis methods still could not be separated on the PCA and PLS-DA plots, as it was shown in Figure 4.4 and discussed in previous section. Thus, the impact of cell lysis methods was relatively small.

In choosing the lysis method for cellular metabolomics, both GB lysis and FT lysis use physical disruption to lyse the cells with no chemical or surfactant added and thus are compatible with the downstream sample processing and analysis in CIL LC-MS. However, based on the LC-UV quantification results, FT lysis gave higher lysis efficiency. In addition, freeze-thaw-cycle is easy to perform, although liquid nitrogen is required for fast processing. We conclude that, if liquid nitrogen is readily available, the FT lysis is preferred for lysis of adherent mammalian cells. If liquid nitrogen is not available, the GB lysis method can be used. It should be noted that for some bacteria cells and yeast cells that have tough cell walls, more aggressive lysis such as ultrasonication lysis⁷ or glass-bead-assisted lysis⁶ should be applied.

4.4 Conclusions

We have examined two cell harvest methods (trypsinization and scraping) and two cell lysis methods (freeze-thaw-cycle and glass-bead-assisted) to evaluate the effects of their combinations on cellular metabolome results. Based on the data obtained from LC-UV measurement of the total concentration of dansyl labeled metabolites in each cell extract and ^{13}C -/ ^{12}C -dansylation LC-MS analysis of the amine/phenol submetabolome, we concluded that the combination of scraping and freeze-thaw-cycle is a simple and efficient method for harvesting and lysing adherent mammalian cells for CIL LC-MS metabolomics. We envisage a wide use of this protocol for cellular metabolomics where comprehensive and quantitative analysis of the chemical-group-based submetabolomes is done using multiple chemical labeling LC-MS.

Chapter 5

Metabolomics of Small Numbers of Cells: Metabolomic Profiling of 100, 1000 and 10000

Human Breast Cancer Cells

5.1 Introduction

Cellular metabolomics involves the study of metabolomic profiles and their associated changes in response to a stimuli or perturbation to a cell (e.g., exposure to a toxin or mutation of a gene). It can be a powerful tool for studying cell biology and looking for potential biomarkers of diseases. In order to increase the number of quantifiable metabolites in a cell extract, multiple analytical techniques, with each often run under several different experimental conditions, are employed,^{220,221} which requires the use of a large number of cells (e.g., millions of cancer cells). However, decreasing the number of cells required for metabolomics would significantly benefit a number of research areas. For biological studies, with a reduced cell number required, one does not need to culture many cells, thereby reducing the overall experimental cost and allowing more biological replicates to be conveniently performed (e.g., no need of pooling cell cultures). In other areas, such as researches on stem cells,²²² circulating tumor cells in blood,²²³ and primary cells from tissues procured using laser capture micro-dissection (LCM),²²⁴ only a limited number of cells are available.

Analysis of metabolites from small numbers of cells, or even single cell, has been attempted by a number of detection techniques including electrochemical detection, vibrational spectrometry, fluorescence-based detection, and mass spectrometry (MS).^{225,226} Among them, only MS has the potential to analyze many metabolites simultaneously with high specificity. For example, matrix-assisted laser desorption ionization (MALDI) can generate ions from a small

sample spot, offering the possibility of detecting cellular components from a few cells²²⁷ or even a single small-size cell (e.g., hemoglobin from a red blood cell²²⁸ and metabolites from HeLa cells²²⁹). However, low MALDI efficiency of metabolites and strong interference of matrix ions in low mass region can limit the number of detectable metabolites. Matrix-free laser desorption ionization from a sample placed onto an active desorption substrate may eliminate matrix interference, but achieving uniformly high ionization efficiency for many metabolites is still a challenge. Some studies have shown the detection of about 100 metabolites and lipids from 1 to 80 cells.²³⁰ An alternative approach of using laser ablation electrospray ionization from a sample spot has shown the possibility of detecting 332 putative metabolite features in 13 *A. Cepa* cells.²³¹ Electrospray ionization (ESI) MS is another sensitive technique that has been shown to be useful to detect metabolites and lipids from a few cells or single plant cell.^{232,233}

The studies noted above only provided a few examples of using MS for analyzing small numbers of cells; excellent reviews on this active research field can be found in the literature.^{225,226} Currently, the major challenges are metabolomic coverage and quantification. Because lipids are major constituents of a cell and, therefore, are in high abundance, MS analysis of small numbers of cells detected more lipids than metabolites, even within the small number of mass spectral features observed. Moreover, cellular metabolomics requires accurate and precise quantification of metabolic changes among comparative cell types (i.e., relative quantification of individual metabolite concentrations among different cells). Without using internal standards, MALDI, ESI and other ionization methods suffers from matrix and ion suppression effects in metabolite quantification. Thus, there is a clear need of developing more sensitive and quantitative tools to perform high-coverage metabolomic profiling of small numbers of cells.

Recognizing that chemical derivatization can improve the sensitivity of metabolite detection in MS, a number of research groups have reported various labeling reagents and chemistries targeting the analysis of metabolites of interest with varying degrees of success.²⁰⁶⁻²¹⁶ We have been involved in developing a high-performance chemical isotope labeling (CIL) liquid chromatography mass spectrometry (LC-MS) platform for quantitative metabolomics.⁷⁹ In our previous studies, we have reported a divide-and-conquer approach of performing deep profiling of the metabolome using four labeling chemistries: $^{12}\text{C}/^{13}\text{C}$ -dansyl labeling for the amine/phenol submetabolome,⁷⁹ $^{12}\text{C}/^{13}\text{C}$ -DmPA labeling for the carboxylic submetabolome,⁸⁰ base-activated $^{12}\text{C}/^{13}\text{C}$ -dansyl labeling for the hydroxyl submetabolome,⁸² and $^{12}\text{C}/^{13}\text{C}$ -dansylhydrazine labeling for the carbonyl submetabolome.²³⁴ These four submetabolomes can cover over 95% of the entire chemical space of the metabolome.²³⁴ These rationally designed labeling methods afford a significant increase in metabolite detectability using reversed phase (RP) LC-MS without the need of changing columns and ionization modes. In addition, using differential isotope labeling, relative quantification of individual metabolites can be carried out with high accuracy and precision.

In this report, we describe a method of performing high-coverage quantitative metabolomics from small numbers of cells using CIL nanoflow LC-MS. To analyze small numbers of cells, previous CIL LC-MS protocols, such as that reported by Luo et al for analyzing 10^8 yeast cells,⁶ cannot be adapted. Because of the need to deal with much smaller amounts of metabolites present in a few cells, compared to analyzing millions of cells, a very sensitive workflow for CIL LC-MS is required. Thus, we focused our research efforts on developing and optimizing each key step from cell lysis to data generation to minimize sample loss during the sample workup and maximize metabolite detection in MS. We also focused on the metabolome profiling of an

analytically more challenging cell type: small-size mammalian cells. This type of cells is far more widely employed in biological studies and biomarker discovery, compared to other types of cells such as yeast cells or large-size cells, thus increasing the overall impact of the analytical workflow for cellular metabolomics research. In this work, we demonstrate the performance of this sensitive workflow in the analysis of 100, 1000 and 10000 MCF-7 breast cancer cells using dansylation labeling for profiling the amine/phenol submetabolome with unprecedented metabolomic coverage.

5.2 Experimental Section

5.2.1 Overall Workflow

Figure 5.1 shows the overall workflow for metabolomic profiling of a small number of cells. MCF-7 breast cancer cells, representative of many different types of mammalian cells commonly used in biological studies, were cultured, harvested, washed, counted, and then aliquoted to separate vials. Cell metabolism was quenched by snap-freezing in liquid nitrogen. The cells were lysed using a glass-bead-assisted lysis method (see below). The cell extracts were separated from glass beads and cell debris by centrifugation, and then dried down. The extracts were re-dissolved in $\text{Na}_2\text{CO}_3/\text{NaHCO}_3$ buffer, aliquoted, and labeled using ^{12}C - and ^{13}C -dansylation separately. The ^{12}C - and ^{13}C -labeled samples were mixed by 1:1 (v/v), and dried down. The dried samples were re-dissolved in 9:1 (v/v) $\text{H}_2\text{O}:\text{ACN}$ and analyzed by LC-MS. To improve detection sensitivity, nanoLC-MS with a Bruker captivespray ionization (CSI) interface was used. CSI uses a non-taped emitter tip which is not easily clogged, allowing robust operation in running complex (and often precious) metabolomic samples. We also employed and optimized a chemical-vapor-assisted technique to further increase MS sensitivity.

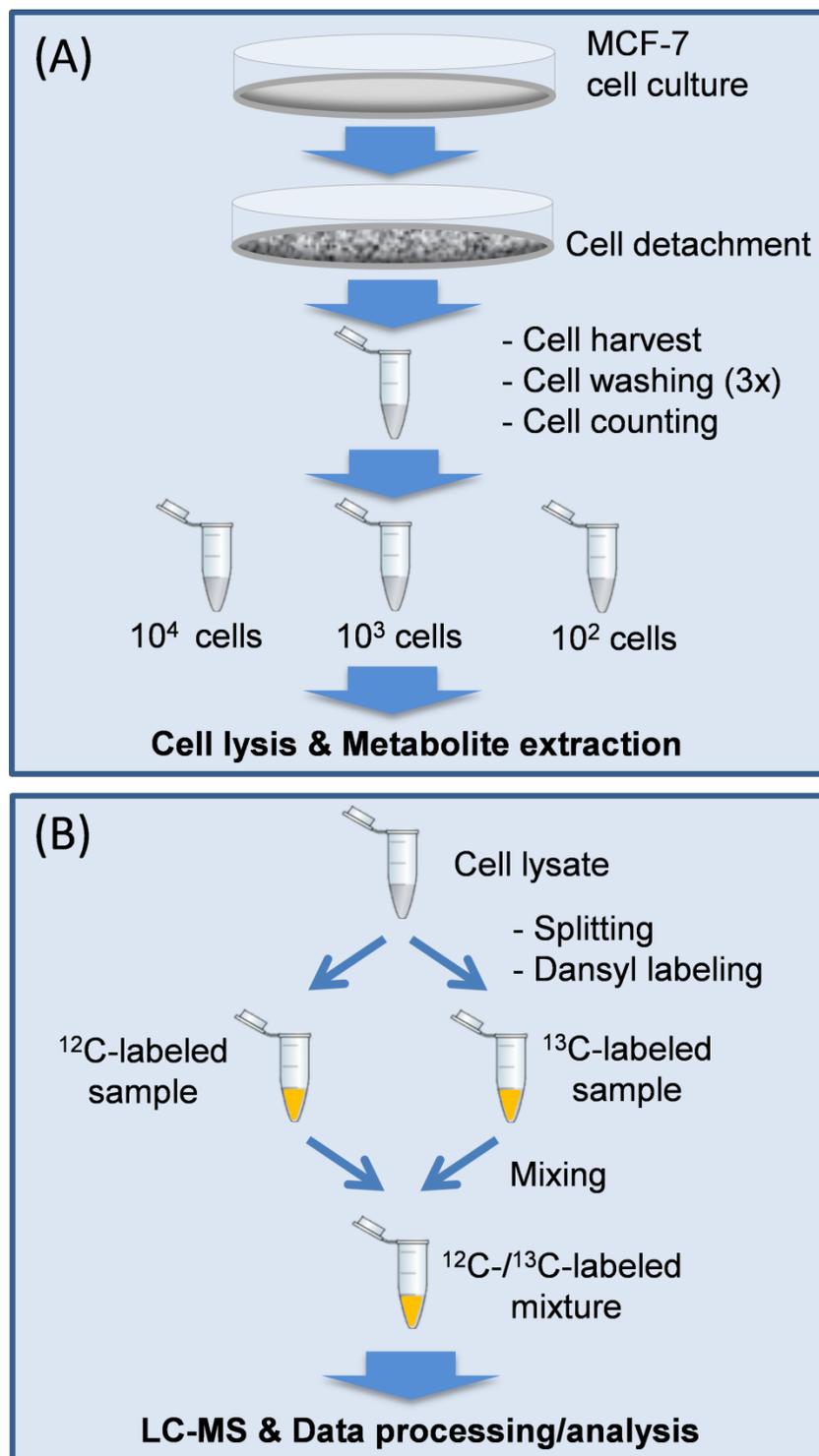


Figure 5.1 Workflow for CIL LC-MS method development.

5.2.2 Cell Culture and Harvest

MCF-7 cells (ATCC HTB-22) were cultured in Hyclone DMEM medium, supplemented with 10% fetal bovine serum (FBS) and 0.01 mg/mL human recombinant insulin, in 10-cm diameter culture dishes at 37°C in a humidified atmosphere with 5% CO₂. The growth medium was renewed every 2 days. For cell harvest, the cells were treated by 0.25% (w/v) trypsin and 0.53 mM EDTA at 37°C. The trypsinization process, monitored under a Zeiss Axiovert 25 inverted microscope (Oberkochen, Germany), was inhibited by adding the growth medium when the rounded cells were in suspension. The trypsin and growth medium were removed by centrifugation at 125g for 5 min at 4°C. The cell pellets were suspended in 1 mL of cold PBS solution and centrifuged at 125g for 5 min at 4°C. After removing PBS, this washing procedure was repeated by two more times. The washed cells in PBS were counted by a hemocytometer, and different numbers of cells were aliquoted into separate vials. The vials were snap-frozen in liquid nitrogen and then stored at -80°C freezer until further use.

5.2.3 Cell Lysis and Metabolite Extraction

Cell lysis was carried out by the glass-bead-assisted method.⁶ For comparison, ultrasonication cell lysis was also examined using a Branson Sonifer 450 Ultrasonic Distrupteror (Danbury, CT). For ultrasonic lysis, the cells were suspended in 1 mL of 50% MeOH, and sonicated on ice-bath for 1 min. For glass-bead lysis, 50 µL of lysis solvent and 0.1 mL of 0.5-mm diameter glass beads (Biospec Products, Bartlesville, OK) were added into the cell vial. The vials were vortexed on a VORTEX-GENIE 2 Mixer holder for 10 min at 4°C. Then additional 400 µL of the same lysis solvent were added and vortexed for 10 min for metabolite extraction. After centrifugation at 16000g for 10 min, the supernatant was transferred to another vial and

dried down in Speed Vac (Savant SC110A). The lysis/extraction solvent examined included 50% (v/v) ACN in water, 50% (v/v) MeOH in water, and a combination solvent of 1:1:1 (v/v/v) ACN: MeOH: H₂O (AMW) (see Results and Discussion).

5.2.4 Dansylation Labeling

For microflow LC-MS analysis of a large number of cells (i.e., 10⁵ cells in this work), a cell extract was re-dissolved in 50 µL of water and labeled using a previously reported protocol.⁶ In brief, a 20-µL aliquot of the extract was taken and mixed with 10 µL of Na₂CO₃/NaHCO₃ buffer and 10 µL of ACN. The solution was spun down and mixed with 20 µL of ¹²C-dansyl chloride (DnsCl) solution (18 mg/mL in ACN) for light labeling. The reaction mixture was incubated at 40°C for 1 h. After 1 h, the mixture was cooled down on ice-water bath, and 4 µL of 250 mM NaOH was added to quench the reaction by consuming the excess DnsCl. The solution was then incubated at 40°C for another 10 min. Finally, 20 µL of 425 mM formic acid (FA) in 1:1 ACN/H₂O was added to consume excess NaOH and to acidify the solution. For heavy labeling using ¹³C-dansyl chloride (available from mcid.chem.ualberta.ca), another 20-µL aliquot of the extract was taken and processed in the same way as ¹²C-labeling. The ¹²C-labeled sample was mixed with the ¹³C-labeled sample in 1:1 (v/v) for microflow LC-MS analysis.

For analyzing small numbers of cells ($\leq 10^4$ cells), a new labeling protocol was developed to handle the small amounts of metabolites in the cell extracts. The cell extract from 100, 1000 or 10000 cells was re-dissolved in 20 µL of Na₂CO₃/NaHCO₃ buffer, and split into two aliquots for labelings. Due to the presence of precipitates in the re-dissolved cell extract, 7.5 µL, instead of 10 µL, was taken into a 0.6 mL vial for ¹²C-labeling and another 7.5 µL was taken for ¹³C-labeling. 7.5 µL of 0.25 mg/mL DnsCl in ACN was added to each vial and the reaction mixture was incubated at 40°C for 1 h, and then quenched by 1 µL of 250 mM NaOH. 5 µL of 425 mM

FA was added to consume the excess NaOH. The ^{12}C - and ^{13}C -labeled samples were mixed in 1:1 (v/v) and dried down in Speed Vac, and re-dissolved in 20 μL of 9:1 (v/v) $\text{H}_2\text{O}:\text{ACN}$ for nanoflow LC-MS analysis.

5.2.5 LC-UV Quantification

The total amount of labeled metabolites was determined by using a step-gradient LC-UV method.¹⁵⁰ 5 μL of a dansyl labeled sample was injected onto a Phenomenex Kinetex C18 column (2.1 mm \times 5 cm, 1.7 μm particle size, 100 \AA pore size) linked to a Waters ACQUITY UPLC system with UV detection at 338 nm (Milford, MA). Mobile phase A was 0.1% (v/v) formic acid in 5% (v/v) ACN, and mobile phase B was 0.1% (v/v) formic acid in ACN. The LC conditions were: t = 0, 0% B; t = 1 min, 0% B; t = 1.1 min, 95% B; t = 2.6 min, 95% B; t = 3.1 min 0% B; t = 6.5 min, 0% B. The flow rate was 0.45 mL/min. The labeled metabolites eluted out together at high organic solvent. Their peak area was compared to that in a linear calibration curve, which was established using known concentrations of labeled amino acid mixtures, for quantifying the concentration of the labeled metabolites in a cell extract.

5.2.6 Microflow LC-MS

For profiling the metabolites in 10^5 cells, Thermo Scientific UltiMate 3000 UHPLC (Sunnyvale, CA) connected to a Bruker Maxis II Quadrupole Time-of-flight (Q-TOF) mass spectrometer (Billerica, MA) was used. The ^{12}C -/ ^{13}C -labeled samples were injected into an Agilent reversed phase Eclipse Plus C18 column (2.1 mm \times 10 cm, 1.8 μm particle size, 95 \AA pore size) for separation. For profiling the metabolites in 10^5 cells, a Thermo Scientific UltiMate 3000 UHPLC system (Sunnyvale, CA) connected to a Bruker Maxis II Quadrupole Time-of-flight (Q-TOF) mass spectrometer (Billerica, MA) was used. The ^{12}C -/ ^{13}C -labeled samples were injected into an Agilent reversed phase Eclipse Plus C18 column (2.1 mm \times 10 cm, 1.8 μm

particle size, 95 Å pore size) for separation. Mobile phase A was 0.1% (v/v) formic acid in 5% (v/v) acetonitrile, and mobile phase B was 0.1% (v/v) formic acid in acetonitrile. The chromatographic conditions were: t = 0 min, 20% B; t = 3.5 min, 35% B; t = 18 min, 65% B; t = 21 min, 99% B; t = 34 min, 99% B. The flow rate was 0.18 mL/min. The ESI-MS conditions used for Q-TOF were as follows: nebulizer, 1.0 bar; dry temperature, 230 °C; dry gas, 8 L/min; capillary voltage, 4500V; end plate offset, 500V; spectra rate, 1.0 Hz.

5.2.7 Nanoflow LC-MS

The analyses of cell extracts from small numbers of cells were performed on a nanoflow LC-MS system. It consisted of Waters NanoAcquity UPLC (Milford, MA) connected to a Bruker Impact HD Q-TOF mass spectrometer (Billerica, MA) equipped with a captivespray nanoBooster ion source (Bruker). Chromatographic separations were performed on a Thermo Scientific Acclaim PepMap 100 trap column (75 µm × 20 mm, 3 µm) and Acclaim PepMap RSLC C18 column (75 µm × 150 mm, 2 µm) (Sunnyvale, CA). The analyses of cell extracts from small numbers of cells were performed on a nanoflow LC-MS system. It consisted of a Waters NanoAcquity UPLC system (Milford, MA) connected to a Bruker Impact HD Q-TOF mass spectrometer (Billerica, MA) equipped with a captivespray nanoBooster ion source (Bruker). Chromatographic separations were performed on a Thermo Scientific Acclaim PepMap 100 trap column (75 µm × 20 mm, 3 µm) and Acclaim PepMap RSLC C18 column (75 µm × 150 mm, 2 µm) (Sunnyvale, CA). Mobile phase A was 0.1% (v/v) FA in water, and mobile phase B was 0.1% (v/v) FA in ACN. A 2-min-trapping procedure was performed prior to sample loading onto the analytical column. The trapping solvent was 99% mobile phase A. The chromatographic conditions were: t = 0 min, 15% B; t = 2.0 min, 15% B; t = 4.0 min, 25% B; t = 24 min, 60% B; t = 28 min, 90% B, t = 45 min, 90% B. The flow rate was 350 nL/min. The

captivespray operation conditions were: dry temperature, 200 °C; dry gas, 3 L/min; capillary voltage, 1400 V; nanoBooster, 0.2 bar, and dopant gas was pure ACN.

5.2.8 Data Processing and Metabolite Identification

The raw LC-MS data were exported as CSV files by Bruker Daltonics Data Analysis 4.3. A software tool, IsoMS,¹⁰⁸ was used to extract the peak pairs from the CSV files, filter the peak pairs by removing redundant peaks such as adduct ions, dimers and multimers to retain only $[M+H]^+$ pairs (i.e., one peak pair corresponds to one unique metabolite), and calculate the peak-pair intensity ratios of individual labeled metabolites.²³⁵ The multiple files generated from different LC-MS runs were aligned together by their accurate mass and retention time, and missing values in aligned files were filled by the Zerofill software.⁹⁷ Metabolite identification was done by searching against the dansyl standard library²³⁶ and MyCompoundID (MCID) libraries (www.mycompoundid.org).¹⁵¹

5.3 Results and Discussion

5.3.1 Cell Lysis and Metabolite Extraction

In cellular metabolomics, efficient cell lysis and metabolite extraction are very important, especially for profiling a small number of cells. In addition, the lysis method should be compatible to downstream sample processing and analysis. Detergent-based cell lysis is often used for cellular proteomics.²⁹ However, for metabolomics, detergent is difficult to separate from the metabolites and may cause interference in chemical labeling and LC-MS. Cell lysis by a physical means is a better option. Ultrasonic cell lysis is perhaps the most widely used method.²³⁷ However, a lot of energy is absorbed in this process, which may cause metabolite degradation. It also has a low throughput, as only one sample can be processed each time using a

conventional ultrasonic tip. In addition, there is a risk of cross-contamination if the tip is not washed thoroughly.

Recently, we reported a workflow for yeast cell metabolomic profiling where a glass-bead-assisted method was optimized for efficient lysis of yeast cells.⁶ The cells were disrupted by shear force generated from glass beads during vortexing. Comparing to ultrasonic lysis, the glass-bead method does not produce much heat, can be performed in parallel for multiple samples when a vortex holder is used, and there is no risk of cross-contamination. However, in this work, our focus was to develop a sensitive workflow for metabolome profiling of small-size mammalian cells, which are prone to metabolite leak during harvesting and cell washing step, compared to yeast cells which have a much stronger cell membrane. Thus, it is much easier to lose metabolites during the sample workup in analyzing mammalian cells. Sample loss may not be a problem if one has a lot of cells to start with; however, when we are forced to deal with small numbers of cells, any sample loss will result in the loss of metabolome information. In order to develop a method to lyse mammalian cells efficiently, we have compared the efficiencies of the ultrasonic and glass-bead methods. Figure 5.2A shows the number of peak pairs detected from microflow LC-MS analysis of ¹²C-/¹³C-labeled cell extracts prepared with a starting material of 10⁵ cells. There were 1599±47 (n=9) peak pairs detected from ultrasonic lysis and 1697±76 (n=9) peak pairs detected from the glass-bead method. These numbers are not significantly different, although the average peak pair number per run is slightly higher in the glass-bead method. Figure 5.2B shows the Venn diagram of the peak pair numbers detected from the two methods. Most of the peak pairs are in common (within a mass tolerance of 10 ppm), but more unique pairs are detected in the ultrasonic method, which may be related to the formation

of degraded metabolites during the sonication process. Based on these comparison results, we chose the glass-bead method for lysing the MCF-7 cells.

