Mass Spectrometry-based Metabolomics: Considerations for Laboratory Testing

by

Paniz Mohajer Jasbi

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved March 2022 by the
Graduate Supervisory Committee:

Carol Johnston, Co-Chair
Haiwei Gu, Co-Chair
Douglas Lake
Karen Sweazea
Natasha Tasevska

ARIZONA STATE UNIVERSITY

May 2022

ABSTRACT

Metabolomics focuses on the study of metabolic changes occurring in various systems and utilizes quantitative and semi-quantitative measurements of multiple metabolites in parallel. Mass spectrometry (MS) is the most ubiquitous platform in this field, as it provides superior sensitivity regarding measurements of complex metabolic profiles in biological systems. When combined with MS, multivariate statistics and advanced machine learning algorithms provide myriad opportunities for bioinformatics insights beyond simple univariate data comparisons. In this dissertation, the application of MS-based metabolomics is introduced with an emphasis on biomarker discovery for human disease detection. To advance disease diagnosis using MS-based metabolomics, numerous statistical techniques have been implemented in this research including principal component analysis, factor analysis, partial least squares-discriminant analysis (PLS-DA), orthogonal PLS-DA, random forest, receiver operating characteristic analysis, as well as functional pathway/enzyme enrichment analyses. These approaches are highly useful for improving classification sensitivity and specificity related to disease-induced biological variation and can help identify useful biomarkers and potential therapeutic targets. It is also shown that MS-based metabolomics can distinguish between clinical and prodromal disease as well as similar diseases with related symptoms, which may assist in clinical staging and differential diagnosis, respectively. Additionally, MS-based metabolomics is shown to be promising for the early and accurate detection of diseases, thereby improving patient outcomes, and advancing clinical care. Herein, the application of MS methods and chemometric statistics to the diagnosis of breast cancer, coccidioidomycosis (Valley fever), and senile dementia (Alzheimer's disease) are presented and discussed. In addition to presenting original research, previous efforts in biomarker discovery will be synthesized and appraised. A Comment will be offered

regarding the state of the science, specifically addressing the inefficient model of repetitive biomarker discovery and the need for increased translational efforts necessary to consolidate metabolomics findings and formalize purported metabolic markers as laboratory developed tests. Various factors impeding the translational throughput of metabolomics findings will be carefully considered with respect to study design, statistical analysis, and regulation of biomedical diagnostics. Importantly, this dissertation will offer critical insights to advance metabolomics from a scientific field to a practical one including targeted detection, enhanced quantitation, and direct-to-consumer considerations.

DEDICATION

This work is dedicated to all those whose lives were lost prematurely to treatable or curable diseases, to those who were not afforded last goodbyes with their loved ones, and those that had insufficient time to settle their affairs before their passing. May this dissertation serve as a small impetus to realizing translational success so that the promise of metabolomics is no longer delayed.

ACKNOWLEDGMENTS

I would like to thank my committee for their scientific guidance and expert opinion; I am incredibly fortunate to have my research activity supervised by such distinguished faculty. I have had the distinct pleasure of successfully collaborating with all of you on various projects and have learned something valuable from each of you. Thank you for having such a profound impact on my scientific development. In particular, I would like to thank my co-chair Dr. Carol Johnston who has routinely advocated on my behalf and overseen my success. Above all, I owe special thanks to my advisor Dr. Haiwei Gu who indoctrinated me in the culture of academia, trained me to think critically about my research, and instilled in me the sincere belief that science can improve the human condition. I am honored to be your student and to carry your academic lineage, and I am eternally indebted to you for your patience, mentorship, and continued support.

I must especially thank Dr. Corrie Whisner for her constant encouragement, her academic mentorship, and her professional friendship. Thank you for showing me the role of kindness and inclusion in academia, I will forever leverage this valuable perspective. I would also like to acknowledge Dr. Christy Alexon for contributing to my professional development; thank you for always being considerate of my teaching time and protecting my ability to do research. Special thanks are owed to Dr. Devina Wadhera for her initial belief in my scientific ability which I have carried with me since my undergraduate studies. I should like to thank my colleague Alex Mohr for his lending me his ears, his eyes, and, from time to time, the great benefit of his mind. Funding from the College of Health Solutions at Arizona State University is also gratefully acknowledged.

Finally, I thank my mother Maryam Ahmadpour for her unwavering love. Especially, I would like to thank my fiancée Cassidy Turner not only for her love, but also her intellectual contributions to much of my work. Cassidy, I adore you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

acetic acid (AcOH), acetonitrile (ACN), Alzheimer's disease (AD), ammonium acetate (NH_4OAc), amyloid β (Aβ), area under curve (AUC), area under receiver operating characteristic (AUROC), Banner Sun Health Research Institute (BSHRI), branched-chain amino acid (BCAA), breast cancer (BC), Centers for Medicare and Medicaid Services (CMS), cerebral amyloid angiopathy (CAA), cerebrospinal fluid (CSF), Clinical Laboratory Improvement Amendments (CLIA), coefficient of variation (CV), community-acquired pneumonias (CAP), complement fixation (CF), computerized tomography (CT), confidence interval (CI), deionized (DI), dried blood spot (DBS), early breast cancer (EBC), electron transport chain (ETC), electrospray ionization (ESI), enzyme immunoassay (EIA), estrogen receptor (ER), exploratory factor analysis (EFA), false discovery rate (FDR), fold change (FC), Food and Drug Administration (FDA), Fred Hutchinson Cancer Research Center (FHCRC), fructose-1,6-bisphosphate (F16BP), gas chromatography (GC), general linear model (GLM), high pathology control (HPC), human epidermal growth factor receptor 2 (HER2), Human Metabolome Database (HMDB), hydrophilic interaction chromatography (HILIC), immunodiffusion (ID), *in vitro* diagnostic (IVD), Ingenuity Pathway Analysis (IPA), inner mitochondrial membrane (IMM), institutional review board (IRB), isopropanol (IPA), laboratory developed test (LDT), leave-one-out cross validation (LOOCV), liquid chromatography (LC), liquid chromatography-tandem mass spectrometry (LC-MS/MS), long-chain fatty acid (LCFA), magnetic resonance imaging (MRI), mass spectrometry (MS), methanol (MeOH), methyl tert-butyl ether (MTBE), microRNA (miRNA), mild cognitive impairment (MCI), Mini-Mental State Examination (MMSE), multiple-reaction-monitoring (MRM), multivariate analysis of variance (MANOVA), N-Methyl-N-(tert-butyldimethylsilyl) trifluoroacetamide (MTBSTFA), Natural Products

