Method Development and Application of Mass Spectrometry for Better Understanding of Disease Mechanisms and Diagnosis

By
Yuan Liu

A dissertation in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Pharmaceutical Sciences)

at the
University of Wisconsin-Madison
2023

Date of final oral examination: May 19th, 2023

The dissertation is approved by the following members of the Final Oral Committee:

Lingjun Li, Professor, School of Pharmacy and Department of Chemistry
Glen Kwon, Professor, School of Pharmacy
Melissa Skala, Professor, Morgridge Institute for Research and Biomedical Engineering
Tim S Bugni, Professor, School of Pharmacy

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr. Lingjun Li, without whom this work would not have been possible. She provided me with the opportunity to rotate in her lab and accepted me as a graduate student. Throughout my academic and professional journey, she has offered me continuous support and freedom, including funding and guidance for experiments, a research assistant position, and opportunities to attend conferences. She also helped me connect with prospective scholars when I sought postdoctoral positions in the late stages of my graduate training and provided me with strong recommendation letters. Her guidance, patience, and hard work have been a constant source of motivation for me throughout my graduate career, and I am grateful for her mentorship as I pursue becoming an independent scientist. Working under her supervision has been a great honor, and I will forever be thankful for this opportunity.

I would also like to express my sincere gratitude to my committee members, Dr. Melissa Skala, Dr. Tim Bugni, and Dr. Glen Kwon, for their invaluable time, insightful feedback, and unwavering support throughout my graduate studies. From the rigorous preparation for my preliminary exam, to the constructive discussions during our annual committee meetings, to the valuable input and guidance on this dissertation, their contributions have been fundamental to my growth as a scientist. I am deeply grateful for their mentorship and encouragement, which have inspired me to pursue excellence in research and teaching.

Furthermore, I would like to acknowledge Dr. Yi Yu, Dr. Xin Liu, and Dr. Yuhui Sun for their invaluable support during my undergraduate years at Wuhan University. Their acceptance into their labs and their strong recommendation letters for my graduate program application encouraged me to pursue a career in scientific research. Their teaching and encouragement ignited

my scientific curiosity and played a crucial role in setting me on the path to becoming a well-rounded scientist.

Throughout my entire graduate career, I have received invaluable assistance from my colleagues. I would like to express my gratitude to several of them who have made significant contributions to my projects. Dr. Xiaofang Zhong introduced me to proteomics during my lab rotation and provided guidance on sample preparation for proteomics. Dr. Jillian Johnson introduced me to MALDI MSI, proposed the cell co-culture project, and taught me how to perform 3D cell culture. Dr. Gongyu Li mentored me to finish my first project and guided me in writing the first manuscript. Dr. Hua Zhang provided invaluable support and expertise for a variety of experiments and issues. Zicong Wang and Dr. Ting-jia Gu were immensely helpful in chemical synthesis and provided valuable insights during discussions. Dr. Zihui Li was a great scientific discussion partner. Jingwei Zhang was wonderful to collaborate with. Additionally, I am grateful to have had the opportunity to collaborate with Dr. Henrik Zetterberg from the University of Gothenburg and Dr. William F. Dove, Dr. Perry J. Pickhardt, and Dr. Mark Reichelderfer, who not only provided valuable samples but also offered insightful comments and suggestions during manuscript preparation.

I would like to express my gratitude to several individuals who have played a significant role in my scientific development by training me on various mass spectrometry instruments and techniques. Dr. Qinying Yu, Caitlin Keller, Dr. Meng Xu, and Dr. Nhu Vu trained me on LTQ MALDI Orbitrap, Dr. Jillian Johnson and Wenxin (Wendy) Wu trained me on Rapiflex, and Dr. Meng Xu trained me on TM-sprayer. Dr. Xiaofang Zhong and Dr. Junfeng Huang trained me on Q Exactive and HILIC-enrichment and Dr. Junfeng Huang also trained me on packing column. Dr. Danqing Wang and Dr. Junfeng Wang trained me on high pH fractionation and glycopeptide data

analysis using Proteome Discoverer and Byonic, respectively. Dr. Dustin Frost trained me on Orbitrap Elite, while Dylan Tabang and Hannah Miles trained me on Q Exactive HF. Dr. Zihui (Jerry) Li trained me on Fusion Lumos and using Maxquant for proteomics data analysis. Dr. Graham Delafield trained me on Agilent Q-TOF and LC-MS on Bruker TIMS TOF (Bruker), while Peng-Hsuan (Timo) Huang and Dr. Hua Zhang trained me on MALDI on TIMS TOF. Additionally, Zicong Wang, Dr. Miyang (Mike) Li, Dr. Ting-Jia (Gutin) Gu, and Peng-Kai Liu provided me with DiLeu tags.

I am also grateful to the following individuals who have provided me with their support and guidance: Dr. Haiyan Lu, who has been very helpful in valuable discussions and experiments, and Haoran Zhang, who is great to work with and helped fix the TM-sprayer. Anne-Sophie Mancha in Dr. Melissa Skala's lab has been helpful in discussions for the co-culture project, and Dr. Xudong (Daniel) Shi has been very helpful and provided technical support for the use of microscopes. Dr. Kellen DeLaney, Dr. Chris Sauer, Dr. Yu Feng, Dr. Min Ma, Dr. Shuling Xu, Dr. Pingli Wei, Dr. Yusi Cui, Dr. Yatao Shi, Dr. Ashley Phetsanthad, Jericha Mill, Dr. Bin (Ben) Wang, Jiabao Guo, Zhijun (Andrew) Zhu, Olga Milian-Martinez, Gaoyuan Lu, Feixuan Wu, Zexin (Peter) Zhu, and Lauren Fields Ching-Yuan (Lily) Yang have been wonderful colleagues and friends, and I am grateful for the chance to work with them in the same lab. I would like to thank Dr. Cameron Scarlett and Molly Pellitteri Hahn of the Analytical Instrumentation Center for maintaining and troubleshooting instruments in the School of Pharmacy. Their help has been invaluable to me throughout my graduate career.

I would like to express my gratitude to Dr. Melgardt M. de Villiers and Bonnie J. Fingerhut, whom I had the honor of working with as a teaching assistant. I am also immensely grateful to Ken Niemeyer, the outstanding graduate program coordinator, for his support and kindness

towards all graduate students in the Pharmaceutical Sciences PhD program. I would also like to extend my thanks to the current graduate program coordinator, Derek Zimmerman, for his assistance with graduation-related matters.

Finally, I would like to express my heartfelt appreciation to my wonderful family. My grandparents, mother, father, and elder sister have been a constant source of support for me, from before the start of my graduate program and throughout the past five years. My girlfriend, Xiaodan Zhang, has been an invaluable pillar of support, providing me with unstoppable encouragement and assistance in all aspects of my graduate studies.

Table of Contents

Method Development and Application of Mass Spectrometry for Better Understanding of Disease Mechanisms and Diagnosis

Yuan Liu
Under the supervision of Professor Lingjun Li
At the University of Wisconsin-Madison

Abstract

Mass spectrometry is a very powerful tool for identifying and quantifying biomolecules of interest, including metabolites, drug compounds, proteins, and post-translational modifications of (PTMs) of proteins. In mass spectrometry (MS)-based analytical techniques, liquid-chromatography coupled with tandem mass spectrometry (LC-MS/MS) is widely used for biomolecule analysis, including relative and absolute quantification of metabolites, proteins and PTMs. Data dependent acquisition (DDA) and data-independent acquisition (DIA) are two main methods for quantification of biomolecules in LC-MS/MS. By integrating isotopic and/or isobaric tags labeling, the DDA method enables precise absolute quantification and high-throughput analysis of biomolecules, facilitating biomarker discovery and elucidation of disease mechanisms. In contrast, DIA has garnered significant interest due to its exceptional reproducibility and depth in identifying and quantifying biomolecules. In this dissertation work, I apply these isotopic/isobaric tag labeling and label-free strategies to investigate potential biomarkers in the serum of Alzheimer's disease (AD) and colorectal cancer (CRC). Additionally, I explore the crosstalk between pancreatic stellate cells and pancreatic cancer cells using multi-faceted MS approaches.

In addition to LC-MS/MS, mass spectrometry imaging (MSI) offers unparalleled insights into the spatial distribution of biomolecules across organs, tissues, and cell cultures, further enhancing our understanding of biological systems. We develop a tagging method to enhance the identification, quantification, and visualization of amine-containing metabolites by matrix-assisted

laser desorption/ionization (MALDI) mass spectrometry. Furthermore, I optimize a workflow for rapid sample preparation and high throughput MSI of biomolecule distribution in a widely used three-dimensional cell culture system - spheroids.

In summary, the integration of various mass spectrometry techniques and method development discussed in this dissertation demonstrates that combining different mass spectrometry approaches can provide a more comprehensive molecular picture of the complex biological system, ultimately enhancing our understanding of disease mechanisms and advancing diagnosis and treatment options.

List of Abbreviations and Acronyms

FA	Formic acid
FDR	FALSE discovery rate
GO	Gene ontology
HCD	Higher-energy collisional dissociation
HILIC	Hydrophilic interaction chromatography
HMDB	Human Metabolome Database
HPLC	High-performance liquid chromatography
IAA	Iodoacetamide
IM-MS	Ion mobility mass spectrometry
IT	Ion trap
iTRAQ	Isobaric tag for relative and absolute quantification
LAESI	Laser ablation electrospray
LC	Liquid chromatography
LMWM	Low-molecular-weight molecule
Lys-C	Endoproteinase lysine C
m/z	Mass-to-charge ratio
MALDI	Matrix-assisted laser desorption
MS	Mass spectrometry
MS1	Precursor mass analysis
MS2	Tandem mass spectrometry
MSI	Mass spectrometry imaging

mTRAQ	Mass differential tags for relative and absolute quantification
NCE	Normalized collision energy
NMR	Nuclear magnetic resonance
nsPCR	Nanosecond photochemical reaction
OC	Optical colonoscopy
OCR	Oxygen consumption rate
PCA	Principal component analysis
PSCs	Pancreatic cancer cells
PDAC	pancreatic ductal adenocarcinoma
POD	Probability of detection
ppm	Parts per million
PSCs	Pancreatic stellate cells
PSM	Peptide-spectrum match
p-tau	Phosphorylated tau
PTM	Post-translational modification
ROC	Receiver operating characteristic
RPLC	Reversed phase liquid chromatography
RT	Retention time
S/N	Signal-to-noise ratio
SCX	Strong cation exchange chromatography
SILAC	Stable isotope labeling by amino acids
TFA	Trifluoroacetic acid

TIC	Total ion chromatogram
TMT	Tandem mass tag
UMAP	Uniform manifold approximation and projection for dimension reduction

Chapter 1

Introduction and Research Summary

Introduction and Research Summary

To enhance our understanding of disease mechanisms and advance diagnosis and treatment options, it proves that it is extremely beneficial to have a comprehensive perspective on genomics, epigenomics, transcriptomics, proteomics, metabolomics, and post-translational modifications of proteins. Mass spectrometry (MS)-based approaches facilitate the identification and quantification of numerous metabolites, proteins, and other biomolecules. Derived from various sources such as organs, tissues, body fluids, and cells, these biomolecules can be analyzed using a comprehensive global approach or a focused targeted strategy. Both methods offer exceptional sensitivity and accuracy, contributing to a deeper understanding of the biological systems. In this dissertation, I applied multiple qualitative and quantitative approaches for untargeted global MS-based metabolomics, proteomics analysis in several biological systems. These applications of MS enable us to profile changes of metabolite and protein levels between disease and normal conditions. Additionally, I develop a tagging strategy for improving the identification and quantification of primary amine-containing metabolites. Furthermore, I optimize a workflow for rapid and high throughput MSI of spheroids. I expect the advancements of technologies presented in this dissertation will lead to a better understanding of disease mechanisms and advancing diagnosis and treatment options.

Chapter 1 is a general introduction to this dissertation and briefly summarizes the major findings of all projects. **Chapter 2** presents the method development of a tagging strategy for enhancing identification, quantification, and visualization of primary amine-containing metabolites. Numerous metabolites, such as amino acids, neurotransmitters, and pharmaceuticals, possess primary amine functional groups. MS analysis of these molecules is crucial for studying cancers and psychogenic diseases.[1-3] However, challenges persist in MS-based detection and

visualization of these bioactive molecules in complex biological systems, primarily due to low ionization efficiency and matrix interference effects. In this study, we present a novel, efficient approach called nanosecond photochemical reaction (nsPCR)-enabled fast chemical derivatization, which facilitates direct MS analysis of primary amine-containing metabolites. This method significantly enhances detection sensitivity for a wide range of metabolites in cell culture media and rat brain sections. Moreover, the nsPCR-based chemical derivatization strategy proves to be a valuable visualization tool, providing improved spatial information for these metabolites. As a result, this technique holds the potential to serve as an alternative approach for gaining new insights into metabolic processes.

In addition to identification and quantification of molecules with low abundance or low ionization efficiency, the tedious sample preparation for MSI also hinder the investigation of biological systems, including drug development and cell-cell interactions. To address this challenge, **Chapter 3** introduces a strategy for rapid sample preparation of multicellular tumor spheroids (MCTS) for high throughput MSI. MCTS are vital for biological studies and cancer research. Mass spectrometry imaging (MSI) is increasingly employed to investigate MCTS, but current sample preparation methods for spheroids often suffer from limitations such as low throughput or extended preparation times.[4, 5] To address these challenges, an optimized sectioning method for preparing MCTS frozen sections was developed. MCTS grown on agarose-based microarrays were coated with gelatin, allowing for simultaneous sectioning of multiple spheroids. This innovative method was utilized to study lipid distribution in spheroids using matrix-assisted laser desorption/ionization MSI. Compared to the ice-assisted sectioning technique, our approach enables high-throughput sectioning of a larger number of spheroids. Additionally, sectioning spheroids alongside the agarose-based microarray improves their visualization and eliminates the

need for staining to locate spheroid positions, thereby simplifying sample preparation procedures. Consequently, this method offers a high-throughput and convenient strategy for preparing spheroid sections for MSI studies.

In this dissertation, we not only develop a strategy for rapid sample preparation for high throughput MSI of pancreatic cancer spheroids, but also investigate the crosstalk between pancreatic stellate cells (PSCs) and pancreatic cancer cells (PCCs), as detailed in **Chapter 4**. In this chapter, I report on the study of crosstalk between PSCs and PCCs at protein levels using data-independent acquisition (DIA)-based label-free quantification methods. PSCs contribute to pancreatic cancer cell survival, proliferation, tumor growth, and an increasingly oxidized redox state in pancreatic cancer cells.[3, 6, 7] However, limited information is available regarding the alterations in protein levels in pancreatic cancer cells as they interact with PSCs, particularly within pancreatic cancer spheroids. Furthermore, there is a knowledge gap about the changes in protein levels of pancreatic cancer cells during different stages. In this chapter, I employ a DIA method to quantify protein levels in pancreatic cancer spheroids, including spheroids cultured alone and co-cultured with PSCs at various time points. This strategy allows us to profile the changes in protein levels within pancreatic cancer spheroids. Over 7,000 proteins were quantified in the samples. Gene ontology (GO) analysis revealed enriched GO terms in extended cell cultures, including response to hypoxia, response to oxygen levels, and ribonucleoprotein complex biogenesis, among others, as part of their biological processes. Additionally, we observed a roughly 24% increase in high-temperature requirement A serine peptidase 1 (HTRA1, or serine protease HTRA1) abundance in PANC1 spheroids co-cultured with PSCs compared to monocultured PANC1 spheroids.

My fascination with uncovering disease mechanisms and improving diagnostic methods extends beyond the realm of pancreatic cancer, encompassing a broader scope of medical conditions and health-related challenges, including biomarker discovery for Alzheimer's disease (AD) and colorectal cancer. **Chapter 5** shows an application of isotopic DiLeu tags for absolute quantification of two proteins – APOE and CLU in the serum of AD and control individuals. In recent decades, considerable efforts have been made towards discovering biomarkers for early diagnosis and treatment of AD. Blood-based biomarkers for AD diagnosis offer cost-effectiveness and reduced invasiveness, making their discovery and validation in AD patient blood samples a popular research topic in Alzheimer's and dementia studies. To date, amyloid β (Aβ), total-Tau, and phosphorylated tau (p-Tau) in blood have demonstrated remarkable accuracy and specificity in diagnosing AD.[8-10] However, the underlying mechanisms of AD pathology remain complex and elusive. In addition to these well-studied proteins, others like CLU and APOE have been linked to AD development, with both implicated in Aβ clearance and deposition. In this study, we applied an isotopic labeling strategy using iDiLeu tags for absolute quantification of CLU and APOE. By measuring the absolute concentrations of these proteins in blood, this study sheds light on their potential roles in AD pathology. Consequently, examining CLU and APOE heterogeneity in AD pathogenesis may facilitate the identification of pathological relevance for specific pathways between different genders, ultimately leading to personalized medicine. In addition, **Chapter 6** presents the application of DiLeu tags for relative quantification of primary/secondary amine-containing metabolites in the serum of colorectal cancer patients and control individuals. A significant challenge in reducing colorectal cancer mortality lies in screening patients using minimally invasive tests.[11] Blood tests offer a high compliance rate and decreased invasiveness. In this study, we employ a multiplex isobaric tag labeling strategy combined with mass

spectrometry to quantify primary and secondary amine-containing metabolites in serum for the purpose of discovering changes in metabolite levels associated with colorectal cancer. We analyze serum samples from patients with different risk levels and colorectal cancer growth statuses. Metabolite identification is based on accurate mass matching and/or retention time of labeled metabolite standards. Across all serum samples, we quantify 40 metabolites, including 18 metabolites validated with standards. We observe significantly reduced levels of threonine and asparagine in patients with growing or high-risk adenomas ($p < 0.05$). Glutamine levels decrease in patients with adenomas of unknown growth status or high-risk adenomas, while arginine levels increase in patients with low-risk adenomas. Based on our findings, we conclude that some of the identified metabolites may help differentiate colorectal patients with growing adenomas from healthy individuals and patients with adenomas of unknown growth status.

Finally, **Chapter 7** offers an overall conclusion of the key findings and methodological advancements from this dissertation work. It also provides some future directions for various ongoing projects.

References

1. Ng, J.; Papandreou, A.; Heales, S. J.; Kurian, M. A., Monoamine neurotransmitter disorders--clinical advances and future perspectives. *Nat Rev Neurol* **2015,** *11* (10), 567-84.

2. Zhang, W. C.; Shyh-Chang, N.; Yang, H.; Rai, A.; Umashankar, S.; Ma, S.; Soh, B. S.; Sun, L. L.; Tai, B. C.; Nga, M. E.; Bhakoo, K. K.; Jayapal, S. R.; Nichane, M.; Yu, Q.; Ahmed, D. A.; Tan, C.; Sing, W. P.; Tam, J.; Thirugananam, A.; Noghabi, M. S.; Pang, Y. H.; Ang, H. S.; Mitchell, W.; Robson, P.; Kaldis, P.; Soo, R. A.; Swarup, S.; Lim, E. H.; Lim, B., Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. *Cell* **2012,** *148* (1-2), 259-72.

3. Sousa, C. M.; Biancur, D. E.; Wang, X.; Halbrook, C. J.; Sherman, M. H.; Zhang, L.; Kremer, D.; Hwang, R. F.; Witkiewicz, A. K.; Ying, H.; Asara, J. M.; Evans, R. M.; Cantley, L. C.; Lyssiotis, C. A.; Kimmelman, A. C., Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. *Nature* **2016,** *536* (7617), 479-83.

4. Johnson, J.; Sharick, J. T.; Skala, M. C.; Li, L., Sample preparation strategies for high-throughput mass spectrometry imaging of primary tumor organoids. *J Mass Spectrom* **2020,** *55* (4), e4452.

5. Xie, P.; Zhao, C.; Liang, X.; Huang, W.; Chen, Y.; Cai, Z., Preparation of Frozen Sections of Multicellular Tumor Spheroids Coated with Ice for Mass Spectrometry Imaging. *Anal Chem* **2020,** *92* (11), 7413-7418.

6. Datta, R.; Sivanand, S.; Lau, A. N.; Florek, L. V.; Barbeau, A. M.; Wyckoff, J.; Skala, M. C.; Vander Heiden, M. G., Interactions with stromal cells promote a more oxidized cancer cell redox state in pancreatic tumors. *Sci Adv* **2022,** *8* (3), eabg6383.

7. Lau, A. N.; Li, Z.; Danai, L. V.; Westermark, A. M.; Darnell, A. M.; Ferreira, R.; Gocheva, V.; Sivanand, S.; Lien, E. C.; Sapp, K. M.; Mayers, J. R.; Biffi, G.; Chin, C. R.; Davidson, S. M.; Tuveson, D. A.; Jacks, T.; Matheson, N. J.; Yilmaz, O.; Vander Heiden, M. G., Dissecting cell-type-specific metabolism in pancreatic ductal adenocarcinoma. *Elife* **2020,** *9*.

8. Polanco, J. C.; Li, C.; Bodea, L. G.; Martinez-Marmol, R.; Meunier, F. A.; Gotz, J., Amyloid-beta and tau complexity - towards improved biomarkers and targeted therapies. *Nat Rev Neurol* **2018,** *14* (1), 22-39.

9. Dujardin, S.; Commins, C.; Lathuiliere, A.; Beerepoot, P.; Fernandes, A. R.; Kamath, T. V.; De Los Santos, M. B.; Klickstein, N.; Corjuc, D. L.; Corjuc, B. T.; Dooley, P. M.; Viode, A.; Oakley, D. H.; Moore, B. D.; Mullin, K.; Jean-Gilles, D.; Clark, R.; Atchison, K.; Moore, R.; Chibnik, L. B.; Tanzi, R. E.; Frosch, M. P.; Serrano-Pozo, A.; Elwood, F.; Steen, J. A.; Kennedy, M. E.; Hyman, B. T., Tau molecular diversity contributes to clinical heterogeneity in Alzheimer's disease. *Nat Med* **2020**.

10. Nakamura, A.; Kaneko, N.; Villemagne, V. L.; Kato, T.; Doecke, J.; Dore, V.; Fowler, C.; Li, Q. X.; Martins, R.; Rowe, C.; Tomita, T.; Matsuzaki, K.; Ishii, K.; Ishii, K.; Arahata, Y.; Iwamoto, S.; Ito, K.; Tanaka, K.; Masters, C. L.; Yanagisawa, K., High performance plasma amyloid-beta biomarkers for Alzheimer's disease. *Nature* **2018,** *554* (7691), 249-254.

11. Issa, I. A.; Noureddine, M., Colorectal cancer screening: An updated review of the available options. *World J Gastroenterol* **2017,** *23* (28), 5086-5096.

Chapter 2

Nanosecond Photochemical Reaction for Enhanced Identification, Quantification, and Visualization of Primary Amine-Containing Metabolites by MALDI-Mass Spectrometry

Adapted from:
Yuan Liu, Gongyu Li, Ting-jia Gu, and Lingjun Li. Nanosecond Photochemical Reaction for Enhanced Identification, Quantification, and Visualization of Primary Amine-Containing Metabolites by MALDI-Mass Spectrometry. (*Anal. Chem. 2022, 94, 3774−3781*)

Author contributions:
L.L. and G.L. designed and supervised the whole study; Y.L. performed experiments, analyzed, interpreted the data, and wrote the manuscript; T.G. performed the experiments. All authors provided feedback.

Abstract

Many metabolites, including amino acids, neurotransmitters, and pharmaceuticals, contain primary amine functional groups. The analysis of these molecules by mass spectrometry (MS) plays an important role in the study of cancers and psychogenic diseases. However, the MS-based detection and visualization of these bioactive metabolites directly from real biological systems still suffer from challenges such as low ionization efficiency and/or matrix interference effects. Here, we introduce a simple and efficient strategy, the nanosecond photochemical reaction (nsPCR)-enabled fast chemical derivatization, enabling direct MS analysis of primary amine-containing metabolites, with enhanced detection sensitivity for numerous metabolites from cell culture medium and rat brain sections. Furthermore, this nsPCR-based chemical derivatization strategy was demonstrated to be a useful visualizing tool that could provide improved spatial information for these metabolites, potentially offering alternative tools for gaining novel insights into metabolic events.

Introduction

As metabolites are the downstream product of a biosystem, metabolomics study is focused on the global analysis of metabolites in a biological specimen and is considered to contain most comprehensive and closely linked biomarker information that reflects the biological alterations.[1] Among the diverse metabolites, amino acids (and the analogues), neurotransmitters, and lipids are the top representative classes that attract the attention of the biological community. Reprogrammed metabolism is a hallmark of cancer cells.[2,3] Amino acids such as glycine and alanine have been shown to be strongly correlated to cancer proliferation and tumorigenesis for some cancers.[3,4] Dysregulation of neurotransmitters is considered a hallmark of various neurological and

psychiatric diseases.[5] Therefore, quantitative studies of small-molecule metabolites and visualization of spatial distributions of metabolites in samples are critical for understanding pathological pathways, improving diagnosis and treatment of diseases.

Nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) are two most widely used analytical tools in metabolite studies.[6] As a noninvasive technique, NMR spectroscopy enables valuable samples kept intact and to be used in the following experiments. Meanwhile, NMR can quantify analytes accurately and reproducibly.[7] However, NMR suffers from poor sensitivity and ubiquitous signal overlap of mixtures of analytes from complex samples such as cells and tissues, limiting its power to identify metabolites. MS-based techniques offer distinct advantages in speed, specificity, sensitivity, and wide detection range compared with other techniques, which enable exciting discoveries in medical research. Mass spectrometry imaging (MSI), with high molecular specificity, sensitivity, and throughput, has provided new dimensionality in the visualization of the metabolite distributions in tissues or cells. A variety of sampling/ionization methods have been employed for MSI, such as secondary ion mass spectrometry, laser ablation electrospray ionization (LAESI), direct liquid extraction and ionization (DLEI), and matrix-assisted laser desorption/ionization (MALDI).[6,8]

MALDI is a commercially available and commonly used soft ionization method. Unlike LC-MS-based analysis, which requires time-consuming and labor-intensive purification process, MALDI-MS-based analysis requires relatively simple sample preparation procedures due to its tolerance to relatively high level of salt and other contaminants. Also, commercial MALDI-MS instruments enable rapid analysis and can achieve high spatial resolution for MSI (5−10 μm or even lower).[9] However, the versatility of MALDI-MS for metabolite studies is still impaired by

inherently low ionization efficiency and concomitant ion suppression, especially for small molecules with lower mass overlapping with MALDI matrix peaks.[10]

To overcome the abovementioned challenges faced by MALDI-MS-based techniques, many chemical derivatization methods have been developed to facilitate small-molecule metabolite analysis.[11,12] Considering the importance and widespread distribution of amine-containing metabolites, including amino acids and neurotransmitters, some chemical derivatization methods focus on improving the detection and visualization of these metabolites specifically.[11,13] A few amine derivatization strategies were described in a review paper.[14] However, most of these strategies require incubation and/or heating to achieve decent derivatization efficiency, whereas these can increase analyte dislocation or degradation, compromising their accuracy in identification, quantification, and visualization of metabolites.

Developing fast and effective derivatization strategies for enhancing metabolite identification, quantification, and visualization is still key to advancing low-molecular-weight molecule (LMWM) analysis with MALDI-MS. In our previous work, we developed a nanosecond photochemical reaction (nsPCR)- based MS strategy for enhancing neuropeptide visualization in brain tissues and protein identification.[15] In this work, we demonstrate the capability of nsPCR-enhanced MALDI-MS analysis of LMWMs with the simple optimization of nsPCR reagents. The derivatization reaction is activated by laser without incubation or heating and can be completed at the nanosecond time scale with high spatial resolution on tissue samples. With similar working principles and derivatization protocols as recently proposed, we herein validated the successful application of the nsPCR-based MS strategy in enhancing identification and quantification of a key class of LMWMs, primary amine-containing metabolites (scheme shown in **Figure 1**).

Facilitated by the nsPCR strategy, the low abundance small molecules as demonstrated by nsPCR assisted MALDI imaging experiments.[16]

Materials and Methods

Chemicals and materials.

Methanol (MeOH), acetonitrile (ACN), and trifluoroacetic acid (TFA) were purchased from Fisher Scientific (Pittsburgh, PA, USA). 2-Nitrobenzaldehyde (NBA) was purchased from Chem-impex int'l Inc (Wood Dale, IL, USA). 4-Bromo-2-nitrobenzaldehyde (4- Br-2-NBA) was purchased from Tokyo Chemical Industry Co; LTD (Tokyo, Japan). 2-Hydroxy-5-nitrobenzaldehyde (2- NBA) and 4-dimethylamino-2-nitrobenzaldehyde (4-N-2- NBA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). 5-Hydroxy-2-nitrobenzaldehyde (5-NBA) was purchased from Acros organics (New Jersey, USA). Dulbecco's Modified Eagle Medium (DMEM, high glucose, GlutaMAX, pyruvate) was purchased from Gibco (Grand Island, NY, USA). Glycine, L-alanine, L-valine, DL-leucine, L-glutamic acid, dopamine, L-histidine, L-phenylalanine, L-arginine, and L-tryptophan were purchased from Sigma-Aldrich (St. Louis, MO, USA). The matrices of α-cyano-4- hydroxycinnamic acid (CHCA) were purchased from Sigma-Aldrich (St. Louis, MO, USA). All solvents used in this study were of high-performance liquid chromatography (HPLC) grade.

Sample preparation for MALDI spots.

Stock solutions of glycine, lysine, L-alanine, L-valine, DL-leucine, L-glutamic acid, dopamine, L-histidine, L-phenylalanine, L-arginine, and Ltryptophan were prepared as 25 or 50 mM in HPLC water, respectively. Then, the stock solution of the metabolites was mixed to get a metabolite mixture with equimolar concentrations (1 mM) for each standard. CHCA was prepared as 10 mg/mL (ACN/H2O/TFA (v/v/v) 49.95:49.95:0.1). Each nsPCR reagent (nsPCR reagents

evaluated in study shown in **Figure 2a**) was first dissolved in the mixed solvent of ACN/ EtOH/H2O/TFA (84/13/2.9/0.1, v/v/v/v). First, 2-NBA and its analogues were prepared at a concentration of 0, 100, and 200 mM. Metabolite mixture and metabolite standard were prepared in serial dilutions (0, 5, 10, 20, 40, 50, 100, 200, and 1000 μM) using 0.1% TFA. For each concentration of nsPCR reagents, it was tested by several concentrations of the metabolite mixture. The optimal concentration of nsPCR was further used to measure the LOD of valine and leucine; the concentrations of the metabolites were optimized based on the results. Glycine in DMEM was quantified by adding serial dilutions of the glycine standard. The standard glycine (0, 5, 10, 20, 40, and 100 μM) as the calibration standard was added to 10-fold diluted DMEM.

A 1:1:1 (v/v/v) mixture of sample, matrix, and nsPCR reagent was deposited onto a stainless-steel MALDI plate for all experiments (workflow shown in **Figure 2b**). A total volume of 2 μL for each spot was deposited on the MALDI plate. At least three replicates for each sample. The samples were dried down for 20 min in a fume hood in the dark before MALDI Orbitrap data acquisition.

Sample preparation for MALDI imaging.

nsPCR reagents were first dissolved in the mixed solvent of ACN/ EtOH/FA/H2O (84/13/0.3/2.7, v/v/v/v). Unless stated otherwise, 5 mg/mL of nsPCR reagent was used for imaging analysis. We complied with all relevant ethical regulations for animal testing and research. Animal experiments were conducted following institutional guidelines approved by the University of Wisconsin−Madison IACUC. The procedure for preparing mouse tissues was described in our previous study.15 Briefly, female C57BL/6J mice were anesthetized, perfused with chilled phosphate-buffered saline, and decapitated, and brains were removed. The brains were cut along midline and 100 mg mL−1 gelatin in water was used to embed each hemisphere, which was then

snap-frozen in dry ice. Tissue samples were stored at −80 °C. An Olympus SZX16 stereo microscope (Olympus, Center Valley, PA) was used to take the optical images of tissue sections with bright-field illumination. Tissue slides were dried for 30 min in a desiccator at room temperature right before matrix and nsPCR reagent application. Thereafter, 5 mg/mL CHCA matrix was deposited via the TM sprayer system using the following conditions: a nozzle temperature of 80 °C, a gas pressure of 10 psi, eight passes, a moving velocity of 1100 mm/min, a drying time of 30 s, and a flow rate of 0.2 mL/min. Then, the nsPCR reagent was sequentially deposited under the gentle condition of 30 °C and a gas pressure of 10 psi. The sprayed slides were stored in a desiccator after drying at room temperature for 30 min. At this point, the tissue slides were ready to be examined with MALDI-MSI experiments (workflow shown in **Figure 2c**).

Data acquisition and analysis.

Here, we used a MALDI LTQ- Orbitrap XL mass spectrometer for spot analysis. This vacuum MALDI source was featured with 60 Hz 337 nm N2 laser. For MALDI setting, plate motion is set as survey CPS, ASF is set as off, and the scan rate is set as 2 microscans/step. For laser setting, laser energy is set as 12 μJ, AGC is set as off, and the number of laser shots is set as 2. MS was operated under the positive polarity mode, with a mass range of *m/z* 70−400 and a mass resolution of 60 000. The number of experiments to acquire was set as 50 for each sample spot. Xcalibur 2.2 (Thermo Scientific, Bremen, Germany) was used to process mass spectra data.

The MALDI imaging experiments were performed using a Thermo QE-HF mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). A SubAP-MALDI source was coupled with the QE-HF orbitrap mass spectrometer for data acquisition. The SubAP-MALDI (ng) UHR (MassTech, Columbia, MD) was used as a laser ion source and a similar laser module can be found in our previous study.15 Notably, this SubAP-MALDI source was featured with a 355

nm high resolution Nd/YAG laser module along with an integrated ion funnel for improved ion collection and transmission. The typical operation source pressure was 3−5 Torr that made the source SubAP-MALDI. The starting laser parameters (unless otherwise specified) were a laser energy of 18.9% and a repetition rate of 1000 Hz. MS was operated under the positive polarity mode, with a mass range of *m/z* 133−2000 and a mass resolution of 60 000 unless otherwise specified. To ensure consistent scan time at each pixel, the maximum injection time was set to 300 ms and automatic gain control target was tuned to 3×10^6. Xcalibur (Thermo Scientific, Bremen, Germany) was used to process QE-HF mass spectra; ImageQuest 1.1.0 (Thermo Scientific, Bremen, Germany) was used for image processing, including image normalization and ion image production. Peaks of protonated ion $[M + H]^+$, corresponding sodiated ion $[M + Na]^+$, and potassium adduct $[M + K]^+$ would be accounted for the same compound. All ions were identified through accurate mass matching from the Human Metabolome Database with less than 10 ppm mass error tolerance.

Results and Discussion

Design and workflow for nsPCR-based chemical derivatization of metabolites.

Upon laser irradiation, the nsPCR reagents, as shown in **Figure 2a**, can selectively react with primary amine groups of metabolites, just like the case of proteins and peptides.[15,16] Similarly, we found that the nsPCR can also improve the MALDI-MS detection of metabolites, especially for those with low molecular weights overlapping with MALDI matrix peaks. On the one hand, the fast and online labeling shifts small metabolites effectively in molecular weights, which allow these metabolites to be separated from overlapping and interfering matrix peaks. On the other hand, the observed apparent ionization efficiency has been greatly enhanced for most of these small molecule metabolites. We thus speculate that the nsPCR for enhanced metabolite analysis by

MALDI-MS benefits from mass shift-associated alleviation of ion suppression effects by MALDI matrix as well as the pH jump induced by nsPCR that basically enriches the available protons during analyte ionization and desorption.[15]

For nsPCR-enhanced MALDI spot analysis, the sample preparation procedure is straightforward and convenient (**Figure 2b**).[16] Unlike most previously reported methods for amine derivatization, we only need to mix metabolites and MALDI matrix with the nsPCR reagent before data acquisition, without incubation or heating for the reaction. Therefore, this strategy enables rapid analysis of metabolites. When the nsPCR strategy is used to enhance MALDI-MS imaging analysis, we only need to apply the nsPCR reagent after and/or during matrix application, followed by conventional data acquisition workflow (**Figure 2c**). Since the derivatization reaction takes place via laser irradiation, samples can be analyzed right after the application of the nsPCR reagent, leaving the localization of metabolites largely undisturbed.

Analytical performance for nsPCR-based chemical derivatization of metabolites.

Here, we evaluated five NBA-derived nsPCR reagents (structure shown in **Figure 2a**), NBA, 2-NBA, 5-NBA, 4-Br-2-NBA, and 4-N-2-NBA, with different functional groups that might affect the reactivity of amine coupling. We chose NBA because it was previously reported for enhancing visualization of peptides through derivatizing the primary amine group of peptides.[15] Also, we speculated that the hydroxy group in 2-NBA and 5-NBA may affect the reaction efficiency between the derivatization reagents and primary amine group. We selected 4-Br-2-NBA because of its relatively high molecular weight (230 Da), which may help avoid matrix interference in the low-molecular weight range. As for 4-N-2-NBA, we speculated that the tertiary amine group could facilitate ionization, improving signal intensities of derivatized molecules. Among the five derivatization reagents, 4-Br-2-NBA and 4-N-2-NBA showed poor performance for the detection

of all tested primary amine-containing molecules. No observable peaks corresponding to derivatized metabolites were shown when 4-N-2-NBA or 4-Br-2-NBA was applied (**Figure S1**).

Instead, NBA, 2-NBA, and 5-NBA can efficiently derivatize some molecules tested in the metabolite standard mixture. Glycine and L-alanine without derivatization were undetectable in the amino acid standard mixture by LTQ MALDI-Orbitrap when the concentration was under 200 μM. Facilitated by derivatization via nsPCR reagents, glycine (**Figure S2**) and alanine (**Figure S3**) were detected with significantly improved signal/noise ratios of up to 68. Also, enhancement of detection for valine (**Figure S4**) and leucine (**Figure 3**) was observed via the nsPCR strategy. In Figure 3, we found that when NBA or 5-NBA was applied, the S/N for derivatized leucine could achieve 414 or more, double of that without nsPCR reagents at the specific concentration. For NBA, the signal intensity of derivatized leucine was 10-fold of underivatized leucine, showing excellent derivatization efficiency (**Figure 3b**). Also, we found that the signal intensity for derivatized leucine was 116-fold of underivatized leucine when 5-NBA was applied, indicating almost complete derivatization (**Figure 3d**). In short, in the case of leucine, the optimal nsPCR reagent was determined to be 5-NBA considering overall labeling efficiency and absolute intensity, while we recommend further optimization using similar workflows when it is applied to other metabolite systems.

As such, we extended the metabolite system using these nsPCR reagents and the derivatization performance evaluation results are shown in **Table S1**. Among these three derivatization reagents, 5-NBA shows the best properties in enhancing the detection of amino acids including glycine, alanine, valine, and leucine (**Table S1**). Moreover, we found that there was a tendency that molecules with lower molecular weight could benefit more from the nsPCR strategy than those with higher molecular weight. The concentration of 5-NBA was then optimized, and

we found that 50 and 100 mM work well for most tested molecules but the optimal concentration of the derivatization reagent for different primary amine containing molecules can be different (representative data for glycine shown in **Figure 4a**). Here, the slope of the calibration curve represented the sensitivity for detecting glycine in the evaluated concentration range. In other words, in the specific concentration range, a higher value of slope indicated a higher sensitivity. The optimal concentration of nsPCR reagents can be used for further experiments. To test the applicability of this derivatization strategy in more complex biological samples, we used 100 mM 5-NBA for calibration curve plotting of glycine in DMEM, and we found good linearity for the quantitation, although the slope of the trendline was significantly smaller than that in the glycine standard (**Figure 4b**).

To determine the improvement of detection of L-valine and DL-leucine, we measured the probability of detection (POD) for these two metabolites (**Figure 5**). We found that after derivatization with 5-NBA, the sensitivity of detecting the L-valine standard was improved by 2.8-fold. For the DL-leucine standard, the sensitivity was improved by 4.5-fold with derivatization.

The POD measures the probability to detect a compound at a specific concentration so that LOD can be calculated based on the fitting curve.[17] Therefore, we prepared a serial dilution of valine and leucine for detection with/without the nsPCR reagent. We measured the POD for these two metabolites based on 8−10 replicates for each concentration (**Figure 5**). Here, the POD50 value represents a specific concentration, where 50% POD can have an S/N equal to or greater than 3. Therefore, POD50 was used here to quantitatively assess the LOD.

The POD50 for valine was calculated to be 9.4 μM without nsPCR reagents, while by adding 5-NBA, the POD50 for derivatized valine decreased to be 3.4 μM (**Figure 5a−b**). For leucine, the POD50 was 3.0 μM without nsPCR reagents, and a decreased POD50 of 0.67 μM was

observed by the application of 5-NBA (**Figure 5c−d**). Thus, after the application of 5-NBA, the sensitivity of nsPCR-enhanced MALDI-MS was improved by 2.8-fold for valine and by 4.5- fold for leucine.

nsPCR-enhanced molecular visualization of metabolite spatial distributions from mouse tissue sections.