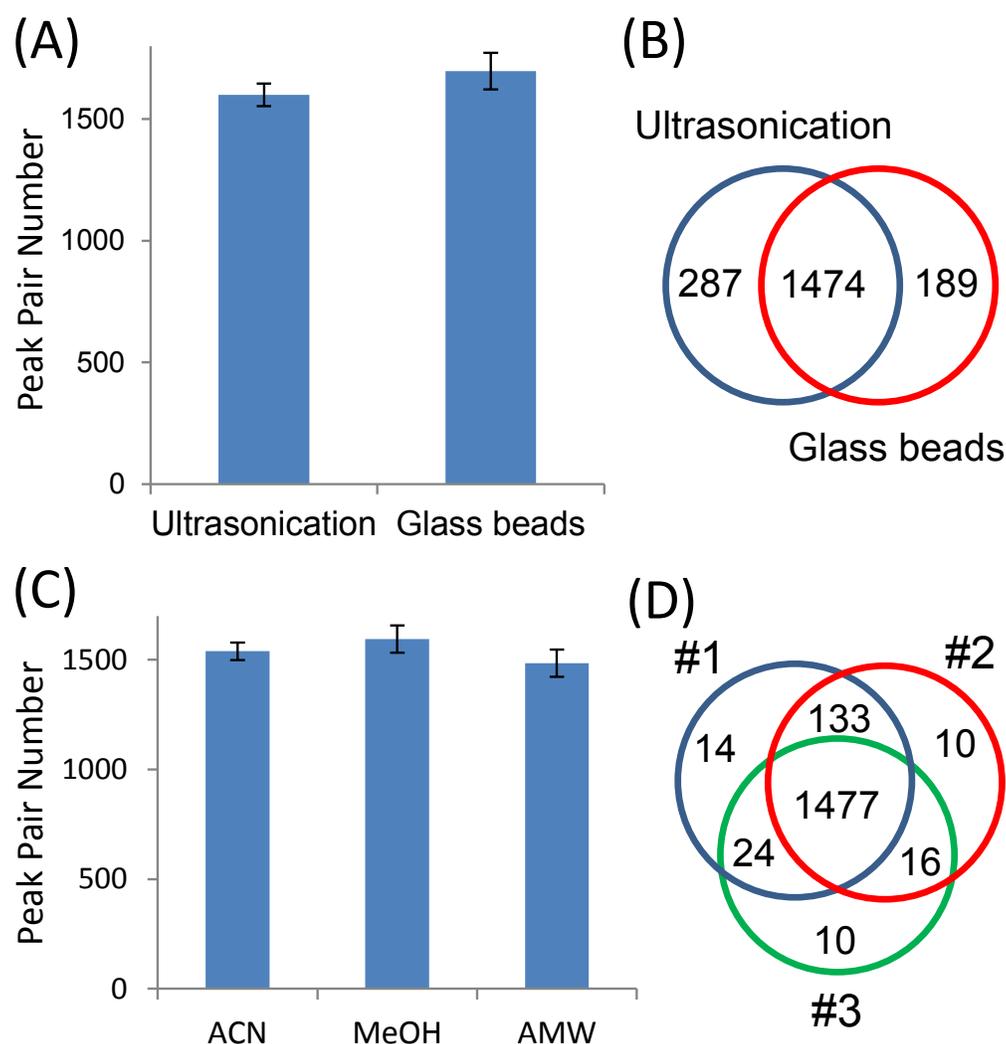


Figure 5.2 (A) Comparison of peak pair numbers detected from the ultrasonic cell lysis method and the glass-bead-assisted cell lysis method. (B) Venn diagram of peak pair numbers from the two methods. (C) Comparison of peak pair numbers detected from different metabolite extraction solvents. (D) Venn diagram of peak pair numbers detected from biological triplicate

analysis using MeOH extraction. Data were from experimental triplicate analysis of three biological replicates (n=9).

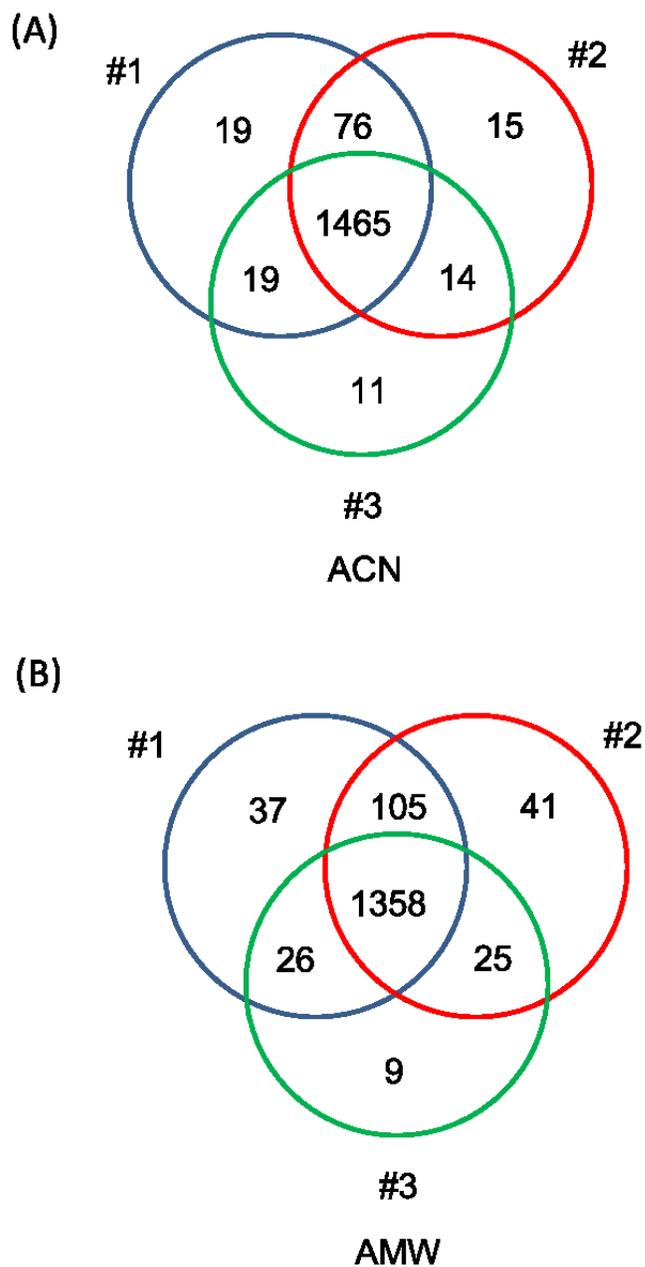


Figure 5.3 Venn diagrams of peak pair numbers detected from biological triplicate analysis using (A) ACN extraction and (B) AMW extraction. For each biological replicate, experimental triplicate analyses were carried out.

Selection of a proper extraction solvent is also important in cellular metabolomics. Based on literature information and our own experience,⁷ we selected and compared three solvent systems: 50% (v/v) ACN in water, 50% (v/v) MeOH in water, and a combination solvent of 1:1:1 (v/v/v) ACN: MeOH: H₂O (AMW). In this case, 10⁵ cells were lysed with the glass-bead method and extracted using one of the three solvents, followed by dansyl labeling and microflow LC-MS analysis. As Figure 5.2C shows, there are 1539±40, 1594±62, and 1484±62 (n=9) peak pairs detected from 50% ACN, 50% MeOH, and AMW extraction, respectively. These numbers are not significantly different, although 50% MeOH extraction gives a slightly larger number. The reproducibility of extraction was found to be excellent. As an example, the Venn diagram of the peak pair numbers detected in triplicate analysis using 50% MeOH extraction (Figure 5.2D) shows that 1477 peak pairs (92.7%) were commonly detected in triplicates (see Figure 5.3 for the ACN and AMW extractions). From the comparison results obtained, we selected 50% MeOH in water as the extraction solvent for the workflow.

5.3.2 Dansylation Protocol

The previously reported dansylation protocol was useful for analyzing samples with a total concentration of labeled metabolites of >1 mM.^{6,79} We found that it was not suitable for labeling a cell extract from a small number of cells. One reason is related to the presence of high concentrations of dansyl dimethylamine and dansyl amine, two major by-products, when using high concentrations of dansyl chloride to label low concentrations of samples. These two products are detected in LC-MS as two high-intensity chromatographic peaks that can interfere with the quantification of other co-eluting labeled metabolites (e.g., signal saturation in MS detection). Another reason is related to the presence of a large amount of dansyl hydroxyl (Dns-OH), a product from the labeling quenching step, in a labeled sample. Dns-OH can suppress

other labeled metabolites in nanoLC-MS analysis. In microflow LC-MS, Dns-OH can be eluted out at the first two minutes of the ion chromatogram. However, in nanoLC-MS, a trapping column is used to capture the labeled metabolites before injecting them into the analytical column for separation. After sample trapping, the relatively hydrophilic Dns-OH is washed away using a high-water-content solvent. Even with the use of an additional washing step, Dns-OH cannot be removed completely. Extension of the washing time or an increase in the number of washings is not ideal, as this will elongate the total analysis time and increase the risk of losing hydrophilic metabolites.²³⁶

To address the above issues, we decreased the DnsCl concentration for the labeling reaction and used a small volume of buffer to re-dissolve a cell lysate to keep the metabolite concentration high. We tested the method blank labeled by different concentrations (0.25, 0.5, 1 mg/mL) of DnsCl in nanoLC-MS. The signals of Dns-OH, Dns-dimethylamine and Dns-amine were significantly reduced when a lower concentration of DnsCl was used (In supplemental information which is available from Dr. Liang Li). Some dansyl labeled background chemicals in the labeled blanks were observed. However, when the DnsCl concentration was decreased to 0.25 mg/mL, these background peaks became very small. Moreover, using this concentration to label a real cell lysate resulted in the detection of thousands of metabolites in nanoLC-MS (see below). Thus, in our workflow, we chose 0.25 mg/mL of DnsCl to label the cell lysates from small numbers of cells.

5.3.3 Captivespray MS

Metabolite detectability can be significantly affected by the MS setup and experimental conditions used. At present, most of the metabolomic analysis experiments are done using a conventional or microflow ESI-MS, while there are only a few reports of using nanoESI-MS.²³⁸⁻

²⁴² The captivespray ion source employed in this work for nanoflow LC-MS uses a gas stream to guide the nanospray-generated ions into the mass spectrometer. Comparing to nanoESI, there is no need for X, Y, Z positioning in CSI, and the non-tapered emitter spray tip avoids being easily clogged for robust operation. There is no report of using captivespray for metabolomic analysis and thus we intended to optimize the CSI setup for CIL LC-MS analysis of cell extracts from small numbers of cells.

We also applied and optimized a chemical-vapor-assisted ESI technique in CSI to improve detection sensitivity.²⁴³ In this technique, a chemical is placed in a container and nitrogen gas flows through the container to carry the chemical vapor to the spray tip chamber. Our group previously demonstrated that this technique could enhance the MS sensitivity in shotgun proteomics, when an appropriate chemical (e.g., butanol) was used.²⁴³ In this study, we evaluated four different dopant gas: ACN, MeOH, isopropanol (IPA), 20% FA in ACN for metabolome analysis using CIL LC-MS. The physical properties of four chemicals can be found in Table 5.1. Dansyl labeled cell lysates were analyzed using CSI MS under different chemical vapors, and several amino acids detected were selected to evaluate the performance. Figure 5.4A shows the signal comparison of labeled amino acids detected in cell lysates. Overall, the use of ACN provided the highest signal enhancement for all the analytes. Although these analytes have different chemical/physical properties, the signal enhancement had the same trend. This can be attributed to the presence of the dansyl tag(s) in each analyte that equalizes the ionization process to some extent and, as a result, different metabolites have similar behaviors in the ionization process. Based on these results, we chose pure ACN as the dopant gas for the subsequent studies. While the exact mechanism for signal enhancement by using ACN dopant is unclear, we speculate that the enhancement was due to enhanced ionization efficiencies for the dansylated

metabolites during the ESI process. It is plausible that ACN vapor molecules surrendering the ESI droplets might reduce the energy barriers required for ejecting the analyte ions from the droplet surfaces to the gas phase.

Table 5.1 Physical properties of chemical vapors tested in this study.

	Boiling Point (°C)	Surface Tension at 25 °C (nN·m ⁻¹)	Proton Affinity (kJ·mol ⁻¹)
Methanol	64.6	22.07	754.3
Isopropanol	82.3	20.93	793.0
Acetonitrile	81.6	28.66	779.2
Formic Acid	101	34.38	742.0

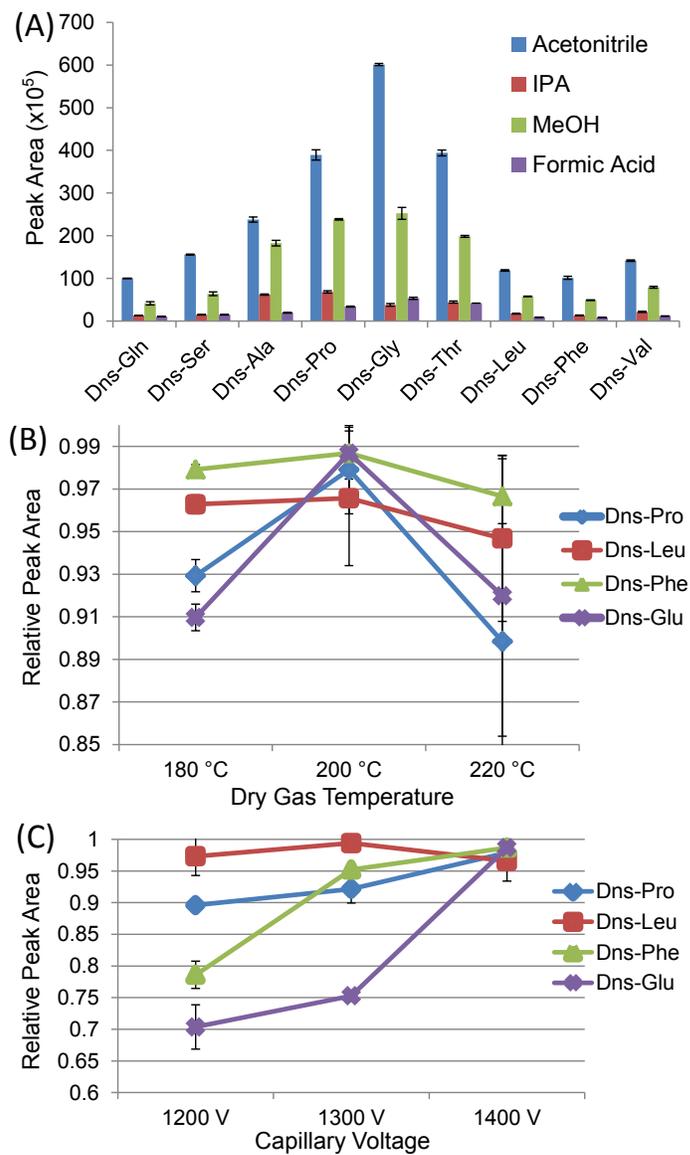


Figure 5.4 (A) Peak areas of molecular ions of 9 dansyl labeled metabolites detected with the use of different dopant chemical vapors. Normalized peak areas of molecular ions of 4 dansyl labeled metabolites detected at (B) different temperatures of dry gas and (C) different capillary voltages. Data were from experimental triplicate analysis (n=3).

Two other adjustable parameters in CSI were optimized. Figures 5.4B and C show the effects of dry gas (nitrogen) temperature and capillary voltage on normalized ion signals of four selected amino acids, respectively. The optimal temperature of 200°C and capillary voltage of 1400 V were chosen.

5.3.4 Injection Amount

The amount of samples injected into LC-MS can also have a significant effect on metabolite detectability. Since the total concentration of labeled metabolites is measured by LC-UV in our workflow, we can readily determine the optimal injection amount that gives the maximal number of peak pairs detectable by LC-MS. Optimization of the injection amount was performed on both microflow LC-MS and nanoLC-MS for comparison. In this case, a series of different volumes of ^{12}C -/ ^{13}C -labeled cell lysates from 10^5 cells were injected into microflow LC-MS, while a labeled cell lysate from 10^4 cells was diluted and then injected into nanoLC-MS. Figures 5.5A and B show the number of peak pairs detected from two systems, respectively. In Figure 5.5A, as the injection amount increases, the peak pair number detected by microflow LC-MS increases and then levels off at 2.04 nmol. On average, 1680 ± 5 ($n=3$) peak pairs were detected by microflow LC-MS. Figure 5.5B shows that, for nanoLC-MS, the maximal number of peak pairs (2301 ± 86) was reached when 11.4 pmol of labeled lysate was injected.

Figure 5.5 shows that, at both optimal injection conditions, nanoLC-MS could detect 37% more metabolites than microflow LC-MS. It should be noted that the two QTOF instruments used gave almost the same detectability of labeled metabolites when microflow LC was linked to both. Thus, the detectability of nanoLC-MS is significantly better than microflow LC-MS. More importantly, for handling small amounts of samples, the sample amount needed to reach the maximal number of detectable peak pairs is about 200-fold less in nanoLC-MS than microflow LC-MS.

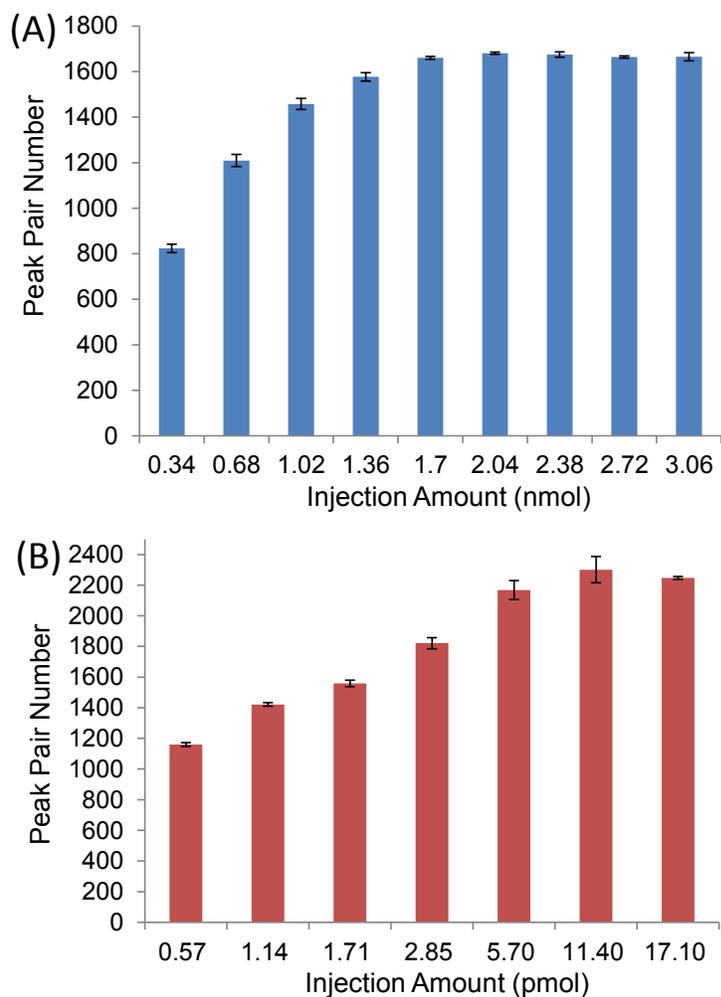


Figure 5.5 Average peak pair numbers detected using (A) microflow LC-MS and (B) nanoLC-MS as a function of injection amount of ^{12}C -/ ^{13}C -labeled cell lysates ($n=3$).

5.3.5 Metabolomic profiling of small numbers of cells

After optimizing sample handling and instrument setting, the CIL nanoLC-MS system was used to profile the amine/phenol submetabolome of 100, 1000 and 10000 cells. In each case, cells were lysed and then split into two aliquots with one aliquot for ^{12}C -labeling and another aliquot for ^{13}C -labeling, followed by mixing the labeled aliquots for nanoLC-MS analysis. Note that, in a metabolomics study of comparing different types of cells (e.g., wild type vs. mutated cells), sample splitting of a cell lysate is needed in order to produce a pooled cell extract from mixing aliquots of all comparative cell lysates. This pooled sample is labeled with a ^{13}C -reagent to serve as a global internal control; an aliquot of ^{13}C -pool is spiked into a ^{12}C -labeled individual cell lysate for relative quantification. In some applications, we could replace the pooled sample with a cell lysate prepared from a large number of similar cells. If this could be done, we would double the sample amount for LC-MS analysis (i.e., the current result of 100 cells would be equivalent to that from a starting material of 50 cells per sample).

For the labeled cell lysates prepared from 10000 cells, the amount of labeled metabolites was found to be ~120 pmol, which is higher than the optimal injection amount (e.g., 11.4 pmol). Thus, only the optimal amount was taken for injection into nanoLC-MS. However, the amount of labeled metabolites from 100 or 1000 cells could not be determined as it was below the detection limit of the current LC-UV setup; for future work, we plan to develop a fluorescence-based detection system for quantifying trace amounts of labeled metabolites. The amount of labeled metabolites was expected to be less than the optimal injection amount and thus all the labeled lysates from 100 or 1000 cells were injected. Figure 5.6A-C shows the representative total ion chromatograms generated from the labeled lysates of 100, 1000, and 10000 cells, while Figure 5.6D-F shows the corresponding extracted ion chromatograms (EIC) of a labeled metabolite

(Dns-uridine) and Figure 5.6G-I shows the molecular ion regions of the detected peak pairs. The signal-to-noise ratios of EICs and mass spectral peaks are slightly lower in the 1000-cell lysate, compared to the 1/10-injection of the 10000-cell lysate, suggesting that sample loss might be more severe in handling 1000 cells than working with 10000 cells. For the 100-cell lysate, the signal intensities are much lower than those from 1000 cells, as expected.

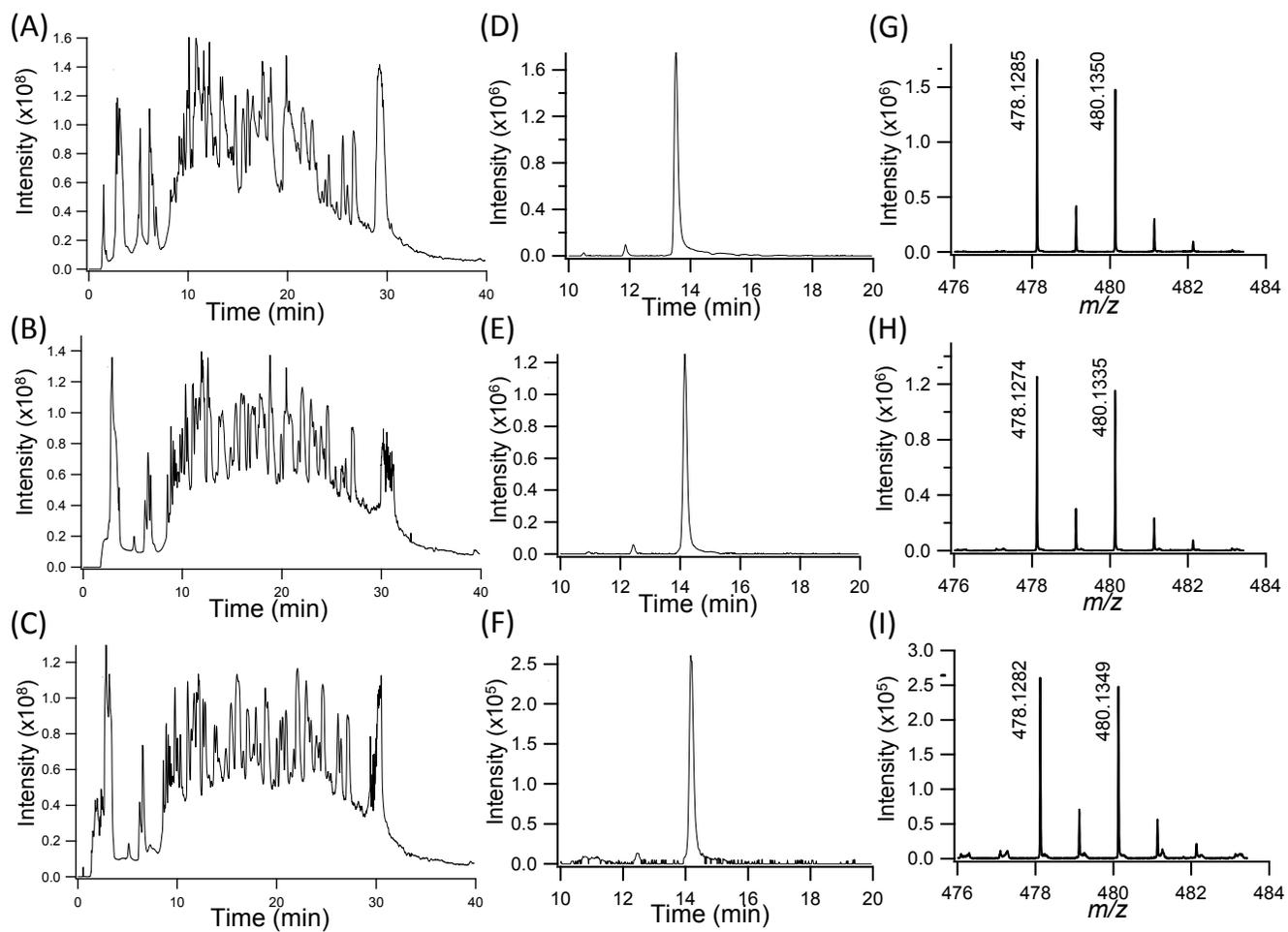


Figure 5.6 (A-C) Total ion chromatograms of labeled cell lysates, (D-F) extracted ion chromatograms of Dns-uridine, and (G-I) molecular ion regions of the ¹²C-/¹³C-Dns-uridine peak

pair obtained from injection of 1/10 of the labeled 10000-cell lysate (top), all labeled 1000-cell lysate (middle), and all labeled 100-cell lysate (bottom).