Magnetic Resonance Database (NP-MRD), normal control (NC), O-methylhydroxylamine hydrochloride (MeOX), orthogonal partial least squares-discriminant analysis (OPLS-DA), outer mitochondrial membrane (OMM), oxidative phosphorylation (OXPHOS), partial least squares-discriminant analysis (PLS-DA), phosphate buffered saline (PBS), positron emission tomography (PET), post-mortem interval (PMI), principal component analysis (PCA), progesterone receptor (PR), proline dehydrogenase (Prodh), quality control (QC), receiver operating characteristic (ROC), retention time (RT), short-chain fatty acid (SCFA), standard operating protocol (SOP), target of rapamycin (TOR), tricarboxylic acid cycle (TCA cycle), Valley fever (VF), variable importance in projection (VIP)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

It is estimated that 92.15% of all human deaths are attributable to communicable disorders and non-communicable diseases, with less than 8% attributable to natural causes, unintended injuries, and suicide (Roth et al., 2018). Metabolomics, the interdisciplinary study of metabolites (T. W. M. Fan & Lane, 2012), has facilitated numerous advances in the field of systems biology and biomedical diagnostics (Patti et al., 2012), which have in turn led to clinical improvements in diagnostic testing (G. N. Gowda et al., 2008), prognostic staging (X. Zhang et al., 2016), drug target identification (Patel & Ahmed, 2015), and therapeutic monitoring of diseases and disorders (Zhanghan Chen et al., 2019). Given its close proximity to the phenotype (Pinu et al., 2019), the metabolome has proven to be a sensitive and reliable indicator of disease (Zhu et al., 2015). As such, mass spectrometry (MS)-based metabolomics has emerged as a powerful tool for the quantitative and semi-quantitative measurement of metabolites for disease classification (Jasbi, Wang, et al., 2019) as well as the elucidation of pathogenic mechanisms (Jasbi et al., 2021). Firmly established as a mainstay testing method for inborn errors of metabolism (Ismail et al., 2019), MS-based metabolomics has expanded to include the testing of cancers (Wei et al., 2021), infectious diseases (Jasbi, Mitchell, et al., 2019), cardiovascular diseases (Shah et al., 2012), digestive disorders (Wilmanski et al., 2019), neurological disorders (Havelund et al., 2017), and even cognitive disorders such as autism (Glinton & Elsea, 2019) and schizophrenia (D. Wang et al., 2019).

Indeed, the field of metabolomics has enjoyed a considerable increase in scientific interest over the last two decades. Regrettably, the number of laboratory developed tests (LDTs) has lagged behind this massive accumulation of metabolomics data and knowledge. As can be seen in **Figure 1.1**, a PubMed search of “metabolomics” and “study” returns 36,904 studies published since 2002; of these, 24,343 were published in just the last five years. In contrast, a similar

PubMed search of “metabolomics” and “laboratory developed test” returns just 691 results in the same time period. Although a plethora of candidate markers for various disorders and diseases have been collectively identified by the field of metabolomics (Xia et al., 2013), less than 5% have been formalized as LDTs (Dias & Koal, 2016). Compared to unvalidated candidate markers, LDTs offer significant advantages to disease assessment including increased accessibility, simple scalability, understandable results, and an enhanced educational effect (Lichtenberg et al., 2021). Nevertheless, translation of preliminary findings of metabolomic biomarker studies to formalized LDTs is prerequisite to the realization of these benefits. Although the U.S. Food and Drug Administration (FDA) has some limited authority in determining the complexity of LDTs and their potential classification as *in vitro* diagnostics (IVDs) and biomedical devices (Sharfstein, 2015), laboratories in the U.S. that develop LDTs are regulated by the Centers for Medicare and Medicaid Services (CMS) under the Clinical Laboratory Improvement Amendments (CLIA). By default, CMS categorizes all LDTs as high complexity tests and requires them to meet the most stringent CLIA standards (Lichtenberg et al., 2021).

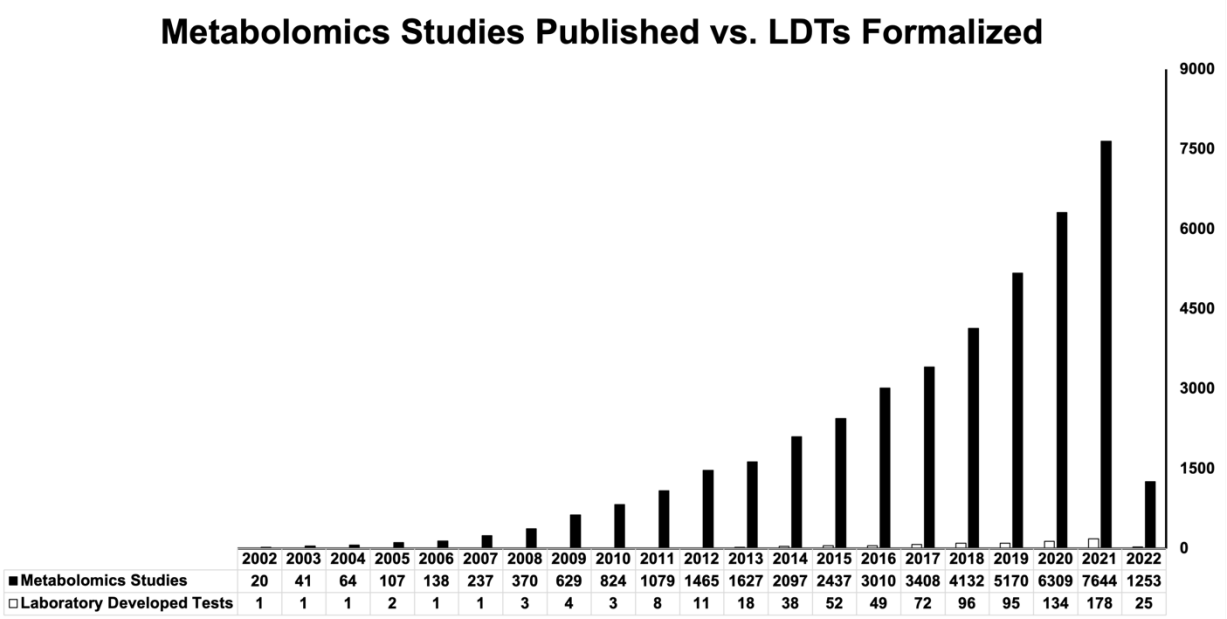


Figure 1.1 Bar graph of published metabolomics studies and formalized LDTs from 2002—2022. X-axis represents year, Y-axis corresponds to number. Data for “Metabolomics Studies” series obtained via PubMed search of “metabolomics” and “study”; data for “Laboratory Developed Tests” series obtained via PubMed search of “metabolomics” and “laboratory developed test”. Search results current as of 2/19/2022. LDT, laboratory developed test.