We also tested the performance of this derivatization strategy for MS imaging in real biological samples using mouse brain tissue section slides (**Figure 6**). Mass spectra obtained from the same region of three consecutive tissue sections showed that a high abundance of metabolites (including lipids) was observed from the 5-NBA group, indicating the effectiveness of 5-NBA in enhancing the detection of metabolites in tissue specimens. A total of 174 metabolites were detected, including 120 metabolites (69% of the total number) detected in the control group, 118 metabolites (68% of the total number) detected in the 2-NBA group, and 145 metabolites (83% of the total number) detected in the 5-NBA group (**Figure 6a**). Among these metabolites, some showed obvious enhanced signal intensity after nsPCR application (representative ion images shown in **Figure 6c−h**). We also observed that the signal intensity for some metabolites/lipids decreased after the application of nsPCR reagents. We speculated that the application of nsPCR reagents altered the ionization of the metabolites, positively or negatively. Overall, we obtained 25 more metabolite IDs in the 5-NBA group than that of the control group, including six primary amine-containing metabolites (cysteamine, 5-aminopentanoic acid, aminoadipic acid, aminocaproic acid, argininic acid, and norepinephrine). The application of nsPCR not only helped profile metabolites undetectable without derivatization (**Figure 6c−e**) but also improved the visualization of metabolites with low signals (**Figure 6f−h**). Also, we found that the imaging contrast of MALDI-MSI ion maps is still maintained and comparable to that of the control group

after the application of nsPCR reagents (**Figure 6c−h**), suggesting that the use of nsPCR reagents has minimal impacts on both the spatial resolution and dislocation of analytes. Notably, all the putative metabolite IDs were obtained through database searching based on accurate mass matching (<10 ppm). Therefore, metabolite standards or MS2 level data would be beneficial for further validating these metabolite IDs. Taken together, notable enhancement of detection sensitivity for metabolites and lipids was achieved via 5-NBA derivatization, encouraging its broad application to on-tissue metabolite profiling via MALDI-MSI.

To compare the performance of nsPCR-enabled MALDI MSI with previous studies for LMWM analysis, **Table 1** lists some representative examples for laser-based, mostly MALDIMSI analysis of LMWMs with various regimes employing different instrument platforms and matrices. For example, previous studies using 9-AA, or 1,5 DAN as a matrix with the negative ion mode enabled direct MSI of some small molecules such as Glu, Gln, Asp, AMP, and so forth as these molecules are easily ionized under the negative ion mode (**Table 1**).[18−20,23] Meanwhile, other materials such as graphene oxide, ZnO nanoparticles (NPs), and so forth were investigated to achieve good coverage for small-molecule detection.[22,24,25] It was reported that by washing away ionization suppression lipids using organic solvent, signal intensities of small metabolites could also be improved.[21] In a recent study, Chen et al. reported the ion images of glycine and alanine in the mouse brain acquired with ZnO NP-assisted LDIMSI.[24] Studies showed that LAESI worked well for MSI of small metabolites including glycine and alanine.[26,27] These prior studies suggest that each technique has varied but unique preferences in visualization of LMWMs. Compared with previous MALDI-MSI protocols applicable to LMWM visualization in mouse/rat brain tissues, nsPCR-equipped MALDIMSI not only preserves most LMWM information and provides some complementary identifications including many lipid species (**Table 1**) but also offers an

opportunity to quantify and visualize some classes of LMWMs with enhanced sensitivity as demonstrated by identification numbers (**Table 1**) and overall coverage compared to prior methods without nsPCR.

Conclusions

In summary, improved detection of metabolites from the metabolite standard and cell culture medium and more sensitive spatial characterization of metabolites from mouse brain tissue samples were achieved by nsPCR derivatization coupled with MALDI-MS. The ionization efficiencies of glycine, alanine, valine, and leucine were notably improved after nsPCR derivatization, especially for glycine and L-alanine, which were not detected without nsPCR derivatization in MALDI LTQ Orbitrap. Improved sensitivity and metabolite coverage were demonstrated by spatial mapping of metabolite distribution in the mouse brain tissue. Furthermore, the nsPCR strategy enabled on-demand derivatization and detection of metabolites, simplifying the derivatization procedure potentially enabling the analysis of time-sensitive samples since there was no incubation time. This study presents a versatile strategy for profiling primary amine-containing metabolites derived from biological samples (e.g., cell culture medium and cell lysates), as well as in situ mapping the spatial distribution of metabolites directly on tissue sections with high speed and high sensitivity. In this study, all the nsPCR reagents were commercially available and inexpensive. These advancements will facilitate investigating metabolic changes related to various biological processes and to discover potential metabolite biomarkers in disease and drug development. It is possible that other analogues of the nsPCR reagents can potentially improve the sensitivity and coverage for metabolite detection. More in-depth study focusing on the optimization of the structures of nsPCR reagents for improved performance is underway in our laboratory.

Acknowledgements

This work was funded in part by NIH (R01DK071801 and P01CA250972) and the United States Department of Agriculture (grant number 2018-67001-28266). G.L. acknowledges funding support by the Fundamental Research Funds for Central Universities (Nankai University, 020/63213057). L.L. would like to acknowledge NIH grant support RF1AG052324, NCRRS10RR029531, and S10OD025084, a Pancreas Cancer Pilot grant from the University of Wisconsin Carbone Cancer Center (233-AAI9632), and a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

1. Srivastava, A.; Creek, D. J., Discovery and Validation of Clinical Biomarkers of Cancer: A Review Combining Metabolomics and Proteomics. *Proteomics* **2019**, 19 (10), e1700448.

2. Hanahan, D.; Weinberg, R. A., Hallmarks of cancer: the next generation. *Cell* **2011**, 144 (5), 646-74.

3. Mohit Jain, R. N., Sonia Sharma, Nikhil Madhusudhan,Toshimori Kitami, Amanda L. Souza, Ran Kafri, Marc W. Kirschner,; Clary B. Clish, V. K. M., Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. *Science* **2012**, 336, 1040 - 1044.

4. Sousa, C. M.; Biancur, D. E.; Wang, X.; Halbrook, C. J.; Sherman, M. H.; Zhang, L.; Kremer, D.; Hwang, R. F.; Witkiewicz, A. K.; Ying, H.; Asara, J. M.; Evans, R. M.; Cantley, L. C.; Lyssiotis, C. A.; Kimmelman, A. C., Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. *Nature* **2016**, 536 (7617), 479-83.

5. Ng, J.; Papandreou, A.; Heales, S. J.; Kurian, M. A., Monoamine neurotransmitter disorders--clinical advances and future perspectives. *Nat Rev Neurol* **2015**, 11 (10), 567-84.

6. Ren, J.-L.; Zhang, A.-H.; Kong, L.; Wang, X.-J., Advances in mass spectrometry-based metabolomics for investigation of metabolites. *RSC Advances* **2018**, 8 (40), 22335-22350.

7. Letertre, M. P. M.; Dervilly, G.; Giraudeau, P., Combined Nuclear Magnetic Resonance Spectroscopy and Mass Spectrometry Approaches for Metabolomics. *Anal Chem* **2021**, 93 (1), 500-518.

8. Otsuka, Y., Direct Liquid Extraction and Ionization Techniques for Understanding Multimolecular Environments in Biological Systems (Secondary Publication). *Mass Spectrom* (Tokyo) **2021**, 10 (1), A0095.

9. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L., Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* **2018**, 90 (1), 240-265.

10. Gemperline, E.; Rawson, S.; Li, L., Optimization and comparison of multiple MALDI matrix application methods for small molecule mass spectrometric imaging. *Anal Chem* **2014**, 86 (20), 10030-5.

11. Manier, M. L.; Spraggins, J. M.; Reyzer, M. L.; Norris, J. L.; Caprioli, R. M., A derivatization and validation strategy for determining the spatial localization of endogenous amine metabolites in tissues using MALDI imaging mass spectrometry. *J Mass Spectrom* **2014**, 49 (8), 665-73.

12. Santa, T., Derivatization reagents in liquid chromatography/electrospray ionization tandem mass spectrometry. *Biomed Chromatogr* **2011**, 25 (1-2), 1-10.

13. Shariatgorji, M.; Nilsson, A.; Fridjonsdottir, E.; Vallianatou, T.; Källback, P.; Katan, L.; Sävmarker, J.; Mantas, I.; Zhang, X.; Bezard, E.; Svenningsson, P.; Odell, L. R.; Andrén, P. E., Comprehensive mapping of neurotransmitter networks by MALDI-MS imaging. *Nat Methods* **2019**, 16 (10), 1021-1028.

14. Plotka-Wasylka, J. M.; Morrison, C.; Biziuk, M.; Namiesnik, J., Chemical derivatization processes applied to amine determination in samples of different matrix composition. *Chem Rev* **2015**, 115 (11), 4693-718.

15. Li, G.; Ma, F.; Cao, Q.; Zheng, Z.; DeLaney, K.; Liu, R.; Li, L., Nanosecond photochemically promoted click chemistry for enhanced neuropeptide visualization and rapid protein labeling. *Nat Commun* **2019**, 10 (1), 4697.

16. Li, G.; Liu, Y.; Li, L., Nanosecond Photochemical Reaction (nsPCR) for Enhanced Mass Spectrometric Identification, Quantification, and Visualization of Metabolites and Neuropeptides. *Methods Mol Biol* **2022**, 2437, 143-157.

17. Li, A.; Zi, Y.; Guo, H.; Wang, Z. L.; Fernandez, F. M., Triboelectric nanogenerators for sensitive nano-coulomb molecular mass spectrometry. *Nat Nanotechnol* **2017**, 12 (5), 481-487.

18. Benabdellah, F.; Touboul, D.; Brunelle, A.; Laprevote, O., In situ primary metabolites localization on a rat brain section by chemical mass spectrometry imaging. *Anal Chem* **2009**, 81 (13), 5557-60.

19. Miura, D.; Fujimura, Y.; Yamato, M.; Hyodo, F.; Utsumi, H.; Tachibana, H.; Wariishi, H., Ultrahighly sensitive in situ metabolomic imaging for visualizing spatiotemporal metabolic behaviors. *Anal Chem* **2010**, 82 (23), 9789-96.

20. Liu, H.; Chen, R.; Wang, J.; Chen, S.; Xiong, C.; Wang, J.; Hou, J.; He, Q.; Zhang, N.; Nie, Z.; Mao, L., 1,5-Diaminonaphthalene hydrochloride assisted laser desorption/ionization mass spectrometry imaging of small molecules in tissues following focal cerebral ischemia. *Anal Chem* **2014**, 86 (20), 10114-21.

21. Yang, H.; Ji, W.; Guan, M.; Li, S.; Zhang, Y.; Zhao, Z.; Mao, L., Organic washes of tissue sections for comprehensive analysis of small molecule metabolites by MALDI MS imaging of rat brain following status epilepticus. *Metabolomics* **2018**, 14 (4), 50.

22. Zhou, D.; Guo, S.; Zhang, M.; Liu, Y.; Chen, T.; Li, Z., Mass spectrometry imaging of small molecules in biological tissues using graphene oxide as a matrix. *Anal Chim Acta* **2017**, 962, 52-59.

23. Rzagalinski, I.; Hainz, N.; Meier, C.; Tschernig, T.; Volmer, D. A., Spatial and molecular changes of mouse brain metabolism in response to immunomodulatory treatment with teriflunomide as visualized by MALDI-MSI. *Anal Bioanal Chem* **2019**, 411 (2), 353-365.

24. Chen, C.; Laviolette, S. R.; Whitehead, S. N.; Renaud, J. B.; Yeung, K. K., Imaging of Neurotransmitters and Small Molecules in Brain Tissues Using Laser Desorption/Ionization Mass Spectrometry Assisted with Zinc Oxide Nanoparticles. *J Am Soc Mass Spectrom* **2021**, 32 (4), 1065-1079.

25. Yagnik, G. B.; Hansen, R. L.; Korte, A. R.; Reichert, M. D.; Vela, J.; Lee, Y. J., Large Scale Nanoparticle Screening for Small Molecule Analysis in Laser Desorption Ionization Mass Spectrometry. *Anal Chem* **2016**, 88 (18), 8926-30.

26. Nemes, P.; Woods, A. S.; Vertes, A., Simultaneous imaging of small metabolites and lipids in rat brain tissues at atmospheric pressure by laser ablation electrospray ionization mass spectrometry. *Anal Chem* **2010**, 82 (3), 982-8.

27. Wang, X.; Hou, Y.; Hou, Z.; Xiong, W.; Huang, G., Mass Spectrometry Imaging of Brain Cholesterol and Metabolites with Trifluoroacetic Acid-Enhanced Desorption Electrospray Ionization. *Anal Chem* **2019**, 91 (4), 2719-2726.

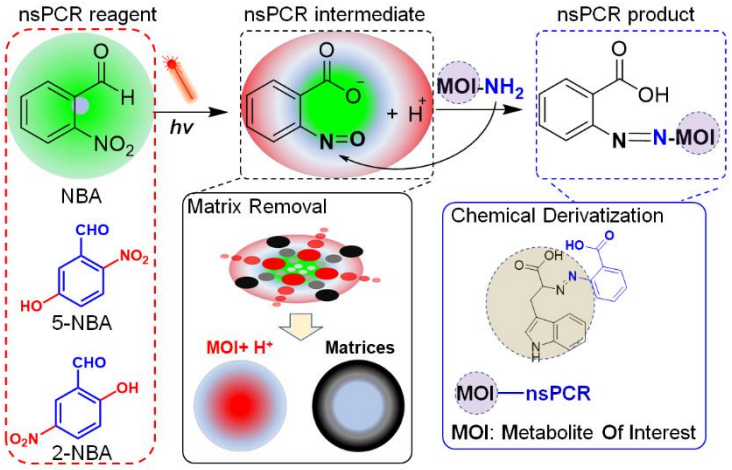


Figure 1. Design and working principle of the nsPCR strategy for on demand matrix removal and chemical derivatization, facilitating the molecular identification and MS visualization of a wide range of biomolecules directly from tissues and other complex biological matrices.

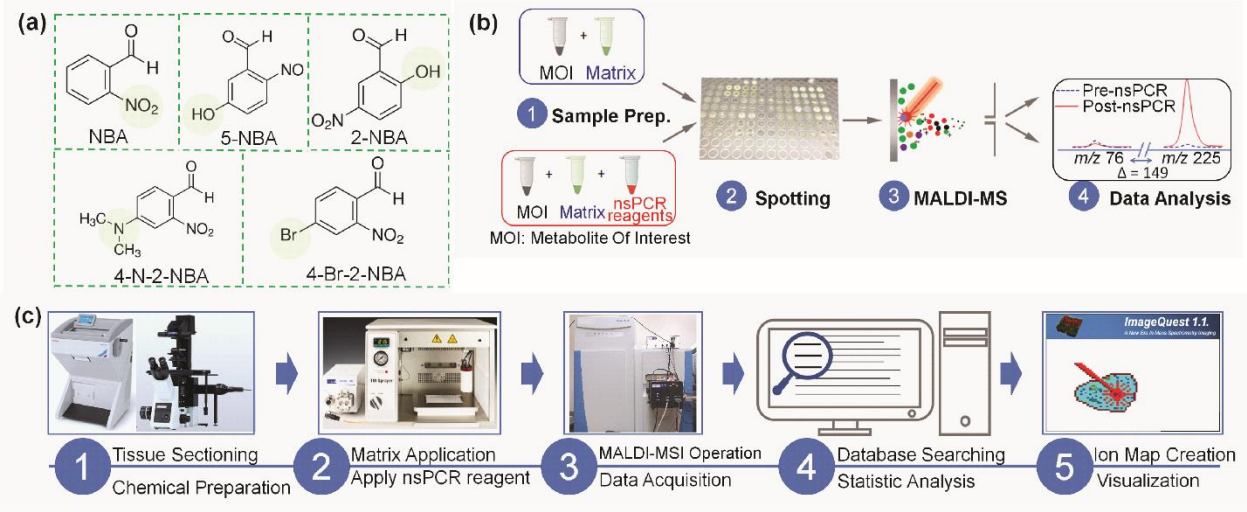


Figure 2. Workflow for nsPCR-enhanced MALDI-MS analysis. (a) Five nsPCR reagents (NBAs) employed in this study. Theoretical accurate mass shifts for NBA, 5-NBA, 2-NBA, 4-N-2-NBA, and 4-Br-2-NBA are 133.0164, 149.0113, 149.0113, 175.0586, and 210.9269 Da per modification, respectively. (b) Generally applicable nsPCR-enhanced MALDI-MS spot analysis workflow. (c) Generally applicable nsPCR-enhanced MALDI-MS imaging workflow. This workflow involves a five-step procedure, step 1: tissue sectioning and chemical preparation; step 2: matrix and nsPCR application to tissue; step 3: MALDI-MS operation and data acquisition; step 4: database searching and statistical analysis; and step 5: ion map creation and visualization.

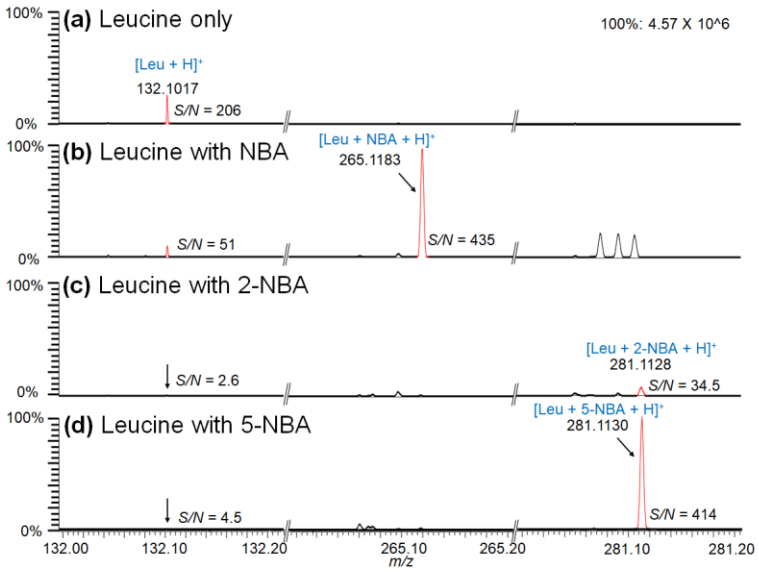


Figure 3. Mass spectra obtained from the leucine standard using the LTQ MALDI Orbitrap MS platform: (a) leucine only, concentration of leucine used here was 50 μM, (b) leucine with NBA, leucine used here was 50 μM and NBA used here was 200 mM, (c) leucine with 2-NBA, leucine used here was 50 μM and 2-NBA used here was 100 mM, and (d) leucine with 5-NBA, leucine used here was 50 μM and 5-NBA used here was 100 mM.

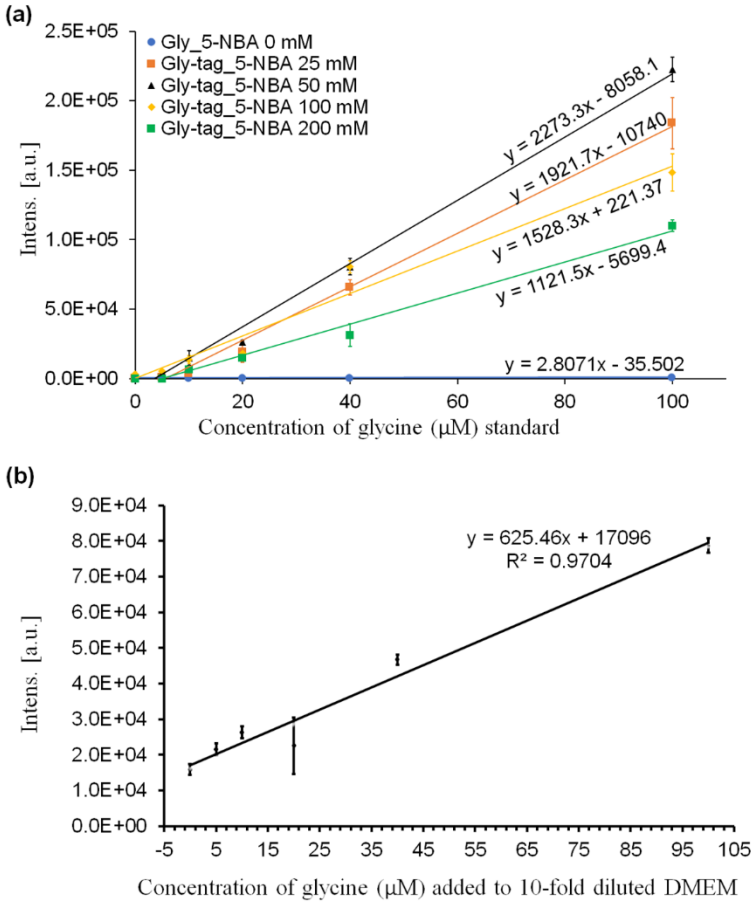


Figure 4. Calibration curve for quantitation of glycine via 5-NBA. (a) Calibration curve for quantitation of glycine in the metabolite standard mixture via 5-NBA with different concentrations. (b) Calibration curve for quantitation of glycine in DMEM via 5-NBA. The medium was 10-fold-diluted, and then, serial dilutions of the glycine standard were added into the medium for quantitation. Concentration of 5-NBA used here was 100 mM.

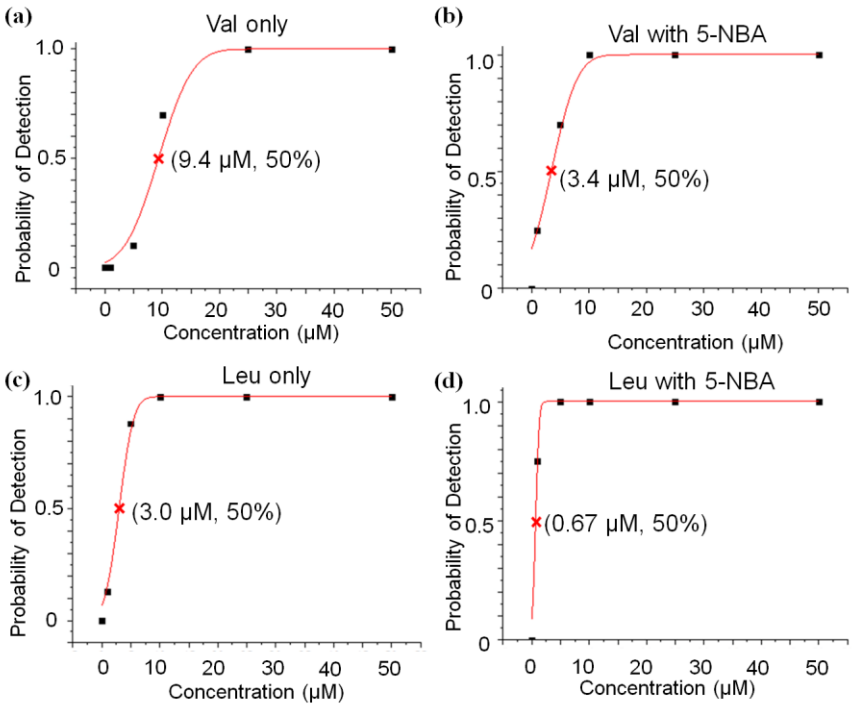


Figure 5. POD for valine and leucine. POD for valine (a) without derivatization and (b) with 5-NBA derivatization. POD for leucine (c) without derivatization and (d) with 5-NBA derivatization. For each concentration of valine or leucine, there were 8−10 replicates. The concentration of 5-NBA used here was 100 mM.

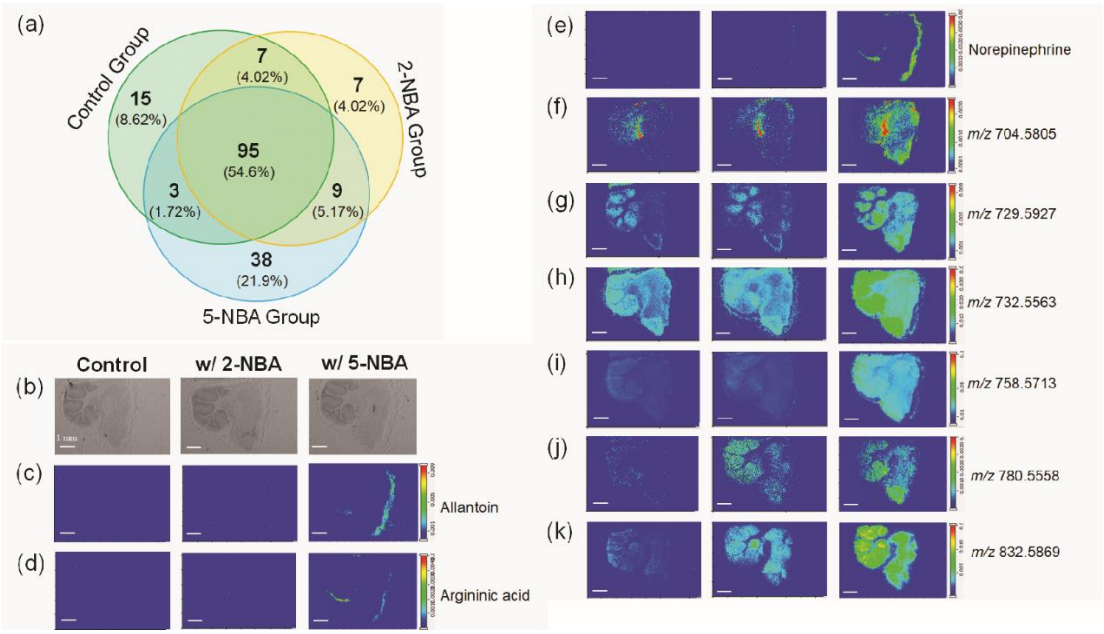


Figure 6. MALDI-MSI ion maps of representative metabolites/lipids in the mouse brain tissue: (a) number of metabolites/lipids identified in the mouse brain tissue, (b) optical images of the mouse brain tissue, and (c–h) (images in the control group were shown in the left, images in the 2- NBA group were shown in the middle, and images in the 5-NBA group were shown in the right) MALDI–MSI maps of the tissue section shown in (c) putative ID: allantoin, $[M + H]^+$, *m/z* 159.0518 (left), $[M + 2\text{-NBA} + NH4]^+$, *m/z* 325.0923 (middle), $[M + 5\text{-NBA} + NH4]^+$, *m/z* 325.0923 (right), (d) putative ID: argininic acid, $[M + H]^+$, *m/z* 176.1035 (left), $[M + 2\text{-NBA} + H]^+$, *m/z* 325.1149 (middle), $[M + 5\text{-NBA} + H]^+$, *m/z* 325.1149 (right), (e) putative ID: norepinephrine, $[M + H]^+$, *m/z* 170.0817 (left), $[M + 2\text{-NBA} + Na]^+$, *m/z* 341.0778 (middle), $[M + 5\text{-NBA} + Na]^+$, *m/z* 341.0778 (right), (f) putative ID: $[DAG + C4H7ON + H]^+$, *m/z* 704.5804, (g) putative ID: $[FAHFA + AMPP]^+$, *m/z* 729.5927, (h) putative ID: $[PC + H]^+$, D16:1-16:0/D14:1-18:0, *m/z* 732.5563, (i) putative ID: $[PC + H]^+$, D16:0-18:2, *m/z* 758.5713, (j) putative ID: $[PC + Na]^+$, D16:0-18:2, *m/z* 780.5558, and (k) putative ID: $[PC + H]^+$, D18:1-22:6/D18:2-22:5, *m/z* 832.5869.

Table 1. Some examples for MALDI-MSI study of low molecular weight molecules comparable to this study.

Tissue type	MS platform	Matrix/NP	Acquisition mode	Metabolites detected	References
Mouse brain	MALDI-Orbitrap MS	CHCA	Positive ion mode	23 LMWM (Arg, AMP, aminoadipic acid, cysteamine, allantoin, etc.) and 122 lipids	This study
Rat brain	MALDI-TOF/TOF MS	9-AA	Negative ion mode	13 primary metabolites (AMP, ADP, ATP, etc.)	Benabdellah, et al., 2009[18]
Mouse brain	MALDI-TOF/TOF-MS	9-AA	Negative-ion mode	More than 30 endogenous metabolites including, Glu, Asp, AMP, ADP, ATP, UMP, GST, etc.	Miura, et al., 2010[19]
Rat brain	MALDI Orbitrap MS	1,5-DAN	Negative ion mode	Asp, Gln, Glu, creatine, adenosine, AMP, GMP, ADP, ATP, etc.	Liu, et al., 2014[20]
Rat brain	MALDI FTICR MS	TiO2	Positive ion mode	34 LMWM (Glu, Gln, GABA, GSH, adenosine, AMP, creatine, etc.)	Yang et al., 2018[21]
Mouse brain	MALDI hybrid Qh-FTICR MS	Graphene oxide	Negative-ion mode	22 of LMWM (inosine, IMP, AMP, GSH etc.) and 190 of lipids	Zhou, et al., 2017[22]
Mouse brain	MALDI FTICR MS	9-AA	Negative ion mode	Asp, Glu, Gln, Tau, GSH, cAMP, AMP, IMP, GMP, ADP, GDP, CMP etc.	Rzagalinski, et al., 2019[23]
Mouse/rat brain	MALDI TOF/TOF MS	ZnO NP	Positive ion mode	Gly, Ala, Gln, Glu, Arg, etc.	Chen, et al., 2021[24]

Supplemental Information

Table S1. Derivatization efficiency of metabolites with different nsPCR reagents and performance of nsPCR in improving signal intensities for different metabolites. Table S1 shows the derivatization efficiency for metabolites at 50 µM and nsPCR reagents at 100 mM. The trendlines were plotted using 3 different concentrations (5 µM, 50 µM and 200 µM) of each metabolite. k1 represents the slope of trendline generated by metabolites without nsPCR reagents, and k2 represents the slope of trendline of derivatized metabolites.

Metabolites (Monoisotopic mass)	Derivatization efficiency (c = 50 µM)			Trendline (slope - k and R^2)					
	NBA	2-NBA	5-NBA	NBA		2-NBA		5-NBA	
Glycine 75.0320	100.00%	100.00%	100.00%	Not applicable k2 =910.96, R^2=0.993	k2/k1 = NA	Not applicable k2 =4207.1, R^2 =0.9008	k2/k1 = NA	Not applicable k2 =32428, R^2 =0.9987	k2/k1 = NA
L-Alanine 89.0477	0.00%	99.62%	99.86%	Not applicable k2 =290.11, R^2=0.9662	k2/k1 = NA	Not applicable k2 =1367.7, R^2=0.9929	k2/k1 = NA	Not applicable k2 =5367.2, R^2=0.9914	k2/k1 = NA
Valine 117.0790	32.26%	83.69%	98.68%	k1 =6289.1, R^2=0.9917 k2 =1195.3, R^2=0.9236	k2/k1 = 0.190	k1 =6289.1, R^2=0.9917 k2 =2559.8, R^2=0.9898	k2/k1 = 0.407	k1 =6289.1, R^2=0.9917 k2 =12036, R^2=0.9586	k2/k1 = 1.91
DL-Leucine 131.0946	20.62%	90.55%	99.18%	k1 =19229, R^2=0.9966 k2 =1331.8, R^2=0.9349	k2/k1 = 0.0692	k1 =19229, R^2=0.9966 k2 =5440.7, R^2=0.9813	k2/k1 = 0.283	k1 =19229, R^2=0.9966 k2 =65441, R^2=0.9771	k2/k1 = 3.40
L-Glutamic acid 146.1055	1.02%	2.58%	28.27%	k1 =13948, R^2=0.9938 k2 =6.0705, R^2=0.6977	k2/k1 = 0.000435	k1 =13948, R^2=0.9938 k2 =401.09, R^2=0.9702	k2/k1 = 0.0288	k1 =13948, R^2=0.9938 k2 =1279.1, R^2=0.9983	k2/k1 = 0.0917
Dopamine 153.0790	68.53%	54.31%	80.42%	k1 =10716, R^2=0.9806 k2 =1733.2, R^2=0.4833	k2/k1 = 0.162	k1 =10716, R^2 =0.9806 k2=1932.3, R^2=0.998	k2/k1 = 0.180	k1 =10716, R^2 =0.9806 k2 =9681.9, R^2 =0.9916	k2/k1 = 0.903
L-Histidine 155.0695	0.00%	0.00%	18.84%	k1 =355113, R^2 =0.9902 Not applicable	k2/k1 = 0	k1 =355113, R^2=0.9902 Not applicable	k2/k1 = 0	k1 =355113, R^2=0.9902 k2 =21477, R^2=0.9923	k2/k1 = 0.0605
L-Phenylalanine 165.0790	10.23%	61.92%	61.99%	k1 =23971, R^2=0.9979 k2 =2426.7, R^2=0.9674	k2/k1 = 0.101	k1 =23971, R^2=0.9979 k2 =4559.6, R^2=0.9999	k2/k1 = 0.190	k1 =23971, R^2=0.9979 k2 =6061.2, R^2=0.9961	k2/k1 = 0.253
L-Arginine 174.1117	0.12%	1.01%	10.03%	k1 =233438, R^2 =0.9907 k2 =589.1, R^2=0.9791	k2/k1 = 0.00252	k1 =233438, R^2=0.9907 k2 =1883.4, R^2=0.6731	k2/k1 = 0.00807	k1 =233438, R^2=0.9907 k2 =61422, R^2=0.9971	k2/k1 = 0.263
L-Tryptophan 204.0899	2.92%	4.73%	3.07%	k1 =8009.6, R^2=0.9976 k2 =456.78, R^2=0.9478	k2/k1 = 0.0570	k1 =8009.6, R^2=0.9976 k2 =166.32, R^2=0.9864	k2/k1 = 0.0208	k1 =8009.6, R^2=0.9976 k2 =366.71, R^2=0.9942	k2/k1 = 0.0458

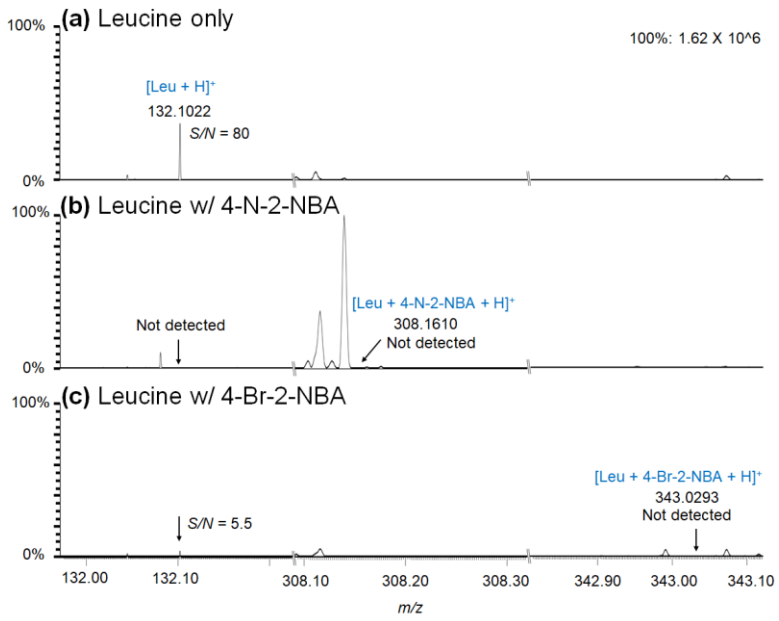


Figure S1. Representative mass spectra obtained from metabolite mixture using LTQ MALDI Orbitrap MS platform: (a) leucine only, concentration of leucine used here was 50 µM, (b) leucine with 4-N-2-NBA, leucine used here was 50 µM and 4-N-2-NBA used here was 100 mM, (c) leucine with 4-Br-2-NBA, leucine used here was 50 µM and 4-Br-2-NBA used here was 100 mM.

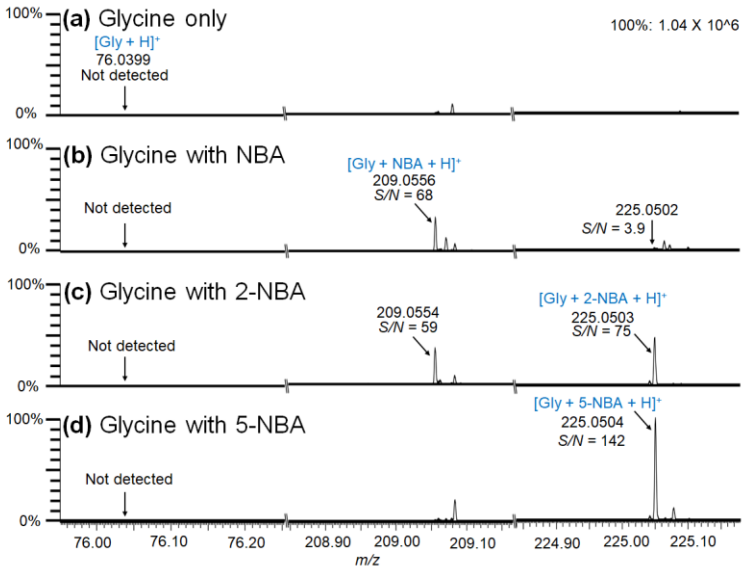


Figure S2. Mass spectra obtained from glycine standard using LTQ MALDI Orbitrap MS platform: (a) glycine only, concentration of glycine used here was 50 µM, (b) glycine with NBA, glycine used here was 50 µM and NBA used here was 200 mM, (c) glycine with 2-NBA, glycine used here was 50 µM and 2-NBA used here was 100 mM, (d) glycine with 5-NBA, glycine used here was 50 µM and 5-NBA used here was 100 mM.

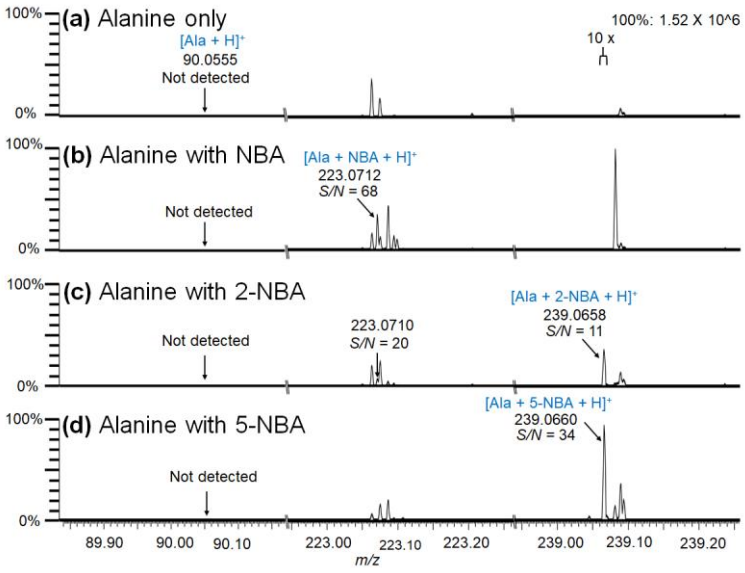


Figure S3. Mass spectra obtained from alanine standard using LTQ MALDI Orbitrap MS platform: (a) alanine only, concentration of alanine used here was 50 µM, (b) alanine with NBA, glycine used here was 50 µM and NBA used here was 200 mM, (c) alanine with 2-NBA, alanine used here was 50 µM and 2-NBA used here was 100 mM, (d) alanine with 5-NBA, alanine used here was 50 µM and 5-NBA used here was 100 mM.

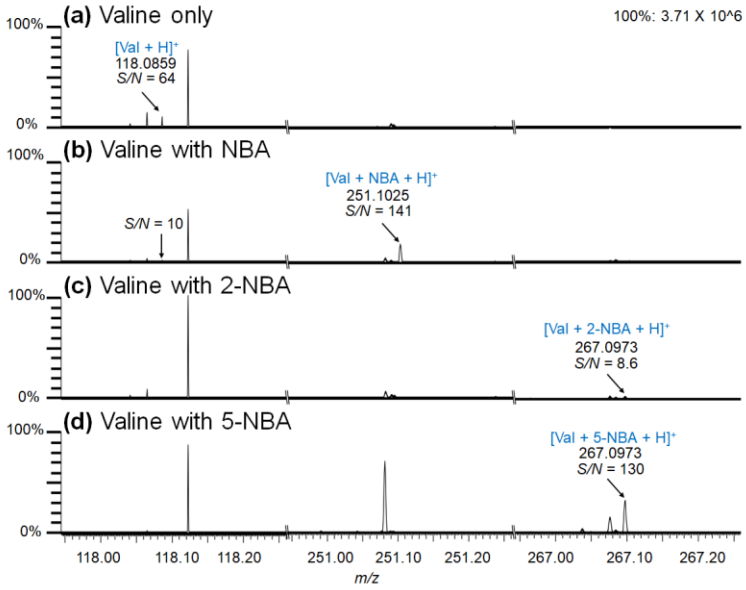


Figure S4. Mass spectra obtained from valine standard using LTQ MALDI Orbitrap MS platform: (a) valine only, concentration of valine used here was 50 µM, (b) valine with NBA, valine used here was 50 µM and NBA used here was 200 mM, (c) valine with 2-NBA, valine used here was 50 µM and 2-NBA used here was 100 mM, (d) valine with 5-NBA, valine used here was 50 µM and 5-NBA used here was 100 mM.

Chapter 3

A Simplified Sample Preparation Strategy for High-Throughput Mass Spectrometry Imaging of Multicellular Tumor Spheroids

Adapted from:
Yuan Liu, Hua Zhang, Jillian Johnson, Penghsuan Huang, Lingjun Li, "A Simplified Sample Preparation Strategy for High-Throughput Mass Spectrometry Imaging of Multicellular Tumor Spheroids", in preparation for submission to *Analytical and Bioanalytical Chemistry*.