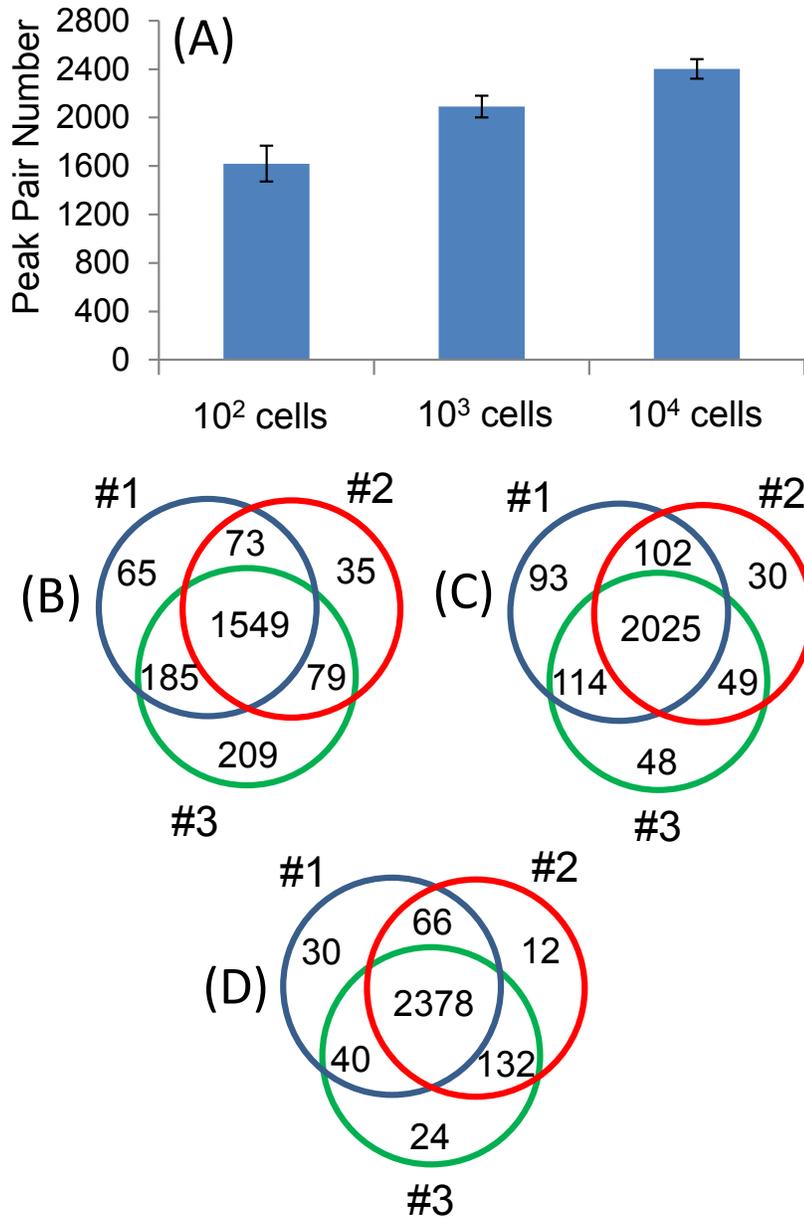


Figure 5.7 (A) Average peak pair numbers detected from ^{12}C -/ ^{13}C -labeled 100-, 1000-, and 10000-cell lysates (duplicate analysis of three biological replicates or $n=6$). Venn diagrams of peak pair numbers detected from biological triplicate analysis of ^{12}C -/ ^{13}C -labeled cell lysates of (B) 100, (C) 1000, and (D) 10000 cells.

Figure 5.7 shows the plots of the number of peak pairs detected from these samples. There are 1620 ± 148 , 2091 ± 89 and 2402 ± 80 ($n=6$) peak pairs detected from the 100-, 1000- and 10000-cell lysates, respectively. Comparing to 2402 peak pairs detected from the 10000-cell lysates, we were still able to detect $\sim 87\%$ peak pairs from 10-fold less cells, and $\sim 67\%$ peak pairs from 100-fold less cells. Figure 5.7B-D shows the Venn diagrams of peak pair numbers detected from experimental duplicate on biological triplicate analysis ($n=6$). Most of peaks pairs could be detected from all the biological triplicate analysis, indicating excellent reproducibility of our workflow.

Many of the peak pairs detected could be identified or mass-matched to human metabolome databases. Using the dansyl standard library consisting of 278 amine/phenol-containing metabolites, we identified 80, 94, 106 metabolites from the 100-, 1000- and 10000-cell lysates, respectively (Tables 5.2-5.4). Based on accurate mass search with a mass tolerance of 10 ppm, we could match 673, 751 and 896 peak pairs to metabolite structures in HMDB (8,021 entries) (In supplemental information which is available from Dr. Liang Li) and additional 369, 474 and 511 peak pairs to the predicted metabolites in MCID (375,809 entries) (In supplemental information which is available from Dr. Liang Li). In total, 1122 (69.3%), 1319 (63.1%) and 1513 (63.1%) peak pairs could be identified and matched in the 100-, 1000- and 10000-cell lysates, respectively.

Table 5.2 List of metabolites detected from 100 cells identified based on accurate mass and retention time matches to the Dns-labeled standard

Peak Pair Information						Identification Result	
Peak Pair #	T _R (min)	Corrected T _R (min)	mz _{light}	mz _{heavy}	monoisotopic mass (Da)	HMDB.No.	Name
1273	8.82	3.48	308.1463	310.1530	74.0879	HMDB00002	1,3-Diaminopropane
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB00020	p-Hydroxyphenylacetic acid
1739	10.21	4.23	501.1538	503.1603	267.0954	HMDB00050	Adenosine
2536	13.59	7.84	323.1081	325.1144	89.0498	HMDB00056	Beta-Alanine
3873	19.63	12.64	363.1355	365.1451	129.0771	HMDB00070	D-Pipecolic acid
4399	21.86	14.36	279.1173	281.1221	45.0590	HMDB00087	Dimethylamine
2553	13.70	7.92	337.1206	339.1269	103.0623	HMDB00112	Gamma-Aminobutyric acid
2296	12.66	7.01	309.0923	311.0990	75.0340	HMDB00123	Glycine
3038	16.38	10.19	363.1003	365.1066	129.0420	HMDB00148_2	L-Glutamic Acid - H ₂ O
2190	12.16	6.54	295.1131	297.1191	61.0547	HMDB00149	Ethanolamine
5079	24.77	16.78	388.0836	390.0889	154.0253	HMDB00152	Gentisic acid
2881	15.52	9.47	370.0962	372.1026	136.0379	HMDB00157_2	Hypoxanthine - multi-tags

2881	15.52	9.47	370.0962	372.1026	136.0379	HMDB00157_3	Hypoxanthine - Isomer
3977	20.03	12.90	399.1390	401.1449	165.0806	HMDB00159	L-Phenylalanine
3054	16.43	10.23	349.1237	351.1300	115.0654	HMDB00162	L-Proline
3012	16.26	10.09	265.1051	267.1120	31.0468	HMDB00164	Methylamine
2075	11.47	5.75	353.1200	355.1257	119.0617	HMDB00167	L-Threonine
1497	9.54	3.87	366.1108	368.1173	132.0525	HMDB00168	L-Asparagine
4048	20.25	13.03	365.1551	367.1614	131.0968	HMDB00172	L-Isoleucine
5684	26.74	18.51	389.1269	391.1333	155.0686	HMDB00177	L-Histidine
1967	10.85	5.07	339.1038	341.1102	105.0455	HMDB00187	L-Serine
2194	12.18	6.56	422.1720	424.1788	188.1137	HMDB00206	N6-Acetyl-L-Lysine
5166	25.00	16.97	300.1020	302.1087	132.0874	HMDB00214	Ornithine
1060	7.58	2.82	359.0718	361.0779	125.0135	HMDB00251	Taurine
3711	19.15	12.25	432.1086	434.1157	198.0503	HMDB00291	Vanillylmandelic acid
2600	14.01	8.19	478.1268	480.1332	244.0685	HMDB00296	Uridine
3517	18.31	11.57	346.0854	348.0919	112.0271	HMDB00300	Uracil
4318	21.33	13.92	372.1007	374.1065	138.0423	HMDB00301	Urocanic acid
5079	24.77	16.78	388.0836	390.0889	154.0253	HMDB00397	2-Pyrocatechuic acid
5195	25.10	17.05	386.1051	388.1114	152.0468	HMDB00440	3-Hydroxyphenylacetic acid
2194	12.18	6.56	422.1720	424.1788	188.1137	HMDB00446	N-Alpha-

							acetyllysine
2981	16.09	9.95	337.1224	339.1291	103.0641	HMDB00452	L-Alpha-aminobutyric acid
5539	26.29	18.10	402.1002	404.1065	168.0418	HMDB00484	Vanillic acid
5545	26.31	18.11	372.0900	374.0964	138.0317	HMDB00500	4-Hydroxybenzoic acid
4070	20.32	13.08	365.1556	367.1609	131.0972	HMDB00557	L-Alloisoleucine
1535	9.67	3.94	380.1285	382.1346	146.0702	HMDB00641	L-Glutamine
2981	16.09	9.95	337.1224	339.1291	103.0641	HMDB00650	D-Alpha-aminobutyric acid
5195	25.10	17.05	386.1051	388.1114	152.0468	HMDB00669	Ortho-Hydroxyphenylacetic acid
3955	19.93	12.85	365.1550	367.1609	131.0967	HMDB00687	L-leucine
3379	17.75	11.12	383.1095	385.1155	149.0511	HMDB00696	L-Methionine
3873	19.63	12.64	363.1355	365.1451	129.0771	HMDB00716	L-Pipecolic acid
1179	8.38	3.25	353.1172	355.1225	119.0589	HMDB00719	L-Homoserine
2503	13.44	7.71	406.1414	408.1490	172.0830	HMDB00721	Glycylproline
2076	11.47	5.76	365.1155	367.1217	131.0572	HMDB00725	Trans-4-Hydroxyl-L-Proline
6958	30.00	21.48	356.0943	358.1002	122.0360	HMDB00750_2	3-Hydroxymandelic acid - COOH
4365	21.62	14.16	416.1142	418.1214	182.0559	HMDB00755	Hydroxyphenyllactic acid
3513	18.28	11.55	422.1726	424.1789	188.1142	HMDB00759	Glycyl-L-Leucine

3431	17.93	11.26	429.1098	431.1156	195.0514	HMDB00840	Salicylic acid
3458	18.00	11.32	351.1402	353.1472	117.0819	HMDB00883	L-Valine
1651	10.01	4.12	409.1526	411.1593	175.0942	HMDB00904	Citrulline
3660	18.94	12.08	438.1473	440.1537	204.0890	HMDB00929	L-Tryptophan
5951	27.51	19.21	428.1136	430.1207	194.0553	HMDB00954	trans-Ferulic acid
5346	25.65	17.52	428.1116	430.1214	194.0533	HMDB00955	Isoferulic acid
1255	8.77	3.46	343.0768	345.0836	109.0185	HMDB00965	Hypotaurine
5613	26.52	18.31	393.1828	395.1885	159.1245	HMDB00991	2-aminooctanoic acid
5060	24.71	16.73	371.1045	373.1105	137.0462	HMDB01123	2-Aminobenzoic acid
2491	13.39	7.66	365.1221	367.1313	131.0638	HMDB01149	5-Aminolevulinic acid
3865	19.61	12.62	401.1193	403.1256	167.0610	HMDB01545	Pyridoxal
4249	21.06	13.69	365.1556	367.1610	131.0973	HMDB01645	L-Norleucine
6018	27.67	19.35	398.1038	400.1105	164.0455	HMDB01713	m-Coumaric acid
1473	9.48	3.84	293.1059	295.1115	59.0475	HMDB01842	Guanidine
5125	24.89	16.88	402.0989	404.1047	168.0406	HMDB01868	5-Methoxysalicylic acid
4847	23.85	16.01	372.0905	374.0965	138.0322	HMDB01895	Salicylic acid
3361	17.69	11.08	365.1517	367.1582	131.0933	HMDB01901	Aminocaproic acid
2553	13.70	7.92	337.1206	339.1269	103.0623	HMDB01906	2-Aminoisobutyric acid

1937	10.73	4.84	399.1037	401.1098	165.0453	HMDB02005_2	Methionine Sulfoxide - Isomer
3547	18.44	11.68	360.1005	362.1068	126.0422	HMDB02024	Imidazoleacetic acid
5306	25.47	17.36	432.1086	434.1162	198.0503	HMDB02085	Syringic acid
5730	26.89	18.64	400.1199	402.1263	166.0615	HMDB02199	Desaminotyrosine
3190	17.00	10.61	468.1678	470.1737	234.1094	HMDB02339	5-Methoxytryptophan
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB02390	3-Cresotinic acid
5315	25.51	17.40	327.1166	329.1233	93.0583	HMDB03012	Aniline
2623	14.17	8.32	351.1340	353.1402	117.0757	HMDB03355	5-Aminopentanoic acid
3361	17.69	11.08	365.1517	367.1582	131.0933	HMDB03640	Beta-Leucine
2806	15.13	9.14	337.1203	339.1274	103.0620	HMDB03911	3-Aminoisobutanoic acid
2100	11.63	5.95	339.1384	341.1448	105.0801	HMDB04437	Diethanolamine
6320	28.39	20.01	386.1039	388.1107	152.0455	HMDB04815	4-Hydroxy-3-methylbenzoic acid
3513	18.28	11.55	422.1726	424.1789	188.1142	HMDB28844	Glycyl-Isoleucine
3639	18.88	12.03	456.1557	458.1625	222.0974	HMDB28848	Glycyl-Phenylalanine
4523	22.49	14.88	402.0989	404.1055	168.0405	HMDB60003	Isovanillic acid

Table 5.3 List of metabolites detected from 1000 cells identified identified based on accurate mass and retention time matches to the DnsCl-labeled standard library.

Peak Pair Information						Identification Result	
Peak Pair #	T _R (min)	Corrected T _R (min)	mz _{light}	mz _{heavy}	monoisotopic mass (Da)	HMDB.No.	Name
1273	8.82	3.48	308.1463	310.1530	74.0879	HMDB00002	1,3-Diaminopropane
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB00020	p-Hydroxyphenylacetic acid
1352	9.12	3.65	501.1512	503.1578	267.0928	HMDB00050	Adenosine
2481	13.34	7.62	323.1086	325.1147	89.0503	HMDB00056	Beta-Alanine
3873	19.63	12.64	363.1355	365.1451	129.0771	HMDB00070	D-Pipecolic acid
2588	13.95	8.14	501.1643	503.1710	267.1060	HMDB00085	Deoxyguanosine
4504	22.39	14.80	279.1180	281.1226	45.0597	HMDB00087	Dimethylamine
2650	14.29	8.42	337.1279	339.1347	103.0696	HMDB00112	Gamma-Aminobutyric acid
2170	12.04	6.41	309.0932	311.0994	75.0349	HMDB00123	Glycine
2034	11.21	5.33	381.1119	383.1181	147.0536	HMDB00148	L-Glutamic Acid
3100	16.56	10.31	363.1012	365.1075	129.0429	HMDB00148_2	L-Glutamic Acid - H ₂ O
2197	12.18	6.57	295.1131	297.1191	61.0548	HMDB00149	Ethanolamine

5079	24.77	16.78	388.0836	390.0889	154.0253	HMDB00152	Gentisic acid
2790	15.06	9.08	370.0972	372.1036	136.0389	HMDB00157_2	Hypoxanthine - multi-tags
3069	16.47	10.26	370.0986	372.1045	136.0403	HMDB00157_3	Hypoxanthine - Isomer
7306	31.33	22.69	324.5953	326.6019	181.0739	HMDB00158	L-Tyrosine
3900	19.73	12.71	399.1386	401.1444	165.0803	HMDB00159	L-Phenylalanine
2319	12.75	7.09	323.1059	325.1123	89.0475	HMDB00161	L-Alanine
3054	16.43	10.23	349.1237	351.1300	115.0654	HMDB00162	L-Proline
3012	16.26	10.09	265.1051	267.1120	31.0468	HMDB00164	Methylamine
2075	11.47	5.75	353.1200	355.1257	119.0617	HMDB00167	L-Threonine
1497	9.54	3.87	366.1108	368.1173	132.0525	HMDB00168	L-Asparagine
3954	19.92	12.84	365.1588	367.1639	131.1005	HMDB00172	L-Isoleucine
5731	26.89	18.64	389.1345	391.1409	155.0761	HMDB00177	L-Histidine
5577	26.41	18.21	307.1150	309.1223	146.1133	HMDB00182	L-Lysine
1967	10.85	5.07	339.1038	341.1102	105.0455	HMDB00187	L-Serine
2080	11.49	5.79	367.0960	369.1024	133.0377	HMDB00191	L-Aspartic Acid
2194	12.18	6.56	422.1720	424.1788	188.1137	HMDB00206	N6-Acetyl-L-Lysine
2632	14.20	8.35	453.1670	455.1736	219.1087	HMDB00210	Pantothenic acid
5166	25.00	16.97	300.1020	302.1087	132.0874	HMDB00214	Ornithine
4369	21.64	14.18	360.0968	362.1029	126.0384	HMDB00262	Thymine
2818	15.19	9.19	386.0978	388.1046	152.0395	HMDB00292	Xanthine

2690	14.51	8.61	478.1374	480.1438	244.0790	HMDB00296	Uridine
3517	18.31	11.57	346.0854	348.0919	112.0271	HMDB00300	Uracil
4318	21.33	13.92	372.1007	374.1065	138.0423	HMDB00301	Urocanic acid
5079	24.77	16.78	388.0836	390.0889	154.0253	HMDB00397	2-Pyrocatechuic acid
5195	25.10	17.05	386.1051	388.1114	152.0468	HMDB00440	3-Hydroxyphenylacetic acid
2482	13.34	7.63	422.1719	424.1757	188.1136	HMDB00446	N-Alpha-acetyllysine
2848	15.31	9.29	337.1278	339.1344	103.0695	HMDB00452	L-Alpha-aminobutyric acid
5539	26.29	18.10	402.1002	404.1065	168.0418	HMDB00484	Vanillic acid
5244	25.24	17.17	372.0913	374.0970	138.0329	HMDB00500	4-Hydroxybenzoic acid
1212	8.57	3.35	408.1687	410.1750	174.1104	HMDB00517	L-Arginine
3953	19.92	12.84	365.1550	367.1610	131.0967	HMDB00557	L-Alloisoleucine
1793	10.35	4.31	380.1293	382.1353	146.0710	HMDB00641	L-Glutamine
2848	15.31	9.29	337.1278	339.1344	103.0695	HMDB00650	D-Alpha-aminobutyric acid
5195	25.10	17.05	386.1051	388.1114	152.0468	HMDB00669	Ortho-Hydroxyphenylacetic acid
4055	20.27	13.04	365.1549	367.1613	131.0966	HMDB00687	L-leucine

3379	17.75	11.12	383.1095	385.1155	149.0511	HMDB00696	L-Methionine
3862	19.60	12.61	363.1436	365.1507	129.0853	HMDB00716	L-Pipecolic acid
1226	8.66	3.40	353.1191	355.1232	119.0608	HMDB00719	L-Homoserine
2503	13.44	7.71	406.1414	408.1490	172.0830	HMDB00721	Glycylproline
2076	11.47	5.76	365.1155	367.1217	131.0572	HMDB00725	Trans-4-Hydroxyl-L-Proline
3776	19.34	12.40	402.0993	404.1060	168.0410	HMDB00750	3-Hydroxymandelic acid
6958	30.00	21.48	356.0943	358.1002	122.0360	HMDB00750_2	3-Hydroxymandelic acid - COOH
4618	22.97	15.28	416.1238	418.1303	182.0655	HMDB00755	Hydroxyphenyllactic acid
3513	18.28	11.55	422.1726	424.1789	188.1142	HMDB00759	Glycyl-L-Leucine
3431	17.93	11.26	429.1098	431.1156	195.0514	HMDB00840	Salicylic acid
3458	18.00	11.32	351.1402	353.1472	117.0819	HMDB00883	L-Valine
1651	10.01	4.12	409.1526	411.1593	175.0942	HMDB00904	Citrulline
3660	18.94	12.08	438.1473	440.1537	204.0890	HMDB00929	L-Tryptophan
5346	25.65	17.52	428.1116	430.1214	194.0533	HMDB00954	trans-Ferulic acid
5463	26.06	17.89	428.1140	430.1224	194.0557	HMDB00955	Isoferulic acid
1136	8.06	3.07	343.0770	345.0828	109.0187	HMDB00965	Hypotaurine
5613	26.52	18.31	393.1828	395.1885	159.1245	HMDB00991	2-aminooctanoic acid

5060	24.71	16.73	371.1045	373.1105	137.0462	HMDB01123	2-Aminobenzoic acid
2340	12.82	7.15	365.1160	367.1227	131.0577	HMDB01149	5-Aminolevulinic acid
2478	13.33	7.62	531.1579	533.1640	297.0996	HMDB01173	5'-Methylthioadenosine
3813	19.44	12.48	401.1145	403.1211	167.0562	HMDB01545	Pyridoxal
4231	21.01	13.65	365.1559	367.1625	131.0975	HMDB01645	L-Norleucine
6018	27.67	19.35	398.1038	400.1105	164.0455	HMDB01713	m-Coumaric acid
955	6.99	2.50	293.1065	295.1104	59.0482	HMDB01842	Guanidine
5125	24.89	16.88	402.0989	404.1047	168.0406	HMDB01868	5-Methoxysalicylic acid
4847	23.85	16.01	372.0905	374.0965	138.0322	HMDB01895	Salicylic acid
2899	15.60	9.53	365.1517	367.1587	131.0934	HMDB01901	Aminocaproic acid
2650	14.29	8.42	337.1279	339.1347	103.0696	HMDB01906	2-Aminoisobutyric acid
1937	10.73	4.84	399.1037	401.1098	165.0453	HMDB02005_2	Methionine Sulfoxide - Isomer
3563	18.54	11.76	360.1080	362.1145	126.0497	HMDB02024	Imidazoleacetic acid
2456	13.24	7.54	364.1681	366.1745	130.1098	HMDB02064	N-Acetylputrescine

5306	25.47	17.36	432.1086	434.1162	198.0503	HMDB02085	Syringic acid
5730	26.89	18.64	400.1199	402.1263	166.0615	HMDB02199	Desaminotyrosine
5195	25.10	17.05	386.1051	388.1114	152.0468	HMDB02390	3-Cresotinic acid
5315	25.51	17.40	327.1166	329.1233	93.0583	HMDB03012	Aniline
2623	14.17	8.32	351.1340	353.1402	117.0757	HMDB03355	5-Aminopentanoic acid
1793	10.35	4.31	380.1293	382.1353	146.0710	HMDB03423	D-Glutamine
3262	17.25	10.78	365.1519	367.1590	131.0935	HMDB03640	Beta-Leucine
2848	15.31	9.29	337.1278	339.1344	103.0695	HMDB03911	3-Aminoisobutanoic acid
2100	11.63	5.95	339.1384	341.1448	105.0801	HMDB04437	Diethanolamine
6320	28.39	20.01	386.1039	388.1107	152.0455	HMDB04815	4-Hydroxy-3-methylbenzoic acid
6229	28.17	19.81	399.1373	401.1437	165.0789	HMDB04992	Benzocaine
7306	31.33	22.69	324.5953	326.6019	181.0739	HMDB06050	o-Tyrosine
3513	18.28	11.55	422.1726	424.1789	188.1142	HMDB28844	Glycyl-Isoleucine
3639	18.88	12.03	456.1557	458.1625	222.0974	HMDB28848	Glycyl-Phenylalanine
3029	16.34	10.16	456.1549	458.1626	222.0966	HMDB28995	Phenylalanyl-Glycine
4523	22.49	14.88	402.0989	404.1055	168.0405	HMDB60003	Isovanillic acid

Table 5.4 List of metabolites detected from 10000 cells identified based on accurate mass and retention time matches to the DnsCl-labeled standard library.