Numerous challenges impede the translation of metabolomics studies to LDTs (Pinu et al., 2019) and, given the disproportionate ratio of generated knowledge to applied knowledge, a careful enumeration of these bottlenecks and a thoughtful discussion of possible solutions is warranted. First and foremost, only vague government guidelines for the assessment of analytical validity and clinical utility are currently offered, rendering an altogether unclear market access pathway and disincentivizing LDT development. Additionally, the results of many metabolomics studies are difficult to reproduce, and the significant cost of metabolomics analysis prohibits replication and validation efforts. Furthermore, the depth and richness of metabolomics data complicates communication of results to the end-user, often rendering results only interpretable to scientists in the field. Nevertheless, the application of metabolomics to medical laboratory

testing offers a promising translational opportunity and represents an excellent return on investment in terms of research and development dollars. In the U.S., 13 billion laboratory tests are performed each year across more than 2,000 CLIA-certified testing sites (Horvath, 2013); even so, laboratory testing comprises just 2.3% of annual healthcare costs yet supports 60-70% of all clinical decisions (Sikaris, 2017). Therefore, given the wealth of translational opportunities and the potential payoff of successful translational efforts, a Comment in the literature offering a critical perspective on the state of the science and considerations for future directions may guide biomarker consolidation and LDT formalization.

This dissertation will present MS-based metabolomics research in addition to synthesizing results as a Comment to the literature. Specifically, metabolomics assessment of breast cancer, coccidioidomycosis (Valley fever), and senile dementia (Alzheimer's disease) will be presented and discussed. In addition to providing methods for the measurement and analysis of disease-related metabolites, this dissertation will compile a detailed catalogue of pre-analytical considerations to study design, spectral data collection, and formal analysis and interpretation that can better facilitate LDT development in future studies. The aim of this Comment will be to identify specific considerations relating to study design and analytics which, if observed, can improve LDT formalization in metabolomics and facilitate enhanced clinical applications.

CHAPTER 2

BREAST CANCER DETECTION USING TARGETED PLASMA METABOLOMICS

(Published in *Journal of Chromatography B*, 2019, 1105: 26—37)

Abstract

Breast cancer (BC) is a major cause of human morbidity and mortality, especially among women. Despite the important role of metabolism in the molecular pathogenesis of cancer, robust metabolic markers to enable enhanced screening and disease monitoring of BC are still critically needed. In this study, a targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS) metabolic profiling approach is presented for the identification of metabolic marker candidates that could enable highly sensitive and specific detection of all-stage as well as early-stage BC. In this targeted approach, 105 metabolites from more than 35 metabolic pathways of potential biological relevance were reliably detected in 201 plasma samples taken from two groups of subjects (102 BC patients and 99 healthy controls). The results of our general linear model and partial least squares-discriminant analysis (PLS-DA) informed the construction of a novel 6-metabolite panel of potential biomarkers. A receiver operating characteristic (ROC) curve generated based on an improved PLS-DA model showed relatively high sensitivity (0.80), specificity (0.75), and area under the receiver-operating characteristic curve (AUROC = 0.89). Similar classification performance of the model was observed for detection of early-stage BC (AUROC = 0.87, sensitivity: 0.86, specificity: 0.75). Bioinformatics analyses revealed significant disturbances in arginine/proline metabolism, tryptophan metabolism, and fatty acid biosynthesis. Our univariate and multivariate results indicate the effectiveness of this metabolomics approach for all-stage as well as early-stage BC diagnosis; our bioinformatics results indicate affected pathways related to tumor growth, metastasis, and immune escape mechanisms. Future studies should validate these results using more samples from different locations.

Introduction

Globally, breast cancer (BC) is the second most common type of cancer and a major cause of human morbidity and mortality, disproportionately affecting women (Siegel et al., 2018). It is reported that BC alone accounts for 25% of all cancer cases and 15% of all cancer deaths among females (Torre et al., 2017). The American Cancer Society (2018) strongly recommends that women with an average risk of developing BC, as determined by a family history-based risk assessment, undergo regular screening mammography beginning at age 45 (Smith et al., 2018). Although mammography has been shown to have high sensitivity (93%) for the detection of symptomatic BC (Jiang et al., 2016), it is far less effective for the detection of early-stage BC. As such, it has been suggested that regular physical examination is of comparable importance, and perhaps the best method of early detection (Shen & Zelen, 2001). Sentinel lymph node biopsy remains the gold standard for detection of BC with lymph node involvement, but the major disadvantages include invasiveness, potential risk of complication, and the inherent inability for detection of early-stage BC (Lyman et al., 2017). The average 5-year BC survival rate is roughly 90%, but can be as high as 99% for those diagnosed and treated with early-stage, localized disease (stages I and II), which regrettably accounts for only 61% of BC patients (DeSantis et al., 2017). Therefore, noninvasive detection methods with high sensitivity and specificity, which would enable the early diagnosis and timely treatment of BC, are still critically needed.

A number of new BC detection methods have been developed at the molecular level, particularly based on genetic/phenotypic testing, computer-aided technology and biomarker identification, such as immunohistochemistry (Onitilo et al., 2009), and serum circulating microRNA (miRNA) profiling (Fong et al., 2015; Hayes et al., 2014), which have shown some favorable evidence of enhanced BC detection. Although efforts to typify ER/PR and HER2 status in order to predict BC pathogenesis and disease progression are promising methods of early detection, they often lead only to the assertion of broad probabilities (Onitilo et al., 2009).