Author Contributions:
L.L. designed, and supervised the whole study; Y.L. performed experiments, analyzed, interpreted the data, and wrote the manuscript; H.Z performed the experiments, analyzed, interpreted the data; Penghsuan Huang analyzed, interpreted the data; All authors provided feedback.

Abstract

Multicellular tumor spheroids (MCTS) play an important role in biological studies and cancer research. An increasing number of studies have applied mass spectrometry imaging (MSI) technology to investigate MCTS. Current methods for sample preparation of spheroids usually suffer from low throughput or increased sample preparation time for MSI. Herein, an optimized sectioning method was developed to prepare MCTS frozen sections to overcome these challenges. MCTS grown on agarose-based microarrays were coated with gelatin for sectioning multiple spheroids simultaneously. The developed method was used for investigating lipid distribution in spheroids by matrix-assisted laser desorption/ionization MSI. Compared with the ice-assisted sectioning method, our method enables high-throughput sectioning of spheroids (81 spheroids or even more). Also, sectioning spheroids together with agarose-based microarray improves the visualization of spheroids and makes the staining of spheroids amenable for locating the position of spheroids, reducing the sample preparation steps and time. Therefore, this method provides a high-throughput and convenient strategy for preparing spheroids sections for MSI studies.

Introduction

Two-dimensional (2D) cell culture is widely used in cancer research and drug development. However, tissues and organs are three dimensional (3D) *in vivo*. Compared to 2D cell culture, 3D cell culture models are usually more physiological in cell morphology and signaling and show better drug metabolism[1]. Multicellular tumor spheroid (MCTS) represents a form of 3D cell culture that gains emerging attention in cancer research and drug evaluation.[2] To study MCTS, matrix-assisted laser desorption/absorption ionization (MALDI) mass spectrometry (MS) is a powerful tool for analyzing biomolecules.[2-4] Furthermore, MALDI MS imaging (MSI) can be used to

visualize the spatial distribution of biomolecules within spheroids.[5-7] This technique enables rapid analysis of drug penetrations and profile changes of some interesting molecules within spheroids.[8-12]

In the study of MCTS using MALDI MSI, sample preparation is a critical step. Currently, gelatin is widely used for embedding during spheroid sample preparation for sectioning.[13, 14] However, because of the limited transparency and the similarity of the color between frozen gelatin solution and spheroids, it is difficult to locate the spheroids sections in gelatin-assisted cryosectioning, especially if the size of spheroids is small. In this case, hematoxylin and eosin staining (H&E staining) technique is used to facilitate the visualization of spheroids sections. To address this issue, a method using ice-coated spheroids was developed to assist sectioning spheroids.[6] In this method, spheroids can be visualized in contrast to ice, avoiding H&E staining during cryosectioning. However, a noted limitation in this work was that the spheroids were sectioned one by one. Another study by Johnson et al. reported on a strategy for high-throughput organoid sectioning using gelatin-based microarray.[5] Although dozens of spheroids could be sectioned and analyzed via MSI, the sample preparation involved transferring samples to gelatin-based microarray, centrifugation, and H&E staining. Additionally, a platform using paper-based cultures was developed for high throughput drug screening.[15] This strategy greatly improved the throughput for MSI of spheroids. However, information regarding the spatial distribution of some molecules including drug distribution may be lost because of the MSI data was acquired on the surface of spheroids as no sectioning was performed.[11] In addition, the equipment for generating wax-patterned paper scaffolds may not be accessible to most biological labs.

Herein, we developed a strategy for simplified sample preparation of spheroid sections for high-throughput MSI analysis. In this strategy, we grow spheroids in agarose-based 3D petri dishes

(microarrays). Spheroids can be embedded in concentration-optimized gelatin solution for cryosectioning directly. The workflow enables MALDI MSI of tens or even hundreds of spheroids on a single glass slide. Also, the workflow is simple without transferring spheroids during sample preparation. Furthermore, all of the materials and equipment for this workflow are easily accessible to most research labs, facilitating rapid implementation of this approach.

Materials and Methods

Chemicals and materials.

Chemicals including acetonitrile (ACN), formic acid (FA), methanol, ethanol, trifluoroacetic acid (TFA), and HPLC grade water were purchased from Fisher Scientific (Pittsburgh, PA, USA). Indium tin oxide (ITO)- coated glass slides (25 mm × 75 mm × 1 mm) were purchased from Delta Technologies (Loveland, CO, USA).

Cell culture and MCTS growth.

The commercially available pancreatic cancer cell line PANC-1 was obtained from American Type Culture Collection (ATCC, Manassas, VA, USA) and was maintained in DMEM (11965, Gibco, Life Technologies, USA) containing 10% fetal bovine serum (FBS) (Gibco, Origin: Mexico) and 1% penicillin-streptomycin solution (Gibco, Life Technologies Corporation, Grand Island, NY, USA). Cells were cultured in a 37°C moisture incubator filled with 5% CO_2. Cells suspensions were prepared by trypsinization at 70%-90% confluence using 0.25% trypsin EDTA solution (Gibco, Life Technologies Corporation, Grand Island, NY, USA). The cell suspension was centrifuged at 300 g for 5 minutes and the medium was discarded. MCTS were prepared by adding PANC-1 cell suspension to 1% agarose-based 3D petri dishes generated using MicroTissues 3D Petri Dish micro-mold (Sigma-Aldrich, St. Louis, MO, USA). Medium was exchanged by half every 2 days.

Preparation of frozen sections of MCTS.

Gelatin solution (50 mg/mL or 5%) in a 50 mL tube was prepared in HPLC grade water, vortexed vigorously, and placed in a water bath at ~60 °C. Gelatin solution was cooled down (less than 37 °C) before use. Medium was removed and MCTS together with 3D petri dishes were washed with PBS 3 times. MCTS carried by 3D petri dish were carefully placed onto plastic disposable base molds and gelatin solution were poured into the molds to cover the 3D petri dish. Then spheroids covered by gelatin solution were placed in a – 20 °C freezer immediately. MCTS and the mold were stored at -20 °C for at least 2 hours or stored at -80 °C for future use. Carboxymethylcellulose (CMC, EMD Millipore Corp., Burlington, MA, USA) were prepared as the procedures described above for 2% solution. Then 2% CMC solution was poured into base molds and was frozen in dry ice.

The procedures of the sectioning spheroids together with agarose-based petri dishes were provided below. The 2% CMC block on a cryostat support was sliced and marked with an oil pen to prepare a surface paralleled with the moving path of the cutting blade (**Figure 1A**). Spheroids on microarrays were placed on base mold and covered by 5% gelatin solution and then frozen at – 20 °C (**Figure 1B**). The frozen sample was taken out of the mold and mounted on CMC block by adding water surrounding the sample and the tissue holder was placed back to the cryostat and secured with the oil pen markers aligned (**Figure 1C**). Then spheroids were cut into slices and transferred to regular glasses or ITO-coated glasses (**Figure 1D**). The sections on the slides were quickly dried in a vacuum prior to the matrix deposition. The sections can also be stored at -80 °C for further use. Matrix was sprayed on spheroids using TM-sprayer (**Figure 1E**).

MALDI-MSI analysis.

Application of matrices were performed based on previous studies with some modifications.[16, 17] Thereafter, DHB (40 mg/mL in 50% MeOH and 0.1% TFA) was deposited via the TM sprayer system using the following conditions: a nozzle temperature of 80 °C, a gas pressure of 10 psi, sixteen passes, a moving velocity of 1250 mm/min, a drying time of 30 s, and a flow rate of 0.1 mL/min.

MALDI-MS imaging experiments were performed on a timsTOF flex mass spectrometer (Bruker Scientific, LLC, Bremen, Germany) coupled with a SmartBeam 3D 10 kHz frequency tripled Nd:YAG laser (355 nm). The laser settings used were 50 μm of diameter circular spot size, with 200 shots per pixel and a raster step size of 50 μm for MCTS section imaging. The laser power was set to 70%. The MS imaging data were collected over a mass range of *m/z* 400-1300 for positive ion detection mode. MSI images were analyzed using SCiLS Lab Pro (Bruker Scientific, LLC, Bremen, Germany) with data normalized to total ion count (TIC). After acquisition of MSI data, H&E staining was performed on MCTS sections.

Lipid extraction and LC-MS/MS analysis.

For lipidomics analysis, PANC1 cellular lipid extracts were subjected to Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer equipped with a Dionex UltiMate 3000 UPLC system (Thermo Scientific, San Jose, CA, USA). The extraction of cellular lipids was based on a modified Folch protocol.[18] Briefly, cell suspensions were centrifuged at 300 ×g for 2 min, and cell pellets were washed with cold PBS three times. For lipids extraction, 1.5 mL of water/methanol/chloroform mixture (1:1:1) was added into the tube with cell pellets. The mixture was vigorously vortexed for 2 min and then sonicated for 15 min in a water bath followed by centrifugation (12000 ×g) at 4 °C for 10 min. After the centrifugation, the chloroform layer was collected to a glass tube. The extraction was repeated once, and the chloroform layers were

combined and dried under a nitrogen stream. All the extracted lipids were stored at −80 °C before analysis.

The HPLC-nanoESI-MS analysis was carried out on Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer equipped with a Dionex UltiMate 3000 UPLC system (Thermo Scientific, San Jose, CA, USA). LC was performed using an in-house fabricated microcapillary column with nanoESI emitter tip (ID of 75 μm). The column was 14 cm packed with Bridged Ethylene Hybrid C18 particles (1.7 μm, 130 Å, Waters). Mobile phase A is composed of ACN/water (40/60, v/v) with 1.0 mM ammonium acetate and mobile phase B is IPA/ACN (90/10, v/v) with 1.0 mM ammonium acetate. A 115-min gradient was used for separation at a flow rate of 0.2 μL min^{-1}: 0-4 min, 5% B; 4-5 min, 5-30% B; 5-10 min, 30-35% B; 10-15 min, 35-51% B; 15-40 min, 51-61% B; 40-50 min, 61-70% B; 50-60 min, 70-99% B; 60-75 min, 99% B; 75-76 min, 99-60% B; 76-79 min, 60% B; 79-80 min, 60-99% B; 80-85 min, 99% B; 85-90 min, 99-5% B; 90-115 min, 5% B. The lipid extracts were dissolved in IPA/ACN/water solution (v/v/v, 45/25/30) for HPLC-MS/MS analysis. Mass spectra were collected over a mass range of *m/z* 200–1200 in both positive and negative ion detection modes. The ionization voltage was ± 2.0 kV and the heated ion transportation capillary was set as 320 °C. Resolving power of MS1 scan was set as 60 k (at *m/z* 200) and AGC target was set as 1×10^5 and maximum injection time was set as 100 ms. For MS2, the mass range was set as *m/z* 100–1200, precursor ions isolation window width was set as 0.8 Da, and normalized collision energy (NCE) was set to 25%. MSDIAL (Version 4.80) and LIPID MAPS database searching was performed for lipids data analysis.

Results and Discussion

Optimization of sectioning method for MCTS.

To section all of spheroids grown on agarose-based microarrays, we optimized the embedding methods. We tried different potential embedding materials including water, CMC, and gelatin.[6, 14, 19] Firstly, we tried ice, 2% CMC and 10% gelatin frozen blocks. We found that 2% CMC and 10% gelatin were sectionable but we were not able to get intact ice sections (**Figure 2A**). The ice tended to break, and it would be difficult to transfer all the spheroids to glass slides without dislocating these spheroids. Even though previous work indicates ice is good in assisting sectioning single spheroids,[6] we excluded ice in this study. Then we picked up CMC and gelatin for further optimization. We tried 1% CMC, 2% CMC, CMC, and gelatin mixture (1% each), 2% gelatin, 5% gelatin and 10% gelatin for embedding and sectioning. We found that it was possible to get intact spheroid sections using 1% CMC or 2% CMC but it was often the case that the sections deformed more or less. The mixture of CMC and gelatin, and 2% gelatin worked better than 1% CMC or 2% CMC but in terms of avoiding deformation, while 5% gelatin and 10% gelatin worked the best for assisting sectioning. The results were shown in **Figure 2**. Gelatin solution of 10% was easy to fold after sectioning. Therefore, we picked up 5% gelatin for further experiments.

For tissue samples or spheroids embedding, we noticed that there were two slightly different ways in sample preparation. One was that fresh samples are frozen and then embedded in embedding solution and frozen again.[5] The other way was that fresh samples were immersed directly in embedding materials and then frozen sequentially.[6] We evaluated the differences between these two methods in preparing intact spheroids sections. We found that it tended to deform when we froze the spheroids before embedding (**Figure S1**). This result suggests that embedding spheroids before freezing works better for sectioning.

Preparation of MCTS sections with different sizes

We began by slicing big spheroids. We grew PNAC1 spheroids on agarose-based microarray (Microtissues, #12-81) and the spheroids were sliced when they were with a diameter about 600~800 μm (**Figure 3A**). The sliced spheroids were mounted on glass slides/ITO-coated glass slides. It could be difficult to locate individual spheroids embedded in gelatin because the color of frozen gelatin (especially when the concentration of gelatin solution used for embedding is high) was very similar to that of spheroids and it was not transparent. In contrast, in this case, spheroids grew on top of agarose and agarose was almost transparent and could be distinguished from spheroids easily even before sectioning. Meanwhile, spheroids could be easily located after sectioning, especially after mounting. The location of spheroids could be further confirmed by hematoxylin staining. The hematoxylin staining demonstrated that spheroids sections obtained by our method were in good shape (**Figure 3B**). We further applied this strategy to sectioning smaller spheroids (300 - 400 μm) and found that it also worked well as shown in microscope images for H&E staining (**Figure 3B**) and scanned images for all the spheroids (**Figure 3C**).

High throughput MALDI MS imaging of metabolites in MCTS.

We performed MALDI-MSI at 50 μm spatial resolution for MCTS sections coated with DHB matrix in the positive ionization mode. The results showed that we could observe high intensities of ion signals in the MCTS sections (**Figure 4A**). We also found that most spheroids could be imaged in one MSI run, showing the high throughput nature of this method for visualizing the distribution of metabolites on spheroids (**Figure 4B**). We also tried smaller spheroids for MSI, and it generated similar results (**Figure S2**).

Sample preparation for sectioning a large number of spheroids/organoids together requires multiple steps of involved procedures. As shown in a previous study, organoids grown in Matrigel needed to be recovered and then transferred to sectionable microwells.[5] These steps involved

centrifugation and transferring of spheroids using pipettes, which might affect the structure of spheroids. Additionally, it was difficult to fill all the wells in microwells with spheroids. Therefore, the throughput could be limited. In this study, we grew spheroids on agarose-based microwells directly, and spheroids could be embedded in embedding materials without centrifugation or transferring of spheroids from somewhere else to microwells, kept the spheroids undisturbed while largely reducing the sample preparation procedures. Furthermore, to section spheroids all together, we cut CMC block to generate a surface that is in parallel with the path of the blade for sectioning. This step would provide a surface that allowed all the spheroids to be sectioned simultaneously.

It is worth mentioning that we did not claim that this agarose-based microwell strategy was designed to replace the previously established gelatin-based microwell strategy for high-throughput sample preparation or ice-assisted sectioning of MCTS. In some cases, the spheroids/organoids are required to grow in Matrigel or other specific materials, therefore, transferring of spheroids/organoids from these materials to microwells would be inevitable. Additionally, compared with ice-assisted sectioning of spheroids, the use of gelatin could bring some interference signals to some extent. We therefore consider that the strategy presented in this chapter provides an attractive alternative for high-throughput MSI of spheroids/organoids, but the selection of which method would be subjected to the actual samples and aims of the experiments.

Conclusions

In summary, we developed a sectioning method using agarose-based microwells as spheroid culture platform and gelatin as embedding material to prepare MCTS that enabled high-throughput MALDI MSI analysis. Compared to the previously reported ice-assisted sectioning, our method enabled sectioning of large quantities of spheroids simultaneously and maintained excellent visibility for small MCTS during cryosectioning, which could greatly increase the

throughput for sample preparation. Compared to the previously reported sample preparation strategy for high-throughput MSI of primary tumor organoids, our method reduces the sample preparation steps and offers a viable alternative for sample preparation for high-throughput MSI. This method can also be utilized to prepare frozen tissue sections for other biological specimens such as organoids and embryos.

Acknowledgements

This work was funded in part by NIH (R01DK071801 and P01CA250972) and the United States Department of Agriculture (grant number 2018-67001-28266). L.L. would like to acknowledge NIH grant support RF1AG052324, NCRRS10RR029531, and S10OD025084, a Pancreas Cancer Pilot grant from the University of Wisconsin Carbone Cancer Center (233-AAI9632), and a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

1. Yamada, K. M.; Cukierman, E., Modeling tissue morphogenesis and cancer in 3D. *Cell* **2007,** *130* (4), 601-10.

2. Wang, Y.; Hummon, A. B., MS imaging of multicellular tumor spheroids and organoids as an emerging tool for personalized medicine and drug discovery. *J Biol Chem* **2021,** *297* (4), 101139.

3. Norris, J. L.; Caprioli, R. M., Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chem Rev* **2013,** *113* (4), 2309-42.

4. Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L., Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal Chem* **2018,** *90* (1), 240-265.

5. Johnson, J.; Sharick, J. T.; Skala, M. C.; Li, L., Sample preparation strategies for high-throughput mass spectrometry imaging of primary tumor organoids. *J Mass Spectrom* **2020,** *55* (4), e4452.

6. Xie, P.; Zhao, C.; Liang, X.; Huang, W.; Chen, Y.; Cai, Z., Preparation of Frozen Sections of Multicellular Tumor Spheroids Coated with Ice for Mass Spectrometry Imaging. *Anal Chem* **2020,** *92* (11), 7413-7418.

7. Tucker, L. H.; Hamm, G. R.; Sargeant, R. J. E.; Goodwin, R. J. A.; Mackay, C. L.; Campbell, C. J.; Clarke, D. J., Untargeted Metabolite Mapping in 3D Cell Culture Models Using High Spectral Resolution FT-ICR Mass Spectrometry Imaging. *Anal Chem* **2019,** *91* (15), 9522-9529.

8. Liu, X.; Flinders, C.; Mumenthaler, S. M.; Hummon, A. B., MALDI Mass Spectrometry Imaging for Evaluation of Therapeutics in Colorectal Tumor Organoids. *J Am Soc Mass Spectrom* **2018,** *29* (3), 516-526.

9. Tian, X.; Zhang, G.; Zou, Z.; Yang, Z., Anticancer Drug Affects Metabolomic Profiles in Multicellular Spheroids: Studies Using Mass Spectrometry Imaging Combined with Machine Learning. *Anal Chem* **2019,** *91* (9), 5802-5809.

10. Cox, M. C.; Reese, L. M.; Bickford, L. R.; Verbridge, S. S., Toward the Broad Adoption of 3D Tumor Models in the Cancer Drug Pipeline. *ACS Biomater Sci Eng* **2015,** *1* (10), 877-894.

11. Feist, P. E.; Sidoli, S.; Liu, X.; Schroll, M. M.; Rahmy, S.; Fujiwara, R.; Garcia, B. A.; Hummon, A. B., Multicellular Tumor Spheroids Combined with Mass Spectrometric Histone Analysis To Evaluate Epigenetic Drugs. *Anal Chem* **2017,** *89* (5), 2773-2781.

12. Zhao, C.; Xie, P.; Yong, T.; Wang, H.; Chung, A. C. K.; Cai, Z., MALDI-MS Imaging Reveals Asymmetric Spatial Distribution of Lipid Metabolites from Bisphenol S-Induced Nephrotoxicity. *Anal Chem* **2018,** *90* (5), 3196-3204.

13. Li, H.; Hummon, A. B., Imaging mass spectrometry of three-dimensional cell culture systems. *Anal Chem* **2011,** *83* (22), 8794-801.

14. Liu, X.; Weaver, E. M.; Hummon, A. B., Evaluation of therapeutics in three-dimensional cell culture systems by MALDI imaging mass spectrometry. *Anal Chem* **2013,** *85* (13), 6295-302.

15. Tobias, F.; McIntosh, J. C.; LaBonia, G. J.; Boyce, M. W.; Lockett, M. R.; Hummon, A. B., Developing a Drug Screening Platform: MALDI-Mass Spectrometry Imaging of Paper-Based Cultures. *Anal Chem* **2019,** *91* (24), 15370-15376.

16. Liu, Y.; Li, G.; Gu, T. J.; Li, L., Nanosecond Photochemical Reaction for Enhanced Identification, Quantification, and Visualization of Primary Amine-Containing Metabolites by MALDI-Mass Spectrometry. *Anal Chem* **2022,** *94* (9), 3774-3781.

17. Gemperline, E.; Rawson, S.; Li, L., Optimization and comparison of multiple MALDI matrix application methods for small molecule mass spectrometric imaging. *Anal Chem* **2014,** *86* (20), 10030-5.

18. Folch, J.; Lees, M.; Sloane Stanley, G. H., A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* **1957,** *226* (1), 497-509.

19. Stutts, W. L.; Knuth, M. M.; Ekelof, M.; Mahapatra, D.; Kullman, S. W.; Muddiman, D. C., Methods for Cryosectioning and Mass Spectrometry Imaging of Whole-Body Zebrafish. *J Am Soc Mass Spectrom* **2020,** *31* (4), 768-772.

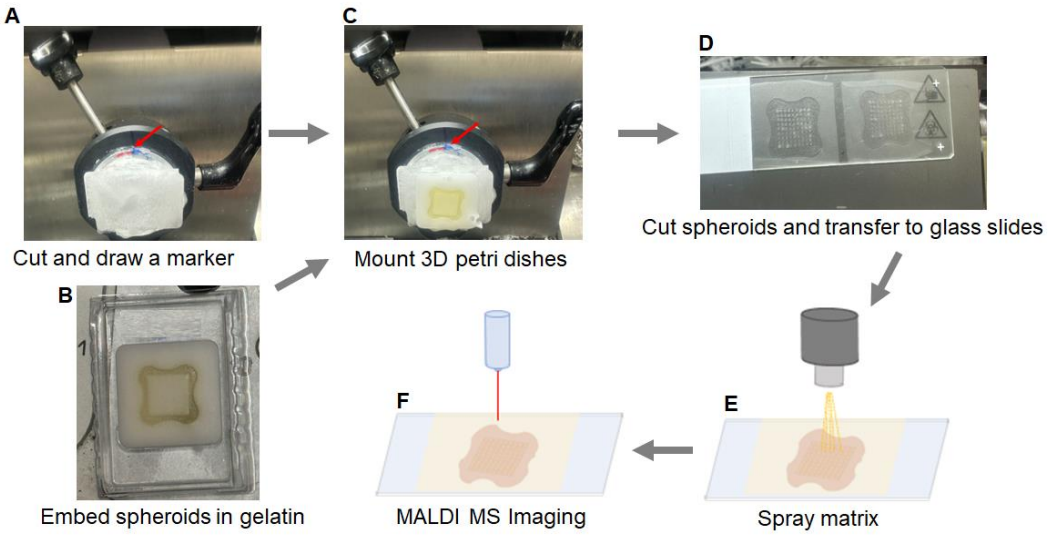


Figure 1. Workflow of the developed method for high-throughput MSI analysis of MCTS: (A) CMC block was cut, and position marked; (B) Spheroids with agarose-based dish embedded in gelatin solution; (C) Spheroids were transferred on CMC block; (D) Spheroids were sectioned and transferred to glass slides; (E) Matrix application; (F) MALDI MSI analysis.

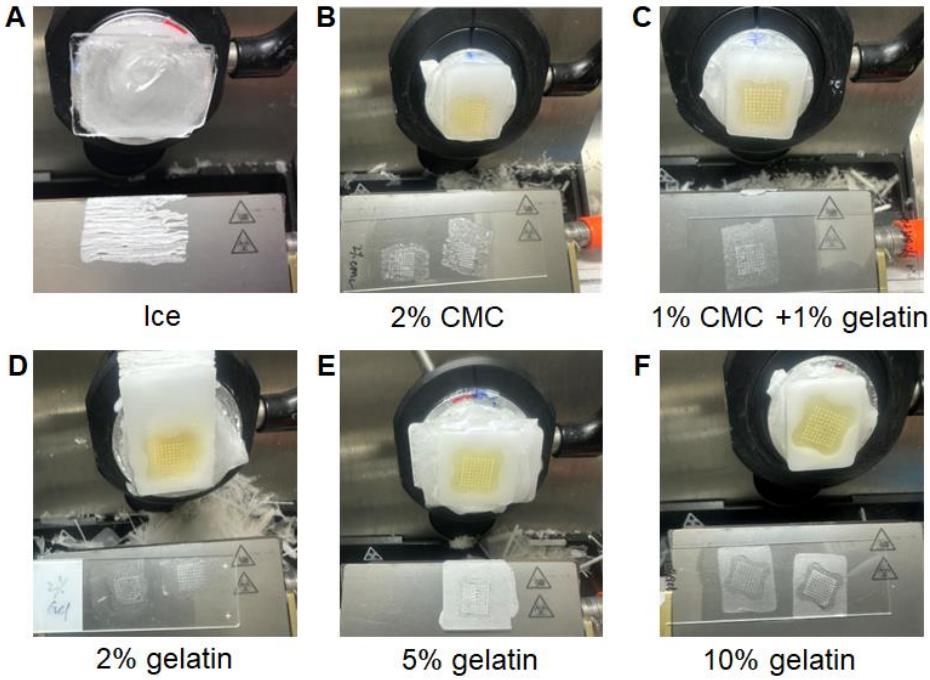


Figure 2. Evaluation of different embedding materials for sectioning. (A) Ice; (B) 2% CMC; (C) 1%CMC + 1% gelatin; (D) 2% gelatin; (E) 5% gelatin; (F) 10% gelatin.

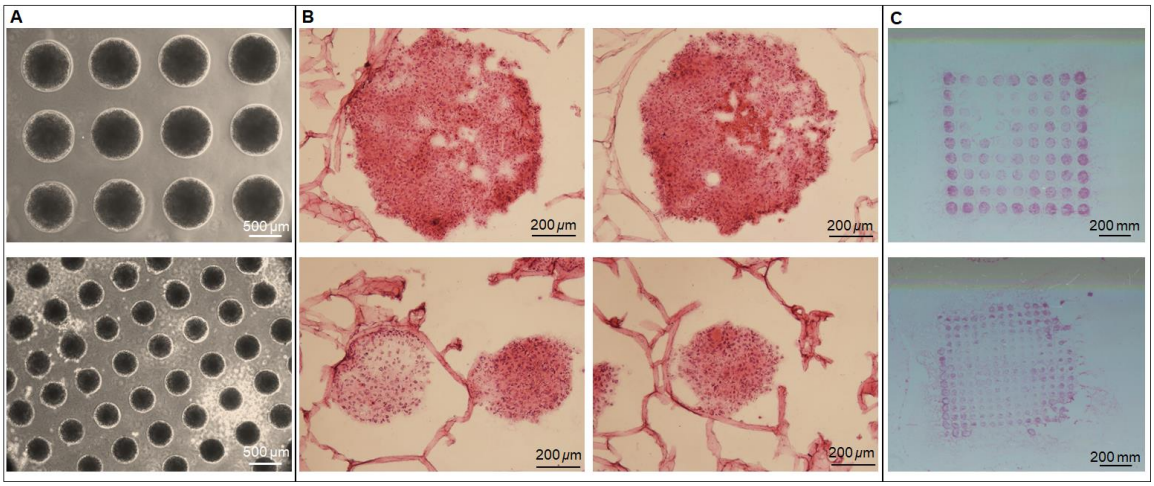


Figure 3. Preparation of MCTS sections with different sizes. (A) Microscope images of spheroids with different sizes cultured in medium. (B) Microscope images of spheroids after H&E staining. (C) Scan of spheroids on glass slides on ITO slides after H&E staining.

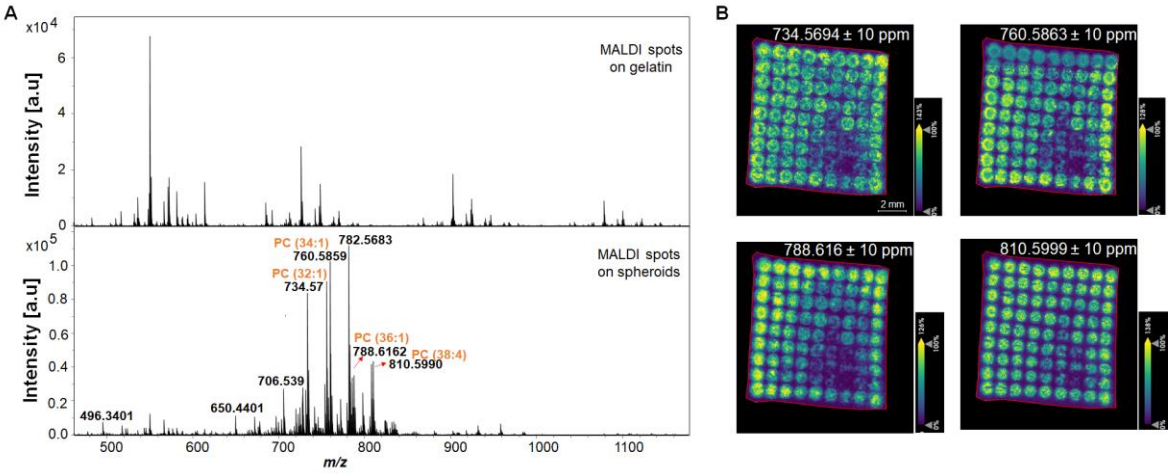
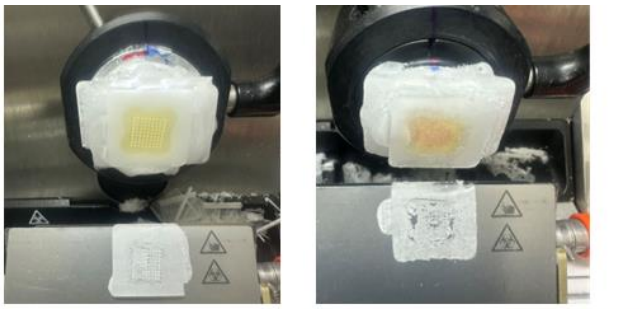
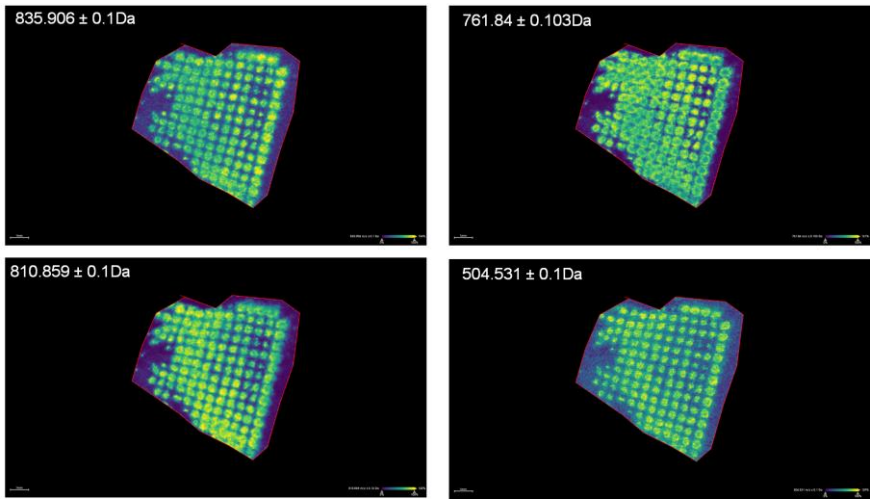


Figure 4. Mass spectra obtained from the PANC1 spheroids: (A) Mass spectra acquired on gelatin (top) and on spheroids (bottom), with several mass spectral peaks assigned and annotated as PC species based on accurate mass matching; (B) Representative MS ion images of selected lipid species with their calculated *m/z* and mass error shown on top of each panel of ion images from the 81 spheroids in one section.

Supplemental Information



Supplementary Figure S1. Evaluate the sample preparation sequence for sectioning. All samples were embedded in 5% gelatin.



Supplementary Figure S2. Representative results of MALDI MS ion images of molecular species from the array of spheroids with a size around 300 – 400 µm.

Chapter 4

Interactions with Stromal Cells Alters the Protein Expression Levels of Pancreatic Cancer Spheroids

Adapted from:
Yuan Liu, Jillian Johnson, Daniel G. Delafield, Sophie Mancha, Jingwei Zhang, Melissa Skala, Lingjun Li, "Interactions with Stromal Cells Alters the Protein Expression Levels of Pancreatic Cancer Spheroids", in preparation for submission to *Journal of Proteome Research*.

Author Contributions:
L.L. designed, and supervised the whole study; Y.L. performed experiments, analyzed, interpreted the data, and wrote the manuscript; J.J performed the experiments; G.D performed the experiments, analyzed the data. All authors provided feedback.

Abstract

Pancreatic stellate cells (PSCs) promote cancer cell survival, proliferation, tumor growth and a more oxidized cancer cell redox state. However, there is little knowledge about changes of protein levels in pancreatic cancer cells when they interact with PSCs, especially for pancreatic cancer spheroids. Also, there is a dearth of knowledge about the changes of protein levels in pancreatic cancer cells during different time points/size of cell culture. Here, we employ a data-independent acquisition method to quantify the protein levels in pancreatic cancer spheroids cultured alone and cocultured with PSCs at different time points and profile the changes of protein levels in pancreatic cancer spheroids. Over 7000 proteins were quantified in the samples. The Gene Ontology analysis reveals enriched GO terms in the long duration of cell cultures for response to hypoxia, response to oxygen levels and ribonucleoprotein complex biogenesis etc. for their biological process (BP). Also, we observe that HTRA1 in the PANC1 spheroids cocultured with PSCs showed about 24% increase in the abundance than that in monoculture of PANC1 spheroids.

Introduction

The stroma of pancreatic ductal adenocarcinoma (PDAC) is composed of a mixture of different kinds of cells, including endothelial cells and cancer-associated fibroblasts (CAFs; derived mainly from activated pancreatic stellate cells (PSCs)),[1] associated with tumor survival, tumor growth and invasiveness.[2-4] Previous studies have shown that stroma can change the patterns of pancreatic cancer epigenome, metabolome, RNA, and proteins in human PDAC,[5-7] and can change the subtypes, differentiation, tumor immunity and treatment responses.[8, 9] As the stroma of PDAC is complex with different cells, information is still lacking in the impact of specific cell type to PDAC in different circumstances. Among these cells, PSCs play a critical role in PDAC

regarding tumor growth and drug resistance. Previous studies find that PSCs secret some metabolites like alanine, deoxycytidine or proteins like fibronectin that can promote the proliferation and survival of pancreatic cancer cells, and even promote the invasiveness and drug-resistance of cancer cells.[10-13] A recent study showed that PSCs promote a more oxidized cancer cell redox state in pancreatic cancer tumors.[14] Still, there is a knowledge gap of how PSCs change the protein expression levels of PDAC cells, especially for PDAC spheroids.

Immortalized cell culture models provide valuable *in vitro* options for research. Among them, two-dimensional (2D) cell culture is widely used because of its relatively low cost, easy manipulation, good reproducibility, and short experimentation duration. Nonetheless, 3D cell culture is emerging as a viable model system and has gained increasingly more attention in cancer research and drug development. Compared with 2D cell culture, 3D cell culture often provides better physiological relevance and shows better drug metabolism and permits longer time of observation of cell growth.[15] Meanwhile, 3D cell culture allows much better real-time and/or fixed microscope imaging than in animal models[15]. It also permits rapid test of drugs compared to animal experiments.[15] 3D cell culture has been used to investigate interactions between stromal cells and PDAC,[14, 16, 17] drug penetrations,[18-20] drug metabolisms.[21, 22] Overall, the significance and usefulness of 3D cell culture in different kinds of research fields is unquestionable. It is a useful model for investigating cellular mechanisms and functions, which can benefit more accurate diagnosis and better therapeutic treatments. Therefore, in this study, we apply the 3D cell culture model for investigating the interactions between PSCs and PDAC.

Given the advantages of 3D cell culture and the importance in answering how PSCs affect PDAC, we utilized data-independent acquisition (DIA)-based proteomic technologies to

investigate the role of PSCs in PDAC spheroids. Also, we analyzed the changes of protein levels at different cell culture time points.

Materials and Methods

Reagents.

Optima LC/MS grade acetonitrile, water, trifluoroacetic acid (TFA), agarose were purchased from Fisher Scientific (Pittsburgh, PA, USA). Formic acid (FA) was purchased from Fluka (Buchs, Switzerland). Dithiothreitol (DTT), iodoacetamide (IAA), calcium chloride (CaCl2) and sodium chloride (NaCl) were acquired from Sigma-Aldrich (St Louis, MO, USA). Complete protease inhibitors cocktail tablets mini (EDTA-free) and phosphatase inhibitors cocktail tablets were acquired from Roche Diagnostics (Mannheim, Germany).

Cell culture.

The commercial pancreatic cancer cell line PANC-1 (ATCC® CRL-1469) was obtained from American Type Culture Collection (ATCC, Manassas, VA, USA). Human primary pancreatic stellate cells (PSCs) were isolated from PDAC tumor specimen resected from patients at the University of Wisconsin Carbone Cancer Center (UWCCC) with informed consent, the procedure of obtaining PSC cell line has been described in previous studies[23, 24]. The cell lines were cultured at 37 °C with 5% CO_2 using DMEM, high glucose (Life Technologies 11965) containing 10% fetal bovine serum (FBS) (Gibco), 1% penicillin–streptomycin solution (Gibco).

Growth of pancreatic cancer spheroids and harvest.

The PANC1 spheroids were formed using MicroTissues 3D Petri dish micro-molds (Sigma-Aldrich, 81 circular wells 800 μm × 800 μm) based on the protocol of manufacture. Briefly, 1% melted agarose was poured on each of the molds and the gelled agarose was pulled out and

placed in 6 well plates. Suspended PANC1 cells were added onto the agarose-based 3D Petri dishes to form spheroids. Seed PANC1 for spheroids. Around 81,000 PANC1 cells were seeded to each Petri dish (~ 1000 cells per well) to form spheroids and cultured for two days before co-culture. About 30,000 PSCs were seeded in each well of 6 well plates (≥80% confluence) one day before co-culture. Medium in 6 well plates were removed, and filters (0.4 μm) were inserted in each well and PANC1 spheroids together with Petri dishes were transferred on the top of filters. In this co-culture setup, PSCs grown on the bottom of 6 well plates and spheroids were separated by filters without direct touch, but nutrients/secreted compounds exchange were allowed. For spheroids cultured alone, no PSCs on the bottom of 6 well plates. Medium (5 mL per well) was changed every two days by half. Spheroids cultured for 1 day, 3 days and 5 days were harvested respectively. For harvesting spheroids, medium was removed, and spheroids were washed using PBS twice and transferred to Eppendorf tubes and washed a third time before the spheroids were frozen at – 80 °C for future use. For each sample, there were 81 spheroids in each 3D petri dish combined for proteomic analyses.

Protein digestion.

Cell pellets of spheroids were resuspended in 200 μL of 8 M urea solution (containing 1× protease inhibitor cocktail) via ultrasonication. Protein concentration was measured using bicinchoninic acid (BCA) assay (ThermoFisher Scientific). DTT was added to reduce disulfide bonds at 5 mM for 1 hour min at room temperature followed by adding 15 mM IAA alkylation at room temperature in dark for 30 min. Redundant IAA was quenched by adding 5 mM DTT. Then trypsin/LysC was added for digesting proteins at an enzyme to protein ratio of 1:50 for 3 hours followed by diluting the 8 M urea to 1 M using 50 mM Tris-HCl for overnight digestion at 37 °C. The digestion reaction was acidified with a final concentration of 0.3% TFA (pH 3), followed by

desalting with Sep-Pak C18 cartridges (Waters Corporation, Milford, Massachusetts, MA, United States). The digested peptides were dried down with a SpeedVac concentrator (Thermo Scientific, Waltham, MA, United States) and resuspended in 0.1% formic acid (FA). Peptide concentration was measured by peptide assay (Thermo Scientific, Rockford, IL, United States). Peptide concentrations were adjusted at 200 ng/μL in 0.1% FA for LC-MS/MS analysis.

NanoLC-MS acquisition.

The HPLC-MS/MS analysis was conducted using a TIMS-TOF MS (Bruker, Germany) equipped with Waters nanoACQUITY UPLC. The separation column was in-house made with an emitter tip and dimensions of 75 μm inner diameter × 16 cm length. The column was packed with 1.7 μm, 150 Å, ethylene-bridged-hybrid (BEH) C18 material (Waters, Milford, MA). The positive ion mode (+ESI) was used, and the standard instrument parameters were set as follows. Mobile phase A was water containing 0.1% formic acid, and mobile phase B was acetonitrile containing 0.1% formic acid. The flow rate was set as 0.3 μL/min, and the LC gradient was 90 min and set as follows: 0−42 min, 2%-13% solvent B; 42−65 min, 13%−20% B; 65 - 70 min, 20%-30% B; 70−75 min, 30% - 85% B; 75-80 min, 85% B; 80-82 min, 85%-2% B. Positive ionization mode was used and full MS scans were acquired from *m/z* 100 to 1700 using dia-PASEF mode, 1/K0 from 0.7 to 1.3 Vs/cm^2. Ramp time and accumulation time was set as 100 ms, duty cycle 100%, ramp rate 9.42 Hz. The parameters for dia-PASEF window were set as mass width 30 Da, mass overlap 1.0 Da, mass steps per cycle 27, mass range *m/z* 300 to 1084. Capillary voltage 1500 V. nebulizer gas pressure, 0.3 bar (N_2); drying gas rate, 3.0 L/min (N_2); drying gas temperature, 180 °C. Collision energy 20 to 59 eV corresponding to 1/K0 from 0.6-1.6 Vs/cm^2.