Peak Pair Information						Identification Result	
Peak Pair #	RT(min)	Corrected RT(min)	mz_light	mz_heavy	monoisotopic mass (Da)	HMDB.No.	Name
927	6.77	2.32	403.1419	405.1482	169.0835	HMDB00001	1-Methylhistidine
1273	8.82	3.48	308.1463	310.1530	74.0879	HMDB00002	1,3-Diaminopropane
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB00020	p-Hydroxyphenylacetic acid
1840	10.44	4.36	501.1552	503.1615	267.0969	HMDB00050	Adenosine
2582	13.93	8.12	323.1097	325.1131	89.0513	HMDB00056	Beta-Alanine
3614	18.77	11.94	363.1434	365.1500	129.0850	HMDB00070	D-Pipecolic acid
1671	10.08	4.16	557.1103	559.1147	323.0520	HMDB00095_2	Cytidine monophosphate - Isomer
2806	15.13	9.14	337.1203	339.1274	103.0620	HMDB00112	Gamma-Aminobutyric acid
4599	22.83	15.17	416.1156	418.1220	182.0573	HMDB00118	Homovanillic acid
2170	12.04	6.41	309.0932	311.0994	75.0349	HMDB00123	Glycine
1324	8.99	3.58	381.1100	383.1159	147.0517	HMDB00148	L-Glutamic Acid
3100	16.56	10.31	363.1012	365.1075	129.0429	HMDB00148_2	L-Glutamic Acid - H2O

2203	12.21	6.60	295.1126	297.1190	61.0543	HMDB00149	Ethanolamine
2790	15.06	9.08	370.0972	372.1036	136.0389	HMDB00157_2	Hypoxanthine - multi-tags
3109	16.59	10.33	370.0984	372.1043	136.0400	HMDB00157_3	Hypoxanthine - Isomer
7306	31.33	22.69	324.5953	326.6019	181.0739	HMDB00158	L-Tyrosine
3462	18.02	11.34	399.1362	401.1431	165.0779	HMDB00159	L-Phenylalanine
2460	13.25	7.54	323.1121	325.1152	89.0537	HMDB00161	L-Alanine
2994	16.18	10.02	349.1240	351.1305	115.0657	HMDB00162	L-Proline
3158	16.89	10.54	265.1000	267.1065	31.0416	HMDB00164	Methylamine
2188	12.14	6.52	353.1175	355.1239	119.0592	HMDB00167	L-Threonine
1497	9.54	3.87	366.1108	368.1173	132.0525	HMDB00168	L-Asparagine
4001	20.12	12.95	365.1550	367.1612	131.0967	HMDB00172	L-Isoleucine
5684	26.74	18.51	389.1269	391.1333	155.0686	HMDB00177	L-Histidine
5317	25.52	17.40	307.1104	309.1169	146.1042	HMDB00182	L-Lysine
1118	7.93	3.00	339.0995	341.1059	105.0412	HMDB00187	L-Serine
2004	11.07	5.15	367.0962	369.1026	133.0379	HMDB00191	L-Aspartic Acid
2352	12.87	7.20	422.1719	424.1771	188.1135	HMDB00206	N6-Acetyl-L-Lysine
2632	14.20	8.35	453.1670	455.1736	219.1087	HMDB00210	Pantothenic acid
4630	23.05	15.35	300.1017	302.1072	132.0868	HMDB00214	Ornithine
1189	8.43	3.27	375.0758	377.0821	141.0175	HMDB00224	O-Phosphoethanola

							mine
1096	7.79	2.93	359.0728	361.0790	125.0145	HMDB00251	Taurine
2570	13.83	8.03	323.1114	325.1132	89.0530	HMDB00271	Sarcosine
1346	9.11	3.64	510.1888	512.1947	276.1304	HMDB00279	Saccharopine
4328	21.37	13.95	432.1086	434.1154	198.0503	HMDB00291	Vanillylmandelic acid
2214	12.28	6.67	478.1255	480.1317	244.0672	HMDB00296	Uridine
3517	18.31	11.57	346.0854	348.0919	112.0271	HMDB00300	Uracil
4318	21.33	13.92	372.1007	374.1065	138.0423	HMDB00301	Urocanic acid
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB00440	3-Hydroxyphenylacetic acid
2482	13.34	7.63	422.1719	424.1757	188.1136	HMDB00446	N-Alpha-acetyllysine
2489	13.38	7.66	337.1222	339.1284	103.0638	HMDB00452	L-Alpha-aminobutyric acid
1185	8.42	3.27	403.1410	405.1477	169.0827	HMDB00479	3-methyl-histidine
5125	24.89	16.88	402.0989	404.1047	168.0406	HMDB00484	Vanillic acid
5244	25.24	17.17	372.0913	374.0970	138.0329	HMDB00500	4-Hydroxybenzoic acid
1212	8.57	3.35	408.1687	410.1750	174.1104	HMDB00517	L-Arginine
4001	20.12	12.95	365.1550	367.1612	131.0967	HMDB00557	L-Alloisoleucine
1793	10.35	4.31	380.1293	382.1353	146.0710	HMDB00641	L-Glutamine
2553	13.70	7.92	337.1206	339.1269	103.0623	HMDB00650	D-Alpha-

							aminobutyric acid
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB00669	Ortho-Hydroxyphenylacetic acid
3657	18.93	12.07	365.1489	367.1578	131.0906	HMDB00687	L-leucine
3457	17.99	11.32	383.1112	385.1171	149.0529	HMDB00696	L-Methionine
3692	19.07	12.19	363.1370	365.1457	129.0787	HMDB00716	L-Pipecolic acid
1916	10.67	4.65	353.1164	355.1227	119.0580	HMDB00719	L-Homoserine
2443	13.18	7.48	406.1409	408.1493	172.0826	HMDB00721	Glycylproline
2076	11.47	5.76	365.1155	367.1217	131.0572	HMDB00725	Trans-4-Hydroxyl-L-Proline
4380	21.73	14.25	402.0993	404.1056	168.0409	HMDB00750	3-Hydroxymandelic acid
6958	30.00	21.48	356.0943	358.1002	122.0360	HMDB00750_2	3-Hydroxymandelic acid - COOH
4365	21.62	14.16	416.1142	418.1214	182.0559	HMDB00755	Hydroxyphenyllactic acid
3513	18.28	11.55	422.1726	424.1789	188.1142	HMDB00759	Glycyl-L-Leucine
7303	31.33	22.69	325.5951	327.6013	183.0736	HMDB00819	Normetanephrine
2923	15.75	9.66	429.1191	431.1230	195.0608	HMDB00840	Salicyluric acid
3458	18.00	11.32	351.1402	353.1472	117.0819	HMDB00883	L-Valine
1651	10.01	4.12	409.1526	411.1593	175.0942	HMDB00904	Citrulline

3418	17.88	11.23	438.1490	440.1555	204.0907	HMDB00929	L-Tryptophan
6174	28.02	19.67	428.1141	430.1209	194.0558	HMDB00954	trans-Ferulic acid
5429	25.93	17.77	428.1214	430.1302	194.0630	HMDB00955	Isoferulic acid
1001	7.26	2.64	343.0787	345.0832	109.0204	HMDB00965	Hypotaurine
1734	10.20	4.22	581.1206	583.1267	347.0623	HMDB01044	2'- Deoxyguanosine 5'-monophosphate
5060	24.71	16.73	371.1045	373.1105	137.0462	HMDB01123	2-Aminobenzoic acid
2340	12.82	7.15	365.1160	367.1227	131.0577	HMDB01149	5-Aminolevulinic acid
2186	12.13	6.51	531.1483	533.1542	297.0900	HMDB01173	5'- Methylthioadenos ine
3594	18.68	11.87	401.1148	403.1218	167.0565	HMDB01545	Pyridoxal
3999	20.11	12.95	365.1557	367.1610	131.0974	HMDB01645	L-Norleucine
6018	27.67	19.35	398.1038	400.1105	164.0455	HMDB01713	m-Coumaric acid
772	5.98	1.53	293.1068	295.1108	59.0485	HMDB01842	Guanidine
5432	25.95	17.79	402.1001	404.1065	168.0418	HMDB01868	5- Methoxysalicylic acid
3361	17.69	11.08	365.1517	367.1582	131.0933	HMDB01901	Aminocaproic acid
2489	13.38	7.66	337.1222	339.1284	103.0638	HMDB01906	2- Aminoisobutyric acid

2009	11.10	5.16	399.1037	401.1098	165.0453	HMDB02005	Methionine Sulfoxide
1704	10.16	4.20	399.1045	401.1106	165.0462	HMDB02005_2	Methionine Sulfoxide - Isomer
3547	18.44	11.68	360.1005	362.1068	126.0422	HMDB02024	Imidazoleacetic acid
2333	12.78	7.12	364.1681	366.1745	130.1098	HMDB02064	N-Acetylputrescine
5306	25.47	17.36	432.1086	434.1162	198.0503	HMDB02085	Syringic acid
2700	14.55	8.65	369.1020	371.1109	135.0437	HMDB02108	Methylcysteine
5730	26.89	18.64	400.1199	402.1263	166.0615	HMDB02199	Desaminotyrosine
5284	25.37	17.28	386.1049	388.1108	152.0466	HMDB02390	3-Cresotinic acid
5066	24.73	16.74	327.1161	329.1229	93.0578	HMDB03012	Aniline
6176	28.02	19.67	395.1042	397.1104	161.0459	HMDB03320	Indole-3-carboxylic acid
1250	8.75	3.45	436.1993	438.2055	202.1410	HMDB03334	Symmetric dimethylarginine
2823	15.21	9.20	540.1330	542.1377	612.1494	HMDB03337	Oxidized glutathione
2916	15.70	9.62	351.1313	353.1405	117.0730	HMDB03355	5-Aminopentanoic acid
1793	10.35	4.31	380.1293	382.1353	146.0710	HMDB03423	D-Glutamine
3657	18.93	12.07	365.1489	367.1578	131.0906	HMDB03640	Beta-Leucine

2806	15.13	9.14	337.1203	339.1274	103.0620	HMDB03911	3-Aminoisobutanoic acid
2100	11.63	5.95	339.1384	341.1448	105.0801	HMDB04437	Diethanolamine
6320	28.39	20.01	386.1039	388.1107	152.0455	HMDB04815	4-Hydroxy-3-methylbenzoic acid
6229	28.17	19.81	399.1373	401.1437	165.0789	HMDB04992	Benzocaine
7306	31.33	22.69	324.5953	326.6019	181.0739	HMDB06050	o-Tyrosine
3513	18.28	11.55	422.1726	424.1789	188.1142	HMDB28844	Glycyl-Isoleucine
3029	16.34	10.16	456.1549	458.1626	222.0966	HMDB28848	Glycyl-Phenylalanine
2507	13.45	7.72	408.1526	410.1630	174.0942	HMDB28854	Glycyl-Valine
4364	21.62	14.16	462.2043	464.2116	228.1460	HMDB28937	Leucyl-Proline
2925	15.77	9.68	456.1540	458.1609	222.0957	HMDB28995	Phenylalanyl-Glycine
2993	16.16	10.01	452.1826	454.1892	218.1242	HMDB29043	Serinyl-Leucine
2752	14.85	8.90	486.1671	488.1738	252.1088	HMDB29046	Serinyl-Phenylalanine
4380	21.73	14.25	402.0993	404.1056	168.0409	HMDB60003	Isovanillic acid

5.4 Conclusions

We have developed a CIL nanoLC-MS method for metabolomic profiling of small numbers of cells and demonstrated the metabolic coverage of this method for analyzing the amine/phenol

submetabolome of 100, 1000 and 10000 MCF-7 breast cancer cells. To our knowledge, there is no other method reported in the literature that could match the performance of the described workflow in both quantification accuracy and coverage for analyzing 100 to 10000 mammalian cells. The potential impact of this work is that, when a bioscience researcher working on mammalian cells wishes to perform high-coverage quantitative metabolomics of 100 to 10000 cells, they now have the option of adapting the method described in this paper to do it.

Our research goal was to achieve the highest possible coverage in order to generate metabolome-wide metabolic information required for in-depth biological and biomarker discovery studies. In the case of 10000 cells, we have shown that only a fraction (10%) of the labeled lysate was needed to reach the optimal sample injection in nanoLC-MS for detecting the maximal number of peak pairs or metabolites. In future work, we will consider splitting a 10000-cell lysate into 4 aliquots to analyze, separately, the four submetabolomes (amines/phenols, carboxyls, hydroxyls, and carbonyls) to produce a very comprehensive profile of the cellular metabolome. We envisage that the CIL nanoLC-MS method can become a routine quantitative platform for cellular metabolomics with a starting material of 10000 cells. For analyzing 1000 or 100 cells, even with the injection of almost all the labeled samples, the coverage was found to be decreased to 2091 ± 89 pairs in the 1000-cell lysate and 1620 ± 148 pairs in the 100-cell lysate ($n=6$), compared to 2402 ± 80 pairs found in the 10000-cell lysate. This level of coverage may find to be sufficient in some areas of applications such as partial mapping of the metabolic network or targeted analysis of detectable metabolites. However, future research in improving sample preparation, separation and MS detection including the use of miniaturized devices is needed to maximize the coverage in metabolomics of 1000, 100, or even a lower number of cells.

Chapter 6

High-Performance Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry for Exosome Metabolomics

6.1 Introduction

Exosomes are small (30-120 nm) extracellular vesicles that play an important role in intercellular communication and transmission of macromolecules between cells^{244,245}. For example, blood exosomes contain miRNAs^{246,247}, mRNAs²⁴⁸, proteins²⁴⁹⁻²⁵² and small molecules which are secreted into the bloodstream and travel throughout the body and can transmit their cargo to other cells²⁵³. Consequently, exosomes are important contributing factors in the development of several diseases, including cancer²⁵⁴⁻²⁵⁶. Blood exosomes can be easily obtained by routine blood draws for liquid biopsies. Therefore, they are one of the most promising sources for discovering cancer biomarkers for early detection and diagnosis as well as therapeutic monitoring²⁵⁷⁻²⁵⁹.

Most blood exosome related studies to date have focused on the analysis of RNA²⁶⁰ and proteins²⁶¹⁻²⁶⁵. Because of the possibility of scale-up in culturing cells, lipids²⁶⁶⁻²⁶⁹ and metabolites²⁷⁰⁻²⁷² have been analyzed in exosomes isolated from cancer cell lines. However, to our knowledge, there is no report of untargeted metabolomics study of blood exosomes with high coverage. The main issue lies in the sensitivity of most analytical platforms which is not sufficient to realize high-coverage profiling of the exosome metabolome from the limited amount of material generally available after isolation. The total amount of exosomes is generally less than 1 μg from several mL of patient serum sample. Most commonly used metabolomics

platforms such as microflow LC-MS, GC-MS and NMR cannot provide sufficient sensitivity for comprehensive metabolomic profiling of exosomes.

In a previous study, we developed a chemical isotope labeling nanoflow liquid chromatography technique coupled with captivespray ionization mass spectrometry (CIL nLC-MS) to profile the metabolome of small numbers of breast cancer cells¹⁵. Even with the use of 100 cells as the starting material, we were able to detect thousands of metabolites. Encouraged by this high performance of analyzing trace amounts of sample, we set out to develop a metabolomics workflow for analyzing a small amount of exosome material obtained from patient serum. In this report, we describe the workflow tailored to exosome isolation and analysis using ultra-centrifugation and CIL nLC-MS, demonstrate the possibility of profiling exosome metabolites with unprecedented coverage, and show a comparative analysis of pancreatic cancer patient samples obtained before and after chemotherapy.

6.2 Experimental Section

6.2.1 Workflow

The overall workflow for exosome isolation and metabolomic profiling is shown in Figure 6.1. Two sets of whole blood samples from pancreatic cancer patients were collected before and after chemotherapy (n=10 for each set). The serum was separated out by centrifugation. After dilution, multiple cycles of ultra-centrifugation were performed to pellet the exosomes from serum. The pellets were cleaned by PBS to remove serum metabolites. The exosomes were then lysed and extracted by 50% MeOH with freeze-thaw cycles (5 times). Each sample was divided into two aliquots: one as sample and another one for generating a pooled sample. An individual sample was labeled by ¹²C-dansyl chloride (DnsCl) (light tag), and the pooled sample was labeled by ¹³C-DnsCl (heavy tag), which served as an internal standard. The ¹²C-dansyl labeled

individual sample and ^{13}C -dansyl labeled pool were mixed together. The mixture was injected onto nanoflow LC-MS. The light and heavy labeled metabolites showed up as peak pairs with a 2.0067 Dalton difference in the mass spectra for single-tag-labeled metabolites. The relative ratio of metabolites could be determined by using the heavy peak as reference. Univariate and multivariate analyses were performed, and metabolites were identified or matched by searching against different libraries in MyCompoundID (MCID) (www.mycompoundid.org)¹⁵¹.

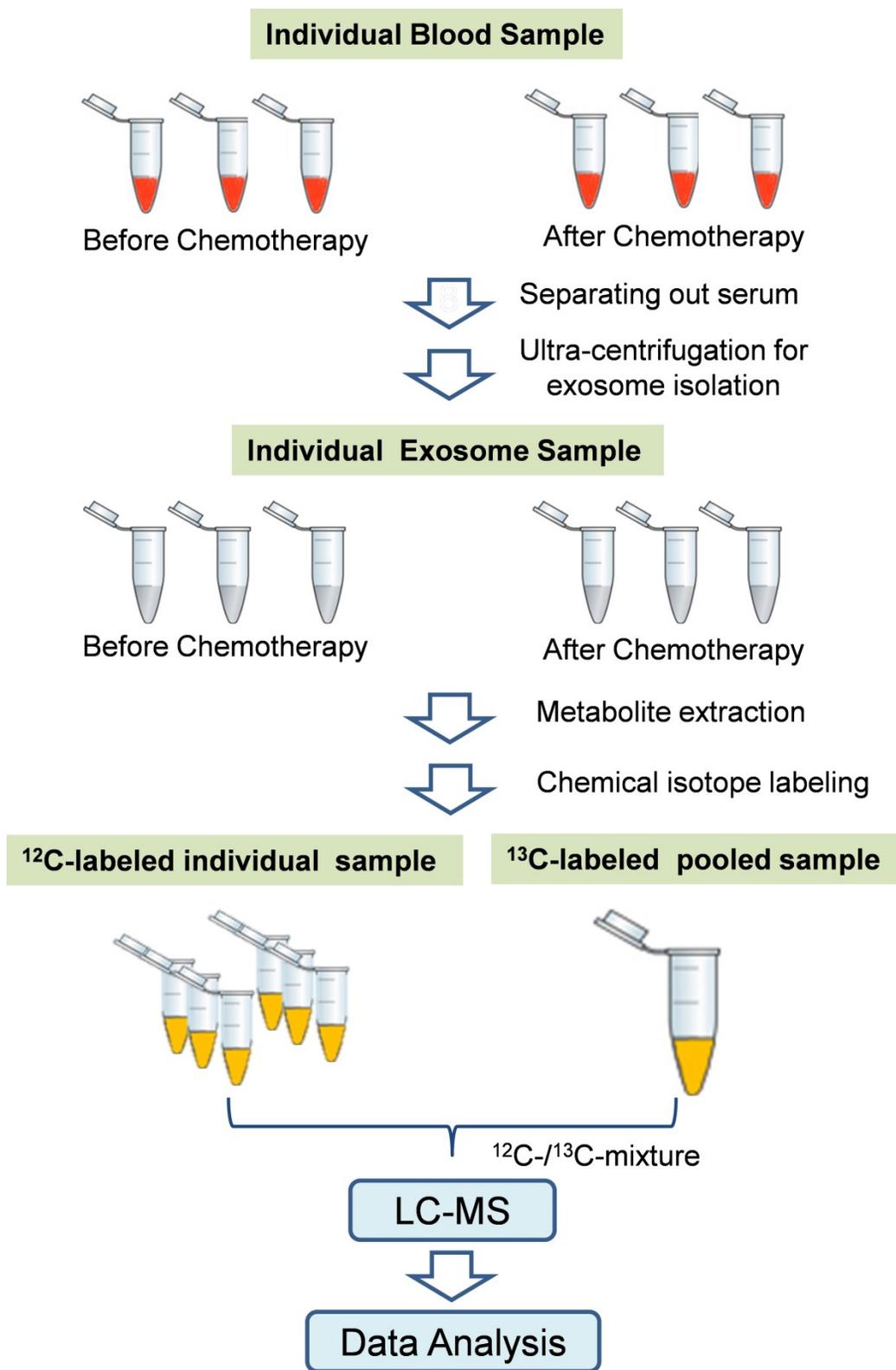


Figure 6.1 Workflow for exosome metabolomics based on CIL nLC-MS.

6.2.2 Serum Samples

Whole blood samples were obtained at the University of Michigan Hospital, Ann Arbor, Michigan, with Institutional Review Board approval. These include samples from ten patients with locally advanced pancreatic cancer prior to treatment and after three-week chemotherapy with gemcitabine in combination with Wee1 inhibitor (AZD1775). Intravenous administration of gemcitabine at a dose of 1000 mg/m² was done on day 1 and day 8 of a three-week cycle. AZD1775 was taken orally on day 1, 2, and day 8, 9 of each three-week cycle. Blood samples were centrifuged at 500×g for 10 min to separate out serum. All serum samples were stored at -80 °C until analysis.

6.2.3 Serum Exosome Isolation

To isolate the exosomes, 2 mL of serum from each patient was used. 2 mL of PBS (AppliChem, St. Louis, MO) was added to a serum sample to decrease the viscosity, followed by centrifugation at 2,000×g for 10 min and then 10,000×g for 30 min at 4 °C to remove cell debris. The supernatant was transferred into an Ultra-Clear™ tube (Beckman Coulter, Indianapolis, IN) and centrifuged at 100,000×g for 120 min at 4 °C using a Beckman Optima XL-70 Ultracentrifuge. Part of the supernatant was removed using a pipette, leaving 2 mL of supernatant remaining above the pellets. To clean the exosomes, 4 mL of PBS was added to the pellets containing exosomes and then centrifuged at 100,000×g for 70 min at 4 °C, followed by partial removal of the supernatant. This step was repeated 4 times to thoroughly clean the exosomes.

6.2.4 Metabolites Extraction and Chemical Isotope Labeling

The exosomes were extracted by 50% MeOH, and freeze-thaw cycles were carried out to assist the releasing of metabolites from exosomes during extraction. 200 μL of 50% MeOH were added into the vials which contained exosome pellets. The vials were placed in liquid nitrogen for 1 min, and then thawed on an ice-bath for 1 min. This procedure was repeated for 5 times. After extraction, the lysates were dried down in a SpeedVac. The lysates were re-dissolved in 17 μL of $\text{Na}_2\text{CO}_3/\text{NaHCO}_3$ buffer (250 mM; pH 9.4). 7.5 μL of lysates were aliquoted into a 0.6 mL vial for ^{12}C -dansyl chloride labeling. Another aliquot of 7.5 μL was taken to generate a pooled sample for ^{13}C -dansyl chloride labeling. For labeling, 7.5 μL of dansyl chloride dissolved in ACN at a concentration of 0.25 mg/mL was added to the lysate vial, followed by incubation for 1 h at 40°C. 1 μL of 250 mM NaOH was added to quench the excess DnsCl, and 5 μL of 425 mM formic acid was added to acidify the reaction mixture.

6.2.5 LC-MS

For the nanoflow LC-MS (nLC-MS) setup, a Waters NanoAcquity ultra-performance liquid chromatograph (UPLC) (Milford, MA) was connected to a Bruker Impact quadruple time-of-flight mass spectrometer (Q-TOF) (Billerica, MA) with a captivespray ion source¹⁵. An Acclaim PepMap 100 trap column (75 $\mu\text{m} \times 20$ mm, 3 μm) (Thermo Scientific, Sunnyvale, CA) was used for trapping the labeled metabolites prior to injection into an Acclaim PepMap RSLC C18 analytical column (75 $\mu\text{m} \times 150$ mm, 2 μm) for analytical separation. The mobile phase A used was 0.1% formic acid in water and the mobile phase B was 0.1% formic acid in acetonitrile. Before injection, the ^{12}C -/ ^{13}C -labeled extracts were combined, dried down, and reconstituted by 9:1 $\text{H}_2\text{O}:\text{ACN}$. A 2-min-trapping procedure was performed prior to sample loading onto the analytical column. The trapping solvent was 99% mobile phase A. The trapping flow rate was 7 $\mu\text{L}/\text{min}$. The chromatographic conditions were: $t = 0$ min, 15% B; $t = 2.0$ min, 15% B; $t = 4.0$

min, 25% B; t =24 min, 60% B; t = 28 min, 90% B, t = 45 min, 90% B. The flow rate was 350 nL/min. The capillary spray operation conditions were: dry temperature, 200 °C; dry gas, 3 L/min; capillary voltage, 1400 V; nanoBooster, 0.2 bar, and dopant gas was pure ACN.

6.3 Results and Discussion

6.3.1 Exosomes Isolation

In order to eliminate serum contamination, five rounds of ultra-centrifugation were applied to purify the exosomes as described previously²⁷³. We chose this UC method because we know from prior work that it is a robust and reliable method for eliminating high abundance proteins and other contaminants from serum for proteomics research²⁷³. Other methods including antibody-based and size-based isolation techniques²⁷⁴ are not evaluated in this study; however, it will be useful, in the future, to compare different exosome isolation methods tailored to metabolomic profiling. In this work, the isolated exosomes were verified by transmission electron microscopy (TEM) and protein markers. The TEM images showed that the diameters of more than 90% of the vesicles were between 30 nm and 100 nm, where the median diameter was around 70 nm²⁷⁵. The exosome markers including CD9, CD63, CD81 and TSG101 were identified in the proteomics mass spectrometry data²⁶¹.

6.3.2 Metabolite Extraction

As in cellular metabolomics, a lysis step is required before extracting metabolites from exosomes. In previous work, we investigated the efficiency and performance of different methods for lysing different types of cells^{6,32,276,277}. We found that for mammalian cells, the freeze-thaw-cycle lysis method was efficient. Considering that exosomes are membrane-bound phospholipid vesicles, this method was therefore selected for exosome lysis in this study. We

also investigated the selection of extraction solvents for cellular metabolomics in previous studies^{6,15,32}. For extraction of prokaryotic microbes (e.g., *E. coli*), eukaryotic microbes (e.g., *S. cerevisiae*) or mammalian cells (e.g., MCF-7 breast cancer cell), we found that 50% MeOH always provided the best extraction efficiency, while 50% ACN resulted in less extraction efficiency than 50% MeOH, and 1:1:1 ACN: MeOH: H₂O resulted in the lowest efficiency. Therefore, 50% MeOH was selected as the extraction solvent in this work.

It should be noted that the presence of residual proteins in a sample does not cause a problem in CIL LC-MS, as residual proteins would precipitate out during the labeling reaction. After centrifugation, the supernatant was used for mixing with a labeled pool and the mixture was then injected into LC-MS. We did not encounter a problem of column clogging in CIL nLC-MS. In this work, there is no desalting step before sample injection. This is because we have already used a very low concentration of the labeling reagent (0.25 mg/mL), and the amount of the buffer used is also small for the labeling reaction. There is a trapping process before the sample is loaded to the analytical column for separation; during trapping, the mobile phase which contains mainly water can remove most of the salts. Overall, the nLC-MS system is robust (we use captivespray source) and the column life time is often more than 2000 injections per column.

6.3.3 nLC-MS

Figure 6.2(A) shows the total ion chromatogram of dansyl labeled exosomes obtained by nLC-MS. A sodium formate peak appears at the very beginning of the chromatogram. Sodium formate was produced when formic acid was used to consume excess NaOH during the labeling step. The cluster peaks of sodium formate were used for mass calibration in each LC run. After the sodium formate peak, a small dansyl hydroxyl (Dns-OH) peak is observed. Dns-OH was

produced from the labeling reaction quenching step. Most of the dansyl labeled metabolites eluted between 7 min to 32 min, where the 45 min gradient was sufficient for CIL nLC-MS. As the ion chromatogram shows, many peaks from labeled metabolites are detected across the entire gradient elution time window.

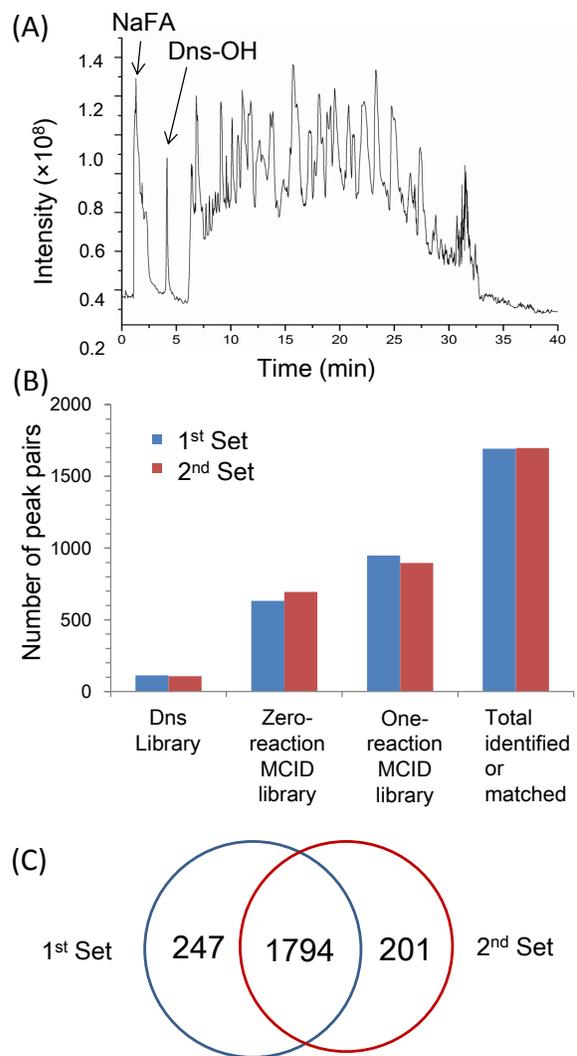


Figure 6.2 (A) Total ion chromatogram of a representative ^{12}C -/ ^{13}C -labeled exosome sample. (B) Number of peak pairs identified or matched in three libraries. (C) Venn diagram showing the commonly detected peak pairs from 1st sample set and 2nd sample set.