A characteristic feature of cancer is its abnormal metabolism. Consequently, recent cancer studies have endeavored to monitor levels of differential metabolites from biologically relevant pathways, resulting in improved BC subtype identification and diagnosis (Y. Fan et al., 2016). Metabolomics, the scientific study of comprehensive sets of metabolites present in biological samples, offers new avenues for advanced disease biomarker discovery (Bain et al., 2009; Dunn, Broadhurst, Atherton, et al., 2011; T. W. M. Fan & Lane, 2011; Y. Fan et al., 2016; Fernie et al., 2004; Fiehn, 2002; G. A. N. Gowda & Raftery, 2013; Griffin et al., 2011; Gu et al., 2012; Hakimi et al., 2016; Halama et al., 2013; Jové et al., 2017; Lindon & Nicholson, 2014; Patti et al., 2012; Reaves & Rabinowitz, 2011; Scalbert et al., 2009; Xia et al., 2013). Mass spectrometry-based metabolic profiling has emerged as a powerful analytical platform for analysis of metabolic alterations caused by various cancers, which has led to substantial advances in cancer diagnosis, pathogenesis clarification, and identification of potential drug targets for clinical treatment (Beger, 2013; Mishra & Ambs, 2015; More et al., 2018). Previous efforts to typify an associated metabolic profile of BC have typically assumed global profiling approaches to differentiate disease patients from healthy controls, employing gas chromatography-mass spectrometry (GC-MS) (Ingram et al., 1997), nuclear magnetic resonance (NMR) spectroscopy (Katz-Brull & Degani, 1996), flow injection analysis-tandem mass spectrometry (FIA-MS/MS) (Halama et al., 2011), and liquid chromatography time-of-flight mass spectrometry (LC-TOF-MS) (Guo et al., 2017). In contrast to global approaches, some metabolomics studies, citing well-known alterations in cancer metabolism such as the Warburg effect and glutamine addiction, have argued for MS-based testing of particular metabolites associated with decreased oxidative phosphorylation and increased glycolysis and lactic acid fermentation (Armitage & Barbas, 2014), while others have emphasized the need to therapeutically target glutamine addiction as breast cancer cells require glutamine that is vital to cancer cell growth and proliferation (D. R. Wise & Thompson, 2010). These metabolic differences, which can be hallmarks of all cancers, also suggest that

metabolomics is a highly promising approach to discover significant risk factors for BC and develop sensitive and specific BC biomarkers.

Although potential applications of metabolomics findings until now are encouraging and their significance is doubtless, limited work has focused on targeted plasma metabolic profiling of BC for accurate diagnosis and pathogenesis clarification. Few metabolomics studies of BC biomarker discovery have, to date, focused on multicenter replication and validation. In this study, a targeted plasma metabolic profiling approach optimized for the detection of over 400 metabolites reflective of more than 35 metabolic pathways of potential biological relevance is reported. A total of 201 plasma samples from BC and control subjects from two different clinical centers were analyzed, and potential biomarkers were selected by means of univariate significance testing and multivariate model construction and validation. In the current study, two sets of control samples taken from two collection sites were analyzed, and only metabolites that were not found to be significantly different between control groups were retained for further comparison of metabolic profiles between breast cancer patients and their healthy counterparts.

Methods

Reagents

Acetonitrile (ACN), methanol (MeOH), acetic acid (AcOH), and ammonium acetate, all LC-MS grade, were purchased from Fisher Scientific (Pittsburgh, PA). Standard compounds (purity >95-99%) corresponding to measured metabolites were purchased from Sigma-Aldrich (Saint Louis, MO) and Fisher Scientific (Pittsburgh, PA). Internal standards (stable ^{13}C -labeled tyrosine and lactate; purity >99%) were purchased from Cambridge Isotope Laboratories (Tewksbury, MA, UK).

Clinical Samples

Clinical samples were purchased from the Fred Hutchinson Cancer Research Center Breast Specimen Repository (FHCRC; Seattle, WA) and Bloodworks Northwest (Seattle, WA). This analysis was deemed IRB exempt, since these samples were commercially purchased. Informed consent was obtained from all participants (both BC patients and healthy controls) before sample collection at the research institutes. All participants were evaluated, and blood samples were obtained after overnight fasting. In total, 201 subject samples were included in the study, among which there were 102 BC patients and 99 healthy controls. The controls were age-matched with BC patients. Among the controls, 31 of 99 samples were allocated from FHCRC (Seattle, WA), while 68 control samples were allocated from Bloodworks Northwest (Seattle, WA). All clinical samples of BC subjects were allocated from FHCRC.

Sample Preparation

The sample preparation protocol was modeled on previous studies (Beckonert et al., 2007; Dunn, Broadhurst, Begley, et al., 2011). Frozen samples were thawed overnight under 4°C, and 50 µL of each plasma sample was placed in a 2 mL Eppendorf vial. Protein precipitation and metabolite extraction were performed by adding 300 µL of methanol. The mixture was then vortexed for 2 min and stored at -20ºC for 30 min, followed by sonication in an ice bath for 10 min and subsequent centrifugation at 14,000 RPM for 20 min at 4°C. The supernatant (150 µL) was collected into a new Eppendorf vial, and dried using a Vacufuge Plus evaporator. The dried samples were reconstituted in 500 µL of 5 mM ammonium acetate in 40% H_2O/60% ACN + 0.2% acetic acid containing 5.13 µM L-tyrosine-$^{13}C_2$ and 22.5 µM sodium-L-lactate-$^{13}C_2$. The two stable isotope-labeled internal standards were added to each sample to monitor system performance. A pooled sample, which was a mixture of plasma from all BC patients and healthy controls, was

extracted using the same procedure as previously described. This sample was used for quality control (QC) purposes and was analyzed once every 10 study samples.

Liquid Chromatography and Mass Spectrometry Conditions

The targeted LC-MS/MS method used here was modeled after that developed and used in a growing number of studies (Carroll et al., 2015; Gu et al., 2015, 2016; Jové et al., 2017; More et al., 2018; Sperber et al., 2015; Zhu et al., 2014). Briefly, LC-MS/MS experiments were performed on a Waters Acquity I-Class UPLC TQS-micro MS system (Milford, MA). Each sample was injected twice, 2 µL and 5 µL for analysis using positive and negative ionization modes, respectively. Chromatographic separation was performed on a Waters Xbridge BEH Amide column (2.5 µm, 2.1 x 150 mm) at 40°C. The flow rate was 0.3 mL/min. For positive mode, the mobile phase was composed of Solvents A (5 mM ammonium acetate in H_2O with 0.1% acetic acid) and B (ACN with 0.1% acetic acid). For negative mode, Solvent A was 10 mM ammonium bicarbonate in H_2O, and Solvent B was ACN. The LC gradient conditions were the same for both positive and negative ionization modes. After an initial 1.5 min isocratic elution of 10% Solvent A, the percentage of Solvent A was increased linearly to 65% at t = 9 min. The percentage of A then remained the same (65%) for 5 min (t = 14 min), after which the percentage of A was reduced to 10% at t = 15 min to prepare for the next injection. The total experimental time for each injection was 30 min. Metabolite identities were confirmed by spiking mixtures of standard compounds into prepared plasma samples. Extracted MRM peaks were integrated using the TargetLynx software (Waters, Milford, MA).