Microscope image analysis.

The optical images of spheroids were taken in bright-field mode under an inverted microscope (TI-S, NIKON, Japan) with a SPOT imaging system (Diagnostic Instruments, Inc., Sterling Heights, MI). The size of spheroids was measured using ImageJ software and analyzed in Origin 2020.

TIMS TOF raw data analysis.

We employed DIA-NN for transforming raw data into precursor and fragment identifications. All data were searched against the reviewed human proteome (UniProt, April 9th, 2023, 20,398 proteins) with trypsin/LysC as digestion enzymes. "FASTA digest for library-free search/library generation" was enabled. "Deep learning-based spectra, RTs and IMs prediction" was enabled. The maximum number of missed cleavages was set to 1. Cysteine carbamidomethylation enabled as a fixed modification. Methionine oxidation, methionine excision at the N terminus were enabled. The maximum variable modification was set as 1. Peptide length set between 7 and 30. Precursor charge was set between 2 and 4. Precursor and fragment ion *m/z* were set between 100 and 1700. Mass and MS1 accuracy were set as 15 ppm. "Match between run" was enabled and "Heuristic protein inference" was enabled. "Robust LC (high precision)" was enabled as the quantification strategy. The proteomics output tables were filtered for a maximum of 1% of q value at both precursor and global protein levels. The "PG.MaxLFQ" column integrated in the DIA-NN output tables was used for quantitative analysis on the protein level[25]. R version 4.2.2 was used for further data analysis. Gene ontology (GO) analysis was performed using the clusterProfiler package in Rstudio.[26]

Results and Discussion

Workflow of relative quantification for proteins in PANC1 spheroids.

The workflow used in this study is shown in **Figure 1.** Briefly, PANC1 cells and PSCs were cultured in 2D cell culture dishes before co-culture. PSCs were seeded in 6 well plates and PANC1 cultured in agarose-based 3D petri dishes were cocultured with PSCs or cultured alone. Pictures were taken to monitor the growth of PANC1 spheroids. Growth of spheroids were analyzed using ImageJ and Origin. Meanwhile, spheroids were harvested and digested to generate peptides and subjected to TIMS TOF for data acquisition. Database searching was performed in DIA-NN.

Growth of PANC1 spheroids.

We monitored the growth of about 70 spheroids since the day of co-culture (representative microscope images were shown in **Figure 2A**). We removed the smallest 5 and biggest 5 spheroids on day 0 (the start of co-culture). The remaining 60 spheroids were used to calculate the average size of spheroids during co-culture or cultured alone. The changes of spheroid size in (diameter) were shown in **Figure 2B**.

Protein quantification accuracy.

Data-independent acquisition (DIA) allows the mass spectrometer to fragment a pre-defined set of precursor isolation windows, therefore unbiasedly fragmenting all the precursors within the isolation window, showing higher reproducibility compared to data-dependent acquisition (DDA)[27]. Here, we prepared 3 biological replicates for each cell culture condition and performed 3 technical replicates in LC-MS/MS runs for each sample. For all the shared quantified proteins identified in DIA-NN, we evaluated the reproducibility of technical replicates and biological replicates (**Figure 3**). Among these 18 samples, there were a total of 7073 proteins quantified in all the samples. And more than 90% of proteins quantified have coefficient variation (CV) less than 15% and about 80% of proteins quantified have CV less than 10% for each of the

samples (**Table S1**). For the six different cell culture groups, there were 7451 proteins quantified in all 6 groups (quantified in at least two out of three samples in each group). Among the quantified proteins, more than 86% of proteins were quantified with a CV less than 15% among the 3 biological replicates in all these groups (**Table S2**). These data suggested that the quantification accuracy is decent in the DIA-based quantification.

Changes of protein levels and GO analysis.

To profile changes of protein levels in different cell culture time points, we quantified cell culture samples from two cell culture conditions at three culture durations with 3 biological replicates for each condition. Firstly, we compared the protein levels across 3 different culture durations within the monocultured PANC1 spheroids and PANC1 spheroids cocultured with PSCs. We identified some proteins that showed up or down regulations (ratio of 1.5 folds and $p < 0.05$) between different culture durations at the same culture condition (**Figure 4**). From the figure we found that some proteins were down or up regulated from day 1 to day 5. However, we found that the protein levels tend to be more stable after longer periods of cell culture based on the numbers of proteins that showed up or down regulation. For example, in the coculture group, we found 69 proteins were down regulated from day 1 to day 3, and 192 proteins were up regulated. However, only 8 proteins were down regulated from day 3 to day 5, and 18 proteins were up regulated in the coculture groups. It was possible that at the early stages, the percentage of cells exposed to high amounts of nutrients was higher than that in the late stages as the spheroids grew bigger. Cells in the inner portion of big spheroids would have very limited nutrients accessible so they grew slower or were inert. GO analysis showed that several pathways were down regulated in the PANC1 spheroids co-cultured in day 5 compared to day 1 (**Figure 5A**), including ribonucleoprotein complex biogenesis, cytoplasmic translation, and ribosome biogenesis. These findings suggested

that cells in day 5 were less active than day 1. Not surprisingly, response to oxygen levels and response to hypoxia and several other metabolic biological processes were up regulated in day 5 compared to day 1 (**Figure 5B**), possibly that more cells were suffered from a shortage of oxygen as the spheroids grew. Similarly, for monoculture of PANC1 spheroids, we observed consistent changes of biological process between day 1 and day 5 (**Figure S1**).

We also compared the changes of protein levels between monoculture of PANC1 spheroids and PANC1 spheroids cocultured with PSCs. We did not find significant changes of protein levels (p-value < 0.05) in day 1 and day 3. However, we found that there was a significant difference of HTRA1 between the monoculture and cocultured PANC1 spheroids in day 5 (**Figure 6**). Protein-protein interaction networks functional enrichment analysis was performed for HTRA1 (**Figure 6B**) in STRING. The database showed several proteins interacting with HTRA1, including CTRC, ARMS2, CFH, ACAN, GALE, CLPP, CLPB, HSPD1, PINK1, MAPT and ACAN.

Clustering analysis.

We further performed hierarchical clustering analysis for selected proteins related to response to hypoxia, ribonucleoprotein complex biogenesis and HTRA1 (**Figure 7A**). A dendrogram was generated using the average linkage method and Euclidean distance to visualize the hierarchical clustering of the protein levels from PANC1 spheroids in different groups (**Figure 7A**). Based on the clustering of groups, we found that PANC1 spheroids showed higher similarities in the same day of coculture regardless of whether it was cultured alone or cocultured with PSCs. Additionally, we performed uniform manifold approximation and projection for dimension reduction (UMAP) analysis for the 6 groups of PANC1 spheroids (18 samples in total). In the UMAP plots, each data point represented a protein profile of HTRA1, NDRG1 and ITPR1. Based on the plots, we found that PANC1 spheroids cultured in different days (day 1, day 3 and day5)

could be separated wells. And for the PANC1 spheroids cultured the same time but cultured alone or cocultured with PSCs, they had similar protein profiles and could be barely separated using HTRA1, NDRG1 and ITPR1. We used the same proteins to do a principal component analysis (PCA) (**Figure 7C**) and found the PCA analysis did not perform as well as the UMAP analysis based on the distance of same and different groups. It is certain that the analysis would benefit from a large data set to draw a more definitive conclusion.

Although the initial focus of this study was on the impact of PSCs on PDAC spheroids, when we compared the different expression levels of PANC1 spheroids between monoculture and coculture collected at the same day, we only observed HTRA1 in cocultured PANC1 spheroids (day 5) expressed at about 24% higher in abundance compared with monoculture of PANC1 spheroids. There are some possible explanations for the results. As there was a gradient of nutrients and waste for PANC1 spheroids, the secreted compounds such as metabolites and proteins could also have limited access to cells in the inner part of spheroids. Therefore, these spheroids could barely be affected by PSCs. In this case, the percentage of cells showing obvious responses to PSCs could be relatively low so that the changes could not be detected. If that was the case, a potential solution could be using a strategy called pulse-chase labeling of spheroids or perform serial trypsin digestion to characterize the changes in the proteome in different regions of spheroids.[28, 29] This could help distinguish proteins at different layers of the spheroids. Another possible explanation is that PANC1 cells may not respond to PSCs sensitively. The clue could be found in a previous publication showing that different pancreatic cancer cell lines reacted to PSC-conditioned medium differently.[10] In their work, the authors found that the increase of oxygen consumption rate (OCR) for different cell lines treated with PSC-conditioned medium showed different levels compared to control group. For example, MiaPaca2 showed 40% increase in OCR,

while PANC1 showed only about 20% increase[10]. It is worth noting that, in this study, the PSCs and PANC1 cells did not have direct contact with each other, and interactions were mediated through secreted molecules. Therefore, the effects of direct contact of PSCs could not be revealed. In addition, it is possible that the interaction of PSCs and PANC1 spheroids could not elicit substantial changes of protein levels, or the changes of protein levels were too subtle to be detected in this study due to their low abundance. It is also possible that the changes of protein levels were relatively low, and we need measure more biological replicates to get significant differences ($p < 0.05$) of different groups. Additionally, the changes may happen at the protein post-translational modification (PTM) levels, such as glycosylation or phosphorylation, or metabolite levels, where we did not investigate in this study.

Some future directions for this research could be increasing the biological replicates and investigating metabolomics and PTMs together with proteomics. To increase the throughput of data acquisition, isobaric tag labeling strategies could be used to relatively quantify the metabolites, proteins and PTMs.[30-33] Also, it would be worth trying to investigate how direct contact between PSCs and PDAC cells affects PDAC.[34, 35]

Conclusions

This study investigates the protein profiles of PANC1 spheroids cultured alone or cocultured with PSCs at different time points. The differences between spheroids at different time points are observable for many proteins, especially for proteins related to response to hypoxia, response to oxygen levels and ribonucleoprotein complex biogenesis etc. These findings are consistent with previous literature about the gradient of nutrients including oxygen, ATP, glucose in the culture of spheroids may change the expression level of metabolites or proteins.[28, 29, 36-38]

And we also discovered that protein HTRA1 showed significant increases (24%, $p < 0.01$) in PANC1 spheroids cocultured with PSCs compared to monoculture.

Acknowledgements

This research was supported in part by UW Carbone Cancer Center pilot grant, the National Institutes of Health (NIH) grants RF1 AG052324, R21AG065728, R01 DK071801, and R01AG078794. Some of the mass spectrometers were acquired using NIH shared instrument grants S10 OD028473, S10 RR029531, and S10 OD025084.

References

1. Apte, M.; Pirola, R. C.; Wilson, J. S., Pancreatic stellate cell: physiologic role, role in fibrosis and cancer. *Curr Opin Gastroenterol* **2015,** *31* (5), 416-23.

2. Apte, M. V.; Wilson, J. S.; Lugea, A.; Pandol, S. J., A starring role for stellate cells in the pancreatic cancer microenvironment. *Gastroenterology* **2013,** *144* (6), 1210-9.

3. Bernard, V.; Semaan, A.; Huang, J.; San Lucas, F. A.; Mulu, F. C.; Stephens, B. M.; Guerrero, P. A.; Huang, Y.; Zhao, J.; Kamyabi, N.; Sen, S.; Scheet, P. A.; Taniguchi, C. M.; Kim, M. P.; Tzeng, C. W.; Katz, M. H.; Singhi, A. D.; Maitra, A.; Alvarez, H. A., Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. *Clin Cancer Res* **2019,** *25* (7), 2194-2205.

4. Gao, Z.; Wang, X.; Wu, K.; Zhao, Y.; Hu, G., Pancreatic stellate cells increase the invasion of human pancreatic cancer cells through the stromal cell-derived factor-1/CXCR4 axis. *Pancreatology* **2010,** *10* (2-3), 186-93.

5. Sherman, M. H.; Yu, R. T.; Tseng, T. W.; Sousa, C. M.; Liu, S.; Truitt, M. L.; He, N.; Ding, N.; Liddle, C.; Atkins, A. R.; Leblanc, M.; Collisson, E. A.; Asara, J. M.; Kimmelman, A. C.; Downes, M.; Evans, R. M., Stromal cues regulate the pancreatic cancer epigenome and metabolome. *Proc Natl Acad Sci U S A* **2017,** *114* (5), 1129-1134.

6. Ligorio, M.; Sil, S.; Malagon-Lopez, J.; Nieman, L. T.; Misale, S.; Di Pilato, M.; Ebright, R. Y.; Karabacak, M. N.; Kulkarni, A. S.; Liu, A.; Vincent Jordan, N.; Franses, J. W.; Philipp, J.; Kreuzer, J.; Desai, N.; Arora, K. S.; Rajurkar, M.; Horwitz, E.; Neyaz, A.; Tai, E.; Magnus, N. K. C.; Vo, K. D.; Yashaswini, C. N.; Marangoni, F.; Boukhali, M.; Fatherree, J. P.; Damon, L. J.; Xega, K.; Desai, R.; Choz, M.; Bersani, F.; Langenbucher, A.; Thapar, V.; Morris, R.;

Wellner, U. F.; Schilling, O.; Lawrence, M. S.; Liss, A. S.; Rivera, M. N.; Deshpande, V.; Benes, C. H.; Maheswaran, S.; Haber, D. A.; Fernandez-Del-Castillo, C.; Ferrone, C. R.; Haas, W.; Aryee, M. J.; Ting, D. T., Stromal Microenvironment Shapes the Intratumoral Architecture of Pancreatic Cancer. *Cell* **2019,** *178* (1), 160-175.e27.

7. Jiang, W.; Qiao, L.; Han, Y.; Zhang, A.; An, H.; Xiao, J.; Ren, L., Pancreatic stellate cells regulate branched-chain amino acid metabolism in pancreatic cancer. *Ann Transl Med* **2021,** *9* (5), 417.

8. Grunwald, B. T.; Devisme, A.; Andrieux, G.; Vyas, F.; Aliar, K.; McCloskey, C. W.; Macklin, A.; Jang, G. H.; Denroche, R.; Romero, J. M.; Bavi, P.; Bronsert, P.; Notta, F.; O'Kane, G.; Wilson, J.; Knox, J.; Tamblyn, L.; Udaskin, M.; Radulovich, N.; Fischer, S. E.; Boerries, M.; Gallinger, S.; Kislinger, T.; Khokha, R., Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell* **2021**.

9. Ogawa, Y.; Masugi, Y.; Abe, T.; Yamazaki, K.; Ueno, A.; Fujii-Nishimura, Y.; Hori, S.; Yagi, H.; Abe, Y.; Kitago, M.; Sakamoto, M., Three Distinct Stroma Types in Human Pancreatic Cancer Identified by Image Analysis of Fibroblast Subpopulations and Collagen. *Clin Cancer Res* **2021,** *27* (1), 107-119.

10. Sousa, C. M.; Biancur, D. E.; Wang, X.; Halbrook, C. J.; Sherman, M. H.; Zhang, L.; Kremer, D.; Hwang, R. F.; Witkiewicz, A. K.; Ying, H.; Asara, J. M.; Evans, R. M.; Cantley, L. C.; Lyssiotis, C. A.; Kimmelman, A. C., Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. *Nature* **2016,** *536* (7617), 479-83.

11. Amrutkar, M.; Aasrum, M.; Verbeke, C. S.; Gladhaug, I. P., Secretion of fibronectin by human pancreatic stellate cells promotes chemoresistance to gemcitabine in pancreatic cancer cells. *BMC Cancer* **2019,** *19* (1), 596.

12. Dalin, S.; Sullivan, M. R.; Lau, A. N.; Grauman-Boss, B.; Mueller, H. S.; Kreidl, E.; Fenoglio, S.; Luengo, A.; Lees, J. A.; Vander Heiden, M. G.; Lauffenburger, D. A.; Hemann, M. T., Deoxycytidine Release from Pancreatic Stellate Cells Promotes Gemcitabine Resistance. *Cancer Res* **2019,** *79* (22), 5723-5733.

13. Richards, K. E.; Zeleniak, A. E.; Fishel, M. L.; Wu, J.; Littlepage, L. E.; Hill, R., Cancer-associated fibroblast exosomes regulate survival and proliferation of pancreatic cancer cells. *Oncogene* **2017,** *36* (13), 1770-1778.

14. Datta, R.; Sivanand, S.; Lau, A. N.; Florek, L. V.; Barbeau, A. M.; Wyckoff, J.; Skala, M. C.; Vander Heiden, M. G., Interactions with stromal cells promote a more oxidized cancer cell redox state in pancreatic tumors. *Sci Adv* **2022,** *8* (3), eabg6383.

15. Yamada, K. M.; Cukierman, E., Modeling tissue morphogenesis and cancer in 3D. *Cell* **2007,** *130* (4), 601-10.

16. Lee, J. H.; Kim, S. K.; Khawar, I. A.; Jeong, S. Y.; Chung, S.; Kuh, H. J., Microfluidic co-culture of pancreatic tumor spheroids with stellate cells as a novel 3D model for investigation of stroma-mediated cell motility and drug resistance. *J Exp Clin Cancer Res* **2018,** *37* (1), 4.

17. Lau, A. N.; Li, Z.; Danai, L. V.; Westermark, A. M.; Darnell, A. M.; Ferreira, R.; Gocheva, V.; Sivanand, S.; Lien, E. C.; Sapp, K. M.; Mayers, J. R.; Biffi, G.; Chin, C. R.; Davidson, S. M.; Tuveson, D. A.; Jacks, T.; Matheson, N. J.; Yilmaz, O.; Vander Heiden, M. G., Dissecting cell-type-specific metabolism in pancreatic ductal adenocarcinoma. *Elife* **2020,** *9*.

18. Johnson, J.; Sharick, J. T.; Skala, M. C.; Li, L., Sample preparation strategies for high-throughput mass spectrometry imaging of primary tumor organoids. *J Mass Spectrom* **2020,** *55* (4), e4452.

19. Liu, X.; Weaver, E. M.; Hummon, A. B., Evaluation of therapeutics in three-dimensional cell culture systems by MALDI imaging mass spectrometry. *Anal Chem* **2013,** *85* (13), 6295-302.

20. Feist, P. E.; Sidoli, S.; Liu, X.; Schroll, M. M.; Rahmy, S.; Fujiwara, R.; Garcia, B. A.; Hummon, A. B., Multicellular Tumor Spheroids Combined with Mass Spectrometric Histone Analysis To Evaluate Epigenetic Drugs. *Anal Chem* **2017,** *89* (5), 2773-2781.

21. Liu, X.; Flinders, C.; Mumenthaler, S. M.; Hummon, A. B., MALDI Mass Spectrometry Imaging for Evaluation of Therapeutics in Colorectal Tumor Organoids. *J Am Soc Mass Spectrom* **2018,** *29* (3), 516-526.

22. LaBonia, G. J.; Lockwood, S. Y.; Heller, A. A.; Spence, D. M.; Hummon, A. B., Drug penetration and metabolism in 3D cell cultures treated in a 3D printed fluidic device: assessment of irinotecan via MALDI imaging mass spectrometry. *Proteomics* **2016,** *16* (11-12), 1814-21.

23. Drifka, C. R.; Eliceiri, K. W.; Weber, S. M.; Kao, W. J., A bioengineered heterotypic stroma-cancer microenvironment model to study pancreatic ductal adenocarcinoma. *Lab Chip* **2013,** *13* (19), 3965-75.

24. Bachem, M. G.; Schneider, E.; Gross, H.; Weidenbach, H.; Schmid, R. M.; Menke, A.; Siech, M.; Beger, H.; Grunert, A.; Adler, G., Identification, culture, and characterization of pancreatic stellate cells in rats and humans. *Gastroenterology* **1998,** *115* (2), 421-32.

25. Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **2014,** *13* (9), 2513-26.

26. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; Fu, X.; Liu, S.; Bo, X.; Yu, G., clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2021,** *2* (3), 100141.

27. Ludwig, C.; Gillet, L.; Rosenberger, G.; Amon, S.; Collins, B. C.; Aebersold, R., Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* **2018,** *14* (8), e8126.

28. McMahon, K. M.; Volpato, M.; Chi, H. Y.; Musiwaro, P.; Poterlowicz, K.; Peng, Y.; Scally, A. J.; Patterson, L. H.; Phillips, R. M.; Sutton, C. W., Characterization of changes in the proteome in different regions of 3D multicell tumor spheroids. *J Proteome Res* **2012,** *11* (5), 2863-75.

29. Beller, N. C.; Lukowski, J. K.; Ludwig, K. R.; Hummon, A. B., Spatial Stable Isotopic Labeling by Amino Acids in Cell Culture: Pulse-Chase Labeling of Three-Dimensional Multicellular Spheroids for Global Proteome Analysis. *Anal Chem* **2021,** *93* (48), 15990-15999.

30. Liu, Y.; Zhang, H.; Dove, W. F.; Wang, Z.; Zhu, Z.; Pickhardt, P. J.; Reichelderfer, M.; Li, L., Quantification of Serum Metabolites in Early Colorectal Adenomas Using Isobaric Labeling Mass Spectrometry. *J Proteome Res* **2023,** *22* (5), 1483-1491.

31. Frost, D. C.; Feng, Y.; Li, L., 21-plex DiLeu Isobaric Tags for High-Throughput Quantitative Proteomics. *Anal Chem* **2020,** *92* (12), 8228-8234.

32. Wang, D.; Ma, M.; Huang, J.; Gu, T. J.; Cui, Y.; Li, M.; Wang, Z.; Zetterberg, H.; Li, L., Boost-DiLeu: Enhanced Isobaric N,N-Dimethyl Leucine Tagging Strategy for a Comprehensive Quantitative Glycoproteomic Analysis. *Anal Chem* **2022,** *94* (34), 11773-11782.

33. Wu, Z.; Xiang, W.; Huang, L.; Li, S.; Zhang, X., Hyperplexing Approaches for up to 45-Plex Quantitative Proteomic Analysis. *Anal Chem* **2023,** *95* (12), 5169-5175.

34. Lee, S. Y.; Park, S. B.; Kim, Y. E.; Yoo, H. M.; Hong, J.; Choi, K. J.; Kim, K. Y.; Kang, D., iTRAQ-Based Quantitative Proteomic Comparison of 2D and 3D Adipocyte Cell

Models Co-cultured with Macrophages Using Online 2D-nanoLC-ESI-MS/MS. *Sci Rep* **2019,** *9* (1), 16746.

35. Lindoso, R. S.; Sandim, V.; Collino, F.; Carvalho, A. B.; Dias, J.; da Costa, M. R.; Zingali, R. B.; Vieyra, A., Proteomics of cell-cell interactions in health and disease. *Proteomics* **2016,** *16* (2), 328-44.

36. Wang, Y.; Hummon, A. B., MS imaging of multicellular tumor spheroids and organoids as an emerging tool for personalized medicine and drug discovery. *J Biol Chem* **2021,** *297* (4), 101139.

37. Zang, Q.; Sun, C.; Chu, X.; Li, L.; Gan, W.; Zhao, Z.; Song, Y.; He, J.; Zhang, R.; Abliz, Z., Spatially resolved metabolomics combined with multicellular tumor spheroids to discover cancer tissue relevant metabolic signatures. *Anal Chim Acta* **2021,** *1155*, 338342.

38. Yue, X.; Lukowski, J. K.; Weaver, E. M.; Skube, S. B.; Hummon, A. B., Quantitative Proteomic and Phosphoproteomic Comparison of 2D and 3D Colon Cancer Cell Culture Models. *J Proteome Res* **2016,** *15* (12), 4265-4276.

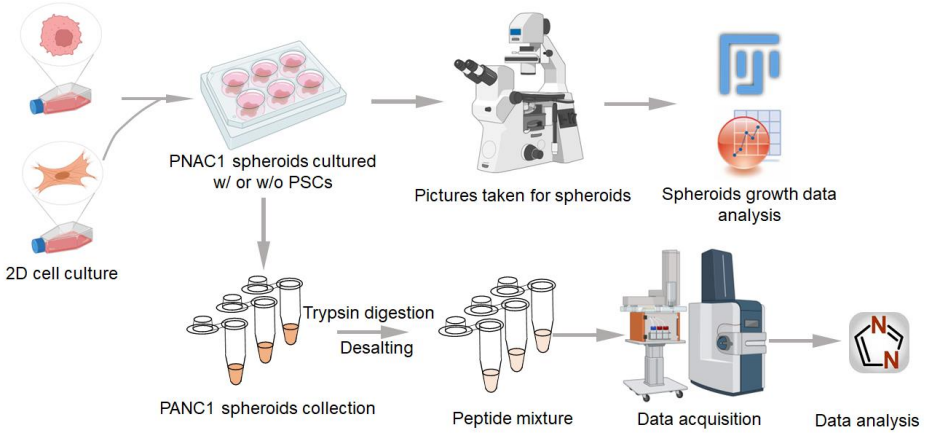


Figure 1. Workflow of study on crosstalk between PCCs and PSCs. PANC1 cells and PSCs were cultured in 2D cell culture dishes. PSCs were seeded in 6 well plates and PANC1 spheroids cultured in agarose-based 3D petri dishes were cocultured with PSCs or cultured alone. Pictures of spheroids were taken to monitor the growth. Growth of spheroids were analyzed using ImageJ and Origin. Spheroids were harvested and digested to generate peptides and subjected to TIMS TOF for data acquisition. Database searching was performed in DIA-NN.

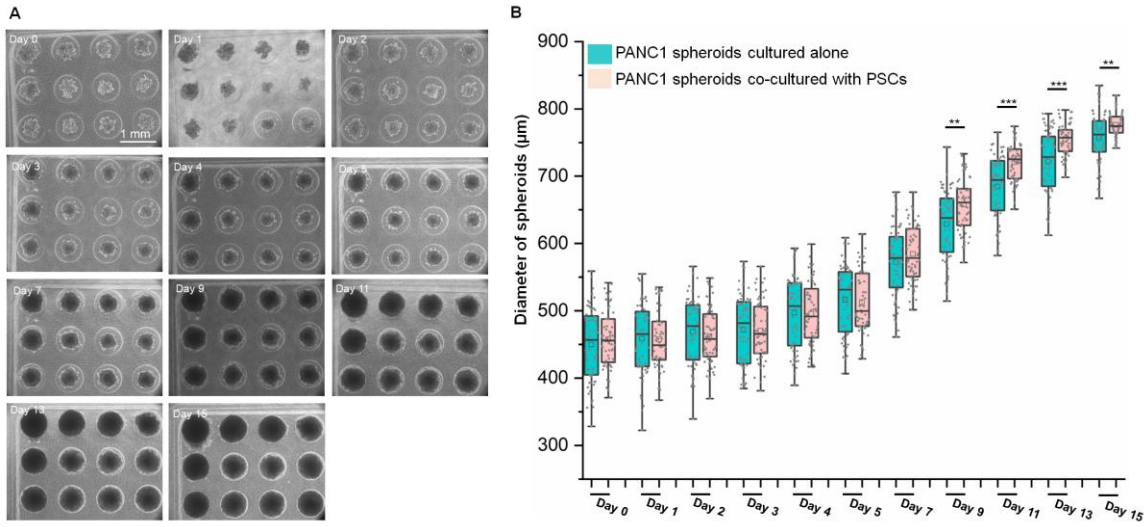


Figure 2. Growth of spheroids. (A) Representative microscope images of spheroids at different time points. (B) Box plot of spheroids size in diameter. The boxes represent the 25% quartile (Q_1), median and 75% quartile (Q_3) and the horizontal line shows the median, and upper whisker limit and lower limit whisker show $Q_3 + 1.5(IQR)$ and $Q_1 - 1.5\ (IQR)$, respectively. Asterisks represent the significance level **P $\leqslant$ 0.01, and ***P $\leqslant$ 0.001) between spheroids cultured alone and co-cultured with PSCs. IQR: Interquartile range.

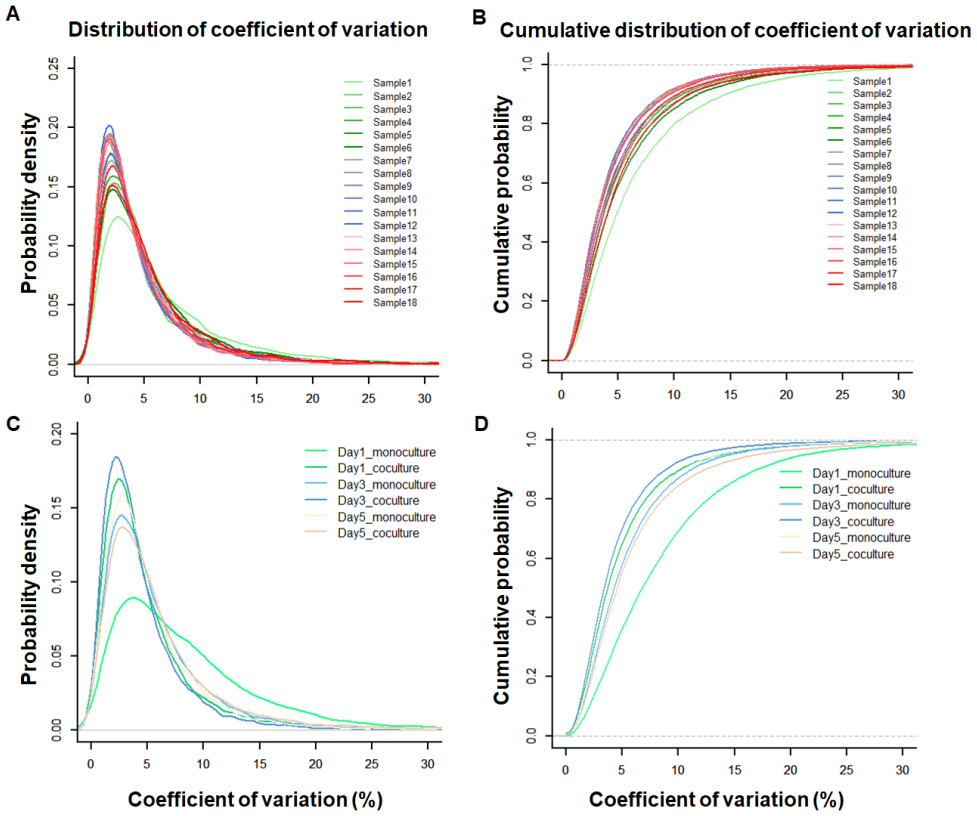


Figure 3. Coefficient of variance for technical replicates and biological replicates. (A) Distribution of coefficient of variance for all 18 samples. Each sample has three technical LC-MS/MS runs. (B) Cumulative distribution of coefficient of variance for all 18 samples. (C) Distribution of coefficient of variance for all 6 cell culture conditions. Each condition has three biological replicates. (D) Cumulative distribution of coefficient of variance for all 6 cell culture conditions.

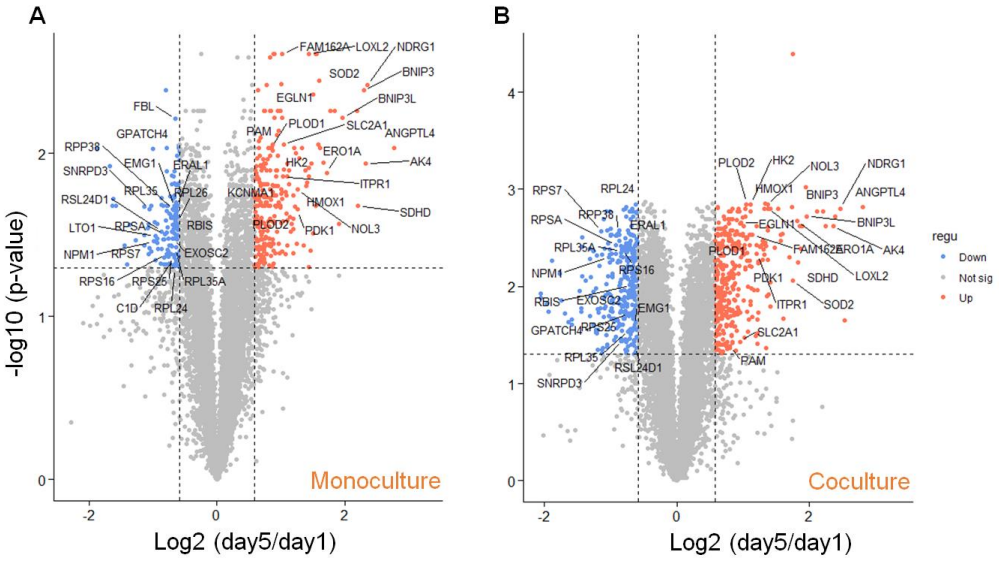


Figure 4. Volcano plots of proteins regulated day 5 compared to day 1 for monoculture of PANC1 spheroids or co-culture of PANC1 spheroids with PSCs. (A) Monoculture on day 5 versus day 1. (B) Coculture on day 5 versus day 1. Two-sided Student's t test, "fold change < 0.667 and p-value < 0.05" = blue, "fold change > 1.5 and p-value < 0.05" = red. Benjamini–Hochberg method was used to adjust the p value. Proteins related to response to hypoxia and proteins and ribonucleoprotein complex biogenesis were selectively labeled.

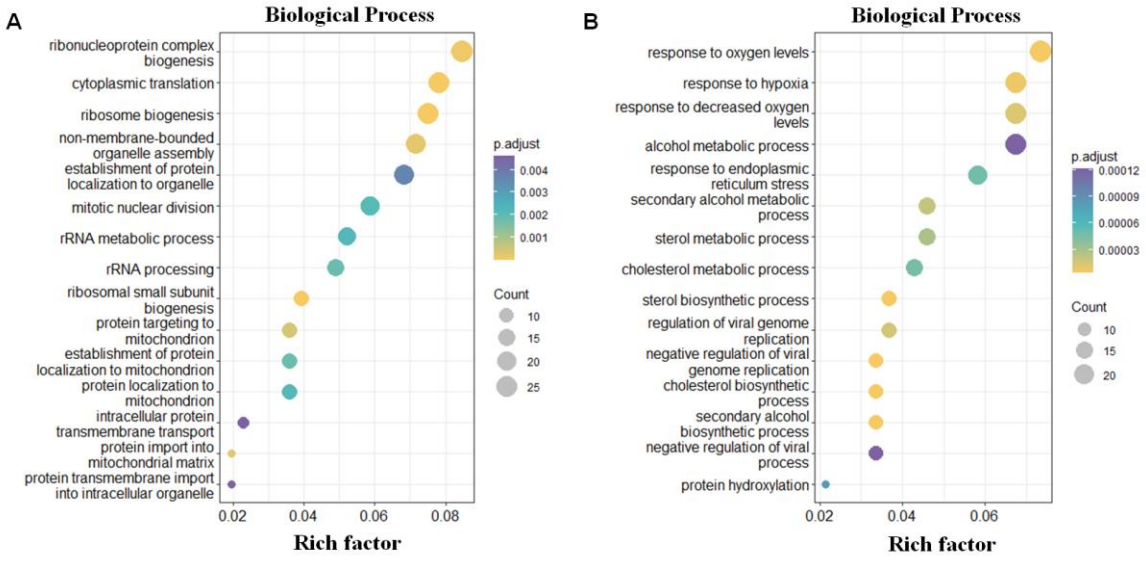


Figure 5. Gene ontology classification of protein changing pattern in PANC1 spheroids cocultured on day 5 and day 1 based on biological processes. (A) Biological processes were down regulated. (B) Biological processes were up regulated.

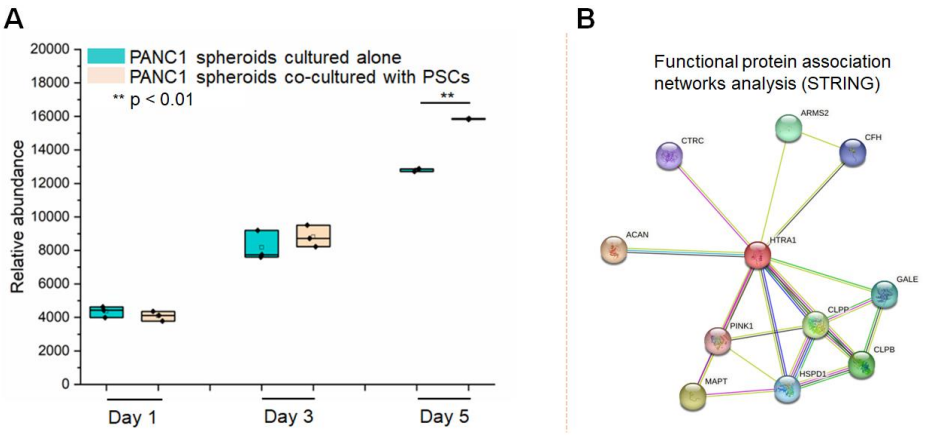


Figure 6. Changes of HTRA1 levels in PANC1 spheroids. (A) Relative abundance of HTRA1 in PANC1 spheroids cultured in different time points and culture conditions. (B) Functional protein association network analysis for HTRA1.

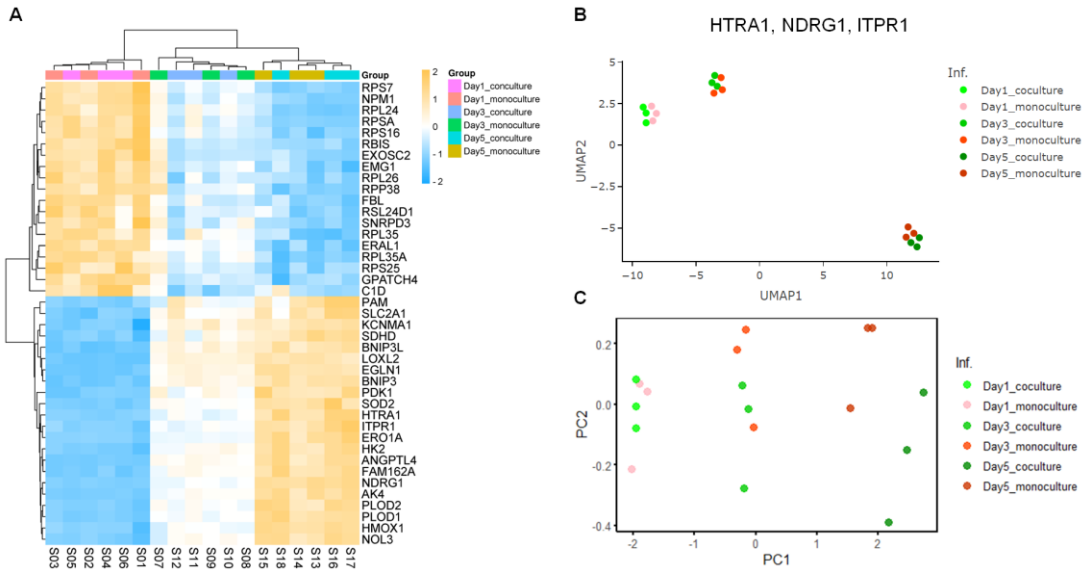


Figure 7. Clustering of PANC1 spheroids. (A) Hierarchical clustering analysis. Proteins showed different expression levels and related to response to hypoxia, ribonucleoprotein complex biogenesis and HTRA1 were used in the analysis. S01 -S03 were monoculture of PANC1 spheroids in day 1, S04 -S06 were coculture of PANC1 spheroids with PSCs in day 1; S07 -S09 were monoculture of PANC1 spheroids in day 3, S10 -S12 were coculture of PANC1 spheroids with PSCs in day 3; S13 -S15 were monoculture of PANC1 spheroids in day 5, S16 -S18 were coculture of PANC1 spheroids with PSCs in day 5; (B) UMAP analysis. Three components were chosen: HTRA1, NDRG1 and ITPR1. (C) PCA analysis. Three components were chosen: HTRA1, NDRG1 and ITPR1.

Supplemental Information

Supplementary Table S1. Number and percentage of quantified proteins at different coefficient variances in all the samples.