6.3.4 Exosome Metabolome

In CIL nLC-MS, the ^{12}C -/ ^{13}C -labeled metabolites are detected as peak pairs, which are different from the singlet peaks arising from background chemical noises. IsoMS¹⁰⁸ was used to pick the peak pairs, remove redundant pairs (e.g., adduct ions, dimers, etc.), align peak pairs from different samples, and determine the intensity ratios of peak pairs for relative quantification. For multiple labeled metabolites, IsoMS can determine the charge number and tag number (e.g., 2 tags with 1 charge) and filter out the redundant peak pairs if any. In general, only one peak pair, $[\text{M}+\text{H}]^+$, was retained for one metabolite and thus the number of peak pairs detected reflects the number of metabolites detected. We analyzed two separate sets of exosome samples to illustrate the technical reproducibility and robustness in terms of the number of peak pairs detectable. The average number of peak pairs detected from the 1st set and 2nd set of exosomes was 1964 ± 60 ($n=10$) and 1948 ± 117 ($n=10$), respectively, with many peak pairs commonly detected in the two sets (see the Venn diagram in Figure 6.2C; only the peak pairs that showed up in more than 50% of the LC-MS runs were included for comparison). The total number of peak pairs detected were 2446 in the 1st sample set ($n=10$) and 2511 in the 2nd sample set ($n=10$). These numbers from the two sets of samples are very consistent, indicating the robustness of the workflow. The large number of metabolites detected from exosomes reflects the complexity of the metabolome contained in exosomes.

Figure 6.2B shows a summary of the number of metabolites identified or matched using different metabolite databases. To identify the metabolite, the peak pair mass and retention time were searched against a dansyl standard library comprised of 304 human endogenous amine/phenol metabolites (the retention time tolerance was set at 1 min and the mass tolerance was set at 10 ppm). There were 101 metabolites identified from the first sample set (Table 6.1),

and 94 metabolites were identified from the second set (Table 6.2); many of them are in common from the two sets. They include most of the common amino acids, some dipeptides, and other metabolites. We also searched the MCID zero-reaction library (8021 known human endogenous metabolites) for putative structure assignments based on accurate mass match. 632 metabolites were matched in the first sample set (In supplemental information which is available from Dr. Liang Li), and 694 metabolites were matched in the second sample set (In supplemental information which is available from Dr. Liang Li). Lastly, we searched the masses against the MCID one-reaction library (i.e., 375,809 predicted human metabolites from one metabolic reaction of known metabolites). In the first sample set, 948 additional metabolites were matched with the predicted metabolites library (In supplemental information which is available from Dr. Liang Li), and 896 metabolites were matched with the predicted metabolites library in the second sample set (In supplemental information which is available from Dr. Liang Li). In total, 1681 out of 2446 peak pairs peak pairs (69%) were identified or matched in the first sample set, and 1675 out of 2511 peak pairs (68%) were identified or matched in the second sample set.

Table 6.1 List of peak pairs detected from the 1st set of exosomes with positive metabolite identification based on accurate mass and retention time search against the dansyl standard library.

Peak Pair Information					Library Information	
Peak Pair #	Corrected T_R (min)	mz_light	mz_heavy	Monoisotopic mass (Da)	HMDB No.	Name
748	2.20	308.1412	310.1475	74.0829	HMDB00002	1,3-Diaminopropane
913	2.46	408.1689	410.1756	174.1106	HMDB00517	L-Arginine
917	2.47	403.1416	405.1479	169.0833	HMDB00001	1-Methylhistidine
917	2.47	403.1416	405.1479	169.0833	HMDB00479	3-Methyl-histidine
1176	3.21	517.1498	519.1564	283.0915	HMDB00133	Guanosine
1181	3.22	293.1057	295.1114	59.0473	HMDB01842	Guanidine
1373	3.60	366.1110	368.1175	132.0526	HMDB00168	L-Asparagine
1453	3.70	409.1152	411.1224	175.0569	HMDB03157	Guanidinosuccinic acid
1554	3.90	380.1275	382.1340	146.0692	HMDB00641	L-Glutamine
1554	3.90	380.1275	382.1340	146.0692	HMDB03423	D-Glutamine
1597	3.99	409.1530	411.1596	175.0947	HMDB00904	Citrulline
1642	4.08	501.1548	503.1614	267.0965	HMDB00050	Adenosine
1688	4.23	353.1163	355.1228	119.0580	HMDB00719	L-Homoserine
1739	4.37	399.1037	401.1101	165.0453	HMDB02005_2	Methionine Sulfoxide - Isomer
1755	4.46	339.1016	341.1083	105.0433	HMDB00187	L-Serine
1838	4.99	581.1211	583.1277	347.0628	HMDB01044	2'-Deoxyguanosine 5'-monophosphate
1853	5.05	381.1114	383.1182	147.0531	HMDB00148	L-Glutamic Acid
1855	5.05	423.1682	425.1749	189.1099	HMDB00679	Homocitrulline
1873	5.06	365.1146	367.1222	131.0563	HMDB00725	Trans-4-Hydroxyl-L-Proline
2033	5.11	422.1747	424.1812	188.1164	HMDB00206	N6-Acetyl-L-Lysine
2056	5.12	339.1379	341.1445	105.0795	HMDB04437	Diethanolamine
2087	5.13	295.1109	297.1176	61.0526	HMDB00149	Ethanolamine
2173	5.17	367.0962	369.1027	133.0379	HMDB00191	L-Aspartic Acid
2216	5.98	353.1169	355.1233	119.0585	HMDB00167	L-Threonine
2236	6.24	395.1274	397.1341	161.0690	HMDB00510	Amino adipic acid
2274	6.54	323.1061	325.1129	89.0478	HMDB00056	Beta-Alanine

2282	6.59	309.0905	311.0982	75.0322	HMDB00123	Glycine
2293	6.69	364.1683	366.1746	130.1100	HMDB02064	N-Acetylputrescine
2328	7.11	406.1420	408.1492	172.0837	HMDB00721	Glycylproline
2383	7.60	323.1071	325.1138	89.0488	HMDB00161	L-Alanine
2405	7.69	337.1225	339.1292	103.0642	HMDB00112	Gamma-Aminobutyric acid
2430	7.78	381.1126	383.1188	147.0543	HMDB02393	N-methyl-D-aspartic acid
2480	7.96	478.1270	480.1338	244.0687	HMDB00296	Uridine
2759	8.71	386.0908	388.0967	152.0325	HMDB00292	Xanthine
2796	8.79	351.1374	353.1440	117.0791	HMDB03355	5-Aminopentanoic acid
2798	8.80	452.1835	454.1898	218.1251	HMDB29043	Serinyl-Leucine
2897	9.06	370.0963	372.1027	136.0380	HMDB00157_2	Hypoxanthine - multi-tags
2947	9.21	265.1025	267.1113	31.0442	HMDB00164	Methylamine
2967	9.27	337.1233	339.1303	103.0650	HMDB00452	L-Alpha-aminobutyric acid
2967	9.27	337.1233	339.1303	103.0650	HMDB00650	D-Alpha-aminobutyric acid
2967	9.27	337.1233	339.1303	103.0650	HMDB01906	2-Aminoisobutyric acid
2967	9.27	337.1233	339.1303	103.0650	HMDB03911	3-Aminoisobutanoic acid
2988	9.43	323.1059	325.1130	89.0476	HMDB00271	Sarcosine
3039	9.58	486.1677	488.1733	252.1094	HMDB29046	Serinyl-Phenylalanine
3132	9.80	337.1222	339.1292	103.0639	HMDB00650	D-Alpha-aminobutyric acid
3154	9.85	456.1571	458.1635	222.0987	HMDB28995	Phenylalanyl-Glycine
3157	9.86	369.0931	371.0996	135.0348	HMDB02108	Methylcysteine
3158	9.86	413.1000	415.1075	179.0417	HMDB00704	Isoxanthopterin
3158	9.86	413.1000	415.1075	179.0417	HMDB00704_2	Isoxanthopterin - Isomer
3202	9.96	422.1730	424.1793	188.1147	HMDB28844	Glycyl-Isoleucine
3211	9.99	370.0970	372.1037	136.0387	HMDB00157_3	Hypoxanthine - Isomer
3246	10.08	466.1991	468.2041	232.1408	HMDB29065	Threoninyl-Leucine
3250	10.09	363.1008	365.1074	129.0425	HMDB00148_2	L-Glutamic Acid - H2O
3295	10.22	349.1243	351.1316	115.0659	HMDB00162	L-Proline
3318	10.29	365.1526	367.1593	131.0943	HMDB01901	Aminocaproic acid
3456	10.66	365.1511	367.1583	131.0928	HMDB03640	Beta-Leucine
3472	10.68	436.1888	438.1951	202.1305	HMDB28691	Alanyl-Leucine
3523	10.81	351.1384	353.1452	117.0801	HMDB00883	L-Valine

3645	10.93	383.1091	385.1158	149.0508	HMDB00696	L-Methionine
3725	11.28	346.0854	348.0920	112.0271	HMDB00300	Uracil
3823	11.52	438.1478	440.1542	204.0895	HMDB00929	L-Tryptophan
4061	12.20	456.1568	458.1624	222.0985	HMDB28848	Glycyl-Phenylalanine
4102	12.27	401.1139	403.1218	167.0555	HMDB01545	Pyridoxal
4108	12.28	363.1361	365.1436	129.0777	HMDB00070	D-Pipecolic acid
4162	12.40	470.1724	472.1785	236.1141	HMDB28694	Alanyl-Phenylalanine
4175	12.43	365.1529	367.1601	131.0945	HMDB00557	L-Alloisoleucine
4193	12.47	363.1374	365.1441	129.0791	HMDB00716	L-Pipecolic acid
4310	12.78	399.1370	401.1435	165.0787	HMDB00159	L-Phenylalanine
4335	12.86	351.1372	353.1435	117.0789	HMDB02141	N-Methyl- α -aminoisobutyric acid
4388	13.05	462.2037	464.2107	228.1454	HMDB28937	Leucyl-Proline
4398	13.07	365.1560	367.1615	131.0977	HMDB00172	L-Isoleucine
4445	13.16	496.1888	498.1953	262.1305	HMDB11177	L-phenylalanyl-L-proline
4679	13.61	365.1551	367.1608	131.0968	HMDB00687	L-leucine
4769	13.82	372.1006	374.1073	138.0423	HMDB00301	Urocanic acid
4873	14.15	365.1540	367.1605	131.0957	HMDB01645	L-Norleucine
4907	14.22	302.0947	304.1028	68.0364	HMDB01525	Imidazole
5336	15.22	416.1147	418.1211	182.0564	HMDB00755	Hydroxyphenyllactici acid
5623	15.76	512.2207	514.2272	278.1624	316	Phenyl-Leucine
5727	16.13	402.0995	404.1059	168.0411	HMDB01868	5-Methoxysalicylic acid
5727	16.13	402.0995	404.1059	168.0411	HMDB60003	Isovanillic acid
5758	16.21	371.1048	373.1120	137.0465	HMDB01123	2-Aminobenzoic acid
5806	16.33	385.1175	387.1271	151.0592	HMDB01859	Acetaminophen
5870	16.43	300.1019	302.1084	132.0871	HMDB00214	Ornithine
5878	16.45	546.2072	548.2138	312.1488	HMDB13302	Phenylalanylphenylalanine
5896	16.48	512.2206	514.2269	278.1623	HMDB13243	Leucyl-phenylalanine
5912	16.52	416.1133	418.1207	182.0549	HMDB00118	Homovanillic acid
6063	16.76	386.1048	388.1111	152.0464	HMDB00020	p-Hydroxyphenylacetic acid
6063	16.76	386.1048	388.1111	152.0464	HMDB00440	3-Hydroxyphenylacetic acid
6063	16.76	386.1048	388.1111	152.0464	HMDB00669	Ortho-Hydroxyphenylacetic acid
6063	16.76	386.1048	388.1111	152.0464	HMDB02390	3-Cresotinic acid
6096	16.82	327.1156	329.1219	93.0573	HMDB03012	Aniline
6281	17.24	372.0893	374.0959	138.0310	HMDB00500	4-Hydroxybenzoic acid

6331	17.32	402.1005	404.1070	168.0421	HMDB00484	Vanillic acid
6411	17.47	307.1107	309.1174	146.1048	HMDB00182	L-Lysine
6417	17.48	428.1143	430.1218	194.0560	HMDB00955	Isoferulic acid
6432	17.54	372.0904	374.0970	138.0321	HMDB00500	4-Hydroxybenzoic acid
6433	17.54	385.1297	387.1401	151.0714	HMDB00239_2	Pyridoxine - H2O
6619	18.09	400.1203	402.1274	166.0620	HMDB02199	Desaminotyrosine
6708	18.37	389.1278	391.1342	155.0695	HMDB00177	L-Histidine
6748	18.50	428.1153	430.1223	194.0570	HMDB00954	trans-Ferulic acid
6766	18.56	397.1458	399.1504	163.0875	HMDB00473	6-Dimethylaminopurine
6776	18.60	460.1775	462.1842	226.1192	HMDB28689	Alanyl-Histidine
6839	18.78	398.1049	400.1121	164.0465	HMDB01713	m-Coumaric acid
7080	19.34	355.1464	357.1519	121.0881	HMDB02017	1-Phenylethylamine
7170	19.58	386.1046	388.1109	152.0463	HMDB04815	4-Hydroxy-3-methylbenzoic acid
7822	21.40	356.0953	358.1020	122.0370	HMDB00750_2	3-Hydroxymandelic acid - COOH
8090	22.05	278.1077	280.1149	88.0987	HMDB01414	1,4-diaminobutane
8203	22.32	285.1153	287.1211	102.1141	HMDB02322	Cadaverine
8268	22.46	373.0840	375.0916	139.0257	HMDB01232	4-Nitrophenol
8345	22.67	324.5945	326.6009	181.0723	HMDB00158	L-Tyrosine
8345	22.67	324.5945	326.6009	181.0723	HMDB06050	o-Tyrosine
8393	22.81	328.0995	330.1058	94.0412	HMDB00228	Phenol

Table 6.2 List of peak pairs detected from the 2nd set of exosomes with positive metabolite identification based on accurate mass and retention time search against the dansyl standard library.

Peak Pair Information					Library Information	
Peak Pair #	Corrected T _R (min)	mz_light	mz_heavy	Monoisotopic mass (Da)	HMDB No.	Name
847	2.36	308.1426	310.1477	74.0843	HMDB00002	1,3-Diaminopropane
913	2.46	408.1689	410.1756	174.1106	HMDB00517	L-Arginine
917	2.47	403.1416	405.1479	169.0833	HMDB00001	1-Methylhistidine
					HMDB00479	3-methyl-histidine
1176	3.21	517.1498	519.1564	283.0915	HMDB00133	Guanosine
1181	3.22	293.1057	295.1114	59.0473	HMDB01842	Guanidine
1373	3.60	366.1110	368.1175	132.0526	HMDB00168	L-Asparagine
1453	3.70	409.1152	411.1224	175.0569	HMDB03157	Guanidinosuccinic acid
1554	3.90	380.1275	382.1340	146.0692	HMDB00641	Glutamine
1597	3.99	409.1530	411.1596	175.0947	HMDB00904	Citrulline
1642	4.08	501.1548	503.1614	267.0965	HMDB00050	Adenosine
1688	4.23	353.1163	355.1228	119.0580	HMDB00719	L-Homoserine
1739	4.37	399.1037	401.1101	165.0453	HMDB02005	Methionine Sulfoxide
1739	4.37	399.1037	401.1101	165.0453	HMDB02005_2	Methionine Sulfoxide - Isomer
1755	4.46	339.1016	341.1083	105.0433	HMDB00187	L-Serine
1838	4.99	581.1211	583.1277	347.0628	HMDB01044	2'-Deoxyguanosine 5'-monophosphate
1853	5.05	381.1114	383.1182	147.0531	HMDB00148	L-Glutamic Acid
1855	5.05	423.1682	425.1749	189.1099	HMDB00679	Homocitrulline
1873	5.06	365.1146	367.1222	131.0563	HMDB00725	Trans-4-Hydroxyl-L-Proline
2033	5.11	422.1747	424.1812	188.1164	HMDB00206	N6-Acetyl-L-Lysine
2056	5.12	339.1379	341.1445	105.0795	HMDB04437	Diethanolamine
2066	5.12	395.1250	397.1322	161.0667	HMDB00510	Amino adipic acid
2087	5.13	295.1109	297.1176	61.0526	HMDB00149	Ethanolamine
2173	5.17	367.0962	369.1027	133.0379	HMDB00191	L-Aspartic Acid
2216	5.98	353.1169	355.1233	119.0585	HMDB00167	L-Threonine

2233	6.21	339.1376	341.1441	105.0793	HMDB04437	Diethanolamine
2236	6.24	395.1274	397.1341	161.0690	HMDB00510	Amino adipic acid
2282	6.59	309.0905	311.0982	75.0322	HMDB00123	Glycine
2293	6.69	364.1683	366.1746	130.1100	HMDB02064	N-Acetylputrescine
2328	7.11	406.1420	408.1492	172.0837	HMDB00721	Glycylproline
2383	7.60	323.1071	325.1138	89.0488	HMDB00056	Beta-Alanine
2383	7.60	323.1071	325.1138	89.0488	HMDB00161	L-Alanine
2405	7.69	337.1225	339.1292	103.0642	HMDB00112	Gamma-Aminobutyric acid
2430	7.78	381.1126	383.1188	147.0543	HMDB02393	N-methyl-D-aspartic acid
2480	7.96	478.1270	480.1338	244.0687	HMDB00296	Uridine
2759	8.71	386.0908	388.0967	152.0325	HMDB00292	Xanthine
2796	8.79	351.1374	353.1440	117.0791	HMDB03355	5-Aminopentanoic acid
2798	8.80	452.1835	454.1898	218.1251	HMDB29043	Serinyl-Leucine
2897	9.06	370.0963	372.1027	136.0380	HMDB00157_2	Hypoxanthine - multi-tags
2947	9.21	265.1025	267.1113	31.0442	HMDB00164	Methylamine
2967	9.27	337.1233	339.1303	103.0650	HMDB00452	L-Alpha-aminobutyric acid
					HMDB00650	D-Alpha-aminobutyric acid
					HMDB01906	2-Aminoisobutyric acid
					HMDB03911	3-Aminoisobutanoic acid
2988	9.43	323.1059	325.1130	89.0476	HMDB00271	Sarcosine
3154	9.85	456.1571	458.1635	222.0987	HMDB28995	Phenylalanyl-Glycine
3157	9.86	369.0931	371.0996	135.0348	HMDB02108	Methylcysteine
3158	9.86	413.1000	415.1075	179.0417	HMDB00704	Isoxanthopterin
					HMDB00704_2	Isoxanthopterin - Isomer
3202	9.96	422.1730	424.1793	188.1147	HMDB28844	Glycyl-Isoleucine
3211	9.99	370.0970	372.1037	136.0387	HMDB00157_3	Hypoxanthine - Isomer
3246	10.08	466.1991	468.2041	232.1408	HMDB29065	Threoninyl-Leucine
3295	10.22	349.1243	351.1316	115.0659	HMDB00162	L-Proline
3318	10.29	365.1526	367.1593	131.0943	HMDB01901	Aminocaproic acid
3456	10.66	365.1511	367.1583	131.0928	HMDB03640	Beta-Leucine
3472	10.68	436.1888	438.1951	202.1305	HMDB28691	Alanyl-Leucine

3523	10.81	351.1384	353.1452	117.0801	HMDB00883	L-Valine
3645	10.93	383.1091	385.1158	149.0508	HMDB00696	L-Methionine
3725	11.28	346.0854	348.0920	112.0271	HMDB00300	Uracil
3823	11.52	438.1478	440.1542	204.0895	HMDB00929	L-Tryptophan
4061	12.20	456.1568	458.1624	222.0985	HMDB28848	Glycyl-Phenylalanine
4102	12.27	401.1139	403.1218	167.0555	HMDB01545	Pyridoxal
4162	12.40	470.1724	472.1785	236.1141	HMDB28694	Alanyl-Phenylalanine
4193	12.47	363.1374	365.1441	129.0791	HMDB00070	Pipelic acid
4310	12.78	399.1370	401.1435	165.0787	HMDB00159	L-Phenylalanine
4335	12.86	351.1372	353.1435	117.0789	HMDB02141	N-Methyl-a-aminoisobutyric acid
4388	13.05	462.2037	464.2107	228.1454	HMDB28937	Leucyl-Proline
4398	13.07	365.1560	367.1615	131.0977	HMDB00172	L-Isoleucine
4445	13.16	496.1888	498.1953	262.1305	HMDB11177	L-phenylalanyl-L-proline
4490	13.26	365.1550	367.1619	131.0966	HMDB00557	L-Alloisoleucine
4519	13.32	365.1561	367.1611	131.0978	HMDB00687	L-leucine
4769	13.82	372.1006	374.1073	138.0423	HMDB00301	Urocanic acid
4873	14.15	365.1540	367.1605	131.0957	HMDB01645	L-Norleucine
4907	14.22	302.0947	304.1028	68.0364	HMDB01525	Imidazole
5336	15.22	416.1147	418.1211	182.0564	HMDB00755	Hydroxyphenyllactici acid
5623	15.76	512.2207	514.2272	278.1624	316	Phenyl-Leucine
5727	16.13	402.0995	404.1059	168.0411	HMDB01868	5-Methoxysalicylic acid
					HMDB60003	Isovanillic acid
5758	16.21	371.1048	373.1120	137.0465	HMDB01123	2-Aminobenzoic acid
5806	16.33	385.1175	387.1271	151.0592	HMDB01859	Acetaminophen
5870	16.43	300.1019	302.1084	132.0871	HMDB00214	Ornithine
5878	16.45	546.2072	548.2138	312.1488	HMDB13302	Phenylalanylphenylalanine
5896	16.48	512.2206	514.2269	278.1623	HMDB13243	Leucyl-phenylalanine
5912	16.52	416.1133	418.1207	182.0549	HMDB00118	Homovanillic acid
6063	16.76	386.1048	388.1111	152.0464	HMDB00020	p-Hydroxyphenylacetic acid
					HMDB00440	3-Hydroxyphenylacetic acid
					HMDB00669	Ortho-Hydroxyphenylacetic acid
					HMDB02390	3-Cresotinic acid

6096	16.82	327.1156	329.1219	93.0573	HMDB03012	Aniline
6281	17.24	372.0893	374.0959	138.0310	HMDB00500	4-Hydroxybenzoic acid
6331	17.32	402.1005	404.1070	168.0421	HMDB00484	Vanillic acid
6411	17.47	307.1107	309.1174	146.1048	HMDB00182	L-Lysine
6417	17.48	428.1143	430.1218	194.0560	HMDB00955	Isoferulic acid
6503	17.72	397.13390 5	399.137580 9	163.0755854	HMDB00473	6-Dimethylaminopurine
6619	18.09	400.1203	402.1274	166.0620	HMDB02199	Desaminotyrosine
6748	18.50	428.1153	430.1223	194.0570	HMDB00954	trans-Ferulic acid
7170	19.58	386.1046	388.1109	152.0463	HMDB04815	4-Hydroxy-3-methylbenzoic acid
7822	21.40	356.0953	358.1020	122.0370	HMDB00750_ 2	3-Hydroxymandelic acid - COOH
8090	22.05	278.1077	280.1149	88.0987	HMDB01414	1,4-diaminobutane
8203	22.32	285.1153	287.1211	102.1141	HMDB02322	Cadaverine
8268	22.46	373.0840	375.0916	139.0257	HMDB01232	4-Nitrophenol
8345	22.67	324.5945	326.6009	181.0723	HMDB00158	L-Tyrosine
					HMDB06050	o-Tyrosine
8393	22.81	328.0995	330.1058	94.0412	HMDB00228	Phenol

The above results indicate that the serum exosome metabolome contains many metabolites with diverse structures and CIL nLC-MS has the sensitivity required to detect a large number of metabolites. To avoid serum metabolite contamination in the exosome metabolome, we used extensive washing (5 rounds), in combination with ultra-centrifugation, to isolate and purify the exosomes from serum. In addition, metabolites were detected after lysis of exosomes using five freeze-thaw cycles. Due to the very small amounts of purified samples available for analysis, nLC-MS with rationally designed chemical isotope labeling of metabolites is needed to achieve high metabolome coverage. We note a report of metabolome analysis of exosomes-like vesicles (ELVs) isolated from human plasma²⁷⁸, where conventional LC-MS using a 2.1-mm-column was used for metabolite detection. Out of 840 features detected in the negative ion mode and 2194

features detected in the positive ion mode, only 6 features were identified (mainly lipids). It was shown that the total ion chromatograms of ELVs isolated from different volumes of human plasma (0.5, 1 mL and 2 mL) were almost the same²⁷⁸, indicating the saturation of the MS signals. If the MS signals were from metabolites, the concentrations of these metabolites must be very high in the injected sample. Thus, the ELVs isolated in their work were very different from the low amount of serum exosomes isolated in our work. Alternatively, the MS signals detected in their work were from background chemicals. In our work of using differential isotope labeling nLC-MS for metabolite detection, background signals can be readily differentiated from those of metabolites; MS signals from background are shown as a singlet peak, while the labeled metabolite is shown as a peak pair.