Data Processing and Statistical Analysis

After exporting from TargetLynx software, data were $\log_{10}$-transformed to approximate normality, and general linear models (GLMs) were used for comparison of metabolite levels

between the sets of control samples from differing centers as well as between BC patients and healthy controls. Age was included as a covariate in all univariate models to control for potential confounding effects. The Benjamini-Hochberg false discovery rate (FDR) control was implemented to correct for multiple comparisons. The FDR *q*-value threshold for significant markers was set at 0.05. Partial correlation analysis was used to calculate correlation coefficients among metabolites. These statistical analyses were performed using SPSS 22.0 (SPSS Inc.; Chicago, IL).

Partial least squares discriminant analysis (PLS-DA) was performed using $\log_{10}$-transformed, Pareto scaled data to construct classification models. For Pareto scaling, data were mean-centered and divided by the square root of the standard deviation of each variable. An internal 7-fold (*n* was automatically selected by the software) cross-validation was carried out to estimate the performance of PLS-DA models. Model validation was also performed using a 300-iteration permutation test. R^2 represents the explanatory capacity of the model, and Q^2 signifies the predictive capacity of the model. The PLS-DA model was constructed using SIMCA-P 14.1 software (Umetrics, Umeå, Sweden). The differential metabolites were obtained based on variable importance in projection values (VIP > 1) taken from an initial PLS-DA model and significant *q*-values ($q < 0.05$) derived from the corrected GLM. These differential metabolites were then selected as a panel of markers to construct a second PLS-DA model for discrimination between BC patients and healthy controls. Area under the receiver operating characteristic curve (AUROC) was then calculated to evaluate the classification performance of the PLS-DA model.

Exploratory factor analysis (EFA) was conducted using Comprehensive Exploratory Factor Analysis (CEFA) software version 2.0 (Columbus, OH) to determine underlying pathways affected in BC patients. We used EFA, a data reduction and analytic technique, to discover patterns of latent variables that could increase the interpretability of the data. Rotation was conducted to achieve simple structure and increase pathway identification. Both analysis and

rotation were unsupervised, and all metabolite names were replaced as variable numbers to maintain unbiased interpretation of factor loadings. Parallel analysis was conducted with a random data matrix of the same order as the experimental data. Factors were retained if and only if they accounted for more variance than the random data, as evidenced by their respective eigenvalues.

Pathway analyses were performed and visualized using both Ingenuity Pathway Analysis (IPA) (Krämer et al., 2014) and the MetaboAnalyst 4.0 software package (Chong et al., 2018).

Results

Targeted Metabolic Profiles of BC versus Healthy Controls

A total of 102 BC patients and 99 healthy controls were included in the study. There was no statistically significant difference in age ($p > 0.05$) between BC patients and healthy controls as calculated by the Mann-Whitney *U* test. The clinical information of BC patients is shown in **Table 2.1**.

Table 2.1 Demographic and Clinical Characteristics of Study Participants

		Breast cancer (*N*=102)	Healthy controls (*N*=99)
Age (years), mean (SD)		54.6 (10.4)	51.6 (12.3)
Menopause status, *n* (%)	Premenopausal	43 (42.2)	
	Perimenopausal	6 (5.9)	
	Postmenopausal	53 (52.0)	
Cancer stage, *n* (%)	I	24 (23.5)	
	II	42 (41.2)	
	III	36 (35.3)	
Cancer type, *n* (%)	Ductal	92 (90.2)	
	Lobular	8 (7.8)	
	Ductal and lobular	2 (2.0)	
ER status, *n* (%)	Negative	7 (6.9)	
	Positive	95 (93.1)	
PR status, *n* (%)	Negative	14 (13.7)	
	Positive	88 (86.3)	
HER2 status, *n* (%)	Negative	69 (67.6)	
	Borderline	13 (12.7)	
	Positive	20 (19.6)	
Molecular subtype, *n* (%)	ER/PR+ Her2+	18 (17.6)	
	ER/PR+ Her2-	65 (63.7)	
	ER/PR- Her2+	2 (2.0)	
	ER/PR- Her2-	4 (3.9)	
	Unknown	13 (12.7)	
Triple-negative (ER-, PR-, HER2-), *n* (%)	Yes	4 (3.9)	
	No	98 (96.1)	

In the current study, we used a large-scale, targeted LC-MS/MS method for reliable and comprehensive plasma metabolite detection. Using this paradigm, targeted analysis of 400+ MRM transitions was achieved for metabolites spanning over 20 different chemical classes (e.g., acyl glycines, bile acids, cyclic amines, etc.) from more than 35 metabolic pathways (e.g., vitamin and cofactor metabolism, citric acid cycle, lipid metabolism, etc.) across positive and negative ionization modes (**Supplementary Table 2.1**). In total, we found that 105 metabolites were detectable in the QC sample with signal-to-noise ratios (S/Ns) > 3. Moreover, 98 of the 105 detected metabolites were observed in >90% of all study samples. Following normalization by averaged values from the QC injection data, relative levels of the 98 metabolites had a median coefficient of variation (CV) value of 5.1%, with ~87% of metabolites having CV < 15% (**Supplementary Figure 2.1**).

Sixty-seven of the 98 metabolites showed statistical significance (FDR $q < 0.05$) between BC patients and controls (**Supplementary Table 2.2**). Since the control samples were from two distinct institutes, we compared the metabolite levels between the two sets of control samples using GLMs with age adjustment. Between the two sets of controls, 68 of 98 metabolites were observed to be significant ($q < 0.05$). Of the 67 metabolites shown to be significantly different between cancer patients and controls, 53 of them were also observed to be statistically significant between the two sets of controls (**Supplementary Table 2.2**). Given this overlap, a real potential for confounding exists and, as such, any true comparison of BC patients and healthy controls necessitates a comparison of metabolites that were not found to be significantly different between the two control cohorts. There were 30 metabolites which exhibited no statistically significant differences between the two groups of control samples ($q > 0.05$) (**Supplementary Figure 2.2**). Thus, these 30 metabolites were used for subsequent analyses. An over-representation analysis was performed to evaluate the scope and depth of the metabolic profile generated from these 30 compounds. As shown in **Supplementary Figure 2.3**, the metabolic profile used for comparison of cancer and control subjects is reflective of 27 pathways, and significantly representative of 7 pathways. Associations among these 30 metabolites are shown in **Figure 2.1**. Most metabolites of BC patients were observed to be more positively correlated with each other than those of controls, and correlations of certain metabolites in the two groups showed inverse relationships **(Figure 2.1)**.