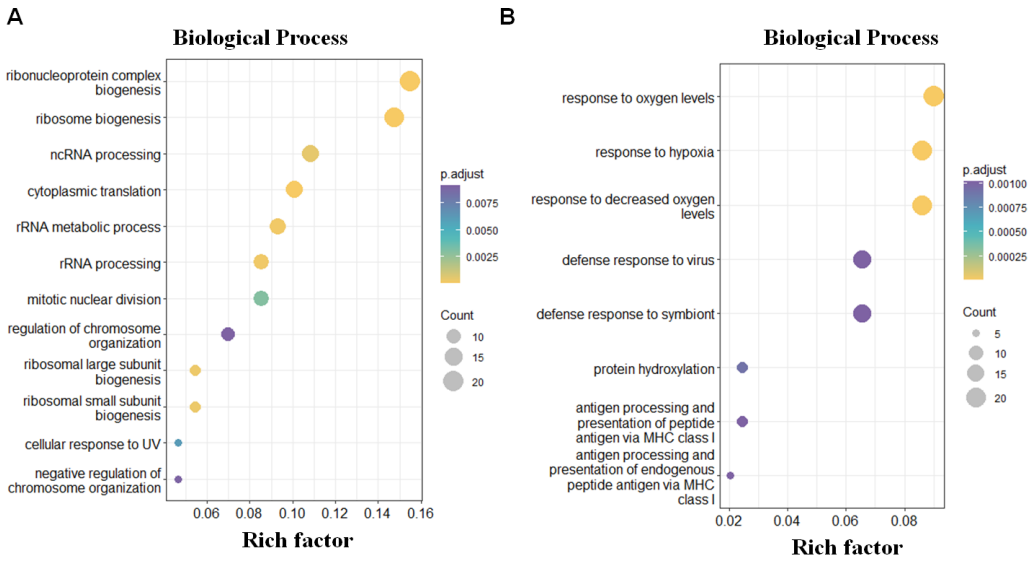
Sample ID	01	02	03	04	05	06	07	08	09
Protein number (CV 15%)	6407	6849	6707	6748	6781	6623	6758	6871	6853
Percentage of proteins (%) (CV 15%)	90.58	96.83	94.83	95.41	95.87	93.64	95.55	97.14	96.89
Protein number (CV 10%)	5656	6489	6171	6257	6294	6022	6296	6471	6443
Percentage of proteins (%) (CV 10%)	79.97	91.74	87.25	88.46	88.99	85.14	89.01	91.49	91.09
Protein number (CV 5%)	3601	5099	4284	4481	4688	4204	4636	5022	4928
Percentage of proteins (%) (CV 5%)	50.91	72.09	60.57	63.35	66.28	59.44	65.55	71.00	69.67

Table S1. Continued.

Sample ID	10	11	12	13	14	15	16	17	18
Protein number (CV 15%)	6851	6785	6863	6771	6845	6861	6843	6784	6684
Percentage of proteins (%) (CV 15%)	96.86	95.93	97.03	95.73	96.78	97.00	96.75	95.91	94.50
Protein number (CV 10%)	6423	6324	6502	6317	6452	6504	6428	6326	6138
Percentage of proteins (%) (CV 10%)	90.81	89.41	91.93	89.31	91.22	91.96	90.88	89.44	86.78
Protein number (CV 5%)	4943	4681	5104	4607	4907	5006	4920	4548	4298
Percentage of proteins (%) (CV 5%)	69.89	66.18	72.16	65.14	69.38	70.78	69.56	64.30	60.77

Supplementary Table S2. Number and percentage of quantified proteins at different coefficient variances in all biological groups.

Groups	Day1 monoculture	Day1 coculture	Day3 monoculture	Day3 coculture	Day5 monoculture
Protein number (CV 15%)	6421	7145	7101	7263	7158
Percentage of proteins (%) (CV 15%)	86.18	95.89	95.30	97.48	96.07
Protein number (CV 10%)	5163	6681	6493	6903	6649
Percentage of proteins (%) (CV 10%)	69.29	89.67	87.14	92.65	89.24
Protein number (CV 5%)	2714	4852	4329	5230	4561
Percentage of proteins (%) (CV 5%)	36.42	65.12	58.10	70.19	61.21



Supplementary Figure S1. Gene ontology classification of protein changing pattern in monoculture of PANC1 spheroids on day 5 and day 1 based on biological processes. (A) Biological processes were down regulated. (B) Biological processes were up regulated.

Chapter 5

Isotopic N,N-dimethyl Leucine Tags for Absolute Quantification of Clusterin and Apolipoprotein E in Alzheimer's Disease

Adapted from:
Yuan Liu, Hua Zhang, Xiaofang Zhong, Zihui Li, Henrik Zetterberg, Lingjun Li. Isotopic N,N-dimethyl Leucine Tags for Absolute Quantification of Clusterin
and Apolipoprotein E in Alzheimer's Disease. (*Journal of Proteomics 257 (2022) 104507*)

Author contributions:
L.L. designed and supervised the whole study; Y.L. performed experiments, analyzed, interpreted the data, and wrote the manuscript; H.Z performed the experiments; X.Z performed the experiments. All authors provided feedback.

Abstract

Alzheimer's disease (AD) is the most common form of dementia and one of the leading causes of death in the United States. In the past decades, extensive efforts have been devoted to biomarker discovery for early diagnosis and treatment of AD. Herein, this study aims to quantify clusterin (CLU) and apolipoprotein E (APOE) in blood samples from AD patients and evaluate these two proteins as potential biomarkers in AD diagnosis. In-house synthesized 5-plex isotopic N,N-dimethyl leucine (iDiLeu) tags were used to label target peptide standards at different concentrations to construct standard curves. Our study revealed that the levels of CLU and APOE exhibited clear differences in male vs. female AD groups but not in male vs. female non-AD groups. In contrast, the levels of serum CLU and APOE did not show statistically significant differences in the AD groups and non-AD groups. Principal component analysis (PCA) with CLU and APOE showed some separation between the AD and non-AD participants. Significance: Dissecting CLU and APOE heterogeneity in AD pathogenesis may therefore facilitate delineating the pathological relevance for sex-related pathways, leading to personalized medicine in the future. Collectively, this study introduces a cost-effective absolute quantitative proteomics strategy for target protein quantitation and lays the foundation for future investigation of CLU and APOE as potential biomarkers for AD. *Significance statement:* As blood-based biomarkers for AD diagnosis are cost-effective and introduce less invasiveness, discovery and validation of biomarkers in the blood samples of AD patients have become a hot topic in Alzheimer's and dementia research. Thus far, amyloid β (Aβ), total-tau and phosphorylated tau (p-tau) in blood show great accuracy and specificity in diagnosis of AD. However, the underlying mechanism of AD pathology remains to be elusive and complex. Besides these well studied proteins, many other proteins, such as clusterin (CLU) and apolipoprotein E (APOE) have also been found to be related to AD development. It

has been implicated that these two proteins are involved in Aβ clearance and deposition. In this study, we measure the absolute concentrations of these two proteins in blood and shed some light on the potential roles of CLU and APOE in AD pathology. Dissecting CLU and APOE heterogeneity in AD pathogenesis may therefore facilitate delineating the pathological relevance for specific pathways between different genders, leading to personalized medicine in the future. Collectively, this study introduces a cost-effective absolute quantitative proteomics strategy for target protein quantitation and lays the foundation for future investigation of CLU and APOE as potential biomarkers for AD.

Introduction

Alzheimer's disease (AD) is a common form of dementia among the elders, which is characterized by the formation of amyloid-β (Aβ) deposition and aggregation of protein tau in the brain.[1] In the past decades, tremendous efforts have been made to understand underlying and validate pathogenesis, identify protective and pathogenic genes, pathogenesis, identify protective and pathogenic genes, and discover and validate cerebrospinal fluid (CSF) and blood-based biomarkers via brain imaging and other techniques. Medical temporal lobe atrophy on magnetic resonance imaging (MRI) and posterior cingulate and temporoparietal hypometabolism on 18fluorodeoxyglucose (18FDG)-positron emission tomography (PET) and cortical Aβ deposition on amyloid- PET imaging are well-validated neuroimaging biomarkers for AD diagnosis.[2] In addition, CSF biomarkers including amyloid β1–42 (Aβ42), amyloid β1–40 (Aβ40), phosphorylated tau181 (p-tau181), p-tau217 and total tau (t-tau) are well-established.[3,4] Compared to imaging biomarkers and CSF-based biomarkers, blood-based biomarkers enable easier access and can reduce the cost and invasiveness in clinical practice. Therefore, blood sample is an appealing resource to discover and validate biomarkers for AD. Thus far, plasma Aβ 42/40 ratio

(Aβ42/40) has been reported to show reduction in AD and can be sensitively measured as biomarker.[5–7] Also, recent studies demonstrated excellent performance of phosphorylated tau 181 and 217 in plasma as diagnostic biomarkers for AD versus other dementias.[8–10] The encouraging and rapid developments in blood-based assays is promising for prescreening in preclinical practice with reduced cost and invasiveness, and once validated, for clinical diagnosis.[11] In addition to Aβ and tau, recent studies in deep proteome profiling of serum, CSF and cortex in AD discovered numerous differentially expressed proteins which need further verifications before clinical practice.[12,13] Among the AD-related proteins, clusterin (CLU) and apolipoprotein E (APOE) play important roles in Aβ clearance and deposition.[14–17]

CLU is a 78–80 kDa heterodimeric glycoprotein that is highly expressed in the central nervous system (CNS) and functions as an extracellular chaperone that can facilitate the transport of Aβ across the blood-brain barrier.[18–20] It has been suggested that CLU can bind to Aβ oligomers to interfere with peptide aggregation and inhibit the growth of extracellular amyloid fibrils.[14–16] There is also significant positive correlation between the concentration of CLU and the regional levels of insoluble Aβ42 in the brain.[14–16] APOE is a 299-amino acid protein with a molecular mass of ~34 kDa that serves important functions in lipoprotein-mediated lipid transportation among organs.[21] The peripheral pool of APOE is mainly expressed by the liver, and in the CNS, APOE is mainly produced by astrocytes and microglia.[22,23] APOE2, APOE3, and APOE4 are the three major APOE isoforms encoded by the Apolipoprotein E2, E3, and E4 alleles, respectively.[22] In general, APOE3 is the most common variant, and APOE4 variant is associated with higher risk for AD, while APOE2 is protective.[22,24] Previous study showed that APOE may negatively affect clearance of soluble Aβ and Aβ deposition at different stages of plaque formation.[17]

However, comparing APOE concentration in patients at different developmental stages of AD showed different and unconvincing results in CSF and plasma samples in different studies.[25–28] Regardless of the APOE genotype, previous studies in our lab observed significant alterations of total APOE levels in CSF of AD patients compared to healthy control.[29,30] As both CLU and APOE are involved in Aβ plaque formation, here, we focused on evaluating the total concentrations of these two proteins in blood samples and to validate their capability to serve as potential biomarkers respectively and together in AD. Enzyme-linked immunosorbent assay (ELISA) is commonly used for measuring protein expression levels in blood and clinical biomarker verification, but the requirement of high-quality antibodies and difficulties when it comes to developing multiplexed assays compromise its applicability to verify novel protein biomarkers.[31] By introducing stable isotope amino acids into proteins metabolically, amino acid coded mass tagging (AACT) or stable isotope labeling by amino acids (SILAC) can be used for large-scale relative quantification of cellular proteins.[32,33] However, this strategy usually is applicable to cell culture and bacteria but not for blood samples. In contrast, absolute quantification (AQUA) strategies using synthetic heavy isotope-encoded peptides as internal standards can be applied to quantifying native peptides in the blood samples.[34] However, the high cost of the isotopically labeled internal peptides limits its usage for verification of a large number of candidate biomarkers. Alternatively, commercially available mass differential tags can be used for relative and absolute quantification such as mass differential tags for relative and absolute quantification (mTRAQ) tag, and iTRAQ.[35–38] Meanwhile, combined precursor isotopic labeling and isobaric tagging (cPILOT) was developed for improved multiplex sample tagging.[39] Previous studies in our lab show that the 5-plex isotopic N,N-dimethyl leucine (iDiLeu) tags are cost-effective and enable improved quantification accuracy.[40] iDiLeu-labeling strategy enables construction of a 4-

point calibration curve and has been adopted previously to quantify analytes in CSF samples.[29] In this study, we applied a 5-plex iDiLeu-labeling strategy to perform absolute quantification for APOE and CLU in serum samples of AD and non-AD participants and investigate their potential to serve as blood-based biomarkers for AD.

Materials and Methods

Serum samples.

The samples were from patients who sought medical advice because of cognitive impairment. CSF samples were collected by lumbar puncture. Blood was sampled by venipuncture at the same time as the CSF and collected in gel separator tubes. Following clotting for 30 min, the sample was centrifuged at 2200 ×g for 10 min with the temperature set at +20◦C. The supernatant was collected, aliquoted and stored at - 80◦C until analysis. Patients were designated as AD and non-AD according to CSF biomarker levels, generated using INNOTEST enzyme-linked immunosorbent assays (Fujirebio, Ghent, Belgium), using cutoffs that are >90% specific for AD: total-tau (T-tau) >350 pg/mL, phospho-tau 181 (P-tau181) >60 pg/mL and Aβ42 < 530 pg/mL (2 out of 3 positive).[41] None of the biochemically normal subjects fulfilled these criteria. The study was approved by the regional ethics committee at the University of Gothenburg.

Serum protein digestion.

Ten μL of serum was diluted in 190 μL water and protein concentrations were measured by protein assay kit (Thermo Scientific Pierce) following the manufacturer's protocols. Serum proteins were dissolved in 6 M urea, 50 mM tris buffer (contains 5 mM CaCl2, 20 mM NaCl, and 1 tablet of EDTA-free protease inhibitor cocktail, pH 8), reduced by dithiothreitol at a final concentration of 5 mM for 1 h. Fifteen mM of iodoacetamide was added for alkylation of cysteines by incubating for 30 min in dark, followed by addition of 5 mM dithiothreitol for quenching

unreacted iodoacetamide. Then the protein mixture was digested with trypsin/LysC (Thermo Scientific, Rockford, IL, United States) at a protein:enzyme ratio of 25:1 at 37◦C for 3 h. The sample was then diluted by 50 mM Tris buffer to a final urea concentration of 1 M, followed by incubation at 37◦C for 15 h. The digestion reaction was quenched by acidification with 10% trifluoroacetic acid (TFA) to a final concentration of 0.3% (pH 3), followed by desalting with Sep-Pak C18 cartridges (Waters Corporation, Milford, Massachusetts, MA, United States). Then the digested peptides were dried down with a SpeedVac concentrator (Thermo Scientific, Waltham, MA, United States) and resuspended in 0.1% formic acid (FA). Peptide concentration was measured by peptide assay (Thermo Scientific, Rockford, IL, United States). Then peptide digestion was aliquot at 50 μg in each tube and was dried in vacuo for further use.

Peptide and protein standards preparation.

Stock solutions of synthetic human peptide (Biomatik, Ontario, Canada) ASSIIDELFQDR (unique peptide of CLU), were prepared at 1 mM. Recombinant human APOE3 (Cat. No. SRP4696) from Sigma- Aldrich (Missouri, USA) protein standard was digested, and the concentrations of digested peptides were measured according to procedures described above. Based on the molecular weight of APOE and concentrations of digested peptides, we could calculate the molar amount. iDiLeu reagents were in-house synthesized as previously described.[40,42] In each peptide aliquot, 4 nmol of CLU peptide standard (about 4 μg) was combined with 400 pmol APOE protein digests (about 14.5 μg), and the mixture was resuspended in 10 μL of 0.5 M triethylammonium bicarbonate (TEAB) buffer. Each mixture was labeled by each iDiLeu tag (d0, d3, d6, d9, and d12) at a ratio of 10:1 (tags-peptide, by weight) or higher with excessive tags. A volume of 40 μL of activated iDiLeu tags was added to each aliquot to 80:20 organic

aqueous solution ratio and shake for 1.5 h at room temperature, respectively. Hydroxylamine (50%) was added to a final concentration of 0.25% to quench the reaction.

Unreacted reagents and by-products of iDiLeu-labeling reaction were removed by SCX TopTips (PolyLC INC., Columbia, MD, United States) according to the manufacturer's protocol. The eluate was dried in vacuo and desalted with Bond Elut OMIX C18 pipette tips (Agilent Technologies, Santa Clara, CA, United States).

Serum protein digests were reconstituted in 10 μL of 0.5 M TEAB and labeled with d0 iDiLeu tag separately according to the peptide standards labeling procedure described above. The other four tag-labeled peptide standards were spiked into each d0-labeled serum sample in a ratio of 1:5:25:100. The combined sample was cleaned up with SCX TopTips and desalted with Bond Elut OMIX C18 pipette tips. For correction factor calculation, four μg of each labeled peptide standard was cleaned up with SCX TopTips and desalted with Bond Elut OMIX C18 pipette tips. All the labeled samples were then dried in vacuo and reconstituted in 50 μL of 0.1% formic acid (FA) in water.

NanoLC-MS acquisition.

LC separation was performed on a Dionex Ultimate 3000 nanoLC system (Thermo Scientific). Capillary column (15 cm length, 75 μm i.d.) was self-fabricated and packed with Bridged Ethylene Hybrid C18 materials (1.7 μm, 130 Å, Waters Corporation). Samples were loaded onto the column in 97% solvent A (water, 0.1% FA) at a flow rate of 0.3 μL/ min. Reverse-phase separation was performed using a linear gradient from 3% to 35% solvent B (ACN, 0.1% FA) for 90 min. Peptide elution was electrosprayed into an Orbitrap Fusion Lumos Tribrid quadrupole ion trap-Orbitrap mass spectrometer (Thermo Scientific). Full MS scan was acquired in profile mode ranging from *m/z* 350 to 950 at a resolution of 60 K. Automatic gain control (AGC)

target was set as 1×10^6, and maximum injection time was 120 ms. The top 20 most abundant precursor ions were selected for stepped higher-energy collisional dissociation fragmentation with a dynamic exclusion for 40 s with 10 ppm. MS2 spectra were acquired at centroid mode. The resolution was set as 30 K, isolation window of 1.6 Th, stepped normalized collision energy (NCE) of 27, 30 and 33, the maximum injection time of 60 ms, AGC target of 1×10^5, and fixed first mass of *m/z* 110. Each sample was acquired in three technical replicates.

Data analysis.

Protein identification was performed using MaxQuant (version 1.5.2.8) against SwissProt human database (February 2020) with 1% false discovery rate (FDR) at peptide and protein level. The first search peptide tolerance for precursor and product ion were 20 ppm and 0.02 Da, respectively. The maximum missed cleavages per peptide were set at 2. Fixed modification was set as carbamidomethylation of cysteine residues (+57.0215 Da) and d0 labeling of N-terminal of peptides and lysine residues (+141.1154 Da). Oxidation of methionine (+15.9949 Da) was selected as variable modification. The intensity obtained from relative quantification in MaxQuant was further analyzed in Excel for advanced downstream analysis. For absolute quantification of CLU and APOE, protein identification was performed using the Peaks Studio Xpro software (Bioinformatics Solutions, Inc., Waterloo, ON, Canada). The data refinement was used to adjust precursor mass by default. All the raw files were searched with UniProt Homo sapiens reviewed database with trypsin as digestion enzyme. The mass error tolerance for precursor ions was 15 ppm using monoisotopic mass and 0.02 Da for product ions. The maximum missed cleavages per peptide were set as two, allowed to be cleaved at both ends of the peptides. Fixed modification was set as carbamidomethylation of cysteine residues (57.0215 Da), iDiLeu labels (+141.1154 Da, +144.1313 Da, +147.1409 Da, +150.1631 Da, and + 153.1644 Da for d0, d3, d6, d9 and d12 tag,

respectively) of peptide N-termini and lysine residues, along with oxidation of methionine (+15.9949 Da) chosen as variable modifications. Peptides with FDR < 1% were considered as unambiguous identifications. Peak areas generated by Genesis peak detection algorithm in Thermo Xcalibur 4.0 software were used for absolute quantification. The precursor ion integration tolerance was 0.015 Da. Retention time of extracted ion chromatogram of 5-plex iDiLeu-labeled peptides was required to be within 2 min. Isotopic interference correction factors were applied to each sample to correct the raw values. A two-sided Student's t-test was used to evaluate the differences of the protein expression levels in the serum samples. A p-value of 0.05 or less is the cutoff for significance.

Results and Discussion

Workflow of a 5-plex iDiLeu-labeling strategy to perform absolute quantification for serum proteins.

A schematic illustration of the workflow used in this study is shown in **Figure 1**. Briefly, serum proteins extracted from male or female AD patients or non-AD cohort were subjected to trypsin digestion. The resultant serum tryptic peptides were labeled with d0 channel of 5-plex iDiLeu tags. Target tryptic peptide standards were labeled with d3, d6, d9, and d12 channels, respectively. The four channels labeled peptide standards were spiked into each serum sample with known molar ratio to enable creating a 4-point calibration curve that allow determination of absolute concentration of target surrogate peptide in serum samples.

Patient demographics.

Ten patients with AD (5 men and 5 women), and 9 healthy non-AD individuals (5 men and 4 women) were included in this study. **Table 1** summarizes the demographic characteristics of the participants. Statistically significant differences ($p < 0.05$) in CSF levels of T-tau, p-tau181 and

Aβ1–42 concentrations were found between AD patients and non- AD individuals. No statistically significant differences ($p > 0.05$) in age were found between AD patients and non-AD individuals (**Supplementary Table S1**).

Quantitative performance of 5-Plex iDiLeu-labeling strategy.

First, label-free and d0 tag-labeled serum protein digests were tested in LC-MS/MS, where CLU and APOE were identified with confidence. A mixture of 5-plex iDiLeu-labeled peptide standard/protein standard digests were then used to evaluate the performance of the iDiLeu-labeling strategy. Impurities of starting materials for the synthesis of iDiLeu tags can compromise the quantification accuracy. When 5-plex iDiLeu reagents are employed to label peptides, mass additions of 141.1154, 144.1313, 147.1409, 150.1631, and 153.1644 Da are added to peptides by d0, d3, d6, d9, and d12 labels, respectively (**Supplementary Table S2**). The mass of five monoisotopic precursors will be 3 Da or 6 Da apart which will cause isotopic interference. To overcome these pitfalls, we firstly determined the correction factors for each of the 5-plex iDiLeu tags to ensure accurate quantification. The correction factors were calculated following previous literature.[40] To evaluate the linearity of the concentration range used in this iDiLeu labeling strategy, unique peptide standards from CLU and APOE proteins were labeled with 5-plex iDiLeu tags and combined at a ratio of 10:1:5:25:100. Two peptides (ASSIIDELFQDR and AATVGSLAGQPLQER) were selected as internal standards. For serum protein digestion, we used trypsin/Lys-C to achieve high-level of complete and specific proteolytic cleavage of protein samples. Nine unique peptides of CLU were identified. These 9 peptides were subjected to criteria described by a previous study, including no miscleavages, between 7 and 30 amino acids residues, no potential modification sites and no ragged end (C-terminal KK, RR, KR or RK).[43] Out of the 9 specific peptides, 3 peptides met the criteria, including SGSGLVGR, EIQNAVNGVK and

ASSIIDELFQDR. However, only peptide ASSIIDELFQDR existed in all of the six isoforms of CLU based on alignment results from Basic Local Alignment Search Tool (BLAST). A previous study showed that ASSIIDELFQDR as surrogate worked well for absolute quantification of CLU in human brain.[44] Therefore, we chose ASSIIDELFQDR as surrogate for CLU. Similarly, we selected peptide AATVGSLAGQPLQER as surrogate for APOE. While several other peptides of APOE also met the criteria, the linearity of calibration curves generated were not as good as peptide AATVGSLAGQPLQER, possibly because of interferences from other peptides with similar molecular weight being eluted at the same time. And some peptides with relatively low ionization efficiency and signal intensity had missing value for d0-labeled peptide standard. Therefore, they were not chosen as surrogates. Surrogate peptide AATVGSLAGQPLQER chosen for APOE was also used in previous study for absolute quantification of APOE in CSF samples.[29] Furthermore, we did not detect miscleavages of the cleavage sites in these two peptides, consistent with previous studies showing that the cleavage sites where glutamic acid (E) or aspartic acid (D) residues located immediately to the N-terminal side of arginine (R) can be efficiently cleaved by Lys-C although they are prone to be miscleaved by trypsin.[45,46] Taken together, these peptides could serve as good surrogates for these proteins. The concentration of stock solution of iDiLeu-labeled CLU peptide (ASSIIDELFQDR) ranged from 50–5000 fmol/μL and then diluted to prepare concentrations of 10–1000 fmol/μL and 2–200 fmol/μL. For the APOE peptides, the concentration of stock solution of iDiLeu-labeled APOE peptide (AATVGSLAGQPLQER) ranged from 5–500 fmol/μL and then diluted twice to prepare concentrations of 1–100 fmol/μL and 0.2–20 fmol/μL. Based on literature, these concentration ranges should cover all the reported concentrations of these two proteins in human serum samples.[28,47,48] Figure 2 displayed normalized peak areas corresponding to the concentrations of iDiLeu-labeled CLU peptides and APOE

peptides across LC–MS analysis in duplicates by normalization to the d12 channel (the normalized peak area of d12 channel was set as 100). Each calibration curve revealed a good correlation with coefficient above 0.99 and a linear concentration range spanning two orders of magnitude. Compared to those which correction factors were not applied (**Supplementary Figure S1**), we could find an improvement of the linearity of the calibration curves in **Figure 2**. For the concentration range used for quantification of CLU in serum, the coefficient of variation (CV) was below 6.5% in the range of 10–1000 fmol/μL. The CV was below 5% in the concentration of 1–100 fmol/μL for APOE. The excellent linearity in the concentration range we used for quantification confirmed that iDiLeu-labeling is a good strategy for absolute quantification of these two proteins.

Absolute quantification of CLU and APOE in Alzheimer's disease.

Here, we used undepleted serum samples because CLU and APOE have relatively high abundance and can be detected by the sensitive instrument platform. Compared with antibody-based protein depletion strategy, which may lead to co-immunoprecipitation and sample loss, the undepleted serum samples may give a more accurate quantification result with lower variation.[49] Here, we set up the data-dependent acquisition (DDA) to do absolute quantification for two peptides of CLU and APOE using 5-plex iDiLeu reagents in both AD participants and non-AD participants. It is worth noting that, considering its sensitivity and reproducibility, data independent acquisition (DIA) may be a more powerful method for the study of ultra-low abundance proteins or posttranslational modifications (PTMs) of proteins.[50,51] Each serum sample was labeled with d0 tag, and standard peptides with a ratio of 1:5:25:100 were labeled with d3, d6, d9, d12, respectively. The labeled peptide standards were then spiked into each serum sample with equal amount at the same ratio. The final concentration of labeled peptide standard ASSIIDELFQDR ranged from 0.5–

50 pmol/50 μg serum protein digests and the final concentration of labeled APOE digests ranged from 0.05–5 pmol/50 μg serum protein digests. iDiLeu-labeled peptides were fragmented under stepped normalized collision energy (NCE) of 27, 30 and 33. **Figure 3** shows representative tandem mass spectra of d0-labeled ASSIIDELFQDR and AATVGSLAGQPLQER in serum sample. iDiLeu d0 tag was fragmented to produce dimethylated immonium reporter ion at *m/z* of 114.1. Rich b- and y-product ions were matched to specific peptide sequence with false discovery rate (FDR) < 1%, which reflected the statistical significance of the peptide-spectrum match.[52] For the data analysis of each sample in PEAKS, the peptide score (-10lgP) threshold was determined by applying FDR of 1%. A peptide with a higher -10lgP score than the score threshold indicates an FDR lower than 1%. Both peptides in all the samples showed a higher score than the peptide score threshold, indicating high confidence of identification (**Supplementary Table S3**). The plot of normalized peak area of each peptide against concentration displayed high linearity, which ensured the accuracy of protein quantification (**Figure 4A and C**). A coefficient of variation (CV) within 12% was demonstrated across the triplicate runs of each sample, indicating a good technical repeatability within the technical replicates.

The average concentrations of CLU were quantified as 287 and 379 μg/mL in male and female of AD, and 299 and 283 μg/mL in serum of male and female of non-AD, respectively (**Figure 4B**). The average concentrations of APOE were 29 and 53 μg/mL in male and female AD participants, and 31 and 35 μg/mL in male and female of non-AD participants, respectively (**Figure 4D**). Based on these measurements, we found that there were no significant differences of APOE or CLU between AD and non-AD participants. Interestingly, the data showed a higher expression level of both CLU and APOE in the female AD participants compared with male AD participants, with a p-value of 0.034 and 0.033 respectively. This may be because that female and

male are impacted by some of the AD risk factors differently and thus have different expression levels of AD related proteins;[53] clearly, more in-depth and larger-scale analyses will be necessary to draw a firm conclusion. Principal component analysis (PCA) with CLU and APOE showed some separation between the AD and non-AD participants (**Supplementary Figure 2**). Previous studies show that the isoforms of APOE and post-translational modification (PTM) patterns of APOE are related to AD risks.[22,54] CLU levels in the blood/CSF/brain and epigenetic regulation of CLU may also play important roles in AD pathogenesis.[55,56] Recent studies in our lab showed that the glycosylation level of CLU decreased in CSF of AD (X. Zhong et al., unpublished data), and the glycosylation patterns of CLU also exhibited differences between AD and healthy control participants.[56] In a recent study, quantitative proteomics map of 95 PTMs on multiple isoforms of Tau was generated to enable a more comprehensive view of Tau isoforms and related PTMs in AD, which may lead to improved understanding of AD pathology and facilitate rational design of therapeutic agents and diagnostics for AD based on Tau pathology.[57] Therefore, further studies focusing on in-depth characterization of the isoforms and PTMs of CLU and APOE in blood samples may help to better elucidate their intriguing roles in AD pathogenesis and develop potential biomarkers for AD diagnosis or treatment strategies.

Conclusions

In summary, we measured the concentration of CLU and APOE proteins in human serum using custom-developed 5-plex iDiLeu tags. Although we did not find significant differences in these two protein concentrations between the AD participants and non-AD participants in the limited cohort we examined, we observed intriguing differences in the levels of these two proteins between female and male AD participants. If replicated with larger cohort, these results might provide new insights into the underlying pathogenesis of AD related to gender differences that

have been less explored. More in-depth and larger cohort studies are necessary to confirm these intriguing results and may lead to future development of more personalized therapeutic approaches. In this study, we showed the effectiveness and robustness of the iDiLeu labeling strategy in absolute quantification of proteins using Orbitrap MS platforms, and this strategy could also be easily adapted on other instrument platforms such as a time-of-flight (TOF) mass spectrometer. A future evaluation of the accuracy and reproducibility of the iDiLeu-based absolute quantification on TOF MS instruments would be very meaningful, and the advantage of the high speed of TOF MS instrument may enable the quantification of a larger sample cohort and facilitate more conclusive results. We envision this cost-effective isotopic labeling strategy could be broadly applied to protein quantification and biomarker verification across many different types of biofluids and various diseases.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jprot.2022.104507. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE[58] partner repository with the dataset identifier PXD030724.

Acknowledgements

This research was supported in part by the National Institutes of Health (NIH) grants RF1 AG052324, R21AG065728, R01 DK071801, and P41GM108538. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) and Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison. L.L. acknowledges a Vilas Distinguished Achievement Professorship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the

Wisconsin Alumni Research Foundation and University of Wisconsin-Madison School of Pharmacy.

References

1. Blennow, K.; de Leon, M. J.; Zetterberg, H., Alzheimer's disease. *The Lancet* **2006**, 368 (9533), 387-403.

2. Scheltens, P.; De Strooper, B.; Kivipelto, M.; Holstege, H.; Chetelat, G.; Teunissen, C. E.; Cummings, J.; van der Flier, W. M., Alzheimer's disease. *Lancet* **2021**, 397 (10284), 1577-1590.

3. Blennow, K.; Hampel, H.; Weiner, M.; Zetterberg, H., Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nat Rev Neurol* **2010**, 6 (3), 131-44.

4. Barthelemy, N. R.; Bateman, R. J.; Hirtz, C.; Marin, P.; Becher, F.; Sato, C.; Gabelle, A.; Lehmann, S., Cerebrospinal fluid phospho-tau T217 outperforms T181 as a biomarker for the differential diagnosis of Alzheimer's disease and PET amyloid-positive patient identification. *Alzheimers Res Ther* **2020**, 12 (1), 26.

5. Nakamura, A.; Kaneko, N.; Villemagne, V. L.; Kato, T.; Doecke, J.; Dore, V.; Fowler, C.; Li, Q. X.; Martins, R.; Rowe, C.; Tomita, T.; Matsuzaki, K.; Ishii, K.; Ishii, K.; Arahata, Y.; Iwamoto, S.; Ito, K.; Tanaka, K.; Masters, C. L.; Yanagisawa, K., High performance plasma amyloid-beta biomarkers for Alzheimer's disease. *Nature* **2018**, 554 (7691), 249-254.

6. Verberk, I. M. W.; Slot, R. E.; Verfaillie, S. C. J.; Heijst, H.; Prins, N. D.; van Berckel, B. N. M.; Scheltens, P.; Teunissen, C. E.; van der Flier, W. M., Plasma Amyloid as Prescreener for the Earliest Alzheimer Pathological Changes. *Ann Neurol* **2018**, 84 (5), 648-658.

7. Nabers, A.; Perna, L.; Lange, J.; Mons, U.; Schartner, J.; Guldenhaupt, J.; Saum, K. U.; Janelidze, S.; Holleczek, B.; Rujescu, D.; Hansson, O.; Gerwert, K.; Brenner, H., Amyloid blood biomarker detects Alzheimer's disease. *EMBO Mol Med* **2018**, 10 (5).

8. Thijssen, E. H.; La Joie, R.; Wolf, A.; Strom, A.; Wang, P.; Iaccarino, L.; Bourakova, V.; Cobigo, Y.; Heuer, H.; Spina, S.; VandeVrede, L.; Chai, X.; Proctor, N. K.; Airey, D. C.; Shcherbinin, S.; Duggan Evans, C.; Sims, J. R.; Zetterberg, H.; Blennow, K.; Karydas, A. M.; Teunissen, C. E.; Kramer, J. H.; Grinberg, L. T.; Seeley, W. W.; Rosen, H.; Boeve, B. F.; Miller, B. L.; Rabinovici, G. D.; Dage, J. L.; Rojas, J. C.; Boxer, A. L.; Advancing, R.; Treatment for Frontotemporal Lobar Degeneration, i., Diagnostic value of plasma phosphorylated tau181 in Alzheimer's disease and frontotemporal lobar degeneration. *Nat Med* **2020**, 26 (3), 387-397.

9. Janelidze, S.; Mattsson, N.; Palmqvist, S.; Smith, R.; Beach, T. G.; Serrano, G. E.; Chai, X.; Proctor, N. K.; Eichenlaub, U.; Zetterberg, H.; Blennow, K.; Reiman, E. M.; Stomrud, E.; Dage, J. L.; Hansson, O., Plasma P-tau181 in Alzheimer's disease: relationship to other biomarkers, differential diagnosis, neuropathology and longitudinal progression to Alzheimer's dementia. *Nat Med* **2020**, 26 (3), 379-386.

10. Palmqvist, S.; Janelidze, S.; Quiroz, Y. T.; Zetterberg, H.; Lopera, F.; Stomrud, E.; Su, Y.; Chen, Y.; Serrano, G. E.; Leuzy, A.; Mattsson-Carlgren, N.; Strandberg, O.; Smith, R.; Villegas, A.; Sepulveda-Falla, D.; Chai, X.; Proctor, N. K.; Beach, T. G.; Blennow, K.; Dage, J. L.; Reiman, E. M.; Hansson, O., Discriminative Accuracy of Plasma Phospho-tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA*. **2020**;324(8):772-781.

11. Mattke, S.; Cho, S. K.; Bittner, T.; Hlavka, J.; Hanson, M., Blood-based biomarkers for Alzheimer's pathology and the diagnostic process for a disease-modifying treatment: Projecting the impact on the cost and wait times. *Alzheimers Dement (Amst)* **2020**, 12 (1), e12081.

12. Dey, K. K.; Wang, H.; Niu, M.; Bai, B.; Wang, X.; Li, Y.; Cho, J. H.; Tan, H.; Mishra, A.; High, A. A.; Chen, P. C.; Wu, Z.; Beach, T. G.; Peng, J., Deep undepleted human serum

proteome profiling toward biomarker discovery for Alzheimer's disease. *Clin Proteomics* **2019**, 16, 16.

13. Wang, H.; Dey, K. K.; Chen, P. C.; Li, Y.; Niu, M.; Cho, J. H.; Wang, X.; Bai, B.; Jiao, Y.; Chepyala, S. R.; Haroutunian, V.; Zhang, B.; Beach, T. G.; Peng, J., Integrated analysis of ultra-deep proteomes in cortex, cerebrospinal fluid and serum reveals a mitochondrial signature in Alzheimer's disease. *Mol Neurodegener* **2020**, 15 (1), 43.

14. Miners, J. S.; Clarke, P.; Love, S., Clusterin levels are increased in Alzheimer's disease and influence the regional distribution of Abeta. *Brain Pathol* **2017**, 27 (3), 305-313.

15. Hatters, D. M.; Wilson, M. R.; Easterbrook-Smith, S. B.; Howlett, G. J., Suppression of apolipoprotein C-II amyloid formation by the extracellular chaperone, clusterin. *Eur J Biochem* **2002**, 269 (11), 2789-94.

16. Beeg, M.; Stravalaci, M.; Romeo, M.; Carra, A. D.; Cagnotto, A.; Rossi, A.; Diomede, L.; Salmona, M.; Gobbi, M., Clusterin Binds to Abeta1-42 Oligomers with High Affinity and Interferes with Peptide Aggregation by Inhibiting Primary and Secondary Nucleation. *J Biol Chem* **2016**, 291 (13), 6958-66.

17. Nielsen, H. M.; Mulder, S. D.; Belien, J. A.; Musters, R. J.; Eikelenboom, P.; Veerhuis, R., Astrocytic A beta 1-42 uptake is determined by A beta-aggregation state and the presence of amyloid-associated proteins. *Glia* **2010**, 58 (10), 1235-46.

18. Poon, S.; Easterbrook-Smith, S. B.; Rybchyn, M. S.; Carver, J. A.; Wilson, M. R., Clusterin is an ATP-independent chaperone with very broad substrate specificity that stabilizes stressed proteins in a folding-competent state. *Biochemistry* **2000**, 39 (51), 15953-60.

19. Roheim, P. S.; Carey, M.; Forte, T.; Vega, G. L., Apolipoproteins in human cerebrospinal fluid. *Proc Natl Acad Sci U S A* **1979**, 76 (9), 4646-9.

20. Zlokovic, B. V.; Martel, C. L.; Matsubara, E.; McComb, J. G.; Zheng, G.; McCluskey, R. T.; Frangione, B.; Ghiso, J., Glycoprotein 330/megalin: probable role in receptor-mediated transport of apolipoprotein J alone and in a complex with Alzheimer disease amyloid beta at the blood-brain and blood-cerebrospinal fluid barriers. *Proc Natl Acad Sci U S A* **1996**, 93 (9), 4229-34.

21. Mahley, R. W.; Rall, S. C., Jr., Apolipoprotein E: far more than a lipid transport protein. *Annu Rev Genomics Hum Genet* **2000**, 1, 507-37.

22. Belloy, M. E.; Napolioni, V.; Greicius, M. D., A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward. *Neuron* **2019**, 101 (5), 820-838.

23. Elshourbagy, N. A.; Liao, W. S.; Mahley, R. W.; Taylor, J. M., Apolipoprotein E mRNA is abundant in the brain and adrenals, as well as in the liver, and is present in other peripheral tissues of rats and marmosets. *Proc Natl Acad Sci U S A* **1985**, 82 (1), 203-7.

24. Farrer, L. A.; Cupples, L. A.; Haines, J. L.; Hyman, B.; Kukull, W. A.; Mayeux, R.; Myers, R. H.; Pericak-Vance, M. A.; Risch, N.; van Duijn, C. M., Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* **1997**, 278 (16), 1349-56.

25. Song, F.; Poljak, A.; Crawford, J.; Kochan, N. A.; Wen, W.; Cameron, B.; Lux, O.; Brodaty, H.; Mather, K.; Smythe, G. A.; Sachdev, P. S., Plasma apolipoprotein levels are associated with cognitive status and decline in a community cohort of older individuals. *PLoS One* **2012**, 7 (6), e34078.

26. Darreh-Shori, T.; Forsberg, A.; Modiri, N.; Andreasen, N.; Blennow, K.; Kamil, C.; Ahmed, H.; Almkvist, O.; Langstrom, B.; Nordberg, A., Differential levels of apolipoprotein E

and butyrylcholinesterase show strong association with pathological signs of Alzheimer's disease in the brain in vivo. *Neurobiol Aging* **2011**, 32 (12), 2320 e15-32.

27. Simon, R.; Girod, M.; Fonbonne, C.; Salvador, A.; Clement, Y.; Lanteri, P.; Amouyel, P.; Lambert, J. C.; Lemoine, J., Total ApoE and ApoE4 isoform assays in an Alzheimer's disease case-control study by targeted mass spectrometry (n=669): a pilot assay for methionine-containing proteotypic peptides. *Mol Cell Proteomics* **2012**, 11 (11), 1389-403.

28. Martinez-Morillo, E.; Hansson, O.; Atagi, Y.; Bu, G.; Minthon, L.; Diamandis, E. P.; Nielsen, H. M., Total apolipoprotein E levels and specific isoform composition in cerebrospinal fluid and plasma from Alzheimer's disease patients and controls. *Acta Neuropathol* **2014**, 127 (5), 633-43.

29. Zhong, X.; Wang, J.; Carlsson, C.; Okonkwo, O.; Zetterberg, H.; Li, L., A Strategy for Discovery and Verification of Candidate Biomarkers in Cerebrospinal Fluid of Preclinical Alzheimer's Disease. *Front Mol Neurosci* **2018**, 11, 483.

30. Zhong, X.; Yu, Q.; Ma, F.; Frost, D. C.; Lu, L.; Chen, Z.; Zetterberg, H.; Carlsson, C.; Okonkwo, O.; Li, L., HOTMAQ: A Multiplexed Absolute Quantification Method for Targeted Proteomics. *Anal Chem* **2019**, 91 (3), 2112-2119.