6.3.5 Comparative Metabolomics of Exosomes

In CIL nLC-MS, the same ¹³C-labeled pool prepared from a sample set was spiked into the individual ¹²C-labeled exosomes samples. Thus, the peak ratio of a ¹²C-labeled metabolite in a sample vs. the ¹³C-labeled same metabolite in the pool reflects its relative concentration to that of the pool. Different ratio values from different samples mixed with the same pool can be used to measure the relative concentration differences among different samples. IsoMS calculates the chromatographic peak area of the ¹²C- and ¹³C-labeled metabolite in a peak pair and then determines the peak area ratio for all the peak pairs detected. The final metabolite-intensity table in CSV file can be uploaded to statistical tools for analysis. Figure 6.3A,B shows the multivariate PCA analysis plots of the 1st and 2nd data set, respectively, while Figure 6.3C,D shows the corresponding PLS-DA plots. In both cases, the healthy and cancer groups show some separation. Univariate analysis using volcano plot was performed. In the 1st dataset, 54 metabolites were up-regulated (fold change or FC>1.5 with p<0.05), and 36 metabolites were down-regulated

(FC<0.67 with p<0.05) (Table 6.3). In the 2nd dataset, 34 metabolites were up-regulated, and 36 metabolites were down-regulated (Table 6.4). Some of the significantly changed metabolites can be positively identified or matched to the MCID libraries. Figure 6.4 shows the box plots of selected metabolites with significant changes (i.e., alanyl-histidine, 6-dimethylaminopurine, leucyl-proline, methionine sulfoxide).

Table 6.3 List of metabolites with significant change in the 1st sample set. ID Level indicates this metabolite was identified or matched to different libraries. Level 1: DnsID Library; Level 2: Zero-Reaction MCID Library; Level 3: One-Reaction MCID predicted library.

Peak Pair #	Monoisotopic mass (Da)	Retention Time (s)	Fold Change	p-value	Positively or Putatively Identified Metabolite	ID Level
5	185.6830	123.7	2.11	0.0497		
37	702.6352	128.7	2.44	0.0459		
70	600.6951	137.0	2.49	0.0414		
92	291.0463	148.2	0.39	0.0336	3,4-Dihydroxymandelic acid with cysteine conjugation	3
101	354.7219	149.5	1.95	0.0303		
117	41.6546	151.3	5.83	0.0045		
193	186.7248	182.5	2.70	0.0076		
199	308.6374	186.6	1.64	0.0362		
212	128.7662	191.3	3.20	0.0040		
221	103.6620	195.1	1.77	0.0479		
230	27.6332	200.3	0.20	0.0043		
235	61.0517	203.9	6.43	0.0338	Ethanolamine	2
240	45.0546	205.0	2.25	0.0237	Dimethylamine	2
253	207.5515	211.5	0.21	0.0140		
303	147.0512	232.9	0.54	0.0340	L-4-Hydroxyglutamate semialdehyde	

308	59.0494	235.9	0.50	0.0053	Guanidine	2
313	75.0308	238.8	0.52	0.0256		
316	142.7952	240.1	1.88	0.0252		
394	434.5790	271.5	0.18	0.0032		
396	99.1142	271.5	0.33	0.0199		
436	406.6455	287.8	0.39	0.0032		
533	89.0470	329.9	7.40	0.0405	Sarcosine	2
565	29.0267	345.3	0.17	0.0035		
651	51.9695	387.6	2.11	0.0447		
707	724.5107	408.5	0.63	0.0292	LysoPC(24:0) with addition of Adenine	3
718	116.6329	413.3	0.46	0.0491		
822	137.9674	447.5	5.21	0.0026		
859	760.5272	458.8	2.16	0.0332	PC(14:0/16:1(9Z)) with glycine conjugation	3
904	29.0254	472.8	0.36	0.0159		
910	148.9321	473.9	2.82	0.0177		
916	582.4077	475.0	0.64	0.0406		
984	61.0534	493.5	3.26	0.0337		
1140	522.3727	535.0	0.18	0.0404	Nonadecanoic acid addition of Thymidine	
1203	75.0308	542.5	0.54	0.0468		
1229	410.2900	546.6	0.24	0.0094		
1237	504.3723	548.3	5.80	0.0082	Ergosterol with addition of Thymine	3
1296	224.3886	560.5	0.33	0.0084		
1323	748.4339	564.9	2.28	0.0351	12-Hydroxy-12-octadecanoylcarnitine with addition of S-Glutathione	3
1469	177.2595	583.9	0.19	0.0240		
1523	395.3281	591.5	0.54	0.0429		
1531	438.3153	594.5	0.50	0.0282		
1534	103.1738	594.7	0.16	0.0015		
1553	604.3591	598.1	3.40	0.0119	Sphingosine with addition of S-Glutathione	3
1780	198.2536	652.8	2.07	0.0285		

1818	236.1926	658.9	0.36	0.0158		
2098	32.9967	737.2	2.07	0.0339		
2178	237.0364	761.7	2.39	0.0414		
2218	282.1389	783.0	4.96	0.0367	5-Hydroxypyrazinamide with addition of carnitine	3
2334	145.1787	820.3	0.31	0.0073		
2413	290.3018	844.8	0.49	0.0492		
2482	47.0644	864.3	6.47	0.0007		
2575	290.3065	882.6	0.29	0.0203		
2864	105.0697	943.9	0.37	0.0103	Diethanolamine	2
3059	101.0829	997.6	1.76	0.0090		
3299	48.1876	1044.2	0.39	0.0367		
3452	194.1127	1075.9	2.35	0.0353	Serotonin with addition of water	3
3865	86.1368	1162.5	4.20	0.0076		
3999	86.1373	1188.9	3.90	0.0079		
4159	86.1378	1217.4	3.54	0.0089		
4202	101.0478	1223.6	0.11	0.0301	3-Aminopropionaldehyde with addition of CO	3
4550	86.1377	1286.5	3.65	0.0111		
4623	119.0853	1297.8	3.23	0.0158	Isovaleric acid with addition of ammonia	3
4726	86.1330	1313.0	4.34	0.0102		
5143	86.1373	1398.4	4.43	0.0074		
5178	392.2796	1403.6	3.71	0.0066		
5332	107.0415	1433.4	2.47	0.0220		
5448	32.0264	1452.3	0.34	0.0294		
5530	107.0453	1464.9	2.73	0.0209		
5635	51.9701	1477.4	2.65	0.0197		
5716	86.1377	1503.2	3.98	0.0146		
5933	196.0707	1540.6	2.15	0.0375	Aniline with cysteine conjugation	3
5978	28.0405	1546.8	2.75	0.0060		
6363	508.4892	1606.9	0.64	0.0067		
6465	298.2716	1621.1	0.24	0.0362	Urea with addition of Palmitic acid	3

6776	226.1192	1671.7	2.38	0.0069	Alanyl-Histidine	1
7136	99.1039	1718.1	1.82	0.0385		
7337	185.1273	1754.8	0.22	0.0067		
7843	238.1161	1820.4	0.51	0.0304		
8029	279.0194	1842.2	3.40	0.0114	Glycerol 3-phosphate with taurine conjugation	3
8595	214.1745	1920.5	2.32	0.0463		
8608	270.3201	1922.8	2.54	0.0062		
11218	572.5735	2268.7	0.25	0.0137		
11507	80.1253	2306.0	1.63	0.0471		
11587	244.7840	2318.9	1.78	0.0041		
12097	81.1249	2390.4	1.84	0.0044		
13100	524.5214	2528.1	0.11	0.0074		
13344	187.1880	2565.0	1.93	0.0017		

Table 6.4 List of metabolites with significant change in the 2nd sample set. ID Level indicates this metabolite was identified or matched to different libraries. Level 1: DnsID Library; Level 2: Zero-Reaction MCID Library; Level 3: One-Reaction MCID predicted library.

Peak Pair #	Monoisotopic mass (Da)	Retention Time (s)	Fold Change	p-value	Positively or Putatively Identified Metabolite	ID Level
42	62.9944	129.2	2.52	0.0279		
90	105.6366	147.4	5.72	0.0032		
186	61.0515	175.7	2.55	0.0211	Ethanolamine	2
514	348.2892	320.4	0.18	0.0003	Palmitic amide addition of Cytosine	3
517	184.8166	320.8	2.06	0.0202		
649	372.3470	386.3	3.18	0.0143		
670	698.4888	394.5	0.41	0.0400		
701	45.0564	406.6	3.22	0.0230	Dimethylamine	2
717	464.3467	413.2	0.18	0.0197		
813	760.5613	445.0	0.15	0.0006		
861	334.2856	458.8	2.58	0.0439		
865	552.3829	459.7	1.78	0.0291	Homophytanic acid with addition of Uridine	3
957	172.1464	487.4	1.89	0.0298		
961	516.3631	488.3	0.23	0.0038	L-phenylalanyl-L-hydroxyproline with addition of Palmitic acid	3
979	438.3483	492.8	6.02	0.0313	Gamma-linolenyl carnitine with addition of ammonia	3
995	45.0559	497.5	3.18	0.0249		
1118	660.3856	531.2	0.39	0.0369	Tetradecanoylcarnitine with addition of Adenine	3
1135	344.3152	533.9	3.19	0.0085		
1150	646.4330	536.0	3.50	0.0111	PE(14:1(9Z)/14:1(9Z)) with addition of NH	3
1168	770.5281	538.1	2.75	0.0324	PE(14:0/20:3(5Z,8Z,11Z)) glycine conjugation	3
1220	504.1660	545.2	3.57	0.0261	Maltotriose	2
1224	662.4519	545.5	0.38	0.0212	(3R,3'R,6'R,9-cis)-Carotene-3,3'-	

					diol with addition of uracil	
1303	504.1705	561.5	3.47	0.0304		
1340	253.9854	568.6	2.73	0.0254		
1382	464.2941	573.7	2.62	0.0460		
1522	596.3626	591.5	2.61	0.0454	L-Urobilinogen	2
1588	114.1336	604.7	0.40	0.0382		
1696	217.1036	628.6	0.32	0.0220	N-a-Acetylcitrulline	2
1707	530.3857	631.5	0.41	0.0310		
1739	165.0453	639.4	0.46	0.0297	Methionine Sulfoxide	1
1816	158.0680	658.4	3.23	0.0151		
1877	227.0849	669.1	4.29	0.0124	Deoxycytidine	
1910	199.0237	686.3	0.44	0.0396	O-Phosphohomoserine	2
1933	366.1429	694.8	3.94	0.0477		
1950	275.1467	699.9	0.53	0.0309	L-a-glutamyl-L-Lysine	2
2104	290.1128	738.2	0.48	0.0457		
2160	141.0181	753.8	0.61	0.0104	O-Phosphoethanolamine	
2259	298.0452	794.7	0.30	0.0158		
2683	186.9884	906.0	0.18	0.0020		
2981	222.9877	976.3	0.53	0.0388		
3276	101.1560	1039.9	1.94	0.0237		
3531	218.1239	1090.2	0.57	0.0213	1-Methylhistamine with cytosine	3
3580	219.9580	1098.8	0.67	0.0406		
3709	125.0141	1129.8	0.62	0.0464	Taurine	2
4067	214.1294	1203.5	0.55	0.0187		
4240	222.9881	1229.5	0.63	0.0173		
4388	228.1454	1260.7	0.54	0.0237	Leucyl-Proline	1
5094	222.9884	1386.1	0.61	0.0032		
5174	133.0350	1403.2	0.62	0.0340	Iminodiacetate	2
5190	141.0181	1405.9	0.62	0.0112		
5258	219.9577	1416.7	0.65	0.0338		
5315	133.0232	1429.7	0.64	0.0486		
5543	135.0936	1466.4	2.58	0.0266		
5855	219.9578	1528.6	0.66	0.0338		
5998	135.0983	1549.8	3.59	0.0489	(R)-Amphetamine	2
6142	135.0991	1570.1	10.69	0.0062		
6503	163.0756	1626.2	3.22	0.0244	6-Dimethylaminopurine	1

6531	112.0530	1631.2	7.43	0.0362	trans-1,2-Dihydrobenzene-1,2-diol	
7997	86.1370	1837.8	0.41	0.0158		
8093	270.1462	1851.8	3.17	0.0477	Anabasine with addition of Thymine	3
9014	295.3408	1976.9	0.28	0.0094		
9551	223.0296	2047.0	2.50	0.0398		
14028	254.1429	2664.1	2.88	0.0166	Homoanserine	2
14501	125.0137	2755.1	0.64	0.0456		
15133	151.1179	2903.5	0.10	0.0261		
15441	133.0229	2976.5	0.57	0.0335		
15800	75.0574	3077.4	1.79	0.0182		
16046	47.0609	3159.5	0.50	0.0397		

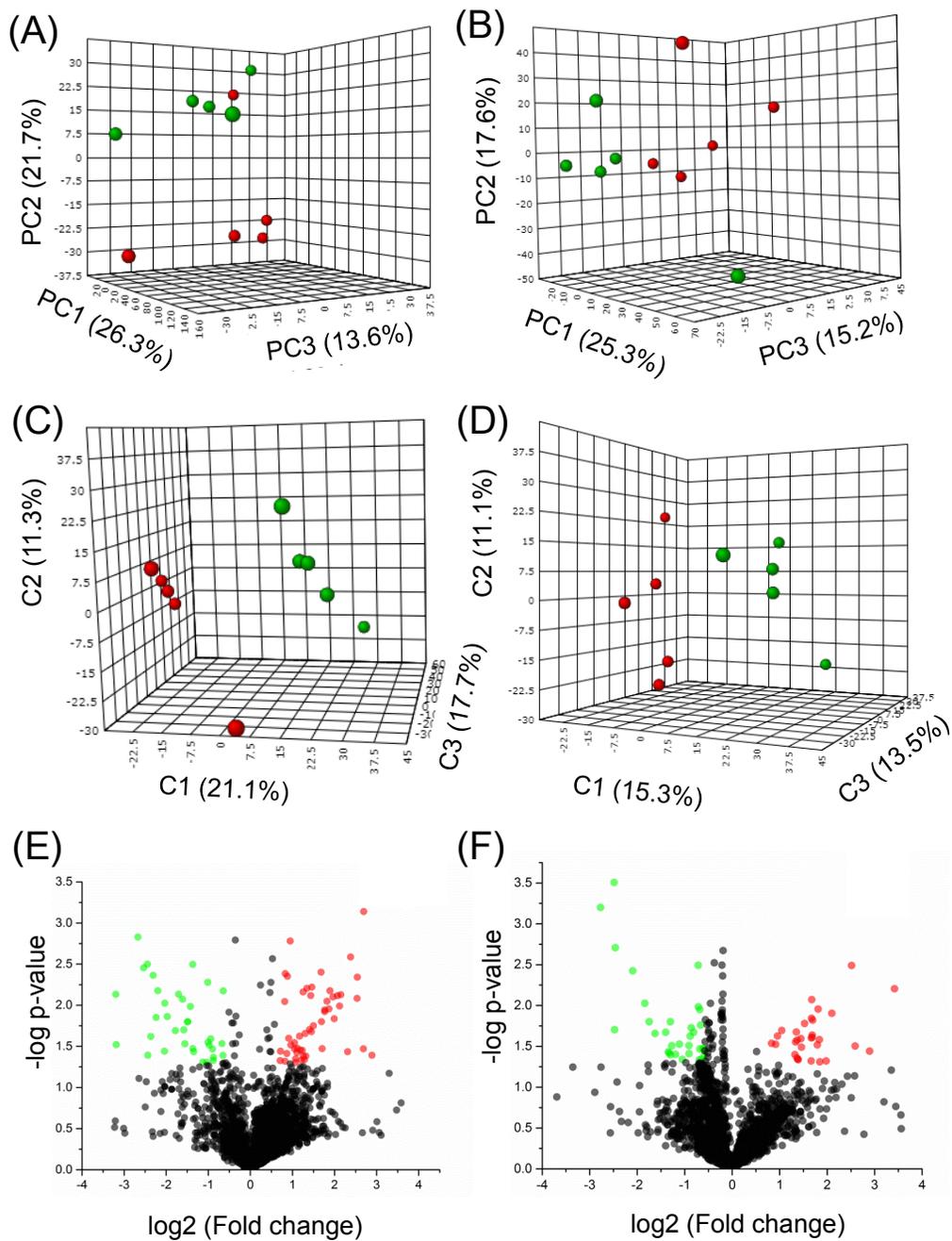


Figure 6.3 (A,B) PCA plots, (C,D) PLS-DA and (E,F) volcano plots of exosome metabolomes from the 1st sample set (n=10) and the 2nd sample set (n=10).

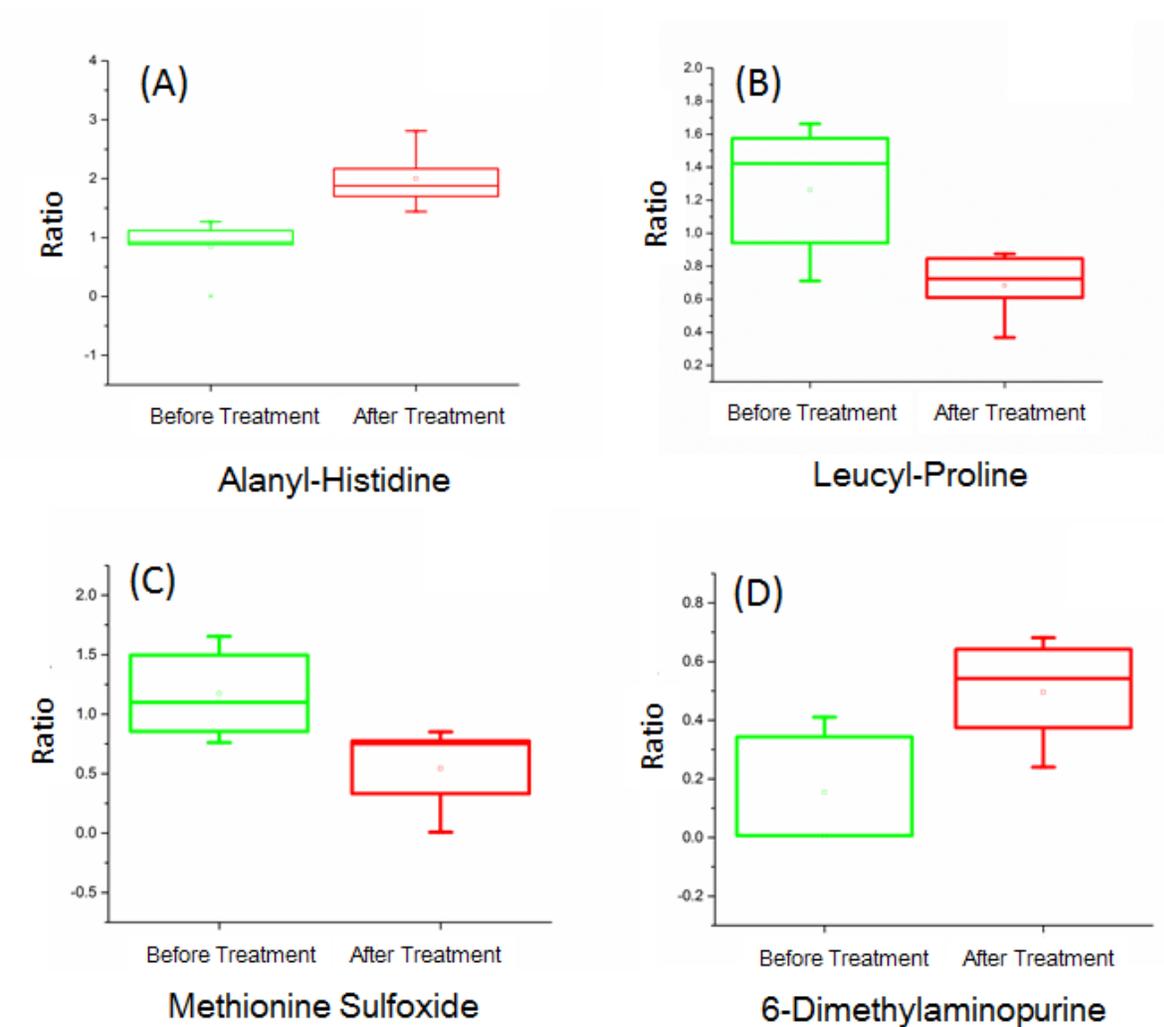


Figure 6.4 Box plots of four identified metabolites with significant changes in exosomes of serum samples collected from cancer patients before and after chemotherapy treatment.

Although the sample size used in this study is too small to draw any biological significance of the changed metabolites, the above results demonstrate that CIL nLC-MS can be used to perform relative quantification of metabolites in the metabolomes of exosomes.

6.4 Conclusions

We have developed a workflow for metabolomic analysis of exosomes isolated from serum that allows the detection of ~1950 metabolites per sample. To our knowledge, this level of detectability is unprecedented. In addition, accurate relative quantification of metabolites using a differentially labeled pooled sample as a control or internal standard can be carried out using CIL nLC-MS. As a proof-of-principle, we demonstrated the application of this workflow to detect significant changes of some metabolites before and after chemotherapy in exosomes isolated from serum of cancer patients. We expect that, after a large number of exosome samples are analyzed in the future, some potential biomarkers for therapeutic monitoring might be discovered. We are also trying to further optimize the coverage of the metabolome by applying other chemical isotope labeling techniques to profile the carbonyl, alcohol, and carboxyl acid sub-metabolomes of exosomes.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

Metabolomics aims to identify or quantify all the metabolites in given biological samples. Due to the great diversity of chemical and physical properties, it is difficult to achieve high-coverage and comprehensive metabolomic profiling. To address this problem, our group has been involved in the development of chemical isotope labeling techniques in recent years. We have developed DnsCl for amine-/phenol submetabolome profiling, DmPA bromide for carboxylic acid submetabolome profiling, base-activated DnsCl for hydroxyl submetabolome profiling, and DnsHz for carbonyl submetabolome profiling. More than 90% of metabolites can be detected if a four-channel labeling technique is performed. The CIL technique has been successfully used in the analysis of different biofluids samples, such as serum, urine, sweat and saliva. However, in cellular metabolomics, specific sample handling and highly sensitive analytical technique is desirable. My thesis work focuses on the establishment of different sample handling methods for different types of cells and the development of highly sensitive nanoflow CIL LC-MS platform for the analysis of small numbers of cells. Brief summaries of each research project are listed below.

Chapter 1 provides an overview of the recent achievements in cellular metabolomics, including sample preparation, analytical technique and data processing.

Chapter 2 describes a workflow that combines the highly efficient glass-bead-assisted lysis with CIL LC-MS for comprehensive metabolomic profiling of *S. cerevisiae*. ^{12}C -/ ^{13}C - DnsCl labeling technique was performed to profile amine and phenol containing metabolites. 5719 peak

pairs or putative metabolites were detected and quantified. 120 metabolites were positively identified by searching against DnsID library and 2980 metabolites were matched based accurate mass search. ^{12}C -/ ^{13}C - DmPA labeling technique was also performed to analyze carboxylic acid containing metabolites. 2286 peak pairs or putative metabolites were detected and quantified. Among these, 33 metabolites were positively identified by searching against DmPA labeled standards library and 1595 metabolites were matched by searching against accurate mass library. The workflow was used to comparatively metabolomic profile *S. cerevisiae* challenged by nitrogen limitation. Metabolomic analysis revealed pantothenate accumulation effect in nitrogen starvation group, and which may have potential applications in metabolic engineering.

Ultrasonication stimulus has been used to improve ethanol production in fermentation engineering, but the mechanism behind is still unclear. In Chapter 3, the technique described in Chapter 2 was employed to comparatively analyze the metabolome of *S. cerevisiae* with or without ultrasonication treatment. Our collaborator also analyzed the transcriptome of *S. cerevisiae* with and without ultrasonication stimulus. There is no significant difference of transcriptome between ultrasonication treated group and untreated group. However, the metabolomic analysis indicates that the ultrasonication treatment has an impact on the core pathways of carbon metabolism.

In Chapter 4, after different cell harvesting and lysis methods were evaluated, we developed a rapid and efficient method for harvesting and lysing mammalian cell cultures tailored to CIL LC-MS metabolomics. Two harvest methods (physical scraping and trypsinization) and two lysis methods (glass-bead-assisted lysis and freeze-thaw-cycle lysis) were compared, and two types of cell cultures (HeLa and MCF-7) were used to cross-validate the results. The total concentration of metabolites of each group was quantified by LC-UV, and the quantification results were used

to evaluate the harvest and lysis efficiency. The physical scrapping combined with freeze-thaw cycle was found to give the highest total metabolite concentration. PCA, PLS-DA and volcano plot were performed to compare the metabolome between different harvesting and lysis groups. Statistical analysis demonstrated that trypsinization can cause more metabolome changes, whereas the metabolome of cells lysed by two different methods were similar. However, the total metabolite concentration of the freeze-thaw-cycle lysis group is higher than that of the glass-beads-assisted lysis group. Thus, physical scraping combined with freeze-thaw-cycle lysis was the optimal method for handling mammalian cells in CIL LC-MS based metabolomics.

In Chapter 5, we described the development of high-performance CIL nanoLC-MS for comprehensive metabolomic profiling of small numbers of cells. A new chemical isotope labeling protocol was established tailored to analyze small amounts of samples. The conditions of nanoLC and captivespray ion source parameters were optimized, and the chemical-vapor-assisted technique was introduced to further improve the sensitivity of the analytical platform. Compared with microflow LC-MS, more metabolites could be detected on nanoLC-MS platform, and the requirement of sample amount was also reduced. On average, 1620 ± 148 , 2091 ± 89 and 2402 ± 80 peak pairs or putative metabolites were detected from 100, 1000, and 10000 MCF-7 breast cancer cells. Among these, more than 60% of metabolites could be positively identified or matched.