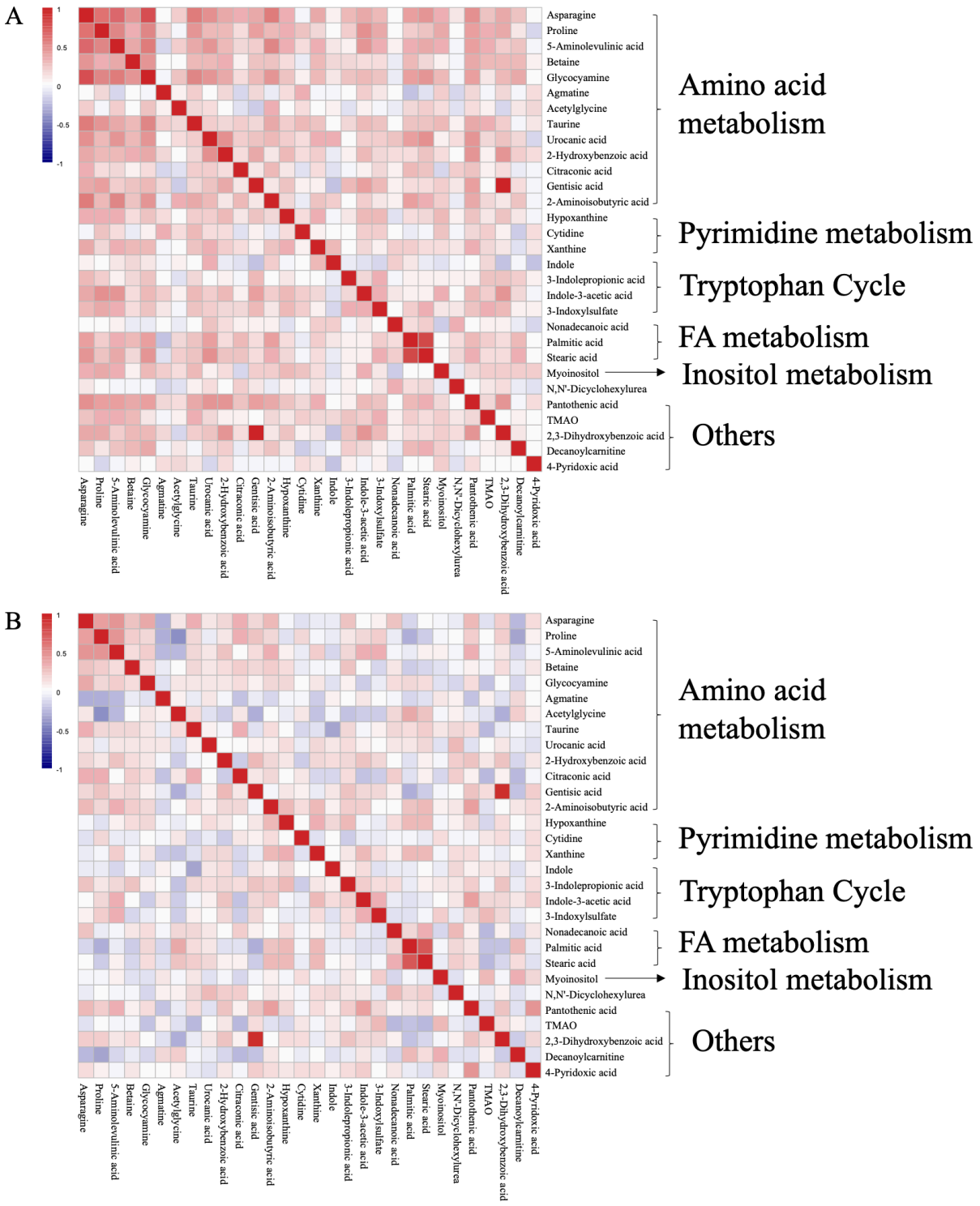


Figure 2.1 Correlation coefficients among 30 metabolites after adjustment for age in (A) BC patients and (B) healthy controls. Blue and red represent negative and positive correlations, respectively. BC, breast cancer; FA, fatty acid, TMAO, trimethylamine N-oxide.

As shown in **Table 2.2**, 18 of these 30 metabolites showed statistical significance between BC patients and healthy controls ($q < 0.05$). With the exception of hypoxanthine, acetylglycine, and three metabolites related to fatty acid metabolism (nonadecanoic acid, palmitic acid, and stearic acid), all other metabolites were decreased in BC patients compared to healthy controls. Considerable fold change values (FC > 1) and significant *p*-values, calculated based on mean ratios and univariate testing respectively, were also observed for 17 of the 18 significant between-group metabolites (**Table 2.3**), when comparing staged cancer patients and healthy controls. Of these, 8 metabolites were observed to be significant between stage I/controls, whereas 9 were significant between stage II/controls. When comparing early stage BC patients (stages I and II) to control subjects, 6 metabolites were found to be significant; of these 6 metabolites, 4 were found to be mutually significant in both stage I/control and stage II/control comparisons. Thirteen metabolites were observed to be significant between stage III patients and controls. Similarly, thirteen metabolites had $p < 0.001$ (with FC ranging from 0.63 to 1.60) when comparing all-stage BC patients to healthy controls (**Table 2.2**).