31. Huttenhain, R.; Malmstrom, J.; Picotti, P.; Aebersold, R., Perspectives of targeted mass spectrometry for protein biomarker verification. *Curr Opin Chem Biol* **2009**, 13 (5-6), 518-25.

32. Gu, S.; Du, Y.; Chen, J.; Liu, Z.; Bradbury, E. M.; Hu, C. A.; Chen, X., Large-scale quantitative proteomic study of PUMA-induced apoptosis using two-dimensional liquid chromatography-mass spectrometry coupled with amino acid-coded mass tagging. *J Proteome Res* **2004**, 3 (6), 1191-200.

33. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **2002**, 1 (5), 376-86.

34. Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A* **2003**, 100 (12), 6940-5.

35. Robinson, R. A.; Evans, A. R., Enhanced sample multiplexing for nitrotyrosine-modified proteins using combined precursor isotopic labeling and isobaric tagging. *Anal Chem* **2012**, 84 (11), 4677-86.

36. Pan, S. Q.; Gu, S.; Bradbury, E. M.; Chen, X., Single peptide-based protein identification in human proteome through MALDI-TOF MS coupled with amino acids coded mass tagging. *Anal Chem* **2003**, 75 (6), 1316-1324.

37. Kang, U. B.; Yeom, J.; Kim, H.; Lee, C., Quantitative analysis of mTRAQ-labeled proteome using full MS scans. *J Proteome Res* **2010**, 9 (7), 3750-8.

38. Rauniyar, N.; Yates, J. R., 3rd, Isobaric labeling-based relative quantification in shotgun proteomics. *J Proteome Res* **2014**, 13 (12), 5293-309.

39. King, C. D.; Robinson, R. A. S., Evaluating Combined Precursor Isotopic Labeling and Isobaric Tagging Performance on Orbitraps To Study the Peripheral Proteome of Alzheimer's Disease. *Anal Chem* **2020**, 92 (4), 2911-2916.

40. Greer, T.; Lietz, C. B.; Xiang, F.; Li, L., Novel isotopic N,N-dimethyl leucine (iDiLeu) reagents enable absolute quantification of peptides and proteins using a standard curve approach. *J Am Soc Mass Spectrom* **2015**, 26 (1), 107-19.

41. Hansson, O.; Zetterberg, H.; Buchhave, P.; Londos, E.; Blennow, K.; Minthon, L., Association between CSF biomarkers and incipient Alzheimer's disease in patients with mild cognitive impairment: a follow-up study. *The Lancet Neurology* **2006**, 5 (3), 228-234.

42. Xiang, F.; Ye, H.; Chen, R.; Fu, Q.; Li, L., N,N-dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics. *Anal Chem* **2010**, 82 (7), 2817-25.

43. Han, B.; Higgs, R. E., Proteomics: from hypothesis to quantitative assay on a single platform. Guidelines for developing MRM assays using ion trap mass spectrometers. *Brief Funct Genomic Proteomic* **2008**, 7 (5), 340-54.

44. Chen, J.; Wang, M.; Turko, I. V., Mass spectrometry quantification of clusterin in the human brain. *Mol Neurodegener* **2012**, 7, 41.

45. Gershon, P. D., Cleaved and missed sites for trypsin, lys-C, and lys-N can be predicted with high confidence on the basis of sequence context. *J Proteome Res* **2014**, 13 (2), 702-9.

46. Slechtova, T.; Gilar, M.; Kalikova, K.; Tesarova, E., Insight into Trypsin Miscleavage: Comparison of Kinetic Constants of Problematic Peptide Sequences. *Anal Chem* **2015**, 87 (15), 7636-43.

47. Martinez-Morillo, E.; Nielsen, H. M.; Batruch, I.; Drabovich, A. P.; Begcevic, I.; Lopez, M. F.; Minthon, L.; Bu, G.; Mattsson, N.; Portelius, E.; Hansson, O.; Diamandis, E. P., Assessment of peptide chemical modifications on the development of an accurate and precise multiplex selected reaction monitoring assay for apolipoprotein e isoforms. *J Proteome Res* **2014**, 13 (2), 1077-87.

48. Yang, C.; Wang, H.; Li, C.; Niu, H.; Luo, S.; Guo, X., Association between clusterin concentration and dementia: a systematic review and meta-analysis. *Metab Brain Dis* **2019**, 34 (1), 129-140.

49. Tu, C.; Rudnick, P. A.; Martinez, M. Y.; Cheek, K. L.; Stein, S. E.; Slebos, R. J.; Liebler, D. C., Depletion of abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res* **2010**, 9 (10), 4982-91.

50. Searle, B. C.; Pino, L. K.; Egertson, J. D.; Ting, Y. S.; Lawrence, R. T.; MacLean, B. X.; Villen, J.; MacCoss, M. J., Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nat Commun* **2018**, 9 (1), 5128.

51. Bern, M.; Finney, G.; Hoopmann, M. R.; Merrihew, G.; Toth, M. J.; MacCoss, M. J., Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. *Anal Chem* **2010**, 82 (3), 833-41.

52. Zhang, J.; Xin, L.; Shan, B.; Chen, W.; Xie, M.; Yuen, D.; Zhang, W.; Zhang, Z.; Lajoie, G. A.; Ma, B., PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* **2012**, 11 (4), M111 010587.

53. Subramaniapillai, S.; Almey, A.; Natasha Rajah, M.; Einstein, G., Sex and gender differences in cognitive and brain reserve: Implications for Alzheimer's disease in women. *Front Neuroendocrinol* **2021**, 60, 100879.

54. Hu, Y.; Meuret, C.; Go, S.; Yassine, H. N.; Nedelkov, D., Simple and Fast Assay for Apolipoprotein E Phenotyping and Glycotyping: Discovering Isoform-Specific Glycosylation in Plasma and Cerebrospinal Fluid. *J Alzheimers* Dis **2020**, 76 (3), 883-893.

55. Yu, J. T.; Tan, L., The role of clusterin in Alzheimer's disease: pathways, pathogenesis, and therapy. *Mol Neurobiol* **2012**, 45 (2), 314-26.

56. Chen, Z.; Yu, Q.; Yu, Q.; Johnson, J.; Shipman, R.; Zhong, X.; Huang, J.; Asthana, S.; Carlsson, C.; Okonkwo, O.; Li, L., In-depth Site-specific Analysis of N-glycoproteome in Human

Cerebrospinal Fluid and Glycosylation Landscape Changes in Alzheimer's Disease. *Mol Cell Proteomics* **2021**, 20, 100081.

57. Wesseling, H.; Mair, W.; Kumar, M.; Schlaffner, C. N.; Tang, S.; Beerepoot, P.; Fatou, B.; Guise, A. J.; Cheng, L.; Takeda, S.; Muntel, J.; Rotunno, M. S.; Dujardin, S.; Davies, P.; Kosik, K. S.; Miller, B. L.; Berretta, S.; Hedreen, J. C.; Grinberg, L. T.; Seeley, W. W.; Hyman, B. T.; Steen, H.; Steen, J. A., Tau PTM Profiles Identify Patient Heterogeneity and Stages of Alzheimer's Disease. *Cell* **2020**, 183 (6), 1699-1713 e13.

58. Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Perez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaino, J. A., The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* **2019**, 47 (D1), D442-D450.

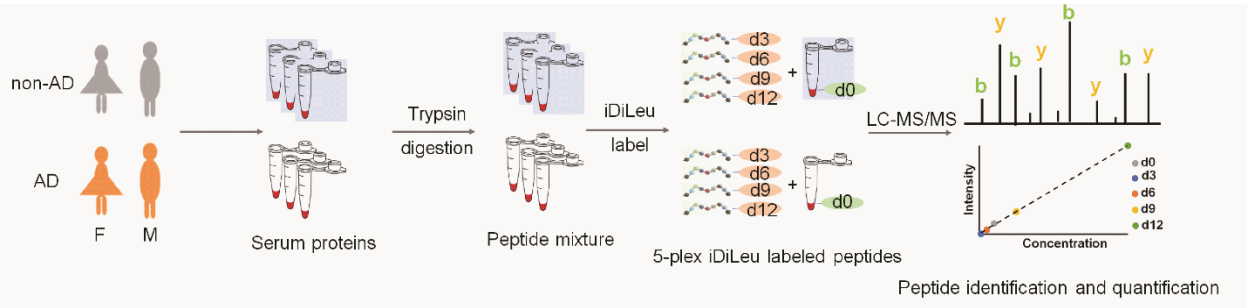


Figure 1. Workflow for absolute protein quantification using iDiLeu-labeling strategy.

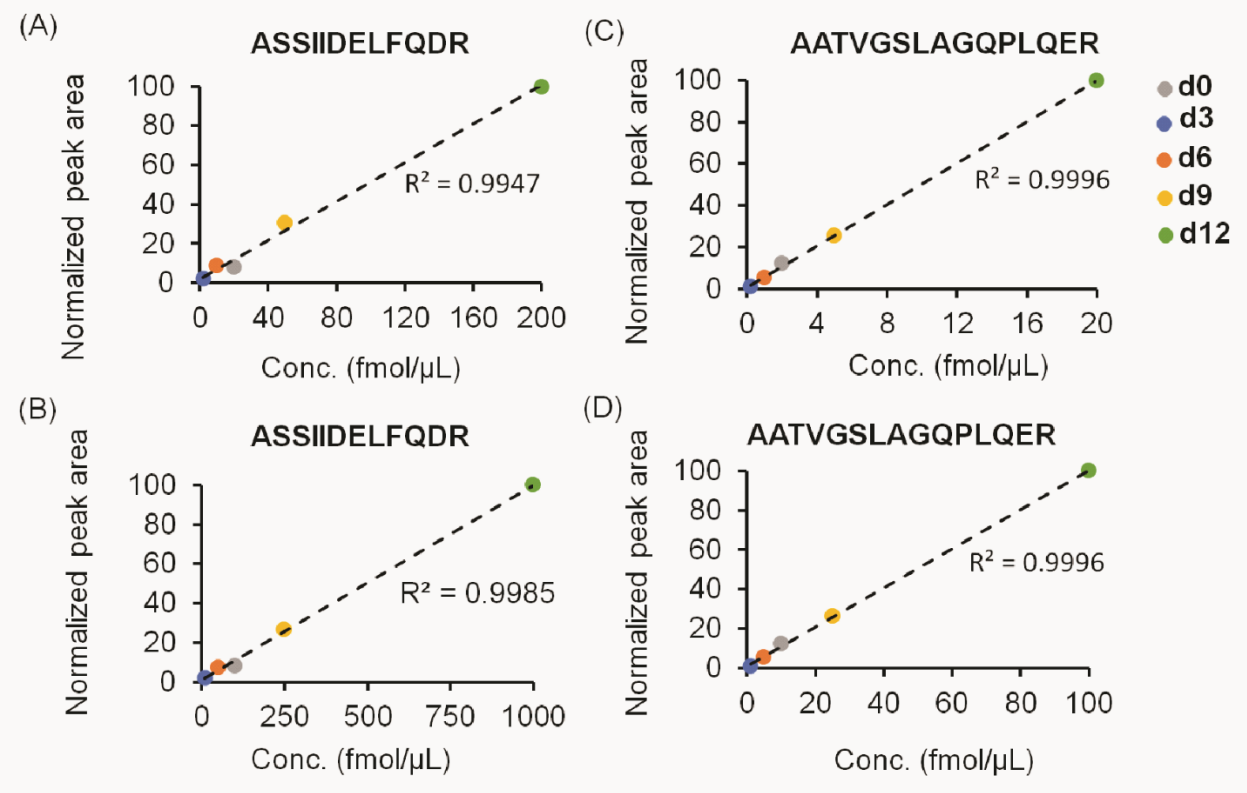


Figure 2. Calibration curves of 5-plex iDiLeu-labeled CLU peptide and APOE peptide. The normalized peak areas of precursor ions were plotted as a function of CLU peptide concentrations at 2–200 fmol/µL (A), 10–1000 fmol/µL (B). The normalized peak areas of precursor ions were plotted as a function of APOE peptide concentrations at 0.20–20 fmol/µL (C), 10–1000 fmol/µL (D). Two µL was injected for each LC-MS/MS run. ASSIIDELFQDR is the peptide of CLU and AATVGSLAGQPLQER is the peptide of APOE. The calibration curves were plotted after correction factors were applied to eliminate the interferences between different tags and isotopic peaks.

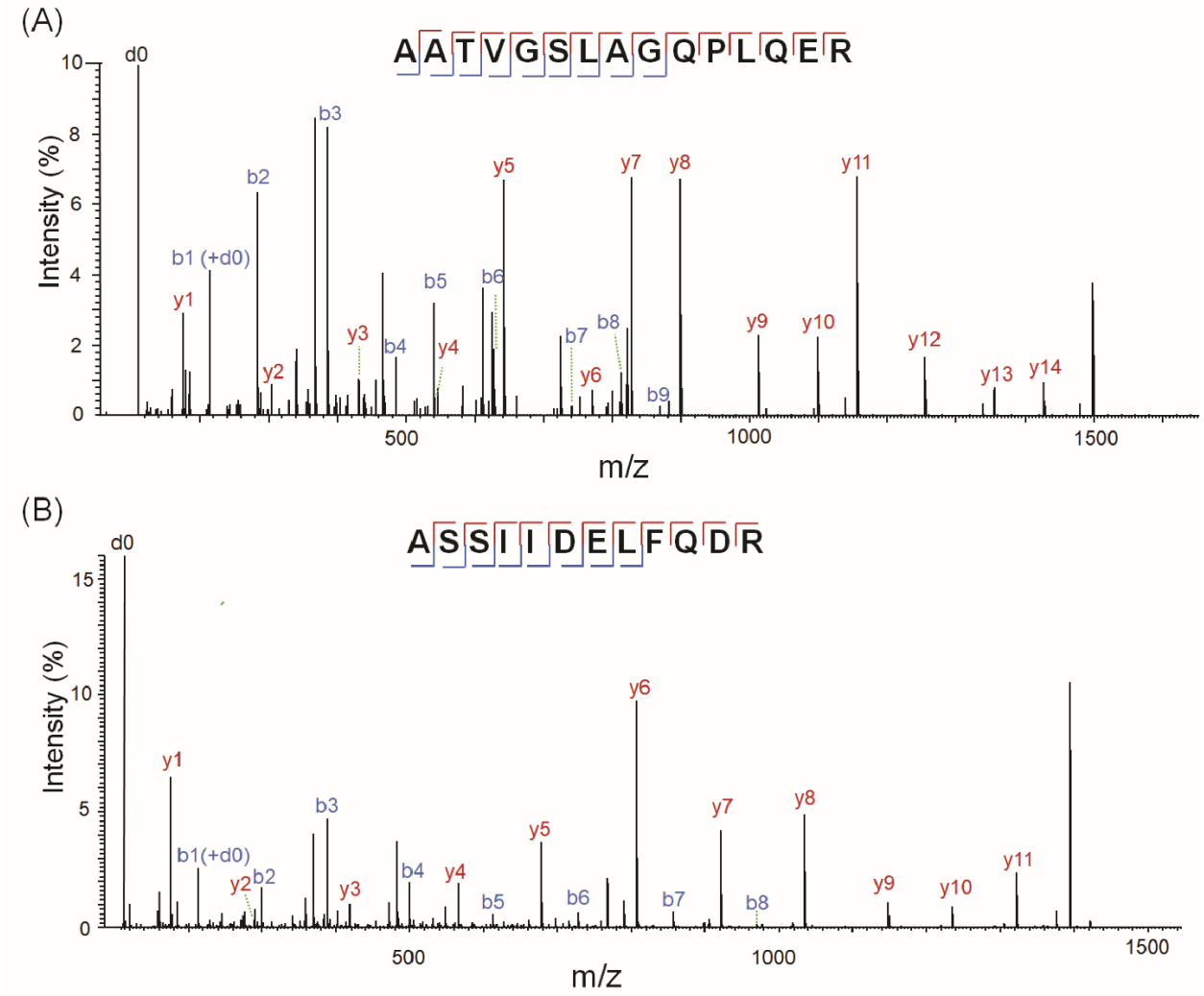


Figure 3. Example tandem mass spectra for iDiLeu d0-labeled CLU peptide (A) and APOE peptide (B) in serum.

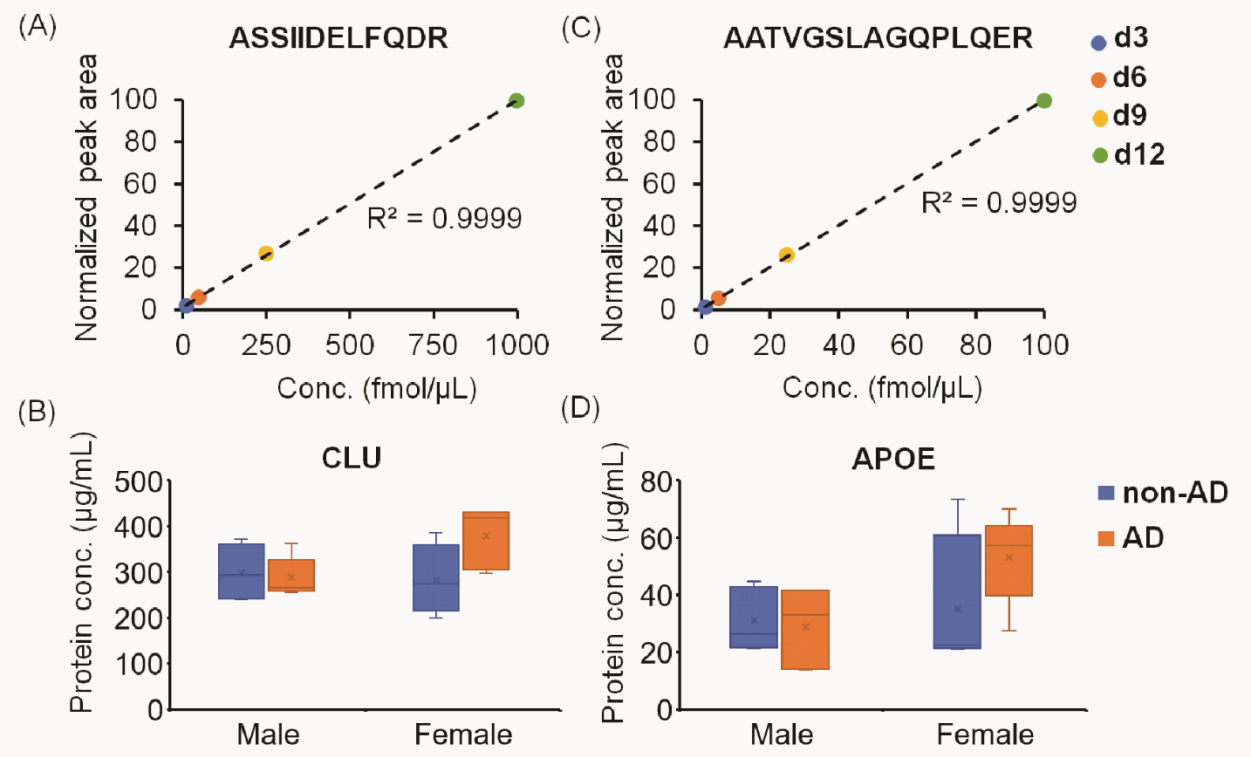


Figure 4. Calibration curves and the determination of absolute concentration of proteins CLU (A, B) and APOE (C, D) in serum samples of male and female in AD and non-AD participants. The boxes represent the 25% quartile, median and 75% quartile and the whiskers show the mean standard deviation (B, D).

Supplemental Information

Supplementary Table S1. More detailed information about the AD/non-AD participants including age, gender, and concentrations of tau, p-tau181, Aβ in CSF.

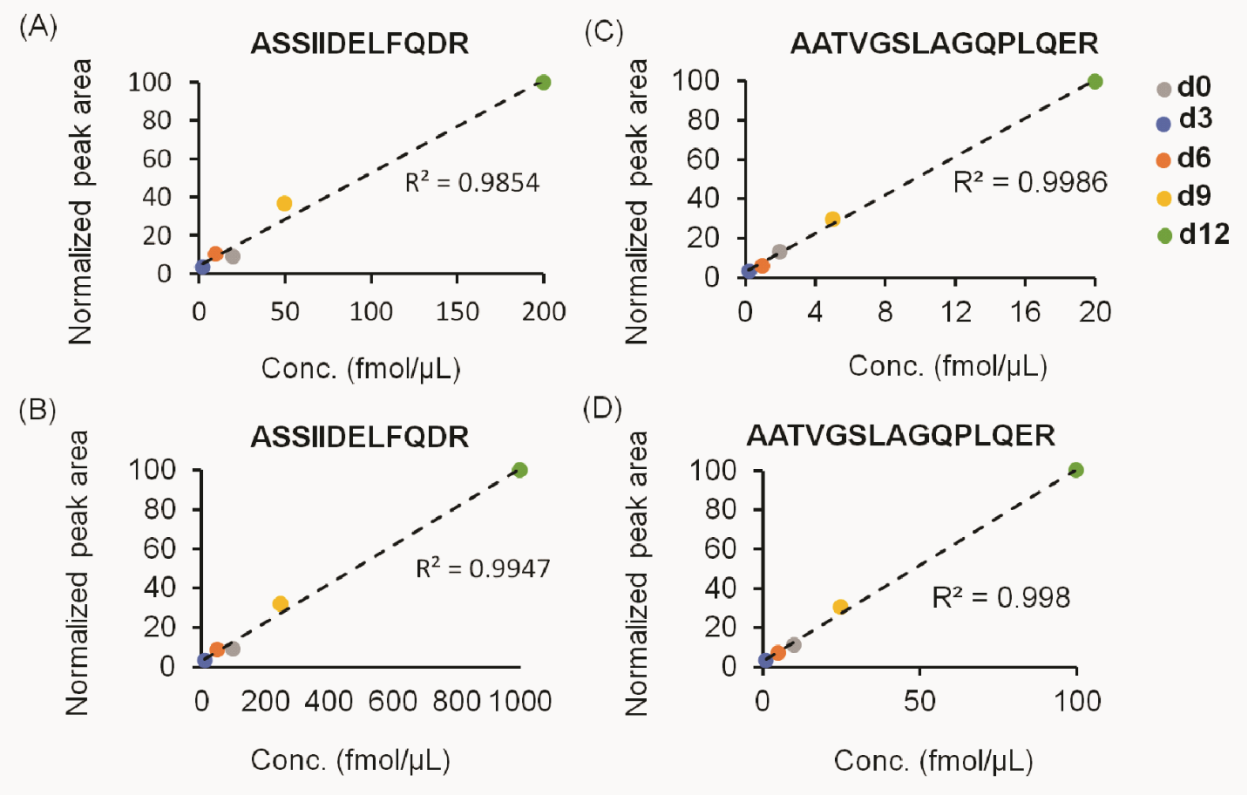
	age	**gender**	**Tau (pg/mL)**	**Aβ42 (pg/mL)**	**p-tau181 (pg/mL)**	**AD stage**
AD1	84	m	667	429	93	AD
AD2	77	m	781	274	94	AD
AD3	75	m	777	382	100	AD
AD4	63	m	881	457	97	AD
AD5	76	m	1160	412	142	AD
AD6	56	f	541	464	80	AD
AD7	73	f	748	483	80	AD
AD8	69	f	1040	264	114	AD
AD9	60	f	642	581	88	AD
AD10	73	f	687	558	80	AD
nonAD1	82	m	386	625	51	non-AD
nonAD2	77	m	192	624	28	non-AD
nonAD3	61	m	248	1010	36	non-AD
nonAD4	65	m	189	931	28	non-AD
nonAD5	61	m	195	860	30	non-AD
nonAD6	76	f	303	877	42	non-AD
nonAD7	65	f	315	703	46	non-AD
nonAD8	67	f	160	876	22	non-AD
nonAD9	72	f	126	603	22	non-AD

Supplementary Table S2. Mass and isotope composition of *N*,*N*-dimethyl leucine (iDiLeu) tags.

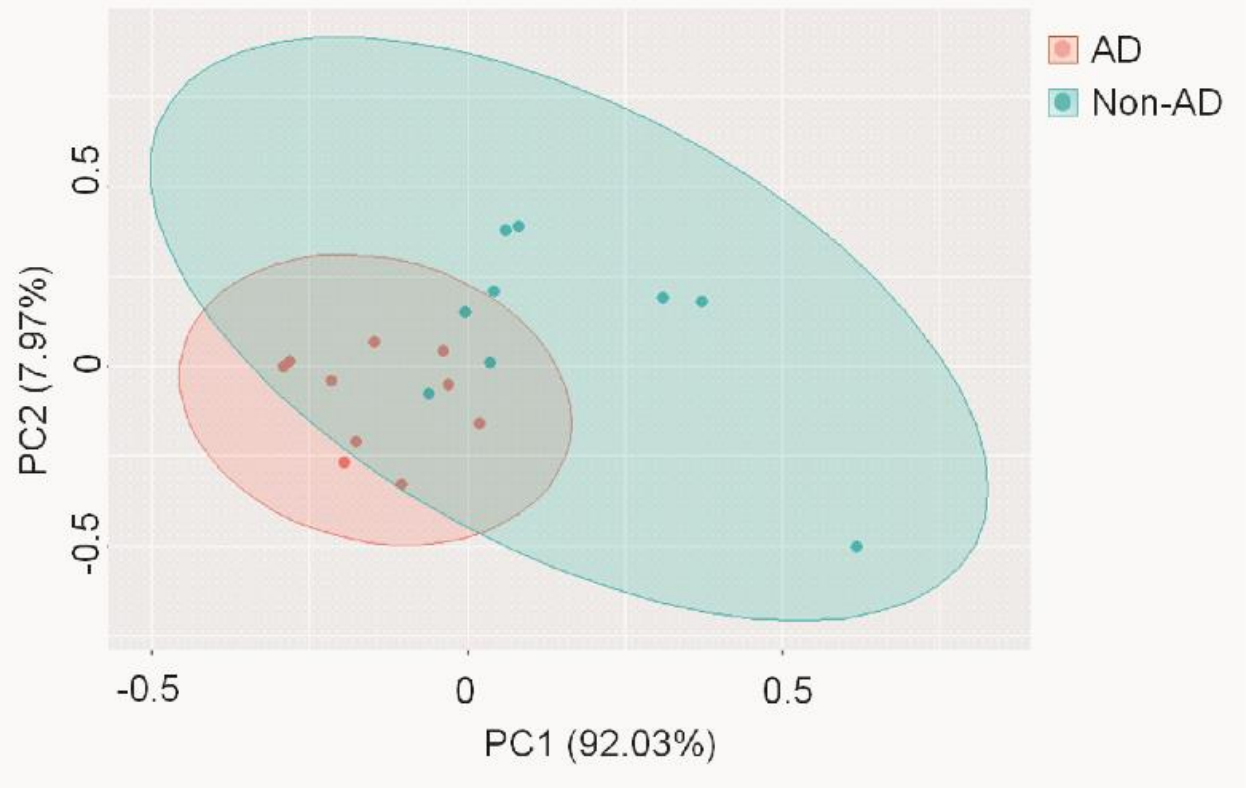
iDiLeu tags	Stable Isotopes				Mass shift (Da)
	^{13}C	^{2}H	^{15}N	^{18}O	
d0	0	0	0	0	141.1154
d3	1	2	0	0	144.1313
d6	1	4	1	0	147.1409
d9	3	6	0	0	150.1631
d12	3	6	1	1	153.1644

Supplementary Table S3. Peptide scores in peptide identification for ASSIIDELFQDR and AATVGSLAGQER in PEAKS data analysis. FDR, false discovery rate.

Sample ID	-10lgP (FDR = 1%)	Score of AATVGSLAGQER	Score of ASSIIDELFQDR
AD-1	25.4	55.85	48.83
AD-2	24.9	61.83	50.96
AD-3	25.2	57.23	49.11
AD-4	39.8	89.68	86.6
AD-5	39.1	93.3	85.68
AD-6	39	94.00	90.16
AD-7	38	90.13	89.88
AD-8	40.5	92.67	88.3
AD-9	38.2	93.87	87.39
AD-10	39.7	93.47	88.62
Non-AD-1	37.7	81.14	82.61
Non-AD-2	41.9	98.58	98.5
Non-AD-3	39.3	87.51	83.39
Non-AD-4	38.8	94.76	85.86
Non-AD-5	39.6	104.35	89.71
Non-AD-6	38.5	94.58	82.35
Non-AD-7	37.8	94.95	82.43
Non-AD-8	39	90.35	78.83
Non-AD-9	38.4	98.32	96.3



Supplementary Figure S1. Calibration curves of 5-plex iDiLeu labeled CLU peptide and APOE peptide before correction factors applied. The normalized peak areas of precursor ions were plotted as a function of CLU peptide concentrations at 2–200 fmol/µL (A), 10–1000 fmol/µL (B). The normalized peak areas of precursor ions were plotted as a function of APOE peptide concentrations at 0.20–20 fmol/µL (C), 10–1000 fmol/µL (D). Two µL was injected for each LC-MS/MS run. ASSIIDELFQDR is the peptide of CLU and AATVGSLAGQPLQER is the peptide of APOE.



Supplementary Figure S2. Principal component analysis from CLU and APOE.

Chapter 6

Quantification of Serum Metabolites in Early Colorectal Adenomas Using Isobaric Labeling Mass Spectrometry

Adapted from:
Yuan Liu, Hua Zhang, William F. Dove, Zicong Wang, Zhijun Zhu, Perry J. Pickhardt, Mark Reichelderfer, and Lingjun Li. Quantification of Serum Metabolites in Early Colorectal Adenomas Using Isobaric Labeling Mass Spectrometry. *J. Proteome Res.* **2023**, 22, 5, 1483–1491

Author contributions:
L.L. designed, and supervised the whole study; Y.L. performed experiments, analyzed, interpreted the data, and wrote the manuscript; H.Z. analyzed and interpreted the data; W.F.D. designed the study, and interpreted the data; Z.W. and Z.Z. performed experiments; P.J.P. and M.R. collected the serum samples. All authors reviewed, revised, and edited the manuscript.

Abstract

A major challenge in reducing the death rate of colorectal cancer is to screen patients using low-invasive testing. A blood test shows a high compliance rate with reduced invasiveness. In this work, a multiplex isobaric tag labeling strategy coupled with mass spectrometry is adopted to quantify primary and secondary amine-containing metabolites in serum for the discovery of metabolite level changes of colorectal cancer. Serum samples from patients at different risk statuses and colorectal cancer growth statuses are studied. Metabolite identification is based on accurate mass matching and/or retention time of labeled metabolite standards. We quantify 40 metabolites across all the serum samples, including 18 metabolites validated with standards. We find significantly decreased levels of threonine and asparagine in the patients with growing adenomas or high-risk adenomas ($p < 0.05$). Glutamine levels decrease in patients with adenomas of unknown growth status or high-risk adenomas. In contrast, arginine levels are elevated in patients with low-risk adenoma. Receiver operating characteristic analysis shows high sensitivity and specificity of these metabolites for detecting growing adenomas. Based on these results, we conclude that a few metabolites identified here might contribute to distinguishing colorectal patients with growing adenomas from normal individuals and patients with unknown growth status of adenomas.

Introduction

Colorectal cancer is the third most diagnosed cancer with an incidence of 10.0% and it is ranked second in mortality (9.4%) in cancer-caused death.[1] Early diagnosis and treatment are critical for reducing the mortality of colorectal cancer. Currently, optical colonoscopy (OC) is the gold standard for the detection of colonic lesions, and computed tomographic colonography (CTC)

is also considered an acceptable alternative.[2] However, these two methods are inaccessible to isolated populations and can be expensive for screening large populations. The identification of low-invasiveness biomarkers for early colorectal cancer detection is challenging. To improve compliance and reduce the invasiveness of detection, stool-based tests have been developed.[3] Although they perform well for frank colorectal cancer, these methods can only detect 11% to 42% of the advanced colonic adenoma.[3]

For colorectal polyps, polypectomy in the early stage of this disease is efficient for long-term prevention.4 Based on longitudinal analysis of colorectal polyps in patients by CTC, it is found that growing adenomas are prone to develop into high-risk adenomas and then become colorectal cancer.[5−7] However, only about one-fourth of (6−9 mm) polyps grow and most colorectal polyps stop growing or regress.6,8 Therefore, overdiagnosis should be avoided to achieve improvements in the prevention of colorectal cancer. Taken together, the longitudinally monitored CTC cohort provides support to determine the degree of correlation of changed levels of serum components with the growing adenoma. In addition, samples from colonoscopy/cancer-free cases classified as screening normal were used as the control. Previous studies reported significant changes in the level of several serum proteins in patients carrying premalignant colonic adenomas.[9,10] Meanwhile, efforts on discovering metabolite biomarkers have been made and a few potential metabolite biomarkers have been reported in blood samples.[11−13, 29] However, metabolite biomarkers associated with early growing colonic adenomas are rarely studied. Mass spectrometry (MS)-based quantification is useful in metabolomics investigations. Typically, by spiking an isotopic analog of the analyte as an internal standard, absolute quantification can be achieved.[14] However, suitable isotopic internal standards for some targeted metabolites may not be available and the cost can be high if many targeted analytes need to be quantified. Alternatively, relative

quantification can be achieved by label-free quantification and chemical derivatization approaches. Compared to label-free quantification, stable isotope labeling is more accurate and offers a high throughput capability for relative quantification. The stable isotope labeling strategy for quantification can be classified as MS1 level-based (mass difference or mass defect) quantification and tandem mass spectrometry (MS/MS) reporter ion-based quantification. Mass difference labeling like mass differential tags for relative and absolute quantification (mTRAQ),[15] formaldehyde dime- thylation,[16] and dansylation[17] could increase mass spectrometry complexity. Strategies like hyperplexing isobaric tags and mass defect-based stable isotope labeling can give more clean spectra for quantification but require high-resolution mass spectrometers.[18,19] In contrast, isobaric labeling strategy can incorporate isotopic tags to analytes with nearly the same precursor ion mass shifts, and MS2 fragmentation can produce reporter ions for relative quantification. Isobaric labeling methods, such as iTRAQ,[20] TMT,[21] and isobaric N,N-dimethyl leucine (DiLeu),[22] provide clean MS1 spectra and only require relatively low-resolution mass spectrometers for quantification. Relative quantification of metabolites using an isobaric labeling strategy has been used in many studies and has demonstrated good performance in discovering metabolite biomarkers.[19,22–24] Therefore, this study employs 4-plex DiLeu tags for the relative quantification of secondary and primary amine-containing metabolites in serum samples. To control the intrinsic variance among patients, serum samples were taken during prepolypectomy and postpolypectomy. Meanwhile, serum samples from screening normal individuals were also analyzed as the control.

Materials and Methods

Human subject protocol.

Detailed information on how the CT method for colorectal cancer screening, derivation of in vivo colorectal polyp volume, and how blood specimens were obtained as described in previous publications.[6,9] The Institutional Review Board at the University of Wisconsin−Madison approved this study.

Chemicals and reagents.

Optima LC/MS grade acetonitrile and water, anhydrous N,N- dimethylformamide (DMF), and 4-(4,6-dimethoxy-1,3,5-tria- zin-2-yl)-4-methylmorpholinium tetrafluoroborate (DMTMM) were purchased from Fisher Scientific (Fair Lawn, NJ). N-Methylmorpholine (NMM) was purchased from TCI America (Tokyo, Japan). Twenty-four metabolite standards (glycine, L-alanine, L-cysteine, L-serine, L-proline, L-valine, L-threonine, L- leucine, L-isoleucine, L-aspartic acid, L-asparagine, L-glutamic acid, L-glutamine, L-methionine, L-dopamine, L-lysine, L-histidine, L-phenylalanine, L-arginine, L-tyrosine, L-tryptophan, serotonin, γ-aminobutyric acid, dopamine, kynurenine), and triethylammonium bicarbonate (TEAB) was purchased from Sigma (St. Louis, MO). Formic acid (FA) was purchased from Fluka (Buchs, Switzerland). SCX Ziptips were purchased from Millipore.

DiLeu reagent synthesis and activation.

The synthesis of 4-plex isobaric DiLeu tags was previously described by Frost et al.[25] DiLeu reagents were activated to DiLeu triazine ester immediately before metabolite labeling. Each 1 mg dried DiLeu tag was activated by mixing with 100 μL of activation solution (1.41 mg of DMTMM and 0.47 μL of NMM in anhydrous DMF) and then vortexed at room temperature for 45 min. The mixture was centrifuged at 10 000 ×g for 1 min, and the supernatant was used immediately for metabolite labeling.

Labeling metabolite standards.

A stock solution containing 1 mM of each metabolite standard was prepared and stored at −20 °C before use. A diluted stock solution containing different amounts of each metabolite was dried with SpeedVac concentrator (Thermo Electron Corporation, West Palm Beach, FL), and redissolved in 10 μL of a pH8 0.5 M TEAB solution to constitute metabolites solution with different concentrations in a ratio of 1:1:1:1 (20 μM each) and a ratio of 1:2:4:8 (5 μM, 10 μM, 20 μM, and 40 μM). Each aliquot was mixed with 40 μL of activated DiLeu tag 115, 116, 117, or 118, respectively. For each labeling reaction, activated DiLeu reagents and the metabolite mixture reacted at a molar ratio higher than 10:1 to ensure efficient and complete labeling under room temperature for 2 h with shaking. Ten microliters of 0.5M TEAB without metabolites was also mixed with 4-plex DiLeu and combined as the negative control. The reaction was then quenched by adding hydroxylamine to 0.25% (v/v). Subsequently, labeled samples were each dried in vacuo and combined at equal ratios. Millipore SCX Ziptips (10 μL) were used to remove residual labeling chemicals, with 0.1% FA in water used as the reconstitution and washing solutions, and 5% $NH_3 \cdot H_2O$ in 30% MeOH as the elution solution. The eluate was dried down and stored at −20 °C until analysis.

Human serum sample preparation.

Serum samples were thawed on ice and centrifuged at 2000g for 10 min to remove particulates and debris. Molecular weight cutoff filters (MWCO, 3 kDa, Millipore Amicon Ultra, Burlington, MA) were prerinsed 3 times with optima water at 14 000g for 20 min. Twenty microliters of serum supernatant from each sample was diluted in 380 μL of water and added to the filter then centrifuged for 20 min at 14 000g. Flowthrough was collected. Then, 400 μL of water was added to the filter to rinse the sample. An additional rinse step was applied. All the flowthrough (about 1.2 mL) was combined and dried down in Speedvac and stored at −20 °C until

labeling. For normalization, a pooled serum sample from screening normal individuals was also prepared. For labeling of serum samples, 40 μL of activated DiLeu reagents was mixed with serum flowthrough dissolved in 10 μL of 0.5MTEAB and shaken for 2 h. After the reaction was quenched, 2.5 μL of each reaction mixture with different DiLeu tags was combined at a 1:1:1:1 ratio and mixed well. To compare the relative abundance of each set of 4-plex DiLeu labeled metabolites, one channel from each 4-plex combination was selected and combined with 118-DiLeu tag labeled pooled serum samples. Ten microliters of the mixture was taken for drying down and desalted using SCX Ziptips. Samples were dissolved in 15 μL of 0.1% FA, and 3 μL was injected into LC- MS for data acquisition with 3 technical replicates.

Liquid chromatography coupled with MS/MS data collection.

Four-plex pooled samples were reconstituted in 0.1% formic acid before injection. The HPLC-MS/MS analysis was conducted using a Waters nanoACQUITY UPLC coupled with Thermo Q Exactive Orbitrap MS. The separation column was in-house made with an emitter tip and dimensions of 75 μm inner diameter × 15 cm length. The column was packed with 1.7 μm, 150 Å, ethylene-bridged-hybrid (BEH) C18 material (Waters, Milford, MA). Mobile phase A was water containing 0.1% formic acid, and mobile phase B was acetonitrile containing 0.1% formic acid. The flow rate was set as 0.3 μL/min, and the LC gradient was 55 min and set as follows: 0−10 min, 3%−30% solvent B; 10−30 min, 30−80% B; 30−30.5 min, 80%−95% B; 30.5−40.5 min, 95% B; 40.5−41 min, 95%−3% B; 41−55 min, 3% B. Positive ionization mode was used and full MS scans were acquired from *m/z* 180 to 800 at a resolution of 60 k, automatic gain control (AGC) was set as 5 × 105, and a maximum injection time as set as 30 ms. The top 20 precursors were selected for normalized collision energy (NCE) dissociation (NCE = 30) with an isolation window

of *m/z* 1, fixed first *m/z* 110, dynamic exclusion of 5 s, charge exclusion of >2, and a resolution of 35 k.

Data analysis.

Raw data files were converted into .mgf format files via the msConvert.[26] Metandem was used to process three technical replicates of each 4-plex DiLeu labeled metabolite for quantification.[27] The average precursor mass shift due to labeling was 145.1273 Da. Data analysis parameters were optimized automatically in Metandem software. Output files with reporter ion information were merged. Among all the features detected in DiLeu labeled pooled serum (1:1:1:1 mixed after labeling), only the ratios of reporter ion intensity (116/115 and 117/115 and 118/115) between 0.67 and 1.5 were kept for further analysis. Then, selected features detected in the serum mixture were compared with DiLeu labeled blanks and features that appeared in the blank along with the same retention time were excluded. The remaining features were used for calculating the monoisotopic molecular weight based on mass shift caused by labeling and then searched against the Human Metabolome Database (HMDB). Only putative metabolites with primary and secondary amines were kept, while putative metabolites without the amine group were excluded. To compare the relative abundance of each set of 4-plex DiLeu labeled metabolites, all the intensities of reporter ions were normalized based on the reporter ion intensities of the shared channel for comparison. A Mann−Whitney U test was performed to compare metabolite levels in different groups. A ROC analysis in MedCalc (version 20.118) was performed to evaluate the sensitivity and specificity of biomarkers for detecting growing adenomas compared with adenomas of unknown growth status or compared with normal controls. Leave-one-out cross-validation was performed for different groups (control group vs. adenoma carrier group, control group vs.

growing adenoma group, and growing adenoma group vs. unknown growth group) via R studio (R version 4.2.2).