Chapter 6 describes the development of nanoLC-MS for metabolomic profiling of circulating exosomes. Two sets of exosomes were isolated by ultracentrifugation from pancreatic cancer patient serum before and after chemotherapy. Exosomes were lysed by freeze-thaw-cycle, and extracted by 50% MeOH. The lysates were labeled by $^{12}\text{C}/^{13}\text{C}$ - DnsCl and analyzed nanoLC-MS for analysis. On average, 1964 ± 60 peak pairs or putative metabolites were detected

from 1st set of sample and 1948±117 peak pairs or putative metabolites were detected from 2nd set. Among these, 101 and 94 metabolites could be positively identified in the 1st and 2nd set, respectively. 1580 and 1590 metabolites could be matched based on accurate mass search. Univariate and multivariate analysis were performed to analysis the dataset. The statistical analysis demonstrated the change of exosome metabolome before and after chemotherapy and some significantly changed metabolites were found.

7.2 Future Work

We have successfully applied the $^{12}\text{C}/^{13}\text{C}$ - dansylation labeling technique to profile the metabolome of bacterial cells, yeast cells and mammalian cells, and small numbers of cells could be analyzed by our highly sensitive CIL nanoLC-MS technique. However, some challenges are still required to be addressed.

Currently, only amine- and phenol- submetabolomes could be analyzed on nanoLC-MS platform after $^{12}\text{C}/^{13}\text{C}$ - dansylation labeling technique. To achieve high-coverage metabolomic analysis of small numbers of cells, we also need to adapt the other three channels of labeling techniques to nanoLC-MS. Compared with $^{12}\text{C}/^{13}\text{C}$ - dansylation labeling, this will be a more challenging task. First, in hydroxyl labeling, additional liquid-liquid extraction is required to remove amine containing metabolites prior to labeling. Sample loss effects becomes significant when handling small numbers of cells, thus, metabolites loss is expected in hydroxyl labeling. Second, chemical background exists in DmPA labeling, and the background mainly comes from unquenched DmPA reagent and byproducts of reaction. In most of the situations, the background could be suppressed by metabolites signal if sample concentration is high enough, but since we aim to analyze small amounts of samples by nanoLC-MS, background will significantly interfere with MS analysis. Third, DnsHz is used for carbonyl submetabolome analysis, but there is no

quenching step in the current labeling protocol, and DnsHz will elute out in the middle of the elution gradient and significantly suppress metabolite signals. As discussed above, these problems need to be resolved before transferring to nanoLC-MS.

Another challenge is sample normalization in the analysis of small numbers of cells. Currently, our group uses a step gradient LC-UV method to determine the total concentration of metabolites. However, the sensitivity of LC-UV cannot meet the requirements for total metabolite quantification of a small number of cells. We are in the process of developing a more sensitive LC fluorescence detection based platform to quantify the total concentration of metabolites from small numbers of cells.

References

- (1) Zhang, A.; Sun, H.; Xu, H.; Qiu, S.; Wang, X. *OMICS : a Journal of Integrative Biology* **2013**, *17*, 495-501.
- (2) Ramirez, T.; Strigun, A.; Verlohner, A.; Huener, H.-A.; Peter, E.; Herold, M.; Bordag, N.; Mellert, W.; Walk, T.; Spitzer, M.; Jiang, X.; Sperber, S.; Hofmann, T.; Hartung, T.; Kamp, H.; van Ravenzwaay, B. *Archives of Toxicology* **2018**, *92*, 893-906.
- (3) Tiziani, S.; Kang, Y.; Choi, J. S.; Roberts, W.; Paternostro, G. *Nature communications* **2011**, *2*, 545-545.
- (4) Cevallos-Cevallos, J. M.; Danyluk, M. D.; Reyes-De-Corcuera, J. I. *Journal of Food Science* **2011**, *76*, M238-M246.
- (5) Alvarez Cubero, M. J.; Lorente, J. A.; Robles-Fernandez, I.; Rodriguez-Martinez, A.; Puche, J. L.; Serrano, M. J. In *Circulating Tumor Cells: Methods and Protocols*, M. Magbanua, M. J.; W. Park, J., Eds.; Springer New York: New York, NY, 2017, pp 283-303.
- (6) Luo, X.; Zhao, S.; Huan, T.; Sun, D.; Friis, R. M. N.; Schultz, M. C.; Li, L. *J. Proteome Res.* **2016**, *15*, 1602-1612.
- (7) Wu, Y. M.; Li, L. *Anal. Chem.* **2013**, *85*, 5755-5763.
- (8) Fu, F. F.; Cheng, V. W. T.; Wu, Y. M.; Tang, Y. A.; Weiner, J. H.; Li, L. *J. Proteome Res.* **2013**, *12*, 4478-4489.
- (9) Xu, Y.-F.; Lu, W.; Rabinowitz, J. D. *Anal. Chem.* **2015**, *87*, 2273-2281.
- (10) Zampieri, M.; Zimmermann, M.; Claassen, M.; Sauer, U. *Cell Reports* **2017**, *19*, 1214-1228.
- (11) Dietmair, S.; Timmins, N. E.; Gray, P. P.; Nielsen, L. K.; Krömer, J. O. *Analytical Biochemistry* **2010**, *404*, 155-164.

- (12) Dettmer, K.; Nürnberger, N.; Kaspar, H.; Gruber, M. A.; Almstetter, M. F.; Oefner, P. J. *Anal. Bioanal. Chem.* **2011**, *399*, 1127-1139.
- (13) García-Cañaveras, J. C.; López, S.; Castell, J. V.; Donato, M. T.; Lahoz, A. *Anal. Bioanal. Chem.* **2016**, *408*, 1217-1230.
- (14) Luo, X.; Gu, X.; Li, L. *Anal. Chim. Acta* **2017**.
- (15) Luo, X.; Li, L. *Anal. Chem.* **2017**, *89*, 11664-11671.
- (16) Perez-Camps, M.; Garcia-Ximenez, F. *Animal* **2008**, *2*, 595-599.
- (17) Hutschenreuther, A.; Kiontke, A.; Birkenmeier, G.; Birkemeyer, C. *Analytical Methods* **2012**, *4*, 1953-1963.
- (18) Myint, K. T.; Uehara, T.; Aoshima, K.; Oda, Y. *Anal. Chem.* **2009**, *81*, 7766-7772.
- (19) Teng, Q.; Huang, W.; Collette, T. W.; Ekman, D. R.; Tan, C. *Metabolomics* **2008**, *5*, 199.
- (20) Lane, A. N.; Fan, T. W. M. *Metabolomics* **2007**, *3*, 79-86.
- (21) Ibáñez, C.; Simó, C.; Valdés, A.; Campone, L.; Piccinelli, A. L.; García-Cañas, V.; Cifuentes, A. *Journal of Pharmaceutical and Biomedical Analysis* **2015**, *110*, 83-92.
- (22) Canelas, A. B.; Ras, C.; ten Pierick, A.; van Dam, J. C.; Heijnen, J. J.; van Gulik, W. M. *Metabolomics* **2008**, *4*, 226-239.
- (23) Bolten, C. J.; Kiefer, P.; Letisse, F.; Portais, J.-C.; Wittmann, C. *Anal. Chem.* **2007**, *79*, 3843-3849.
- (24) Castrillo, J. I.; Hayes, A.; Mohammed, S.; Gaskell, S. J.; Oliver, S. G. *Phytochemistry* **2003**, *62*, 929-937.
- (25) Taymaz-Nikerel, H.; de Mey, M.; Ras, C.; ten Pierick, A.; Seifar, R. M.; van Dam, J. C.; Heijnen, J. J.; van Gulik, W. M. *Analytical Biochemistry* **2009**, *386*, 9-19.

- (26) Kapoore, R. V.; Coyle, R.; Staton, C. A.; Brown, N. J.; Vaidyanathan, S. *Analyst* **2017**, *142*, 2038-2049.
- (27) Garcia-Canaveras, J. C.; Lopez, S.; Castell, J. V.; Donato, M. T.; Lahoz, A. *Anal. Bioanal. Chem.* **2016**, *408*, 1217-1230.
- (28) Lorenzo Tejedor, M.; Mizuno, H.; Tsuyama, N.; Harada, T.; Masujima, T. *Anal. Chem.* **2012**, *84*, 5221-5228.
- (29) Wang, N.; Xu, M.; Wang, P.; Li, L. *Anal. Chem.* **2010**, *82*, 2262-2271.
- (30) Zheng, H.; Yin, J.; Gao, Z.; Huang, H.; Ji, X.; Dou, C. *Applied Biochemistry and Biotechnology* **2011**, *164*, 1215-1224.
- (31) Patterson, A. D.; Li, H.; Eichler, G. S.; Krausz, K. W.; Weinstein, J. N.; Fornace, A. J.; Gonzalez, F. J.; Idle, J. R. *Anal. Chem.* **2008**, *80*, 665-674.
- (32) Wu, Y.; Li, L. *Anal. Chem.* **2013**, *85*, 5755-5763.
- (33) Brown, R. B.; Audet, J. *Journal of the Royal Society Interface* **2008**, *5*, S131-S138.
- (34) Najafpour, G. D. In *Biochemical Engineering and Biotechnology (Second Edition)*, Najafpour, G. D., Ed.; Elsevier: Amsterdam, 2015, pp 495-526.
- (35) Cha, J.-O.; Talha, A. F. S. M.; Lim, C. W.; Kim, B. *Experimental Parasitology* **2014**, *138*, 18-24.
- (36) Ron, E. Z.; Kohler, R. E.; Davis, B. D. *Science* **1966**, *153*, 1119-&.
- (37) Tsai, T. S. *Journal of Biochemical and Biophysical Methods* **1986**, *13*, 343-346.
- (38) Melendez, J.; Kiang, D.; Santaus, T.; Geddes, C. *Biophysical Journal* **2015**, *108*, 315a.
- (39) Ritter, J. B.; Genzel, Y.; Reichl, U. *Analytical Biochemistry* **2008**, *373*, 349-369.
- (40) Hans, M.; Heinzle, E.; Wittmann, C. *Appl. Microbiol. Biotechnol.* **2001**, *56*, 776-779.

- (41) Luttik, M. A. H.; Vuralhan, Z.; Suir, E.; Braus, G. H.; Pronk, J. T.; Daran, J. M. *Metab. Eng.* **2008**, *10*, 141-153.
- (42) Rabinowitz, J. D.; Kimball, E. *Anal. Chem.* **2007**, *79*, 6167-6173.
- (43) Prasad Maharjan, R.; Ferenci, T. *Analytical Biochemistry* **2003**, *313*, 145-154.
- (44) Gielisch, I.; Meierhofer, D. *J. Proteome Res.* **2015**, *14*, 224-235.
- (45) Feng, J.; Li, J.; Wu, H.; Chen, Z. *Metabolomics* **2013**, *9*, 874-886.
- (46) Meissen, J. K.; Hirahatake, K. M.; Adams, S. H.; Fiehn, O. *Metabolomics* **2015**, *11*, 707-721.
- (47) Teahan, O.; Bevan, C. L.; Waxman, J.; Keun, H. C. *The International Journal of Biochemistry & Cell Biology* **2011**, *43*, 1002-1009.
- (48) Muschet, C.; Möller, G.; Prehn, C.; de Angelis, M. H.; Adamski, J.; Tokarz, J. *Metabolomics* **2016**, *12*, 151.
- (49) Watanabe, M.; Sheriff, S.; Lewis, K. B.; Cho, J.; Tinch, S. L.; Balasubramaniam, A.; Kennedy, M. A. *Journal of molecular biomarkers & diagnosis* **2012**, *3*, S3-002.
- (50) Ye, Y.; Zhang, L.; Hao, F.; Zhang, J.; Wang, Y.; Tang, H. *J. Proteome Res.* **2012**, *11*, 2559-2566.
- (51) Fleisher, T. A.; de Oliveira, J. B. In *Clinical Immunology (Third Edition)*, Rich, R. R.; Fleisher, T. A.; Shearer, W. T.; Schroeder, H. W.; Frew, A. J.; Weyand, C. M., Eds.; Mosby: Edinburgh, 2008, pp 1435-1446.
- (52) Groves, J. D.; Falson, P.; le Maire, M.; Tanner, M. J. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, *93*, 12245-12250.
- (53) Sezonov, G.; Joseleau-Petit, D.; D'Ari, R. *Journal of Bacteriology* **2007**, *189*, 8746-8749.

- (54) K., F.; F., D.; A.R., S.; A.H., G.; I., C.; J.F., V. I.; J., D. *Journal of Applied Microbiology* **2005**, *99*, 1503-1515.
- (55) Cao, B.; Aa, J.; Wang, G.; Wu, X.; Liu, L.; Li, M.; Shi, J.; Wang, X.; Zhao, C.; Zheng, T.; Guo, S.; Duan, J. *Anal. Bioanal. Chem.* **2011**, *400*, 2983-2993.
- (56) Silva, L. P.; Lorenzi, P. L.; Purwaha, P.; Yong, V.; Hawke, D. H.; Weinstein, J. N. *Anal. Chem.* **2013**, *85*, 9536-9542.
- (57) Wu, Y.; Li, L. *Anal. Chem.* **2012**, *84*, 10723-10731.
- (58) Wu, Y.; Li, L. *Anal. Chem.* **2014**, *86*, 9428-9433.
- (59) Mung, D.; Li, L. *Anal. Chem.* **2017**, *89*, 4435-4443.
- (60) Hooton, K.; Li, L. *Anal. Chem.* **2017**, *89*, 7847-7851.
- (61) Goldansaz, S. A.; Guo, A. C.; Sajed, T.; Steele, M. A.; Plastow, G. S.; Wishart, D. S. *PLoS One* **2017**, *12*, 26.
- (62) Zulak, K. G.; Weljie, A. M.; Vogel, H. J.; Facchini, P. J. *BMC Plant Biology* **2008**, *8*, 5.
- (63) Markley, J. L.; Brüschweiler, R.; Edison, A. S.; Eghbalnia, H. R.; Powers, R.; Raftery, D.; Wishart, D. S. *Curr. Opin. Biotechnol.* **2017**, *43*, 34-40.
- (64) Castiglione, F.; Ferro, M.; Mavroudakos, E.; Pellitteri, R.; Bossolasco, P.; Zaccheo, D.; Morbidelli, M.; Silani, V.; Mele, A.; Moscatelli, D.; Cova, L. *Scientific Reports* **2017**, *7*, 15808.
- (65) Guennec, A. L.; Giraudeau, P.; Caldarelli, S. *Anal. Chem.* **2014**, *86*, 5946-5954.
- (66) Féraud, B.; Govaerts, B.; Verleysen, M.; de Tullio, P. *Metabolomics* **2015**, *11*, 1756-1768.
- (67) Kostidis, S.; Addie, R. D.; Morreau, H.; Mayboroda, O. A.; Giera, M. *Anal. Chim. Acta* **2017**, *980*, 1-24.
- (68) Bingol, K.; Bruschiweiler-Li, L.; Li, D.; Zhang, B.; Xie, M.; Brüschweiler, R. *Bioanalysis* **2016**, *8*, 557-573.

- (69) Fiehn, O. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]* **2016**, *114*, 30.34.31-30.34.32.
- (70) Koek, M. M.; Jellema, R. H.; van der Greef, J.; Tas, A. C.; Hankemeier, T. *Metabolomics* **2011**, *7*, 307-328.
- (71) Papadimitropoulos, M.-E. P.; Vasilopoulou, C. G.; Maga-Nteve, C.; Klapa, M. I. In *Metabolic Profiling: Methods and Protocols*, Theodoridis, G. A.; Gika, H. G.; Wilson, I. D., Eds.; Springer New York: New York, NY, 2018, pp 133-147.
- (72) Fancy, S.-A.; Rumpel, K. In *Biomarker Methods in Drug Discovery and Development*, Wang, F., Ed.; Humana Press: Totowa, NJ, 2008, pp 317-340.
- (73) Winnike, J. H.; Wei, X.; Knagge, K. J.; Colman, S. D.; Gregory, S. G.; Zhang, X. *J. Proteome Res.* **2015**, *14*, 1810-1817.
- (74) Yu, Z.; Huang, H.; Reim, A.; Charles, P. D.; Northage, A.; Jackson, D.; Parry, I.; Kessler, B. *M. Talanta* **2017**, *165*, 685-691.
- (75) Ibáñez, C.; Simó, C.; Palazoglu, M.; Cifuentes, A. *Anal. Chim. Acta* **2017**, *986*, 48-56.
- (76) He, Y.; Zhang, Z. M.; Ma, P.; Ji, H. C.; Lu, H. M. *Analytical Methods* **2018**, *10*, 1266-1274.
- (77) Han, X.; Jiang, X. *European journal of lipid science and technology : EJLST* **2009**, *111*, 39-52.
- (78) Xiao, J. F.; Zhou, B.; Ressom, H. W. *TrAC Trends in Analytical Chemistry* **2012**, *32*, 1-14.
- (79) Guo, K.; Li, L. *Anal. Chem.* **2009**, *81*, 3919-3932.
- (80) Guo, K.; Li, L. *Anal. Chem.* **2010**, *82*, 8789-8793.
- (81) Zhao, S.; Dawe, M.; Guo, K.; Li, L. *Anal. Chem.* **2017**, *89*, 6758-6765.
- (82) Zhao, S.; Luo, X.; Li, L. *Anal. Chem.* **2016**, *88*, 10617-10623.

- (83) Yuan, B.-F.; Zhu, Q.-F.; Guo, N.; Zheng, S.-J.; Wang, Y.-L.; Wang, J.; Xu, J.; Liu, S.-J.; He, K.; Hu, T.; Zheng, Y.-W.; Xu, F.-Q.; Feng, Y.-Q. *Anal. Chem.* **2018**, *90*, 3512-3520.
- (84) Li, Z.; Tatlay, J.; Li, L. *Anal. Chem.* **2015**, *87*, 11468-11474.
- (85) Wilm, M.; Mann, M. *Anal. Chem.* **1996**, *68*, 1-8.
- (86) Marginean, I.; Tang, K.; Smith, R. D.; Kelly, R. T. *Journal of the American Society for Mass Spectrometry* **2014**, *25*, 30-36.
- (87) Chetwynd, A. J.; David, A. *Talanta* **2018**, *182*, 380-390.
- (88) Navarro-Reig, M.; Jaumot, J.; Baglai, A.; Vivó-Truyols, G.; Schoenmakers, P. J.; Tauler, R. *Anal. Chem.* **2017**, *89*, 7675-7683.
- (89) Edwards, J. L.; Edwards, R. L.; Reid, K. R.; Kennedy, R. T. *J. Chromatogr. A* **2007**, *1172*, 127-134.
- (90) Ramautar, R.; Somsen Govert, W.; Jong Gerhardus, J. *Electrophoresis* **2016**, *38*, 190-202.
- (91) Acunha, T.; Simó, C.; Ibáñez, C.; Gallardo, A.; Cifuentes, A. *J. Chromatogr. A* **2016**, *1428*, 326-335.
- (92) Sugimoto, M.; Kikuchi, S.; Arita, M.; Soga, T.; Nishioka, T.; Tomita, M. *Anal. Chem.* **2005**, *77*, 78-84.
- (93) Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. *J. Proteome Res.* **2003**, *2*, 488-494.
- (94) Liu, J.-X.; Aerts, J. T.; Rubakhin, S. S.; Zhang, X.-X.; Sweedler, J. V. *The Analyst* **2014**, *139*, 5835-5842.
- (95) Lapainis, T.; Rubakhin, S. S.; Sweedler, J. V. *Anal. Chem.* **2009**, *81*, 5858-5864.
- (96) Zhou, B.; Xiao, J. F.; Tuli, L.; Ransom, H. W. *Mol. Biosyst.* **2012**, *8*, 470-481.
- (97) Huan, T.; Li, L. *Anal. Chem.* **2015**, *87*, 1306-1313.

- (98) Jin, Z.; Kang, J.; Yu, T. *Bioinformatics* **2018**, *34*, 1555-1561.
- (99) Wei, R.; Wang, J.; Su, M.; Jia, E.; Chen, S.; Chen, T.; Ni, Y. *Scientific Reports* **2018**, *8*, 663.
- (100) Worley, B.; Powers, R. *Current Metabolomics* **2013**, *1*, 92-107.
- (101) Chen, T.; Cao, Y.; Zhang, Y.; Liu, J.; Bao, Y.; Wang, C.; Jia, W.; Zhao, A. *Evidence-Based Complementary and Alternative Medicine* **2013**, *2013*, 11.
- (102) Bartel, J.; Krumsiek, J.; Theis, F. J. *Computational and Structural Biotechnology Journal* **2013**, *4*, e201301009.
- (103) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78*, 779-787.
- (104) Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G. *Anal. Chem.* **2012**, *84*, 5035-5039.
- (105) Chong, J.; Soufan, O.; Li, C.; Caraus, I.; Li, S.; Bourque, G.; Wishart, D. S.; Xia, J. *Nucleic Acids Res.* **2018**, *46*, W486-W494.
- (106) Wanichthanarak, K.; Fan, S.; Grapov, D.; Barupal, D. K.; Fiehn, O. *PLoS One* **2017**, *12*, 14.
- (107) Broeckling, C. D.; Reddy, I. R.; Duran, A. L.; Zhao, X.; Sumner, L. W. *Anal. Chem.* **2006**, *78*, 4334-4341.
- (108) Zhou, R.; Tseng, C. L.; Huan, T.; Li, L. *Anal. Chem.* **2014**, *86*, 4675-4679.
- (109) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A. *Nucleic Acids Res.* **2018**, *46*, D608-D617.

- (110) Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; Wolan, D. W.; Spilker, M. E.; Benton, H. P.; Siuzdak, G. *Anal. Chem.* **2018**, *90*, 3156-3164.
- (111) Li, L.; Li, R.; Zhou, J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y.; Huan, T.; Zheng, J.; Shi, Y.; Wishart, D. S.; Lin, G. *Anal. Chem.* **2013**, *85*, 3401-3408.
- (112) Huan, T.; Wu, Y.; Tang, C.; Lin, G.; Li, L. *Anal. Chem.* **2015**, *87*, 9838-9845.
- (113) Zheng, S.-J.; Liu, S.-J.; Zhu, Q.-F.; Guo, N.; Wang, Y.-L.; Yuan, B.-F.; Feng, Y.-Q. *Anal. Chem.* **2018**, *90*, 8412-8420.
- (114) Lei, Z.; Jing, L.; Qiu, F.; Zhang, H.; Huhman, D.; Zhou, Z.; Sumner, L. W. *Anal. Chem.* **2015**, *87*, 7373-7381.
- (115) Cao, M.; Fraser, K.; Huege, J.; Featonby, T.; Rasmussen, S.; Jones, C. *Metabolomics* **2015**, *11*, 696-706.
- (116) Botstein, D.; Fink, G. R. *Genetics* **2011**, *189*, 695-704.
- (117) Legras, J. L.; Merdinoglu, D.; Cornuet, J. M.; Karst, F. *Mol. Ecol.* **2007**, *16*, 2091-2102.
- (118) Dequin, S. *Appl. Microbiol. Biotechnol.* **2001**, *56*, 577-588.
- (119) Oud, B.; van Maris, A. J. A.; Daran, J. M.; Pronk, J. T. *FEMS Yeast Res.* **2012**, *12*, 183-196.
- (120) Mattanovich, D.; Sauer, M.; Gasser, B. *Microb. Cell. Fact.* **2014**, *13*, 5.
- (121) Nielsen, J. *Science* **2015**, *349*, 1050-1051.
- (122) Ellis, D. I.; Goodacre, R. *Curr. Opin. Biotechnol.* **2012**, *23*, 22-28.
- (123) Dromms, R. A.; Styczynski, M. P. *Metabolites* **2012**, *2*, 1090-1122.
- (124) Joyner, P. M.; Matheke, R. M.; Smith, L. M.; Cichewicz, R. H. *J. Proteome Res.* **2010**, *9*, 404-412.

- (125) Yu, S. L.; An, Y. J.; Yang, H. J.; Kang, M. S.; Kim, H. Y.; Wen, H.; Jin, X.; Kwon, H. N.; Min, K. J.; Lee, S. K.; Park, S. *J. Proteome Res.* **2013**, *12*, 1619-1627.
- (126) Airoidi, C.; Tripodi, F.; Guzzi, C.; Nicastro, R.; Coccetti, P. *Mol. Biosyst.* **2015**, *11*, 379-383.
- (127) Kondo, E.; Marriott, P. J.; Parker, R. M.; Kouremenos, K. A.; Morrison, P.; Adams, M. *Anal. Chim. Acta* **2014**, *807*, 135-142.
- (128) Carneiro, S.; Pereira, R.; Rocha, I. In *Yeast Metabolic Engineering: Methods and Protocols*, Mapelli, V., Ed.; Humana Press Inc: Totowa, 2014, pp 197-207.
- (129) Klavins, K.; Drexler, H.; Hann, S.; Koellensperger, G. *Anal. Chem.* **2014**, *86*, 4145-4150.
- (130) Farres, M.; Pina, B.; Tauler, R. *Metabolomics* **2015**, *11*, 210-224.
- (131) Li, K. F.; Wang, X.; Pidatala, V. R.; Chang, C. P.; Cao, X. H. *J. Proteome Res.* **2014**, *13*, 5879-5887.
- (132) Ito, Y.; Hirasawa, T.; Shimizu, H. *Biosci. Biotechnol. Biochem.* **2014**, *78*, 151-159.
- (133) Venter, L.; Lindeque, Z.; van Rensburg, P. J.; van der Westhuizen, F.; Smuts, I.; Louw, R. *Metabolomics* **2015**, *11*, 111-121.
- (134) Magasanik, B.; Kaiser, C. A. *Gene* **2002**, *290*, 1-18.
- (135) Ljungdahl, P. O.; Daignan-Fornier, B. *Genetics* **2012**, *190*, 885-929.
- (136) Saldanha, A. J.; Brauer, M. J.; Botstein, D. *Molecular Biology of the Cell* **2004**, *15*, 4089-4104.
- (137) Rungtaphan, W.; Keasling, J. D. *Metab. Eng.* **2014**, *21*, 103-113.
- (138) Kamisaka, Y.; Noda, N.; Tomita, N.; Kimura, K.; Kodaki, T.; Hosaka, K. *Biosci. Biotechnol. Biochem.* **2006**, *70*, 646-653.