Table 2.2 Significant Metabolites for Comparison of BC Patients and Healthy Controls

Metabolite	AUC	*p*-value[a]	FDR *q*-value	Fold Change (BC/Control)	VIP[b]
Proline	0.76	1.10E-12	3.31E-11	0.73	1.10
Myoinositol	0.71	1.60E-08	2.40E-07	0.78	1.02
2-Hydroxybenzoic acid	0.70	3.94E-08	3.94E-07	0.63	1.60
Gentisic acid	0.66	2.10E-06	1.58E-05	0.68	1.09
Hypoxanthine	0.68	1.64E-04	6.16E-04	1.60	1.46
2,3-Dihydroxybenzoic acid	0.60	1.97E-04	6.55E-04	0.72	1.05
Palmitic acid	0.66	1.19E-05	7.14E-05	1.29	0.61
5-Aminolevulinic acid	0.65	6.41E-05	3.20E-04	0.81	0.61
Pantothenic acid	0.63	1.30E-04	5.55E-04	0.69	0.80
Cytidine	0.66	2.26E-04	6.77E-04	0.77	0.91
Stearic acid	0.64	3.34E-04	9.10E-04	1.18	0.51
4-Pyridoxic acid	0.61	3.73E-04	9.33E-04	0.75	0.82
TMAO	0.62	9.61E-04	2.22E-03	0.73	0.83
Agmatine	0.64	2.09E-03	4.49E-03	0.84	0.55
Indole-3-acetic acid	0.56	7.28E-03	1.46E-02	0.88	0.75
Indole	0.53	7.61E-03	1.43E-02	0.85	0.71
Acetylglycine	0.60	9.64E-03	1.70E-02	1.40	0.49
Nonadecanoic acid	0.62	1.23E-02	2.05E-02	1.17	0.52
Betaine	0.61	7.39E-02	1.17E-01	0.93	0.19
Asparagine	0.53	1.01E-01	1.52E-01	1.02	0.42
Glycocyamine	0.55	1.23E-01	1.76E-01	0.97	0.26
Citraconic acid	0.58	1.82E-01	2.48E-01	1.14	0.43
2-Aminoisobutyric acid	0.59	2.01E-01	2.62E-01	1.19	0.43
3-Indolepropionic acid	0.56	2.60E-01	3.25E-01	1.17	0.91
N,N'-Dicyclohexylurea	0.50	2.60E-01	3.12E-01	1.04	0.23
Taurine	0.51	3.07E-01	3.54E-01	1.01	0.19
Decanoylcarnitine	0.55	3.71E-01	4.12E-01	1.03	0.23
Xanthine	0.51	4.46E-01	4.78E-01	0.99	0.20
3-Indoxylsulfate	0.58	6.80E-01	7.04E-01	0.99	0.11
Urocanic acid	0.53	9.10E-01	9.10E-01	1.04	0.31

[a]*p*-values calculated from univariate GLM testing.
[b]VIP values obtained from the age-enhanced PLS-DA model (see **Figure 2.2**).

Table 2.3 Significant Metabolites for Comparison of Staged BC Patients and Controls

	stage I *vs.* controls		stage II *vs.* controls		stage I & II *vs.* controls[a]		stage III *vs.* controls	
	p	FC	*p*	FC	*p*	FC	*p*	FC
2-Hydroxybenzoic acid	1.8E-03	0.55	3.6E-04	0.64	2.7E-03	0.60	2.5E-04	0.65
Myoinositol	1.9E-04	0.69	1.5E-03	0.82	2.1E-04	0.76	1.0E-04	0.80
Proline	1.0E-05	0.75	3.2E-05	0.81	2.9E-04	0.73	5.5E-08	0.68
Palmitic acid	1.1E-02	1.40	6.8E-03	1.24			4.4E-02	1.23
Hypoxanthine	7.1E-04	2.67	2.3E-03	1.55	1.8E-04	2.21		
Indole	1.7E-02	0.78						
Gentisic acid	1.9E-02	0.63			3.0E-03	0.59	6.3E-04	0.64
5-Aminolevulinic acid	2.3E-02	0.88					1.1E-02	0.78
4-Pyridoxic acid			6.1E-03	0.68			2.8E-02	0.69
Cytidine			1.6E-02	0.77			7.8E-04	0.78
Nonadecanoic acid			2.1E-02	1.15				
Stearic acid			2.4E-03	1.27				
Agmatine							8.0E-04	0.76
Indole-3-acetic acid							8.8E-03	0.73
Pantothenic acid							1.1E-02	0.61
2,3-Dihydroxybenzoic acid					1.8E-02	0.66	1.7E-02	0.77
Glycocyamine							3.3E-02	0.88

[a]Early-stage BC is regarded as stages I and II.

We compared patients with different molecular subtypes (ER/PR$^+$, HER2$^+$ *vs.* ER/PR$^+$, HER2$^-$), and compared triple negative (ER$^-$, PR$^-$, and HER2$^-$) with non-triple negative patients. These results are shown in **Table 2.4**. No significant difference in metabolites was observed between these groups (all FDR $q > 0.05$). Interestingly, 15 metabolites were observed to have *p*-values less than 0.05 when comparing ER, PR, HER2 status, and cancer stage among BC patients, although no significant differences in 30 metabolites were observed between the above-mentioned comparison groups after multiple controlled comparisons (FDR) (**Table 2.4**).

Table 2.4 Differences in Metabolites of Patients between Cancer Stages, and Different ER, PR, HER2 Status

Metabolites	Cancer stages				ER status			PR status			HER2 status		
	p	*q*	FC (II/I)	FC (III/I)	*p*	*q*	FC (ER+/ER-)	*p*	*q*	FC (PR+/PR-)	*p*	*q*	FC (HER2+/HER2-)
3-Indolepropionic acid	0.026	0.811	1.41	1.88									
Betaine					0.010	0.105	1.08	0.033	0.346	0.91			
Palmitic acid					0.013	0.099	1.21						
Asparagine					0.015	0.093	1.05	0.035	0.271	0.94			
2,3-Dihydroxybenzoic acid					0.021	0.109	1.49						
3-Indoxylsulfate					0.028	0.123	1.94						
Indole-3-acetic acid					0.034	0.130	1.34				0.046	0.479	1.26
Hypoxanthine					0.041	0.143	1.94						
Gentisic acid					0.042	0.130	1.86						
Stearic acid					0.047	0.133	1.17						
Taurine					0.049	0.126	1.34	0.020	0.303	0.95			
5-Aminolevulinic acid					0.049	0.118	1.06						
Proline								0.005	0.142	0.92	0.030	0.462	1.13
Pantothenic acid								0.049	0.303	0.77			
Cytidine											0.014	0.439	1.28