Results and Discussion

Patient demographics.

We studied 43 patients in total here. Twenty patients screened normal by OC were set as the control. Meanwhile, 23 patients with adenomas discovered by CTC got their blood drawn before and after polypectomy (**Figure 1**). Among the 23 patients carrying adenomas, 15 patients were classified histologically as high-risk and the remaining 8 patients were classified as low-risk cases. In the high-risk cases, 10 were classified as growing by longitudinal CTC analysis, one as static, and 4 as unknown growth status.6,9 In these low-risk cases, 4 were classified as growing, 2 as static, and 2 as unknown status (Figure 1). More details about the patients and screening normal controls can be found in **Table S1** and **Table S2** or previous studies.[9,10]

Quantitative performance of 4-plex labeling strategy.

The structure of theDiLeu isobaric tag is composed of a reporter group, a balance group, and an amine-reactive group (**Figure S1A**). The tag can react with the amine group of metabolites (**Figure S1B**) and introduce a mass shift of 145.1263 or 145.1283 Da (**Figure S1C**). Though 12-plex and even 21-plex DiLeu tags have been developed for relative quantification of peptides and/or amine-containing metabolites,[18,23] we only selected four tags that had the smallest mass value differences (only 0.002 Da). This could facilitate accurate mass matching and enhance confidence in identifying features of those unavailable metabolite standards. It is noted that the metabolite identifications mainly relied on metabolite standards and accurate mass matching. To evaluate the accuracy and reproducibility of relative quantitation of metabolite standards using the 4-plex DiLeu tags, two sets of metabolite standard mixtures with different molar ratios (1:1:1:1

and 1:2:4:8) were labeled by 4-plex DiLeu and combined. Combined labeled metabolites were purified and analyzed by nanoLC ESI-MS/MS analysis. Relative quantification of the metabolite was achieved by calculating the intensity ratios of DiLeu reporter ions (*m/z* 115.1, 116.1, 117.1, 118.1) (**Figure S1C**) generated in MS2 fragmentation. The results are summarized in the box plots in **Figures 2A,B**. To further evaluate the accuracy and reproducibility of the relative quantitation of metabolites in serum samples, two sets of serum flowthrough with different volumes of starting material (20:20:20:20 μL and 5:10:20:40 μL) were labeled by 4-plex DiLeu. The results are summarized in the box plots in **Figures 2C,D**. The median ratios measured among 21 metabolite standards were 1:1.08:0.98:1.03 with a theoretical ratio of 1:1:1:1, and 1:2.52:4.93:7.56 with a theoretical ratio of 1:2:4:8, respectively. The median ratio measured among 40 metabolites in pooled serum were 1:1.08:1.07:1.09 with a theoretical ratio of 1:1:1:1 and 1:2.06:4.01:6.90 with a theoretical ratio of 1:2:4:8, respectively. Both groups for labeling of serum metabolites showed satisfactory accuracy (within 14% error) for relative quantification. This result suggested that the amount of tags used here was suitable for labeling serum samples with starting volumes ranging from 5 to 40 μL. Therefore, the accuracy for relative quantification of metabolites in 20 μL serum was confirmed.

Serum metabolite identification and quantification.

A schematic illustration of the workflow used in this study is shown in **Figure 3**. Briefly, serum metabolites from patients or controls were collected after serum was passed through a 3 kDa MWCO filter. The 4-plex DiLeu tags were used for labeling secondary and primary amine-containing metabolites from each individual and then combined at a ratio of 1:1:1:1. The relative abundance of each metabolite can be determined by comparing the relative intensity of reporter ions. Among all the features detected in the DiLeu labeled pooled serum sample (1:1:1:1 labeled

and combined), we applied the following criteria to filter: 1) the *m/z* values were not shown in DiLeu labeled blank samples or features had different retention time (>1 min) or had the same retention time, but the peak intensity was 100-fold lower than that in the serum samples; 2) the MS2 spectra of the features contained all the four reporter ions; 3) the ratio of reporter ions (116/115, 117/115, 118/115) ranged from 0.67 to 1.5; 4) the calculated mass of the features can match a secondary or primary metabolite in Human Metabolome Database (HMDB) (mass differences <10 ppm). When we calculated the masses of unlabeled metabolites, we subtracted precursor masses by 145.1273 Da (for one-tag- labeled) or 290.2546 (two-tag-labeled), then the masses were searched against a human metabolite database (HMDB). Among hundreds of features detected in the serum samples, a total of 59 passed the filtering criteria. Furthermore, 40 features were detected in all the serum samples, including 18 features validated using metabolite standards (**Table S3**). The 40 metabolites were assessed for statistical significance in detecting colorectal cancer from cancer-free cases. The identified metabolites were compared quantitatively among adenoma-free people ("screening normal"), people carrying adenomas of unknown growth, people carrying growing adenoma, and people carrying adenoma (including both unknown growth and growing adenoma). Among the validated metabolites, threonine, asparagine, and glutamine in patients with high-risk adenomas showed significantly decreased serum levels compared with screening normal cases (p-value <0.05 with Benjaminin−Hochberg to control the false discovery rate q at <0.05) (**Figure 4**). Consistently, threonine and glutamine also showed decreased serum levels in growing adenoma compared with screening normal. Glutamine showed a significant decrease in sera of patients with unknown growth adenoma compared with normal controls (**Figure 4**). Significant level changes of asparagine, glutamine, and arginine were observed in adenoma carriers compared to screening normal controls (**Figure 4**). In contrast,

arginine showed significantly enhanced levels in the sera of patients with low-risk adenoma compared with normal controls (**Figure 4**). Cysteine in the blood is not stable and can be oxidized to form cystine and levels of cystine may be affected during sample preparation. Therefore, cystine was excluded from discussion here. We also looked into the comparison of the ratios of post/pre-polypectomy between different groups, including a comparison between unknown growth and growing adenoma, and a comparison between low-risk and high-risk adenoma, but we did not find significant differences between them.

We then explored whether the levels of these altered metabolites reflected the total adenoma volume in a patient. Previous studies showed that levels of serum protein biomarkers F5, ITIH4, LRG1, and VTN in the ApcPirc$^{/+}$ rat positively correlated with the number of colonic adenomas, while no significant correlation between the total volume of adenomas and the prepolypectomy levels of the four serum markers.[9,28] In this study, we found no significant correlation between the four metabolites (threonine, asparagine, glutamine, arginine) and adenoma volume (**Table S4**) by a Spearman test. Furthermore, we asked if there were any correlations between these metabolites and protein biomarkers identified in previous work (**Table 2** in ref 8).We generated a correlation heatmap for metabolites identified in this study and previously identified protein biomarkers (**Figure 5A**). Interestingly, we found that there was a strong positive correlation between ITIH4 and F5 ($\rho > 0.7$) (**Figure 5B**). A moderate positive correlation was found between CRP and VTN ($\rho > 0.4$) (**Figure 5C**). In contrast, there was a strong negative correlation between LRG1 and asparagine ($\rho < -0.7$) (**Figure 5D**).

We assessed quantitatively the sensitivity and specificity for using these serum metabolic markers to detect growing adenomas compared with normal controls. Receiver operating characteristic (ROC) analysis showed that the nominal area under curve (AUC) values for these

individual markers ranged from 0.829 (glutamine and threonine) to 0.871 (glutamine, threonine, asparagine, and arginine) (**Figure 6A**).When patients with growing adenomas were compared with patients with unknown growth of adenomas, the AUC values of the markers ranged from 0.905 to 0.925 (**Figure 6B**).

The identification of low-invasiveness biomarkers for early colorectal cancer detection is challenging. Although minimally invasive tests for occult fecal blood and tumor-derived DNA have been widely used for the detection of colorectal cancer, performance for detecting advanced colonic adenoma needs to be further improved.[3] To increase the sensitivity and specificity in predicting colorectal cancer and improve the understanding of the pathological mechanism of colorectal cancer development, efforts have been made to discover protein and metabolite biomarkers. Recent studies have reported that potential metabolite biomarkers show dysregulated levels in the blood of colorectal cancer patients compared with healthy controls. However, most of these studies only focus on identifying biomarkers between colorectal cancer and healthy controls.[11-13] One of the studies grouped the patients based on the risk of colorectal cancer precursors (including conventional adenomas and serrated polyps) and found that lipid metabolism and metabolites may be involved in the early stage of serrated pathway-related colorectal carcinogenesis.[13] In this study, we focused on investigating potential serum metabolite biomarkers that can distinguish healthy controls, and adenoma carriers with different growth rates or risks. This study is a follow-up of a previous work aiming at discovering serum protein biomarkers related to growing early colorectal adenomas.[9] A more comprehensive profiling of serum content (including proteins and metabolites and even post-translational modifications of proteins) may enable enhanced detection of colorectal cancer and offer deeper insights into the pathological mechanisms of colorectal cancer. It would be interesting to investigate what power can emerge

from using two or more modalities (metabolomics and proteomics) and then from combining different analytes from the same modality or from different modalities. **Figure 6** shows cases of enhanced ROC scores by combining metabolite markers that were positively correlated with each other. A combination of orthogonal markers that were each correlated with a tumor phenotype but not with each other (with an absolute value of Spearman ρ less than 0.4) also showed enhanced ROC scores compared with each individual marker (**Figure S2**). In the ROC analysis performed in MedCalc, we used all the data set to generate ROC curves. We also evaluated cross-validated area under the ROC curve (AUC) using leave- one-out validation in R studio (**Table S5**).[30] Compared with unvalidated AUCs conducted using MedCalc, we saw a general decrease of the value in the cross-validated AUCs. We could find that some metabolites still showed decent performance in distinguishing different groups of patients in this case. For example, a combination of threonine and asparagine got a cross- validated AUC of 0.8 for control group and growing adenomas group. We also noticed that ROC analysis for growing adenoma group and unknown growth did not work well (**Table S5**). This result could be caused by the relatively small number in these two groups, 14 samples and 6 samples, respectively. Furthermore, paired prepolypectomy and postpolypectomy serum samples from colorectal cancer patients enabled us to investigate the metabolite profile changes before and after removing polyps.

In this work, we applied our in-house synthesized isobaric tags for the relative quantification of primary and secondary amine- containing metabolites. Among the 40 quantified metabolites, we found that threonine, asparagine, glutamine, and arginine exhibited significant differences between the healthy controls and at least one colorectal cancer group (**Figure 4**). ROC analysis combining these metabolic markers showed decent specificity and sensitivity in detecting colorectal cancer. When we compared the metabolic markers with previously identified potential

protein biomarkers in the shared patients, we found that LRG1 showed a strong negative correlation with asparagine. It would be interesting to validate if a correlation exists using a bigger sample size and investigate if any biological pathways are involved in this correlation once validated. A combination of protein biomarkers with metabolite biomarkers should improve the sensitivity and specificity for detecting early growing adenomas of colorectal cancer.

Here, we putatively quantified 40 amine-containing metabolites in the serum, which is a little bit too low compared with some other labeling strategies like dansylation labeling.[31] Some reasons for this situation can be that 1) we excluded hundreds of *m/z* values shown in the DiLeu-labeled blank tube, inevitably removing some potential metabolites; 2) This method was based on MS2 quantification, so that it required the presence of higher concentrations of analytes for quantification since relative quantitation is achieved through comparison of reporter ions in tandem MS compared to MS1 quantification; 3) Because labeled metabolites are mainly in relative low mass range (*m/z* 200−500), coisolation of some labeled metabolites for tandem mass fragmentation affected quantification accuracy, then we excluded these features; 4) Desalting tips used here could hold only very limited amounts of metabolites, therefore the total injection amount of metabolites for LC-MS/MS analysis was small. Though it is not the focus of this study, there is still some room for improving the DiLeu-based relative quantification method for metabolite quantification in serum samples. Although the performance of the 4-plex DiLeu tags utilized in this study cannot compete with dansylation labeling in identified metabolite numbers, the DiLeu tags are superior in their high throughput, especially when 12-plex or even 21-plex DiLeu tags are applied for targeted primary/secondary amine- containing metabolites where metabolite standards are available.[18,23] Additionally, because the tags only label primary/secondary amines, there is no doubt that other metabolites without the amine group cannot be identified here. Therefore, a

combination of other types of chemical tags targeting different classes of metabolites or detecting sera metabolites using a label- free quantification strategy could expand the coverage of metabolites identified/quantified.

Enhancing prevention of colorectal cancer must be balanced by minimizing overdiagnosis. As only about 5% of all screening adults contain large polyps and the lifetime cancer risk,5 identifying the subset of early adenomas that are growing would be a key step to attenuate overdiagnosis. It is worth mentioning that this is a small cohort study with limited controls, so the confidence to claim the outcome is compromised. Although ROC analysis shows that some metabolites quantified here can distinguish samples from screening normal and adenoma carriers, there is still a lack of confidence to claim that these metabolites can be potential biomarkers for colorectal cancer because the sample size is small. Therefore, this study would benefit from further validation using a larger cohort of colorectal cancer patients with a classification of growth status. For example, with an estimated diagnostic accuracy (AUC) of 0.85 and marginal errors of 0.05 with 95% confidence level, the calculated sample sizes for each group of control and adenoma cases are 184.[32,33] Although there is a long way to go before we know whether these metabolites can be validated as biomarkers by independent validation studies, this study illustrates a small step toward the direction of developing early detection in colorectal cancer. Hopefully, this work and strategies could improve early diagnosis of colorectal cancer and finally lead to a reduced death rate of colorectal cancer.

Acknowledgements

This research was supported in part by R01 DK071801, RF1 AG052324, R01AG078794. The Orbitrap instruments were purchased through the support of an NIH shared instrument grant (NIH-NCRR S10RR029531) and Office of the Vice Chancellor for Research and Graduate

Education at the University of Wisconsin−Madison. L.L. acknowledges funding support of NIH shared instrument grants (S10OD028473 and S10OD025084), NIH grants P01CA250972 and R21AG065728, a Pancreas Cancer Pilot grant from the University of Wisconsin Carbone Cancer Center (233- AAI9632), and a Vilas Distinguished Achievement Professor- ship and Charles Melbourne Johnson Distinguished Chair Professorship with funding provided by the Wisconsin Alumni Research Foundation and University of Wisconsin−Madison School of Pharmacy.

References

1. Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **2021**, 71 3., 209-249.

2. Pickhardt, P. J. CT colonography: does it satisfy the necessary criteria for a colorectal screening test? *Expert Rev Gastroenterol Hepatol* **2014**, 8 3., 211-213.

3. Issa, I. A.; Noureddine, M. Colorectal cancer screening: An updated review of the available options. *World J Gastroenterol* **2017**, 23 28., 5086-5096.

4. Zauber, A. G.; Winawer, S. J.; O'Brien, M. J.; Lansdorp-Vogelaar, I.; van Ballegooijen, M.; Hankey, B. F.; Shi, W.; Bond, J. H.; Schapiro, M.; Panish, J. F.; et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med* **2012**, 366 8., 687-696.

5. Pickhardt, P. J.; Pooler, B. D.; Kim, D. H.; Hassan, C.; Matkowskyj, K. A.; Halberg, R. B. The Natural History of Colorectal Polyps: Overview of Predictive Static and Dynamic Features. *Gastroenterol Clin North Am* **2018**, 47 3., 515-536.

6. Pickhardt, P. J.; Kim, D. H.; Pooler, B. D.; Hinshaw, J. L.; Barlow, D.; Jensen, D.; Reichelderfer, M.; Cash, B. D. Assessment of volumetric growth rates of small colorectal polyps with CT colonography: a longitudinal study of natural history. *Lancet Oncol* **2013**, 14 8., 711-720.

7. Lieberman, D.; Moravec, M.; Holub, J.; Michaels, L.; Eisen, G. Polyp size and advanced histology in patients undergoing colonoscopy screening: implications for CT colonography. *Gastroenterology* **2008**, 135 4., 1100-1105.

8. Tutein Nolthenius, C. J.; Boellaard, T. N.; de Haan, M. C.; Nio, C. Y.; Thomeer, M. G.; Bipat, S.; Montauban van Swijndregt, A. D.; van de Vijver, M. J.; Biermann, K.; Kuipers, E. J.; et al. Evolution of Screen-Detected Small 6-9 mm. Polyps After a 3-Year Surveillance Interval:

Assessment of Growth With CT Colonography Compared With Histopathology. *Am J Gastroenterol* **2015**, 110 12., 1682-1690.

9. Ivancic, M. M.; Anson, L. W.; Pickhardt, P. J.; Megna, B.; Pooler, B. D.; Clipson, L.; Reichelderfer, M.; Sussman, M. R.; Dove, W. F. Conserved serum protein biomarkers associated with growing early colorectal adenomas. *Proc Natl Acad Sci U S A* **2019**, 116 17., 8471-8480.

10. Ivancic, M. M.; Megna, B. W.; Sverchkov, Y.; Craven, M.; Reichelderfer, M.; Pickhardt, P. J.; Sussman, M. R.; Kennedy, G. D. Noninvasive Detection of Colorectal Carcinomas Using Serum Protein Biomarkers. *J Surg Res* **2020**, 246, 160-169.

11. Papadimitriou, N.; Gunter, M. J.; Murphy, N.; Gicquiau, A.; Achaintre, D.; Brezina, S.; Gumpenberger, T.; Baierl, A.; Ose, J.; Geijsen, A.; et al. Circulating tryptophan metabolites and risk of colon cancer: Results from case-control and prospective cohort studies. *Int J Cancer* **2021**, 149 9., 1659-1669.

12. Zhang, C.; Zhou, S.; Chang, H.; Zhuang, F.; Shi, Y.; Chang, L.; Ai, W.; Du, J.; Liu, W.; Liu, H.; et al. Metabolomic Profiling Identified Serum Metabolite Biomarkers and Related Metabolic Pathways of Colorectal Cancer. *Dis Markers* 2021, **2021**, 6858809.

13. Hang, D.; Zeleznik, O. A.; Lu, J.; Joshi, A. D.; Wu, K.; Hu, Z.; Shen, H.; Clish, C. B.; Liang, L.; Eliassen, A. H.; et al. Plasma metabolomic profiles for colorectal cancer precursors in women. *Eur J Epidemiol* **2022**.

14. Bennett, B. D.; Yuan, J.; Kimball, E. H.; Rabinowitz, J. D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat Protoc* **2008**, 3 8., 1299-1311.

15. Kang, U. B.; Yeom, J.; Kim, H.; Lee, C. Quantitative analysis of mTRAQ-labeled proteome using full MS scans. *J Proteome Res* **2010**, 9 7., 3750-3758.

16. Wang, J.; Zhang, Y.; Xiang, F.; Zhang, Z.; Li, L. Combining capillary electrophoresis matrix-assisted laser desorption/ionization mass spectrometry and stable isotopic labeling techniques for comparative crustacean peptidomics. *J Chromatogr* A **2010**, 1217 26., 4463-4470.

17. Guo, K.; Li, L. Differential 12C-/13C-isotope dansylation labeling and fast liquid chromatography/mass spectrometry for absolute and relative quantification of the metabolome. *Anal Chem* **2009**, 81 10., 3919-3932.

18. Frost, D. C.; Feng, Y.; Li, L. 21-plex DiLeu Isobaric Tags for High-Throughput Quantitative Proteomics. *Anal Chem* **2020**, 92 12., 8228-8234.

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

20. Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **2004**, 3 12., 1154-1169.

21. Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **2003**, 75 8., 1895-1904.

22. Hao, L.; Zhong, X.; Greer, T.; Ye, H.; Li, L. Relative quantification of amine-containing metabolites using isobaric N,N-dimethyl leucine DiLeu. reagents via LC-ESI-MS/MS and CE-ESI-MS/MS. *Analyst* **2015**, 140 2., 467-475.

23. Wei, P.; Hao, L.; Thomas, S.; Buchberger, A. R.; Steinke, L.; Marker, P. C.; Ricke, W. A.; Li, L. Urinary Amine Metabolomics Characterization with Custom 12-Plex Isobaric DiLeu Labeling. *J Am Soc Mass Spectrom* **2020**, 31 9., 1854-1860.

24. Murphy, J. P.; Everley, R. A.; Coloff, J. L.; Gygi, S. P. Combining amine metabolomics and quantitative proteomics of cancer cells using derivatization with isobaric tags. *Anal Chem* **2014**, 86 7., 3585-3593.

25. Frost, D. C.; Greer, T.; Li, L. High-resolution enabled 12-plex DiLeu isobaric tags for quantitative proteomics. *Anal Chem* **2015**, 87 3., 1646-1654.

26. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* **2012**, 30 10., 918-920.

27. Hao, L.; Zhu, Y.; Wei, P.; Johnson, J.; Buchberger, A.; Frost, D.; Kao, W. J.; Li, L. Metandem: An online software tool for mass spectrometry-based isobaric labeling metabolomics. *Anal Chim Acta* **2019**, 1088, 99-106.

28. Ivancic, M. M.; Irving, A. A.; Jonakin, K. G.; Dove, W. F.; Sussman, M. R. The concentrations of EGFR, LRG1, ITIH4, and F5 in serum correlate with the number of colonic adenomas in ApcPirc/+ rats. *Cancer Prev Res Phila.* **2014**, 7 11., 1160-1169.

29. Wang, Z.; Cui, B.; Zhang, F.; Yang, Y.; Shen, X.; Li, Z.; Zhao, W.; Zhang, Y.; Deng, K.; Rong, Z.; et al. Development of a Correlative Strategy To Discover Colorectal Tumor Tissue Derived Metabolite Biomarkers in Plasma Using Untargeted Metabolomics. *Anal Chem* **2019**, 91 3., 2401-2408.

30. LeDell, E.; Petersen, M.; van der Laan, M. Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates. Electron J Stat **2015**, 9 1., 1583-1607. Hao,

L.; Wang, J.; Page, D.; Asthana, S.; Zetterberg, H.; Carlsson, C.; Okonkwo, O. C.; Li, L. Comparative Evaluation of MS-based Metabolomics Software and Its Application to Preclinical Alzheimer's Disease. *Sci Rep* **2018**, 8 1., 9291.

31. Chen, D.; Han, W.; Su, X.; Li, L.; Li, L. Overcoming Sample Matrix Effect in Quantitative Blood Metabolomics Using Chemical Isotope Labeling Liquid Chromatography Mass Spectrometry. *Anal Chem* **2017**, 89 17., 9424-9431.

32. Hajian-Tilaki, K. Sample size estimation in diagnostic test studies of biomedical informatics. *J Biomed Inform* **2014**, 48, 193-204.

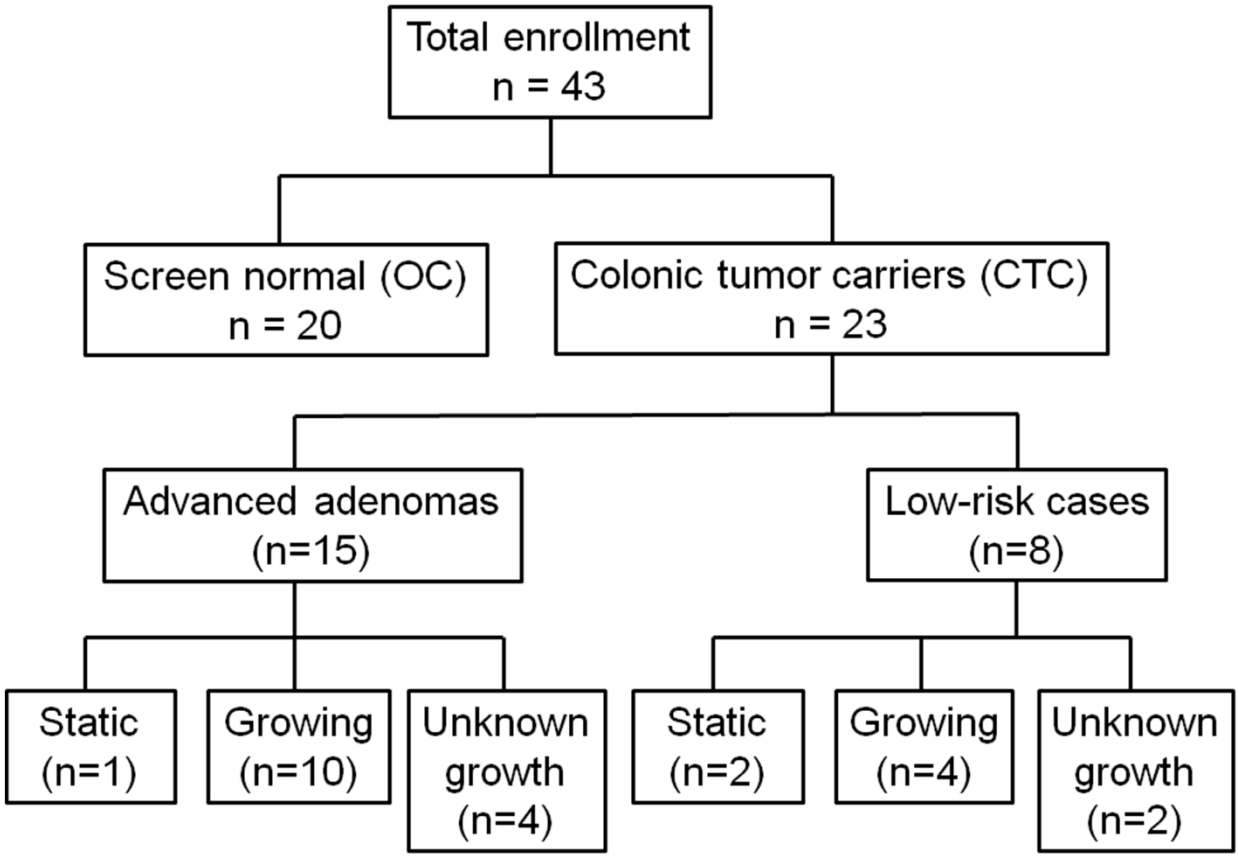


Figure 1. A summary of patient cases enrolled into this study. For patients carrying polyps, the polyps were examined and classified by standard histopathologic criteria as low-risk adenomas or high-risk (advanced) adenomas. A blood draw was performed before and after polypectomy for patients carrying tumor in this study. When available, tumors were classified as growing and static independently based on the longitudinal size profiles. The level of a biomarker of interest was compared between prepolypectomy and postpolypectomy sera.

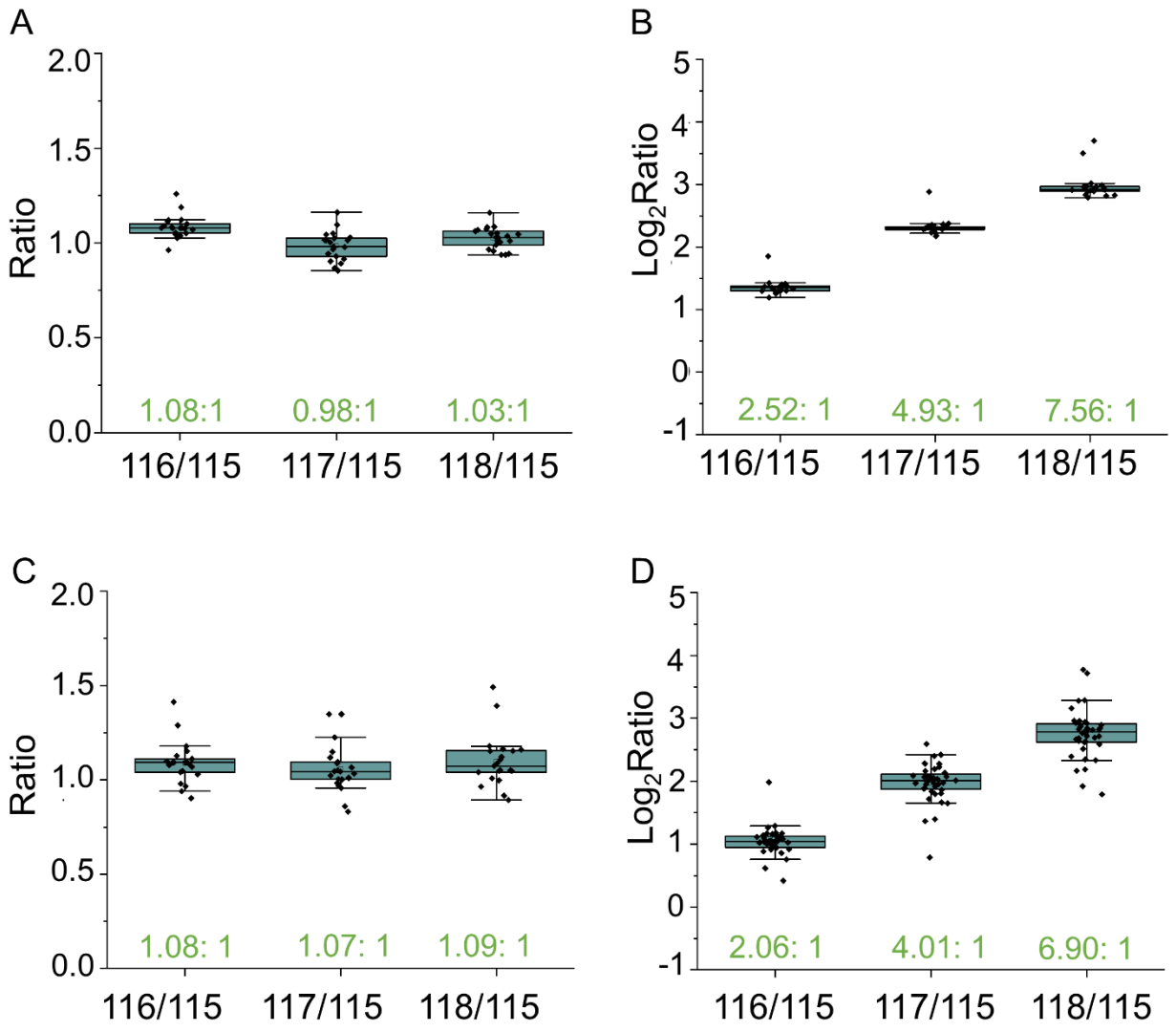


Figure 2. Box plots of reporter ion ratios of DiLeu labeled metabolite standards at a theoretical 1:1:1:1 ratio (A) and 1:2:4:8 ratios (B), and box plots of reporter ion ratios of DiLeu labeled serum metabolites at a theoretical 1:1:1:1 ratio (C) and 1:2:4:8 ratios (D). Each box contains 21 data points (average of 3 replicates) from 21 metabolite standards for metabolite standards, and 40 data points for serum metabolites, obtained from LC-ESI-MS/MS analysis. Box denotes 25th and 75th percentiles; line within box denotes 50th percentile; whiskers denote standard deviation.

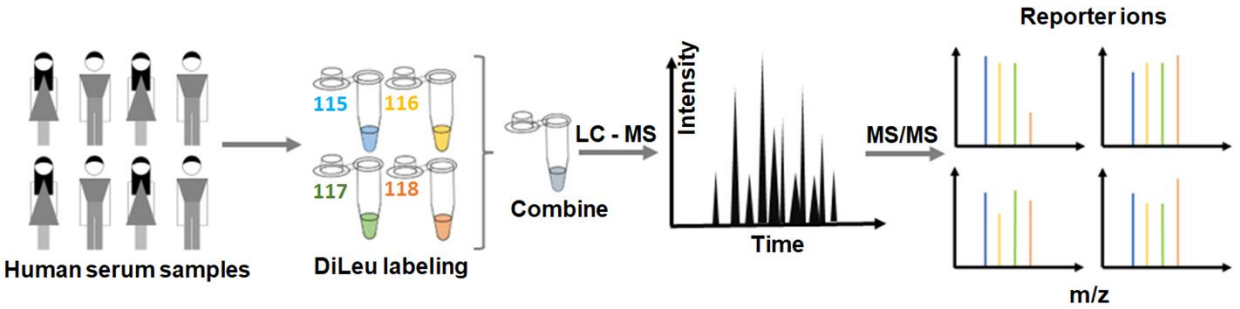


Figure 3. Workflow for relative quantification of serum metabolites using the DiLeu-labeling strategy.

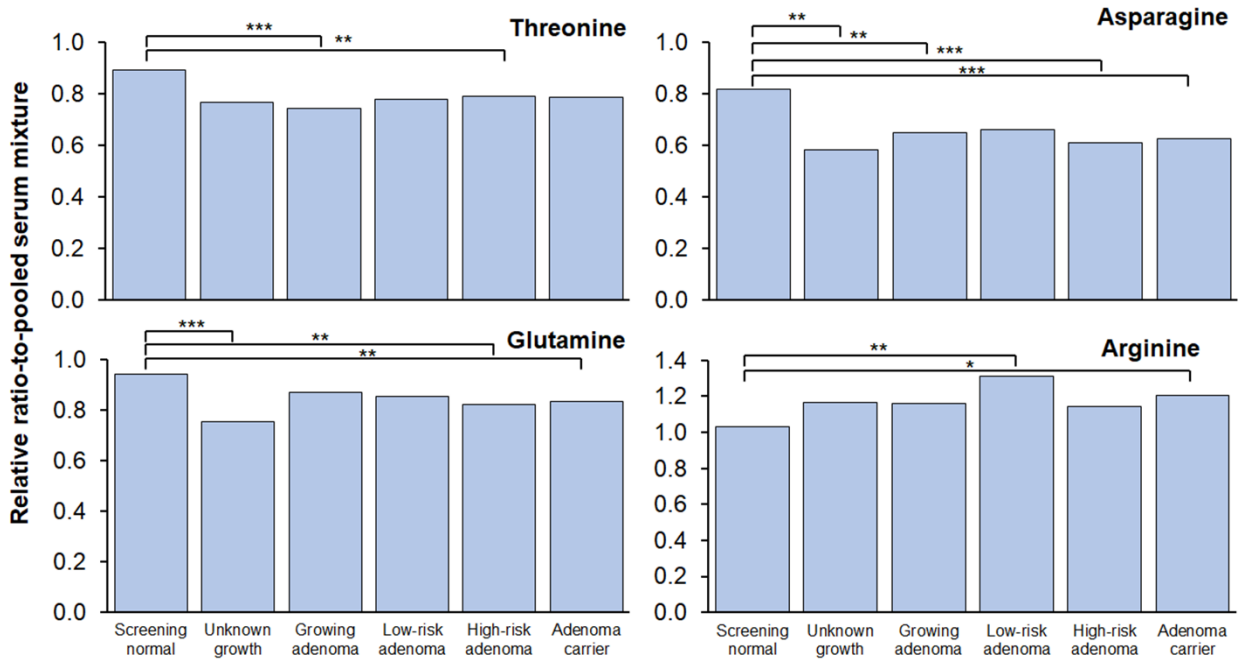


Figure 4. Relative ratio-to-pooled serum mixture for six comparison groups in four metabolites that showed statistically significant tumor-associated level changes. A blood draw was performed prepolypectomy. A Mann−Whitney U test was performed. Asterisks represent the significance level (*P ⩽ 0.05, **P ⩽ 0.01, and ***P ⩽ 0.001) across the different adenoma groups with Benjaminin−Hochberg to control the false discovery rate q at <0.05.

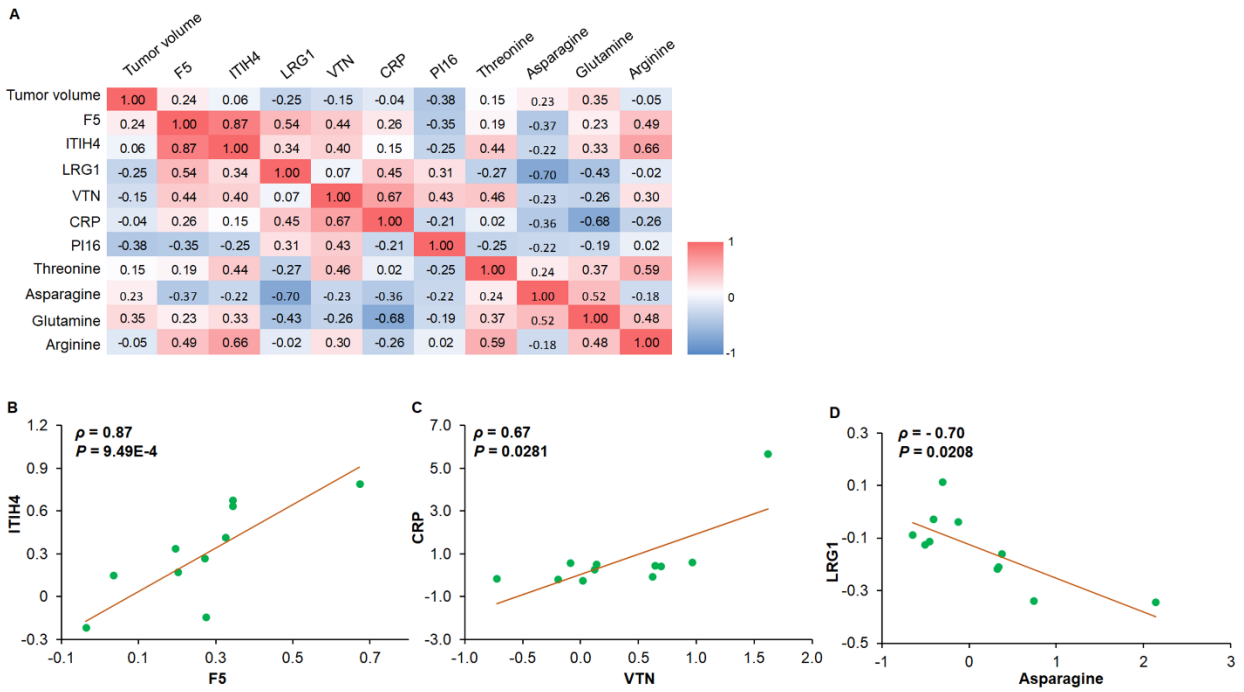


Figure 5. Correlation between tumor volume and differential marker level. Correlation heatmap for tumor volume, protein biomarkers, and metabolite markers (A). Correlation between ITIH4 and F5 (B). Correlation between CRP and VTN (C). Correlation between LRG1 and asparagine (D). A Spearman test was carried out for correlation between the monotonic rank orders of tumor volume, protein biomarker level, and metabolite biomarker. The Spearman ρ-values and their P-values indicated the strength of the correlation between different biomarkers. The corresponding levels of each of the biomarkers were determined, relative to the pooled serum samples from normal control samples of each metabolite.

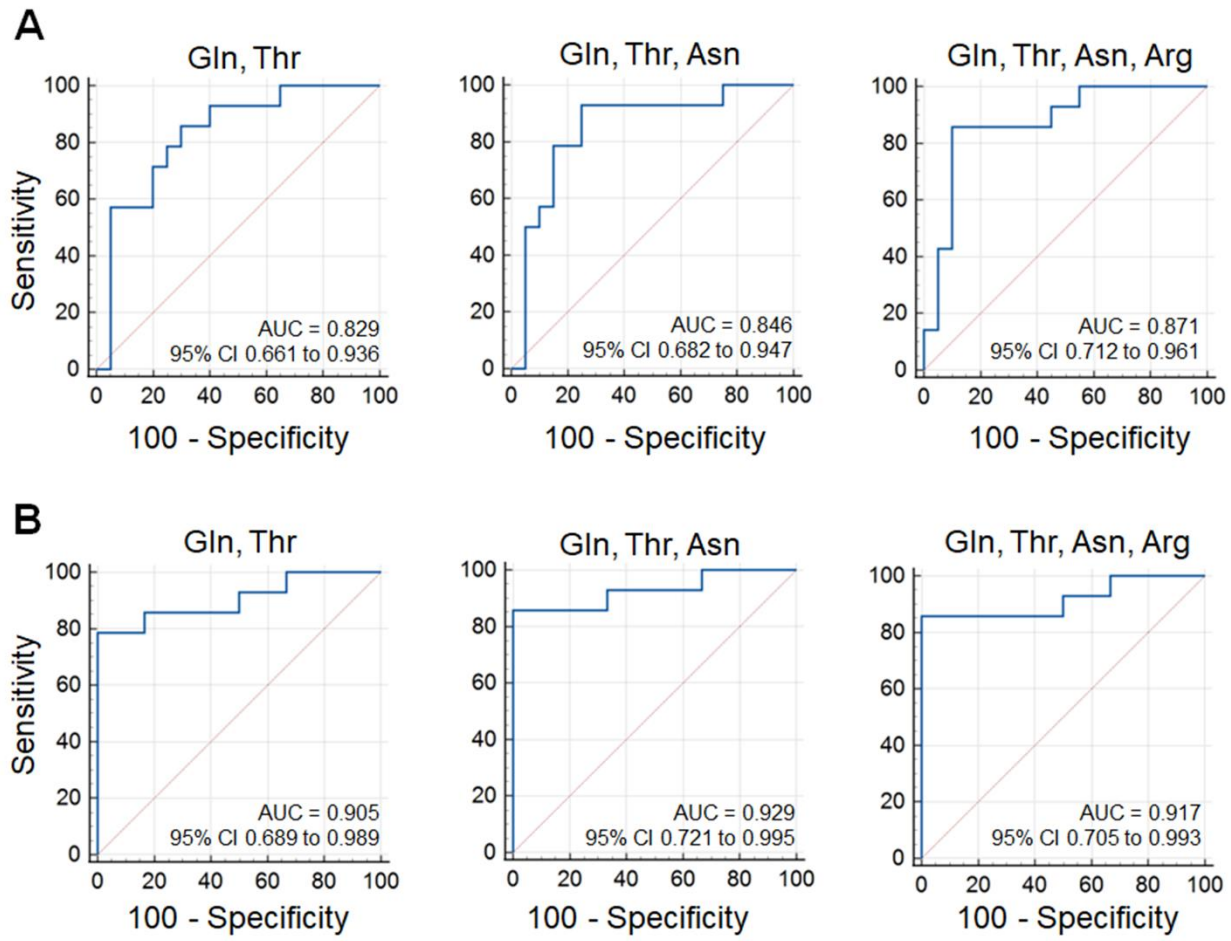


Figure 6. ROC analysis showing sensitivity and specificity of panels of potential metabolite biomarkers for detecting growing adenomas compared with normal controls (A) or compared with adenomas of unknown growth status (B).