- (139) Treco, D. A.; Lundblad, V. In *Current Protocols in Molecular Biology*, Ausubel F. M, B. R., Kingston R. E. , Moore D. D., Seidman J. G., Ed.; John Wiley and Sons: New York, 1993, pp 13.11.11–13.11.17.
- (140) Brachmann, C. B.; Davies, A.; Cost, G. J.; Caputo, E.; Li, J. C.; Hieter, P.; Boeke, J. D. *Yeast* **1998**, *14*, 115-132.
- (141) Steinle, A.; Bergander, K.; Steinbuchel, A. *Appl. Environ. Microbiol.* **2009**, *75*, 3437-3446.
- (142) Chen, B.; Lee, D.-Y.; Chang, M. W. *Metab. Eng.* **2015**, *31*, 53-61.
- (143) Reyes, L. H.; Gomez, J. M.; Kao, K. C. *Metab. Eng.* **2014**, *21*, 26-33.
- (144) Wang, P. P.; Wei, Y. J.; Fan, Y.; Liu, Q. F.; Wei, W.; Yang, C. S.; Zhang, L.; Zhao, G. P.; Yue, J. M.; Yan, X.; Zhou, Z. H. *Metab. Eng.* **2015**, *29*, 97-105.
- (145) Steen, E.; Chan, R.; Prasad, N.; Myers, S.; Petzold, C.; Redding, A.; Ouellet, M.; Keasling, J. *Microb. Cell. Fact.* **2008**, *7*, 36.
- (146) Hanscho, M.; Ruckerbauer, D. E.; Chauhan, N.; Hofbauer, H. F.; Krahulec, S.; Nidetzky, B.; Kohlwein, S. D.; Zanghellini, J.; Natter, K. *FEMS Yeast Res.* **2012**, *12*, 796-808.
- (147) Cohen, R.; Engelberg, D. *FEMS Microbiol. Lett.* **2007**, *273*, 239-243.
- (148) Yanes, O.; Tautenhahn, R.; Patti, G. J.; Siuzdak, G. *Anal. Chem.* **2011**, *83*, 2152-2161.
- (149) Peng, J.; Li, L. *Anal. Chim. Acta* **2013**, *803*, 97-105.
- (150) Wu, Y. M.; Li, L. *Anal. Chem.* **2012**, *84*, 10723-10731.
- (151) Li, L.; Li, R. H.; Zhou, J. J.; Zuniga, A.; Stanislaus, A. E.; Wu, Y. M.; Huan, T.; Zheng, J. M.; Shi, Y.; Wishart, D. S.; Lin, G. H. *Anal. Chem.* **2013**, *85*, 3401-3408.
- (152) Han, J.; Gagnon, S.; Eckle, T.; Borchers, C. H. *Electrophoresis* **2013**, *34*, 2891-2900.
- (153) Cardaci, S.; Ciriolo, M. R. *International Journal of Cell Biology* **2012**, *2012*, 9.

- (154) Brauer, M. J.; Yuan, J.; Bennett, B. D.; Lu, W.; Kimball, E.; Botstein, D.; Rabinowitz, J. D. *Proc Natl Acad Sci U S A* **2006**, *103*, 19302-19307.
- (155) Liu, B. S.; Sutton, A.; Sternglanz, R. *J. Biol. Chem.* **2005**, *280*, 16659-16664.
- (156) Mandal, S.; Mandal, A.; Johansson, H. E.; Orjalo, A. V.; Park, M. H. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, *110*, 2169-2174.
- (157) Olzhausen, J.; Schubbe, S.; Schuller, H. J. *Current genetics* **2009**, *55*, 163-173.
- (158) Chen, Y.; Siewers, V.; Nielsen, J. *PLoS One* **2012**, *7*, e42475.
- (159) Nielsen, J. *mBio* **2014**, *5*, 3.
- (160) Zhou, Y. J.; Buijs, N. A.; Siewers, V.; Nielsen, J. *Frontiers in Bioengineering and Biotechnology* **2014**, *2*, 32.
- (161) Kozak, B. U.; van Rossum, H. M.; Luttkik, M. A. H.; Akeroyd, M.; Benjamin, K. R.; Wu, L.; de Vries, S.; Daran, J. M.; Pronk, J. T.; van Maris, A. J. A. *mBio* **2014**, *5*, 11.
- (162) Lian, J. Z.; Si, T.; Nair, N. U.; Zhao, H. M. *Metab. Eng.* **2014**, *24*, 139-149.
- (163) Terrance, G. C. *Cold Spring Harbor Monograph Archive; Volume 11B (1982): The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression* **1982**.
- (164) Patist, A.; Bates, D. *Ultrasonic innovations in the food industry: From the laboratory to commercial production*, 2008; Vol. 9, p 147-154.
- (165) Kumar, K. S.; Melissa, M.; J., v. L.; Gowrishankar, S.; David, G. *Biotechnology and Bioengineering* **2007**, *98*, 978-985.
- (166) Pandey, A. *Handbook of plant-based biofuels*; CRC Press, 2008.
- (167) Stathopoulos, P. B.; Scholz, G. A.; Hwang, Y.-M.; Rumfeldt, J. A. O.; Lepock, J. R.; Meiering, E. M. *Protein Science : A Publication of the Protein Society* **2004**, *13*, 3017-3027.
- (168) Nyborg, W. L. *Ultrasound Med. Biol.* **2001**, *27*, 301-333.

- (169) Xu, P.; Gul-Uludag, H.; Ang, W. T.; Yang, X. Y.; Huang, M.; Marquez-Curtis, L.; McGann, L.; Janowska-Wieczorek, A.; Xing, J.; Swanson, E.; Chen, J. *Biotechnology Letters* **2012**, *34*, 1965-1973.
- (170) Zhou, S. X.; Schmelz, A.; Seufferlein, T.; Li, Y. P.; Zhao, J. S.; Bachem, M. G. *J. Biol. Chem.* **2004**, *279*, 54463-54469.
- (171) Min, B. H.; Woo, J. I.; Cho, H. S.; Choi, B. H.; Park, S. J.; Choi, M. J.; Park, S. R. *Scand. J. Rheumatol.* **2006**, *35*, 305-311.
- (172) Sun, J. S.; Hong, R. C.; Chang, W. H. S.; Chen, L. T.; Lin, F. H.; Liu, H. C. *J. Biomed. Mater. Res.* **2001**, *57*, 449-456.
- (173) Ashokkumar, M. *Ultrason. Sonochem.* **2015**, *25*, 17-23.
- (174) Wood, B. E.; Aldrich, H. C.; Ingram, L. O. *Biotechnol. Prog.* **1997**, *13*, 232-237.
- (175) Savchenko, O.; Xing, J. D.; Yang, X. Y.; Gu, Q. R.; Shaheen, M.; Huang, M.; Yu, X. J.; Burrell, R.; Patra, P.; Chen, J. *Scientific Reports* **2017**, *7*, 13.
- (176) Martin, E. *Ultrasound* **2009**, *17*, 214-219.
- (177) Furusawa, Y.; Hassan, M. A.; Zhao, Q. L.; Ogawa, R.; Tabuchi, Y.; Kondo, T. *Ultrason. Sonochem.* **2014**, *21*, 2061-2068.
- (178) Shindo, T.; Ito, K.; Ogata, T.; Hatanaka, K.; Kurosawa, R.; Eguchi, K.; Kagaya, Y.; Hanawa, K.; Aizawa, K.; Shiroto, T.; Kasukabe, S.; Miyata, S.; Taki, H.; Hasegawa, H.; Kanai, H.; Shimokawa, H. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2016**, *36*, 1220.
- (179) Zhao, Y.; Ang, W. T.; Xing, J.; Zhang, J.; Chen, J. *Ultrasound Med. Biol.* **2012**, *38*, 1582-1588.
- (180) Shaheen, M.; Choi, M.; Ang, W.; Zhao, Y.; Xing, J.; Yang, R.; Xing, J.; Zhang, J.; Chen, J. *Renewable Energy* **2013**, *57*, 462-468.

- (181) Lindén, T.; Peetre, J.; Hahn-Hägerdal, B. *Appl. Environ. Microbiol.* **1992**, *58*, 1661-1669.
- (182) Stanley, D.; Bandara, A.; Fraser, S.; Chambers, P. J.; Stanley, G. A. *Journal of Applied Microbiology* **2010**, *109*, 13-24.
- (183) Xia, J.; Mandal, R.; Sinelnikov, I. V.; Broadhurst, D.; Wishart, D. S. *Nucleic Acids Res.* **2012**, *40*, W127-W133.
- (184) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorn Dahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. *Nucleic Acids Res.* **2013**, *41*, D801-D807.
- (185) Bretkreutz, A.; Boucher, L.; Bretkreutz, B.-J.; Sultan, M.; Jurisica, I.; Tyers, M. *Genetics* **2003**, *165*, 997.
- (186) Carvalho-Netto, O. V.; Carazzolle, M. F.; Mofatto, L. S.; Teixeira, P. J. P. L.; Noronha, M. F.; Calderón, L. A. L.; Mieczkowski, P. A.; Argueso, J. L.; Pereira, G. A. G. *Microb. Cell. Fact.* **2015**, *14*, 13.
- (187) Xia, J.; Wishart, D. S. *Nat. Protoc.* **2011**, *6*, 743.
- (188) Xu, W.; Chen, D.; Wang, N.; Zhang, T.; Zhou, R.; Huan, T.; Lu, Y.; Su, X.; Xie, Q.; Li, L.; Li, L. *Anal. Chem.* **2015**, *87*, 829-836.
- (189) Slavov, N.; Budnik, B. A.; Schwab, D.; Airoidi, E. M.; van Oudenaarden, A. *Cell reports* **2014**, *7*, 705-714.
- (190) Snyder, N. W.; Tomblin, G.; Worth, A. J.; Parry, R. C.; Silvers, J. A.; Gillespie, K. P.; Basu, S. S.; Millen, J.; Goldfarb, D. S.; Blair, I. A. *Analytical Biochemistry* **2015**, *474*, 59-65.
- (191) White, W. H.; Gunyuzlu, P. L.; Toyn, J. H. *J. Biol. Chem.* **2001**, *276*, 10794-10800.

- (192) Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. *Nucleic Acids Res.* **2016**, *44*, D457-D462.
- (193) Zaman, S.; Lippman, S. I.; Zhao, X.; Broach, J. R. *Annual Review of Genetics* **2008**, *42*, 27-81.
- (194) Patti, G. J.; Yanes, O.; Siuzdak, G. *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 263-269.
- (195) Teng, Q.; Huang, W. L.; Collette, T. W.; Ekman, D. R.; Tan, C. *Metabolomics* **2009**, *5*, 199-208.
- (196) Danielsson, A. P. H.; Moritz, T.; Mulder, H.; Spegel, P. *Anal. Biochem.* **2010**, *404*, 30-39.
- (197) Dettmer, K.; Nurnberger, N.; Kaspar, H.; Gruber, M. A.; Almstetter, M. F.; Oefner, P. J. *Anal. Bioanal. Chem.* **2011**, *399*, 1127-1139.
- (198) Lorenz, M. A.; Burant, C. F.; Kennedy, R. T. *Anal. Chem.* **2011**, *83*, 3406-3414.
- (199) Martineau, E.; Tea, I.; Loac, G.; Giraudeau, P.; Akoka, S. *Anal. Bioanal. Chem.* **2011**, *401*, 2133-2142.
- (200) Bi, H. C.; Krausz, K. W.; Manna, S. K.; Li, F.; Johnson, C. H.; Gonzalez, F. J. *Anal. Bioanal. Chem.* **2013**, *405*, 5279-5289.
- (201) Hounoum, B. M.; Blasco, H.; Nadal-Desbarats, L.; Dieme, B.; Montigny, F.; Andres, C. R.; Emond, P.; Mavel, S. *Anal. Bioanal. Chem.* **2015**, *407*, 8861-8872.
- (202) Kapoore, R. V.; Coyle, R.; Staton, C. A.; Brown, N. J.; Vaidyanathan, S. *Metabolomics* **2015**, *11*, 1743-1755.
- (203) Muschet, C.; Moller, G.; Prehn, C.; de Angelis, M. H.; Adamski, J.; Tokarz, J. *Metabolomics* **2016**, *12*, 12.
- (204) Kostidis, S.; Addie, R. D.; Morreau, H.; Mayboroda, O. A.; Giera, M. *Anal. Chim. Acta.* **2017**, *980*, 1-24.

- (205) Ibanez, C.; Simo, C.; Palazoglu, M.; Cifuentes, A. *Anal. Chim. Acta.* **2017**, *986*, 48-56.
- (206) Yang, W. C.; Regnier, F. E.; Jiang, Q.; Adamec, J. *J. Chromatogr. A* **2010**, *1217*, 667-675.
- (207) Dai, W. D.; Huang, Q.; Yin, P. Y.; Li, J.; Zhou, J.; Kong, H. W.; Zhao, C. X.; Lu, X.; Xu, G. W. *Anal. Chem.* **2012**, *84*, 10245-10251.
- (208) Leng, J. P.; Wang, H. Y.; Zhang, L.; Zhang, J.; Wang, H.; Guo, Y. L. *Anal. Chim. Acta* **2013**, *758*, 114-121.
- (209) Tayyari, F.; Gowda, G. A. N.; Gu, H. W.; Raftery, D. *Anal. Chem.* **2013**, *85*, 8715-8721.
- (210) Yuan, W.; Li, S. W.; Edwards, J. L. *Anal. Chem.* **2015**, *87*, 7660-7666.
- (211) Mochizuki, T.; Todoroki, K.; Inoue, K.; Min, J. Z.; Toyo'oka, T. *Anal. Chim. Acta* **2014**, *811*, 51-59.
- (212) Chu, J. M.; Qi, C. B.; Huang, Y. Q.; Jiang, H. P.; Hao, Y. H.; Yuan, B. F.; Feng, Y. Q. *Anal. Chem.* **2015**, *87*, 7364-7372.
- (213) Arrivault, S.; Guenther, M.; Fry, S. C.; Fuenfgeld, M.; Veyel, D.; Mettler-Altmann, T.; Stitt, M.; Lunn, J. E. *Anal. Chem.* **2015**, *87*, 6896-6904.
- (214) Wong, J. M. T.; Malec, P. A.; Mabrouk, O. S.; Ro, J.; Dus, M.; Kennedy, R. T. *J. Chromatogr. A* **2016**, *1446*, 78-90.
- (215) Yu, L.; Ding, J.; Wang, Y. L.; Liu, P.; Feng, Y. Q. *Anal. Chem.* **2016**, *88*, 1286-1293.
- (216) Hao, L.; Johnson, J.; Lietz, C. B.; Buchberger, A.; Frost, D.; Kao, W. J.; Li, L. J. *Anal. Chem.* **2017**, *89*, 1138-1146.
- (217) Luo, X.; Li, L. *Anal. Chem.* **2017**, *89*, 11664-11671.
- (218) Xia, J.; Sinelnikov, I. V.; Han, B.; Wishart, D. S. *Nucleic Acids Res.* **2015**, *43*, W251-W257.
- (219) Halliwell, B. *FEBS Lett.* **2003**, *540*, 3-6.

- (220) Zamboni, N.; Saghatelian, A.; Patti, G. J. *Mol. Cell* **2015**, *58*, 699-706.
- (221) Hounoum, B. M.; Blasco, H.; Emond, P.; Mavel, S. *Trac-Trends Anal. Chem.* **2016**, *75*, 118-128.
- (222) Shyh-Chang, N.; Ng, H. H. *Genes Dev.* **2017**, *31*, 336-346.
- (223) Jackson, J. M.; Witek, M. A.; Kamande, J. W.; Soper, S. A. *Chem. Soc. Rev.* **2017**, *46*, 4245-4280.
- (224) Datta, S.; Malhotra, L.; Dickerson, R.; Chaffee, S.; Sen, C. K.; Roy, S. *Histol. Histopath.* **2015**, *30*, 1255-1269.
- (225) Comi, T. J.; Do, T. D.; Rubakhin, S. S.; Sweedler, J. V. *J. Am. Chem. Soc.* **2017**, *139*, 3920-3929.
- (226) Zenobi, R. *Science* **2013**, *342*.
- (227) Amantonico, A.; Urban, P. L.; Zenobi, R. *Anal. Bioanal. Chem.* **2010**, *398*, 2493-2504.
- (228) Li, L.; Golding, R. E.; Whittall, R. M. *J. Am. Chem. Soc.* **1996**, *118*, 11662-11663.
- (229) Guillaume-Gentil, O.; Rey, T.; Kiefer, P.; Ibanez, A. J.; Steinhoff, R.; Bronnimann, R.; Dorwling-Carter, L.; Zambelli, T.; Zenobi, R.; Vorholt, J. A. *Anal. Chem.* **2017**, *89*, 5017-5023.
- (230) Walker, B. N.; Antonakos, C.; Retterer, S. T.; Vertes, A. *Angew. Chem. Int. Ed.* **2013**, *52*, 3650-3653.
- (231) Shrestha, B.; Vertes, A. *Anal. Chem.* **2009**, *81*, 8265-8271.
- (232) Fujii, T.; Matsuda, S.; Tejedor, M. L.; Esaki, T.; Sakane, I.; Mizuno, H.; Tsuyama, N.; Masujima, T. *Nat. Protoc.* **2015**, *10*, 1445-1456.
- (233) Gong, X. Y.; Zhao, Y. Y.; Cai, S. Q.; Fu, S. J.; Yang, C. D.; Zhang, S. C.; Zhang, X. R. *Anal. Chem.* **2014**, *86*, 3809-3816.
- (234) Zhao, S.; Dawe, M.; Guo, K.; Li, L. *Anal. Chem.* **2017**, *89*, 6758-6765.

- (235) Huan, T.; Li, L. *Anal. Chem.* **2015**, *87*, 7011-7016.
- (236) Huan, T.; Wu, Y. M.; Tang, C. Q.; Lin, G. H.; Li, L. *Anal. Chem.* **2015**, *87*, 9838-9845.
- (237) Brown, R. B.; Audet, J. *J. R. Soc. Interface* **2008**, *5*, S131-S138.
- (238) Uehara, T.; Yokoi, A.; Aoshima, K.; Tanaka, S.; Kadowaki, T.; Tanaka, M.; Oda, Y. *Anal. Chem.* **2009**, *81*, 3836-3842.
- (239) Yuan, W.; Anderson, K. W.; Li, S. W.; Edwards, J. L. *Anal. Chem.* **2012**, *84*, 2892-2899.
- (240) Jones, D. R.; Wu, Z. P.; Chauhan, D.; Anderson, K. C.; Peng, J. M. *Anal. Chem.* **2014**, *86*, 3667-3675.
- (241) Chetwynd, A. J.; Abdul-Sada, A.; Hill, E. M. *Anal. Chem.* **2015**, *87*, 1158-1165.
- (242) Li, Z. D.; Tatlay, J.; Li, L. *Anal. Chem.* **2015**, *87*, 11468-11474.
- (243) Li, Z.; Li, L. *Anal. Chem.* **2014**, *86*, 331-335.
- (244) van Niel, G.; D'Angelo, G.; Raposo, G. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 213-228.
- (245) Cocucci, E.; Meldolesi, J. *Trends Cell Biol.* **2015**, *25*, 364-372.
- (246) Lai, X.; Wang, M.; McElyea, S. D.; Sherman, S.; House, M.; Korc, M. *Cancer Lett.* **2017**, *393*, 86-93.
- (247) Malla, B.; Zaugg, K.; Vassella, E.; Aebersold, D. M.; Dal Pra, A. *Int. J. Radiat. Oncol. Biol. Phys.* **2017**, *98*, 982-995.
- (248) Mathivanan, S.; Simpson, R. J. *Proteomics* **2009**, *9*, 4997-5000.
- (249) Simpson, R. J.; Jensen, S. S.; Lim, J. W. *Proteomics* **2008**, *8*, 4083-4099.
- (250) Simpson, R. J.; Lim, J. W.; Moritz, R. L.; Mathivanan, S. *Expert Rev. Proteomics* **2009**, *6*, 267-283.
- (251) Chen, Y.; Xie, Y.; Xu, L.; Zhan, S.; Xiao, Y.; Gao, Y.; Wu, B.; Ge, W. *Int. J. Cancer* **2017**, *140*, 900-913.

- (252) Fontana, S.; Saieva, L.; Taverna, S.; Alessandro, R. *Proteomics* **2013**, *13*, 1581-1594.
- (253) Franzen, C. A.; Blackwell, R. H.; Foreman, K. E.; Kuo, P. C.; Flanigan, R. C.; Gupta, G. N. *J. Urol.* **2016**, *195*, 1331-1339.
- (254) Azmi, A. S.; Bao, B.; Sarkar, F. H. *Cancer Metastasis Rev.* **2013**, *32*, 10.1007/s10555-10013-19441-10559.
- (255) Drake, R. R.; Kislinger, T. *Expert Rev. Proteomics* **2014**, *11*, 167-177.
- (256) Henderson, M. C.; Azorsa, D. O. *Front Oncol* **2012**, *2*, 38.
- (257) Lu, L.; Risch, H. A. *Future Oncol.* **2016**, *12*, 1081-1090.
- (258) Nilsson, J.; Skog, J.; Nordstrand, A.; Baranov, V.; Mincheva-Nilsson, L.; Breakefield, X. O.; Widmark, A. *Br. J. Cancer* **2009**, *100*, 1603-1607.
- (259) Whiteside, T. L. *Adv. Clin. Chem.* **2016**, *74*, 103-141.
- (260) Li, M.; Zeringer, E.; Barta, T.; Schageman, J.; Cheng, A.; Vlassov, A. V. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **2014**, *369*, 20130502.
- (261) An, M.; Lohse, I.; Tan, Z.; Zhu, J.; Wu, J.; Kurapati, H.; Morgan, M. A.; Lawrence, T. S.; Cuneo, K. C.; Lubman, D. M. *J. Proteome Res.* **2017**, *16*, 1763-1772.
- (262) Sun, Y.; Liu, S.; Qiao, Z.; Shang, Z.; Xia, Z.; Niu, X.; Qian, L.; Zhang, Y.; Fan, L.; Cao, C. X.; Xiao, H. *Anal. Chim. Acta* **2017**, *982*, 84-95.
- (263) Gonzalez-Begne, M.; Lu, B.; Han, X.; Hagen, F. K.; Hand, A. R.; Melvin, J. E.; Yates, J. R. *J. Proteome Res.* **2009**, *8*, 1304-1314.
- (264) Olver, C.; Vidal, M. *Subcell. Biochem.* **2007**, *43*, 99-131.
- (265) Principe, S.; Jones, E. E.; Kim, Y.; Sinha, A.; Nyalwidhe, J. O.; Brooks, J.; Semmes, O. J.; Troyer, D. A.; Lance, R. S.; Kislinger, T.; Drake, R. R. *Proteomics* **2013**, *13*, 1667-1671.

- (266) Llorente, A.; Skotland, T.; Sylvanne, T.; Kauhanen, D.; Rog, T.; Orłowski, A.; Vattulainen, I.; Ekroos, K.; Sandvig, K. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2013**, *1831*, 1302-1309.
- (267) Buratta, S.; Urbanelli, L.; Sagini, K.; Giovagnoli, S.; Caponi, S.; Fioretto, D.; Mitro, N.; Caruso, D.; Emiliani, C. *PLoS One* **2017**, *12*, 23.
- (268) Roberg-Larsen, H.; Lund, K.; Seterdal, K. E.; Solheim, S.; Vehus, T.; Solberg, N.; Krauss, S.; Lundanes, E.; Wilson, S. R. *J. Steroid Biochem. Mol. Biol.* **2017**, *169*, 22-28.
- (269) Skotland, T.; Sandvig, K.; Llorente, A. *Prog. Lipid Res.* **2017**, *66*, 30-41.
- (270) Zhao, H.; Yang, L.; Baddour, J.; Achreja, A.; Bernard, V.; Moss, T.; Marini, J. C.; Tudawe, T.; Seviour, E. G.; San Lucas, F. A.; Alvarez, H.; Gupta, S.; Maiti, S. N.; Cooper, L.; Peehl, D.; Ram, P. T.; Maitra, A.; Nagrath, D. *eLife* **2016**, *5*, e10250.
- (271) Puhka, M.; Takatalo, M.; Nordberg, M.-E.; Valkonen, S.; Nandania, J.; Aatonen, M.; Yliperttula, M.; Laitinen, S.; Velagapudi, V.; Mirtti, T.; Kallioniemi, O.; Rannikko, A.; Siljander, P. R. M.; af Hällström, T. M. *Theranostics* **2017**, *7*, 3824-3841.
- (272) Iraci, N.; Gaude, E.; Leonardi, T.; Costa, A. S. H.; Cossetti, C.; Peruzzotti-Jametti, L.; Bernstock, J. D.; Saini, H. K.; Gelati, M.; Vescovi, A. L.; Bastos, C.; Faria, N.; Occhipinti, L. G.; Enright, A. J.; Frezza, C.; Pluchino, S. *Nat. Chem. Biol.* **2017**, *13*, 951-+.
- (273) Kim, J.; Tan, Z.; Lubman, D. M. *Electrophoresis* **2015**, *36*, 2017-2026.
- (274) Ko, J.; Carpenter, E.; Issadore, D. *Analyst* **2016**, *141*, 450-460.
- (275) An, M.; Zhu, J.; Wu, J.; Cuneo, K. C.; Lubman, D. M. *J. Proteome Res.* **2018**, *17*, 1690-1699.
- (276) Schultz, M. C.; Zhang, J.; Luo, X.; Savchenko, O.; Li, L.; Deyholos, M.; Chen, J. *J. Proteome Res.* **2017**, *16*, 2975-2982.

(277) Luo, X.; Gu, X.; Li, L. *Anal. Chim. Acta* **2018**, *in press*
(<https://doi.org/10.1016/j.aca.2017.11.054>).

(278) Altadill, T.; Campoy, I.; Lanau, L.; Gill, K.; Rigau, M.; Gil-Moreno, A.; Reventos, J.; Byers, S.; Colas, E.; Cheema, A. K. *PLoS One* **2016**, *11*, 17.