Supplemental Information

Table S1. An overview of CTC patient data.

CTC case #	Sex	Path report segment	Path report diagnosis	Length of time on study (years)	Linear size at final CTC (mm)	First/Final volume measure-ment (mm^3)	Percent growth by volume	Growth Classify-cation	Overall growth classify-cation	Risk Classify-cation
01	M	Rectosigmoid	Tubular adenoma	5.10	9.3	87/88	1.1	Static	Static	Low-Risk
02	M	Sigmoid	Tubular adenoma	5.10	8.2	28/83	196.4	Growing	Growing	Low-Risk
03	F	Transverse	Tubular adenoma, lymphoid follicle	5.33	6.7	36/45	25.0	Static	Static	Low-Risk
05	F	Ascending	Hyperplastic polyp	5.33	20.8	38/227	497.4	Growing	Growing	High-risk
		Rectum	Tubular adenoma	N/A	N/A	N/A		Unknown		
07	M	Cecum	Tubulovillous Adenoma	5.61	13.7	33/900	2627.3	Growing	Growing	High-risk
		Distal Sigmoid	Inflammatory Polyp	2.01	11.0	300/393	31.0	Growing		
08	F	Proximal Sigmoid	Tubular adenoma	N/A	N/A	N/A		Unknown	Growing	High-risk
		Mid Sigmoid	Tubular adenoma	N/A	N/A	N/A		Unknown		
09	F	Ascending	Tubular adenoma	5.02	7.7	12/19	58.3		Growing	Low-Risk
10	F	Ascending	Inflammatory	7.18	8.7	--/266		Unknown	Unknown	High-risk
11	M	Hepatic flexure	Tubular adenoma	N/A	N/A	N/A		Unknown	Unknown	Low-Risk
		Sigmoid colon	Tubular adenoma	N/A	N/A	N/A		Unknown		
12	F	Cecum	Tubular adenoma	N/A	N/A	N/A		Regressing	Unknown	High-risk
		Cecum	Tubular adenoma	1.05	5.6	42/26	-38.1	Regressing		
		Transverse	Tubular adenoma	4.84	4.0	13/7	-46.2	Regressing		
15	F	Ascending	Tubular adenoma	4.84	5.0	60/13	-78.3	Regressing	Growing	High-risk
		(Missing)	Tubular adenoma	4.84	7.7	60/91	51.7	Growing		
		Hepatic flexure	Tubular adenoma	3.18	9.1	174/257	47.7	Growing		
18	M	Rectum	Tubular adenoma and hyperplastic polyp	N/A	N/A	N/A		Unknown	Growing	Low-Risk
		Sigmoid	Tubular adenoma	7.47	5.6	8/40	400.0	Growing		
19	M	Sigmoid	Tubular adenoma	7.47	5.5	36/41	13.9	Static	Growing	High-risk
		Descending	Tubular adenoma	N/A	N/A	N/A		Unknown		
20	M	Sigmoid	Tubulovillous Adenoma	4.12	9.2	89/100	12.4	Static	Static	High-risk
21	M	Descending	Tubulovillous Adenoma	6.00	10.2	178/113	-36.5	Growing	Unknown	High-risk
		Cecum/Ascend-ing	Tubular adenoma	N/A	N/A	N/A		Unknown		
		Sigmoid	Tubulovillous Adenoma	4.20	7.0	104/161	54.8	Growing		
23	F	Ascending	Lymphoid Aggregate	N/A	N/A	N/A		Unknown	Growing	High-risk
		Rectum	Hyperplastic polyp	N/A	N/A	N/A		Unknown		
24	M	Rectum	Tubulovillous Adenoma	3.25	11.4	134/211	57.5	Growing	Growing	High-risk
26	M	Rectum	Hyperplastic polyp	3.11	6.1	28/31	10.7	Static	Unknown	Low-Risk
		Sigmoid	Hyperplastic polyp	N/A	N/A	N/A		Unknown		
		Sigmoid	Tubular adenoma and hyperplastic polyp	3.40	6.1	16/25	56.3	Growing		
28	F	Rectum	Tubular adenoma	3.40	9.0	56/94	67.9	Growing		
		Ascending	Tubular adenoma	N/A	N/A	N/A		Unknown	Growing	High-risk

CTC case #	Sex	Path report segment	Path report diagnosis	Length of time on study (years)	Linear size at final CTC (mm)	First/Final volume measure-ment (mm^3)	Percent growth by volume	Growth Classify-cation	Overall growth classify-cation	Risk Classify-cation
29	F	Cecum	Sessile Serrated Adenoma	5.02	7.8	40/35	-12.5	Static	Unknown	High-risk
		Transverse	Sessile Serrated Adenoma	N/A	N/A	N/A		Unknown		
32	M	Descending	Tubular adenoma	5.27	10.4	161/321	99.4	Growing	Growing	Low-Risk
33	F	Ascending	Sessile Serrated Adenoma	7.98	6.2	11/17	54.5	Growing	Growing	High-risk
		Transverse	Tubular adenoma	7.98	7.0	--/21		Unknown		

Growth classification is based on volume (+/- 30% cutoff), risk classification is based on standard-of-care criteria. This table was adopted and modified from a previous study.[1]

Table S2. Colonoscopy/Cancer Free Cases.

Case #	Specimen ID	Sex	Age	Race/Eth.	Polyp Count	Polyp Count BIN	Pathology Rating	Polyp Location (Colon vs. Rectum)	Polyp Location effect (Ascending, Transverse, Descending, Rectal)	Pathology Summary
N01	22848	M	67	White	0	0	Low	Colon	Normal	Screening normal
N02	21684	M	64	White	0	0	Normal	Normal	Normal	Screening normal
N03	21685	F	50	White	0	0	Normal	Normal	Normal	Screening normal
N04	21755	F	50	White	0	0	Normal	Normal	Normal	Screening normal
N05	22847	M	50	White	0	0	Normal	Normal	Normal	Screening normal
N06	23032	M	51	White	0	0	Normal	Normal	Normal	Screening normal
N07	31089	F	56	White	0	0	Normal	Normal	Normal	Screening normal
N08	31090	F	50	White	0	0	Normal	Normal	Normal	Screening normal
N09	31528	M	72	White	0	0	Normal	Normal	Normal	Screening normal
N10	31571	M	50	White	0	0	Normal	Normal	Normal	Screening normal
N11	31634	F	51	White	0	0	Low	Colon	Normal	Screening normal
N12	32312	M	54	White	0	0	Normal	Normal	Normal	Screening normal
N13	32491	F	53	White	0	0	Normal	Normal	Normal	Screening normal
N14	33029	F	62	White	0	0	Normal	Normal	Normal	Screening normal
N15	33072	F	56	White	0	0	Normal	Normal	Normal	Screening normal
N16	33287	M	55	White	0	0	Normal	Normal	Normal	Screening normal
N17	34188	M	50	White	0	0	Normal	Normal	Normal	Screening normal
N18	34312	F	56	White	0	0	Normal	Normal	Normal	Screening normal
N19	34411	F	57	White	0	0	Normal	Normal	Normal	Screening normal
N20	34412	M	51	White	0	0	Normal	Normal	Normal	Screening normal

This table was adopted from a previous study.[2]

Table S3. Primary or secondary amine-containing metabolites that were labeled and putatively identified from serum samples using the LC-MS/MS platform.

Compound	Formula	Molecular weight (theoretical)	Molecular weight (calculated)	Mass of labeled metabolite *m/z* (detected)	Mass of labeled metabolite	Retention time (min)	Δppm
Ornithine	$C_5H_{12}N_2O_2$	132.0899	132.0898	212.1795	422.3433	17.31	-1.11
Alanine[a]	$C_3H_7NO_2$	89.0477	89.0471	235.1822	234.1744	16.3	-7.06
Creatinine	$C_4H_7N_3O$	113.0589	113.0585	259.1936	258.1858	18.2	-3.89
Proline[a] *	$C_5H_9NO_2$	115.0633	115.0628	261.1979	260.1901	16.9	-5.03
Valine[a]	$C_5H_{11}NO_2$	117.0790	117.0785	263.2136	262.2058	17.1	-4.52
Threonine[a]	$C_4H_9NO_3$	119.0582	119.0579	265.193	264.1852	15.2	-3.30
Cystine[a]	$C_6H_{12}N_2O_4S_2$	240.0239	240.0239	266.1463	530.2769	17.4	-6.46
Creatine	$C_4H_9N_3O_2$	131.0695	131.0689	277.2040	276.1962	18	-4.65
Isoleucine[a]	$C_6H_{13}NO_2$	131.0947	131.0942	277.2293	276.2215	18.2	-3.65
Leucine[a]	$C_6H_{13}NO_2$	131.0946	131.0942	277.2293	276.2215	18.8	-3.65
Asparagine[a]	$C_4H_8N_2O_3$	132.0535	132.0531	278.1882	277.1804	3.5	-3.35
Glutamine[a]	$C_5H_{10}N_2O_3$	146.0691	146.0687	292.2038	291.1960	4.4	-3.37
Glutamic acid[a]	$C_5H_9NO_4$	147.0532	147.0525	293.1876	292.1798	15.8	-4.81
LysoPE(0:0/15:0)	$C_{20}H_{42}NO_7P$	439.2699	439.2685	293.2057	584.3958	17	-3.29
Methionine[a]	$C_5H_{11}NO_2S$	149.0510	149.0497	295.1848	294.1770	17.8	-9.39
LysoPE(0:0/16:0)	$C_{21}H_{44}NO_7P$	453.2855	453.2843	300.2136	598.4116	17.3, 17.9	-2.74
Histidine[a]	$C_6H_9N_3O_2$	155.0695	155.0692	301.2043	300.1965	2.3	-2.11
Aminoadipic acid	$C_6H_{11}NO_4$	161.0688	161.0682	307.2033	306.1955	17.8	-3.79
Amino (2S)-2,5-diamino-5-oxopentanoate	$C_5H_{11}N_3O_3$	161.0800	161.0795	307.2146	306.2068	16.3	-3.28
2-Hydroxy-L-methionine	$C_5H_{11}NO_3S$	165.0460	165.0454	311.1806	310.1727	16.3	-3.58
Phenylalanine[a]	$C_9H_{11}NO_2$	165.0790	165.0786	311.2137	310.2059	19.4	-2.60
1-Methylhistidine	$C_7H_{11}N_3O_2$	169.0851	169.0839	315.2190	314.2112	26.2	-7.00
Arginine[a]	$C_6H_{14}N_4O_2$	174.1117	174.1107	320.2458	319.2380	2.5	-5.89
Citrulline	$C_6H_{13}N_3O_3$	175.0957	175.0953	321.2304	320.2226	17.6	-2.55
Tyrosine[a]	$C_9H_{11}NO_3$	181.0739	181.0734	327.2085	326.2007	17.4	-3.00
Tryptophan[a]	$C_{11}H_{12}N_2O_2$	204.0899	204.0898	350.2249	349.2171	19.8	-0.63
Kynurenine[a]	$C_{10}H_{12}N_2O_3$	208.0848	208.0854	354.2205	353.2127	19.4	2.69

5-Hydroxy-L-tryptophan	$C_{11}H_{12}N_2O_3$	220.0848	220.0845	366.2196	365.2118	19.9	-1.39
Nitrotyrosine	$C_9H_{10}N_2O_5$	226.059	226.0588	372.1939	371.1861	20.5	-1.02
Uridine*	$C_9H_{12}N_2O_6$	244.0695	244.0689	390.2041	389.1962	18.2	-2.40
Phenylalanyl-Serine	$C_{12}H_{16}N_2O_4$	252.111	252.1107	398.2458	397.2380	28.8	-1.21
Glutamylleucine	$C_{11}H_{20}N_2O_5$	260.1372	260.1369	406.2720	405.2642	19.75	-1.22
Norophthalmic acid	$C_{10}H_{17}N_3O_6$	275.1117	275.1109	421.2461	420.2382	16.5	-2.80
Lysine[a]	$C_6H_{14}N_2O_2$	146.1055	146.1053	437.3677	436.3599	17.1	-1.90
Phenylalanylglutamic acid	$C_{14}H_{18}N_2O_5$	294.1216	294.1212	440.2564	439.2485	20.5	-1.25
Fructosyl-lysine	$C_{12}H_{24}N_2O_7$	308.1584	308.1573	454.2924	453.2846	17.4	-3.59
Gamma-Glutamyltyrosine	$C_{14}H_{18}N_2O_6$	310.1165	310.1163	456.2515	455.2436	18.9	-0.64
Phenylalanylphenylalanine	$C_{18}H_{20}N_2O_3$	312.1474	312.1472	458.2824	457.2745	22.6	-0.59
N-Fructosyl phenylalanine*	$C_{15}H_{21}NO_7$	327.1318	327.1313	473.2664	472.2586	20.1	-1.67
N-(1-Deoxy-1-fructosyl)tyrosine	$C_{15}H_{21}NO_8$	343.1267	343.1265	489.2617	488.2538	18.8	-0.44

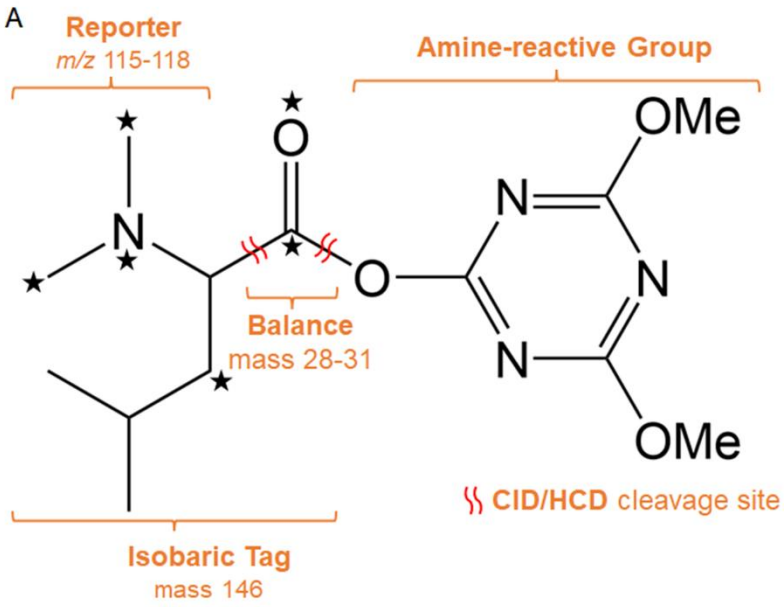
[a] Metabolites validated using standards. *Secondary amine containing metabolites.

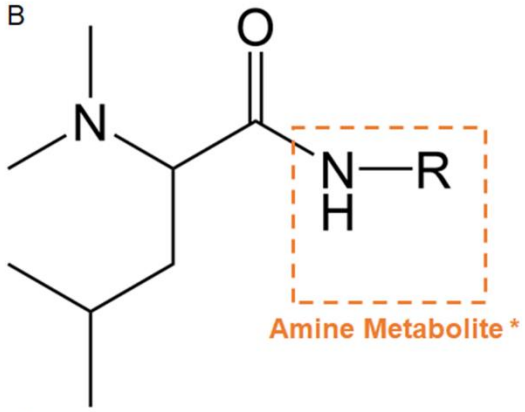
Some *m/z* features show several peaks in the chromatograph extraction, they might be isomers.

Table S4. Test for correlation between tumor volume and differential biomarker level.

Patient #	Total volume of adenomas	Prepolypectomy value for patient minus average of tumor-free controls			
		Threonine	Asparagine	Glutamine	Arginine
1	88	-0.0014	-0.1136	-0.1103	0.4839
2	83	-0.0701	-0.0399	0.0968	0.3757
3	45	0.0592	-0.3451	-0.2558	0.3497
7	900	0.0575	0.1136	0.0605	0.1324
9	19	-0.2767	-0.2188	-0.1633	0.1537
10	266	-0.1453	-0.0889	-0.1326	-0.0783
15	111	-0.1998	-0.1614	-0.1780	0.0199
20	100	0.3416	-0.1251	0.0048	0.5422
24	211	-0.2717	-0.3400	-0.1439	0.0350
32	231	-0.1254	-0.2092	0.1427	0.4856
33	38	-0.1948	-0.0291	-0.1124	-0.2712
Spearman	ρ	0.15	0.23	0.35	-0.05
	Two-sided *P* value	0.65	0.50	0.29	0.88

The corresponding levels of each of the conserved markers (threonine, kynurenine, cystine and ornithine) were determined, relative to the pooled serum samples from control of each marker. A Spearman test was performed for correlation between the monotonic rank orders of tumor volume and metabolite biomarker level. The Spearman ρ values and their *P* values show a lack of correlation between tumor volume and the level of each informative serum biomarker. The tumor volume data was adopted from previous study.[1]





C

DiLeu Label	Stable Isotopes		Reporter Mass	Mass Shift
	Reporter	Balance		
115	●	●●	115.1311	145.1263
116	●●	●	116.1344	145.1263
117	●●●	●	117.1373	145.1283
118	●●●●		118.1407	145.1283
	● ^{13}C ● ^{2}H ● ^{15}N ● ^{18}O			

Figure S1. General structures of 4-plex DiLeu isobaric tags (A), DiLeu labeled amine metabolite (B) and stable isotopes (^{13}C, ^{2}H, ^{15}N and ^{18}O) distribution and mass shift caused by DiLeu labeling (C).

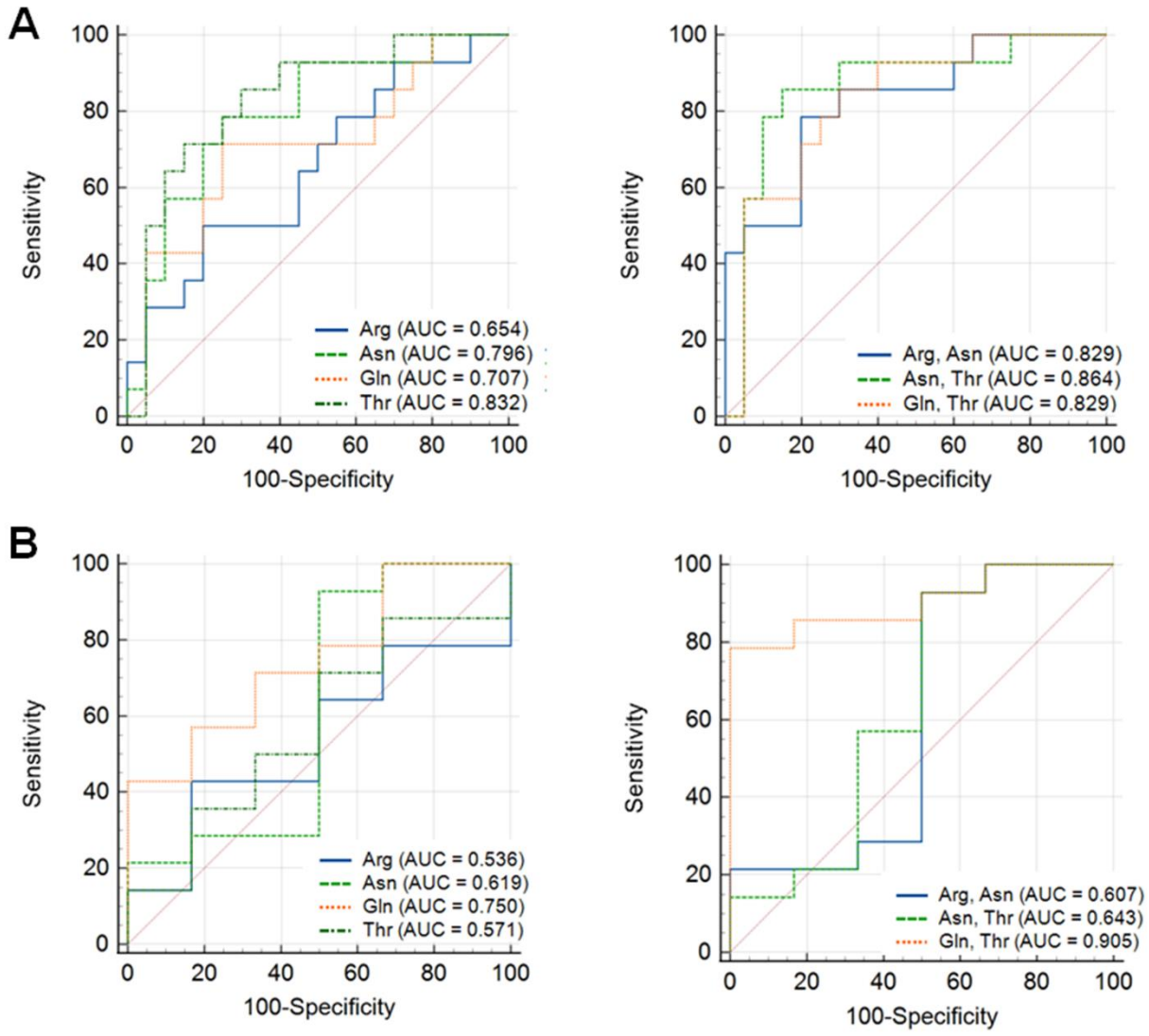


Figure S2. ROC analysis showing sensitivity and specificity of potential metabolite biomarkers for detecting growing adenomas compared with normal controls (A) or compared with adenomas of unknown growth status (B).

Table S5. Cross-validated area under the ROC curve between different patient groups.

Patient Groups and metabolites	cvAUC	SE	CI	Confidence
control/adenoma (Arg)	0.66	0.08	0.50 - 0.82	0.95
control/adenoma (Asp)	0.80	0.07	0.66 - 0.93	0.95
control/adenoma (Glu)	0.72	0.08	0.56 - 0.88	0.95
control/adenoma (Thr)	0.70	0.08	0.55 - 0.86	0.95
control/adenoma (Asp+Arg)	0.80	0.07	0.67 - 0.94	0.95
control/adenoma (Asp+Gln)	0.78	0.07	0.63 - 0.92	0.95
control/adenoma (Glu+Arg)	0.77	0.07	0.62 - 0.92	0.95
control/adenoma (Thr+Arg)	0.75	0.08	0.60 - 0.90	0.95
control/adenoma (Thr+Asp)	0.78	0.07	0.64 - 0.93	0.95
control/adenoma (Thr+Gln)	0.75	0.08	0.59 - 0.90	0.95
control/adenoma (Thr+Gln+Asp)	0.75	0.08	0.61 - 0.90	0.95
control/adenoma (Thr+Gln+Arg)	0.79	0.07	0.65 - 0.94	0.95
control/adenoma (Thr+Asp+Arg)	0.80	0.07	0.66 - 0.94	0.95
control/adenoma (Gln+Asp+Arg)	0.79	0.07	0.64 - 0.93	0.95
control/adenoma (Thr+Gln+Asp+Arg)	0.78	0.07	0.64 - 0.93	0.95
control/growing (Arg)	0.57	0.10	0.37 - 0.77	0.95
control/growing (Asp)	0.76	0.09	0.59 - 0.93	0.95
control/growing (Glu)	0.61	0.10	0.41 - 0.81	0.95
control/growing (Thr)	0.79	0.08	0.64 - 0.95	0.95
control/growing (Asp+Arg)	0.74	0.09	0.56 - 0.91	0.95
control/growing (Asp+Gln)	0.72	0.09	0.54 - 0.89	0.95
control/growing (Glu+Arg)	0.68	0.09	0.50 - 0.86	0.95
control/growing (Thr+Arg)	0.79	0.08	0.63 - 0.95	0.95
control/growing (Thr+Asp)	0.80	0.08	0.64 - 0.96	0.95
control/growing (Thr+Gln)	0.76	0.08	0.59 - 0.92	0.95
control/growing (Thr+Gln+Asp)	0.78	0.09	0.61 - 0.95	0.95
control/growing (Thr+Gln+Arg)	0.78	0.08	0.62 - 0.93	0.95
control/growing (Thr+Asp+Arg)	0.77	0.08	0.61 - 0.93	0.95
control/growing (Gln+Asp+Arg)	0.70	0.09	0.51 - 0.88	0.95
control/growing (Thr+Gln+Asp+Arg)	0.75	0.09	0.58 - 0.92	0.95
NA/growing (Arg)	0.00	0.00	0 - 0	0.95
NA/growing (Asp)	0.38	0.15	0.09 - 0.67	0.95
NA/growing (Glu)	0.67	0.13	0.41 - 0.92	0.95
NA/growing (Thr)	0.24	0.11	0.02 - 0.45	0.95
NA/growing (Asp+Arg)	0.29	0.12	0.04 - 0.53	0.95
NA/growing (Asp+Gln)	0.58	0.13	0.33 - 0.84	0.95
NA/growing (Glu+Arg)	0.68	0.13	0.42 - 0.93	0.95
NA/growing (Thr+Arg)	0.05	0.04	0.00 - 0.13	0.95
NA/growing (Thr+Asp)	0.37	0.15	0.07 - 0.66	0.95
NA/growing (Thr+Gln)	0.79	0.10	0.59 - 0.98	0.95
NA/growing (Thr+Gln+Asp)	0.74	0.12	0.51 - 0.97	0.95
NA/growing (Thr+Gln+Arg)	0.73	0.12	0.49 - 0.96	0.95
NA/growing (Thr+Asp+Arg)	0.37	0.16	0.05 - 0.69	0.95
NA/growing (Gln+Asp+Arg)	0.60	0.14	0.33 - 0.86	0.95

cvAUC: cross-validated area under curve; SE: standard errors; CI: confidential interval; control: control group (or screening normal group); adenoma: adenoma carrier group; growing: growing adenoma group; NA: unknown growth. Cross-validated area under the ROC curve was generated using leave-out-one cross validation performed in R.

References

1. Ivancic, M. M.; Anson, L. W.; Pickhardt, P. J.; Megna, B.; Pooler, B. D.; Clipson, L.; Reichelderfer, M.; Sussman, M. R.; Dove, W. F., Conserved serum protein biomarkers associated with growing early colorectal adenomas. *Proc Natl Acad Sci U S A* **2019**, 116 (17), 8471-8480.

2. Ivancic, M. M.; Megna, B. W.; Sverchkov, Y.; Craven, M.; Reichelderfer, M.; Pickhardt, P. J.; Sussman, M. R.; Kennedy, G. D., Noninvasive Detection of Colorectal Carcinomas Using Serum Protein Biomarkers. *J Surg Res* **2020**, 246, 160-169.

Chapter 7

Conclusions and Future Directions

Conclusions and Future Directions

In this dissertation, I developed a derivatization strategy for improving the identification, quantification, and visualization of primary amine-containing metabolites. This strategy may be applied to various tissue samples to visualize the spatial distribution of low molecular weight metabolites. Initially, we hoped to develop this method for rapid analysis of primary amine-containing metabolites in cell culture medium for monitoring the dynamic changes of metabolite profiles. Furthermore, we expected that we could apply this strategy to mapping primary amine-containing metabolites of PANC1 spheroids, hoping to add insights to the crosstalk between pancreatic cancer cells and pancreatic stellate cells (PSCs). However, as we described in **Chapter 2**, this strategy can improve the detection of some metabolites but not all the metabolites. Additionally, when we applied the NBA reagents to more complex samples, such as DMEM, we found that the performance of this strategy for improving the detection of metabolites was compromised, making it not suitable for this purpose. Although further optimization of reaction conditions, such as pH, may help to solve this issue to some extent. Another potential direction could be exploring more analogues of NBA reagents. As shown in **Chapter 2**, we only tried 5 commercially available NBA and its analogues, but there should be more choices out there. We could even synthesize more analogues of NBA by ourselves, like work reported in another study.[1] Meanwhile, as we used the LTQ MALDI Orbitrap to optimize the reaction conditions of metabolite and NBA reagents, the optimized conditions may not be applicable to other MS instruments. The LTQ MALDI Orbitrap was an old instrument, and its performance was not as good as some high-end instrument such as Thermo QE-HF orbitrap mass spectrometer or Orbitrap Fusion™ Lumos™ Tribrid™ mass spectrometer. We previously tested the SubAP-MALDI source coupled with the QE-HF Orbitrap mass spectrometer for spots analysis, and we found that this instrument could

detect alanine standard easily at a concentration of 100 μM but the LTQ MALDI Orbitrap could not. Therefore, in **Chapter 2**, we only presented some proof-of-principle results indicating that NBA reagents may have the potential to be applied for improving detection of primary amine-containing metabolites in complex samples.

Although the performance of NBA reagents was not as satisfactory as expected, we moved on to investigate crosstalk between PSCs and PNAC1 spheroids. To investigate the crosstalk between PANC1 spheroids and PSCs, we mainly utilized two techniques – one was MALDI MSI and the other one was nano-LC-ESI MS/MS. Nano-LC-ESI enabled us to perform relative quantification of proteins in PANC1 spheroids, profiling the changes of protein levels. MALDI MSI enabled us to study the spatial distribution of biomolecules in PANC1 spheroids. To interrogate the impact of PSCs on PANC1 spheroids, we optimized a sample preparation workflow for high throughput MSI of spheroids, as presented in **Chapter 3**. In this sample preparation strategy, 81 spheroids or even more than one hundred spheroids could be imaged in a simplified workflow. As the reproducibility of MALDI MS was not as good as LC-MS/MS, it would be very meaningful to increase the replicates of spheroids if we wanted to compare spheroids treated under different conditions. Increasing the replicates of spheroids being imaged could help us draw a more confident conclusion about the significant differences in abundance of biomolecules between different treatment conditions. Further work for this direction could be doing MSI for PANC1 spheroids cultured alone or cocultured with PSCs. One pitfall for this direction could be that we might not be able to detect biomolecules related to the crosstalk between pancreatic cancer cells and PSCs because of the relatively low S/N ratio of MALDI MS signal leading to less comprehensive coverage of metabolites. Another option could be detecting metabolites using LC-MS/MS techniques for discovery of changes of metabolite levels followed by targeted and

untargeted MSI of spheroids. In parallel, we applied the DIA method for relative quantification of proteins in PANC1 spheroids cultured alone and co-cultured with PSCs. We quantified over 7000 proteins in each sample and found some significant changes of protein levels from day 1 to day 3 to day 5 as detailed in **Chapter 4**. However, we only found one protein, HTRA1, that showed significant differences in the expression level between monocultured PANC1 spheroids and cocultured spheroids. Though we found some evidence suggested that HTRA1 may be related to tumor growth and migration,[2, 3] based on limited information about HTRA1 in pancreatic cancer, I was cautious to draw any conclusions here. For future directions, we could investigate protein post-translational modifications, such as phosphorylation and glycosylation. This may help reveal the interactions between pancreatic cancer cells and PSCs. Additionally, in my current coculture setup, there was no direct contact/physical interaction between PANC1 cells and PSCs, this may also affect the interactions between these two kinds of cells. We could co-culture these two types of cells together and investigate the changes of protein levels. In that case, to quantify the protein levels of each type of cell, it would be better to separate PANC1 cells from PSCs. Otherwise, it would be difficult to determine whether the proteins were from PANC1 cells or PSCs if both cells expressed the same proteins. Therefore, utilization of fluorescence labeled PANC1 cells (commercially available in ATCC) and/or labeled PSCs could address this issue. The labeled cells could be sorted based on the fluorescence label by flow cytometry and then proteins from each type of cells could be quantified.

In addition to method development for enhancing the detection of metabolites and investigation of crosstalk, I also applied isotopic/isobaric tags for absolute/relative quantification of proteins/metabolites in serum sample from AD patients/colorectal cancer patients as shown in **Chapter 5** and **Chapter 6**, respectively. We observed some differences of proteins/metabolites

between different groups, these might help reveal the disease mechanisms and advancing diagnostic options. For these projects, further research could be performed to investigate the metabolome, proteome, and protein PTMs in each of the samples. This could add multimodality information for each patient and potentially lead to better understanding of disease mechanisms and more accurate diagnosis. Furthermore, as both projects focused on application of the isotopic/isobaric labeling strategies, limited method development/optimization was pursued. However, I noticed that the usage of SCX for removing excess DiLeu tags in labeled metabolites encountered some problems. For example, SCX cleanup suffered from substantial sample losses. Although this did not affect the quantitation accuracy, it could limit the detection of metabolites. We may try more options for desalting, such HLB column.

In summary, the work described in this dissertation has demonstrated powerful MS platforms that can facilitate identification and quantitation of metabolites, proteins and peptides from tissues and cells. We developed a derivatization strategy for enhancing the detection of metabolites and developed a simplified workflow for high throughput MSI of spheroids and investigated the crosstalk between stromal cells and pancreatic cancer 3D cell culture. Also, we applied our customized isobaric/isotopic tag toolkit for biomarker discovery. These methods or techniques can be readily adopted to investigate other biological systems and will hopefully assist other researchers in their exploration. Hopefully, the analytical methodology described in this thesis, and experimental data obtained here can add some value to future research in analytical chemistry, cancer research, pharmaceutical research, and drug development.

References

(1) Guo, A. D.; Wei, D.; Nie, H. J.; Hu, H.; Peng, C.; Li, S. T.; Yan, K. N.; Zhou, B. S.; Feng, L.; Fang, C.; et al. Light-induced primary amines and o-nitrobenzyl alcohols cyclization as a versatile photoclick reaction for modular conjugation. *Nat Commun* **2020**, *11* (1), 5472.

(2) Klose, R.; Adam, M. G.; Weis, E. M.; Moll, I.; Wustehube-Lausch, J.; Tetzlaff, F.; Oka, C.; Ehrmann, M.; Fischer, A. Inactivation of the serine protease HTRA1 inhibits tumor growth by deregulating angiogenesis. *Oncogene* **2018**, *37* (31), 4260-4272.

(3) Chien, J.; Ota, T.; Aletti, G.; Shridhar, R.; Boccellino, M.; Quagliuolo, L.; Baldi, A.; Shridhar, V. Serine protease HtrA1 associates with microtubules and inhibits cell migration. *Mol Cell Biol* **2009**, *29* (15), 4177-4187.

Appendix I

Publications and Presentations

Publications

1. **Y. Liu**; H. Zhang; WF. Dove; Z. Wang; Z. Zhu; PJ. Pickhardt; M. Reichelderfer; L. Li. Quantification of Serum Metabolites in Early Colorectal Adenomas Using Isobaric Labeling Mass Spectrometry. *J. Proteome Res. 2023, 22, 5, 1483–1491*. 10.1021/acs.jproteome.3c00006

2. **Y. Liu**; G. Li*; T. Gu; L. Li*. Nanosecond Photochemical Reaction for Enhanced Identification, Quantification, and Visualization of Primary Amine-Containing Metabolites by MALDI-Mass Spectrometry. *Anal Chem.* **2022** Mar 8;94(9):3774-3781. doi: 10.1021/acs.analchem.1c03840.

3. **Y. Liu**; H. Zhang; X. Zhong; Z. Li; H. Zetterberg; L. Li. Isotopic N,N-dimethyl Leucine Tags for Absolute Quantification of Clusterin and Apolipoprotein E in Alzheimer's Disease. *J Proteomics.* **2022** Apr 15;257:104507. doi: 10.1016/j.jprot.2022.104507.

4. **Y. Liu**; H. Zhang; J. Johnson; P. Huang; L. Li. A Simplified Sample Preparation Strategy for High-Throughput Mass Spectrometry Imaging of Multicellular Tumor Spheroids. *Manuscript in preparation.*

5. **Y. Liu**; JK. Johnson; DG. Delafield; S. Mancha; J. Zhang; M. Skala; L. Li. Interactions with Stromal Cells Alters the Protein Expression Levels of Pancreatic Cancer Spheroids. *Manuscript in preparation.*

6. G. Li; **Y. Liu**; L. Li. Nanosecond Photochemical Reaction (nsPCR) for Enhanced Mass Spectrometric Identification, Quantification, and Visualization of Metabolites and Neuropeptides. *Methods Mol Biol.* **2022**;2437:143-157. doi: 10.1007/978-1-0716-2030-4_10. (Book chapter)

7. DG. Delafield; HN. Miles; **Y. Liu**; WA. Ricke; L Li. Complementary proteome and glycoproteome access revealed through comparative analysis of reversed phase and porous graphitic carbon chromatography. *Anal Bioanal Chem.* **2022** Jul;414(18):5461-5472. doi: 10.1007/s00216-022-03934-7.

8. H. Zhang; X. Shi; **Y. Liu**; B. Wang; M. Xu; NV. Welham; L Li. On-tissue amidation of sialic acid with aniline for sensitive imaging of sialylated N-glycans from FFPE tissue sections via MALDI mass spectrometry. *Anal Bioanal Chem.* **2022** Jul;414(18):5263-5274. doi: 10.1007/s00216-022-03894-y.

9. J. Huang; D. Wang; RD. Shipman; Z. Zhu; **Y. Liu**; L. Li. Simultaneous enrichment and separation of neutral and sialyl glycopeptides of SARS-CoV-2 spike protein enabled by dual-functionalized Ti-IMAC material. *Anal Bioanal Chem.* **2021** Dec;413(29):7295-7303. doi: 10.1007/s00216-021-03433-1.

10. H. Zhang; M. Xu; X. Shi; **Y. Liu**; Z. Li; JC. Jagodinsky; M. Ma; NV. Welham; ZS. Morris; L Li. Quantification and molecular imaging of fatty acid isomers from complex biological samples by mass spectrometry. *Chem Sci.* **2021** Jun 16; 12(23): 8115–8122. doi: 10.1039/d1sc01614h.

11. H. Zhang; **Y. Liu**; X. Shi; P. Huang; H. Lu; NV. Welham; L. Li. Single-cell lipidomics enabled by high-throughput mass spectrometry imaging at subcellular resolution. Under revision.

12. H. Lu; B. Wang; **Y. Liu**; D. Wang; L. Fields; H. Zhang; M. Li; X. Shi; H. Zetterberg; L. Li. DiLeu isobaric labeling coupled with limited proteolysis mass spectrometry for high-throughput profiling of protein structural changes in Alzheimer's disease. Under review in *Anal Chem*.

13. T. Gu; PK. Liu; YW. Wang; MT. Flowers; S. Xu; **Y. Liu**; DB. Davis; L. Li. Diazobutanone-assisted isobaric labeling of phospholipids and sulfated glycolipids enables multiplexed quantitative lipidomics using LC-MS/MS. Under revision.

Presentations

1. *19th Annual US HUPO Conference on Enabling Technology for Biological Discover, Illinois, USA:* Isobaric Labelling-based Relative Quantification for Discovery of Serum Metabolite Biomarkers Associated with Growing Early Colorectal Adenomas. Y. Liu; H. Zhang; W.F. Dove; Z. Wang; Z, Zhu; P. Pickhardt; M. Reichelderfer; L. Li. 2023. (Poster)

2. *The 2nd CASMS conference on Advancing biological and Pharmaceutical Mass Spectrometry, virtually:* Isobaric Labeling-based Relative Quantification for Discovery of Serum Metabolite Biomarkers Associated with Growing Early Colorectal Adenomas. Y. Liu; H. Zhang; W.F. Dove; Z. Wang; P. Pickhardt; M. Reichelderfer; L. Li. 2022. (Poster)

3. *70th ASMS Conference on Mass Spectrometry and Allied Topics, Minnesota, USA:* Quantification of Serum Proteins in Alzheimer's Disease for Biomarker Discovery. Y. Liu; H. Zhang; X. Zhong; Z. Li; H. Zetterberg; L. Li. 2022. (Poster)

4. *69th ASMS Conference on Mass Spectrometry and Allied Topics, Pennsylvania, USA:* Mass Spectrometry Reveals Molecular Crosstalk between Pancreatic Cancer Cells and Pancreatic Stellate Cells. Y. Liu; J. Johnson; L. Li. 2021. (Poster)

5. *68th ASMS Conference on Mass Spectrometry and Allied Topics:* Nanosecond Photochemical Reaction (nsPCR) for Enhanced Identification and Visualization of Metabolites. Y. Liu; G. Li; L. Li. 2020. (Poster)

6. *35th ASMS Asilomar Conference on Mass Spectrometry, California, USA:* Imaging Mass Spectrometry Reveals Crosstalk between PANC1 and Pancreatic Stellate cells. Y. Liu; J. Johnson; L. Li. 2019. (Poster)