Biomarker Selection and Evaluation of Classification Performance

To further explore potential biomarkers for discrimination between BC patients and healthy controls, levels of the 30 comparative metabolites were selected to establish an initial PLS-DA model. As can be seen in **Supplementary Figure 2.4**, a separation trend was observed in the initial PLS-DA score plot [R^2X (cum) = 0.291, R^2Y (cum) = 0.398, Q^2 (cum) = 0.312]. Clinical factors (i.e., gender, age, and medication) have often been incorporated in the building of predictive or diagnostic clinical models, and such variables have recently been used to enhance metabolite biomarker models (Rhee et al., 2013). To enhance the VIP metabolite model, age was included as a clinical factor. The enhanced metabolite model (**Figure 2.2A**) showed a distinct separation trend between the two groups [R^2X (cum) = 0.709, R^2Y (cum) = 0.481, Q^2 (cum) = 0.417], which indicated better performance than the initial PLS-DA model (**Supplementary Figure 2.4**). To validate the reliability of the enhanced prediction model, a permutation test (n = 300) was conducted (**Figure 2.2B**). The Q^2-intercept value (-0.158) of the predictive model was lower than 0.05, indicating the model to be statistically sound. VIP values were obtained as listed in **Table 2.2**.

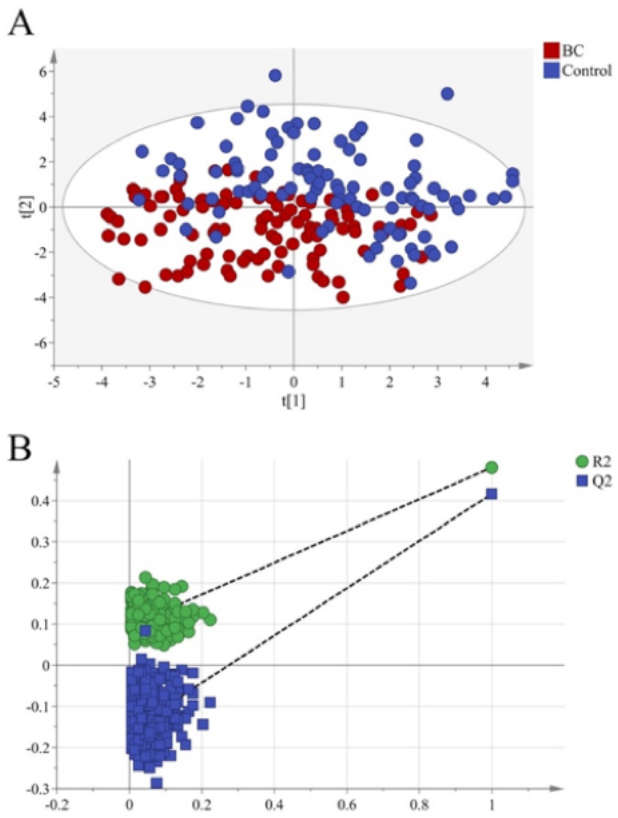


Figure 2.2 The enhanced PLS-DA model for discrimination between breast cancer patients and healthy controls [R^2X (cum) = 0.709, R^2Y (cum) = 0.481, Q^2 (cum) = 0.417]. (A) Score plot of the PLS-DA model: data were log_{10}-transformed, and Pareto scaled. Thirty metabolites and age were used to construct this model. (B) Statistical validation of the PLS-DA model by permutation testing (n = 300) [Q^2-intercept = -0.158]. PLS-DA, partial least squares-discriminant analysis.

According to the VIP values from the enhanced PLS-DA model (VIP > 1) and the *q*-values from FDR-controlled comparisons ($q < 0.05$), 6 statistically significant, highly predictive features were retained for further analysis (**Figure 2.3**). A third PLS-DA model was built using the 6 differential metabolites and subject age to distinguish BC patients from healthy controls. As can be seen in **Figure 2.4**, the resulting PLS-DA model proved to be powerful in distinguishing BC patients from healthy controls, with an AUROC of 0.89 (95% CI: 0.85-0.93, sensitivity = 0.80,

specificity = 0.75), which was more explanatory than that of each individual metabolite (see **Table 2.2**).

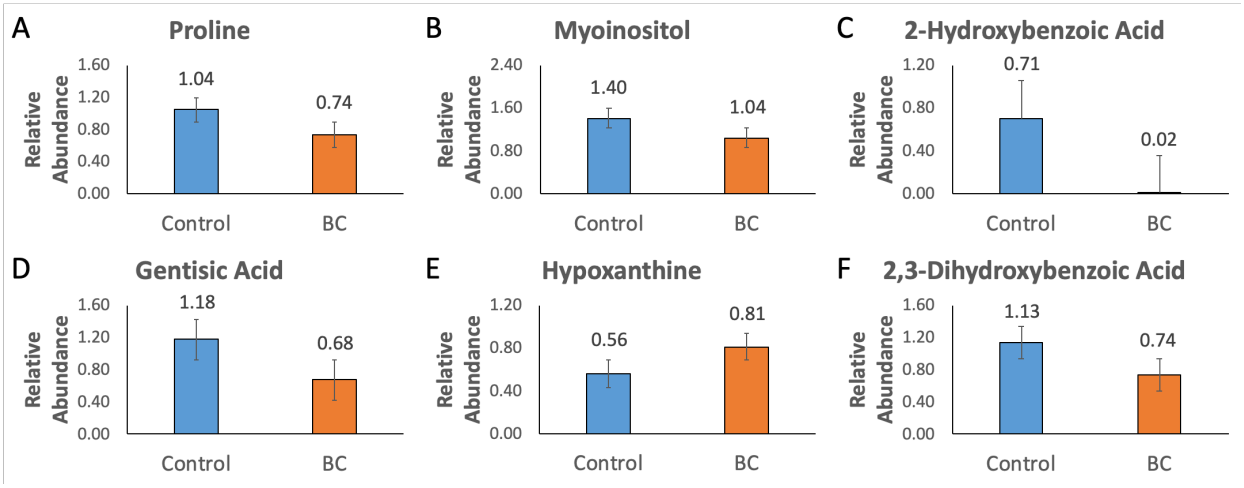


Figure 2.3 Box plots of 6 differential metabolites with both $q < 0.05$ and VIP > 1 for comparison between BC patients and healthy controls: (A) Proline, (B) Myoinositol, (C) 2-Hydroxybenzoic acid, (D) Gentisic acid, (E) Hypoxanthine, and (F) 2,3-Dihydroxybenzoic acid. BC, breast cancer; VIP, variable projection in importance.

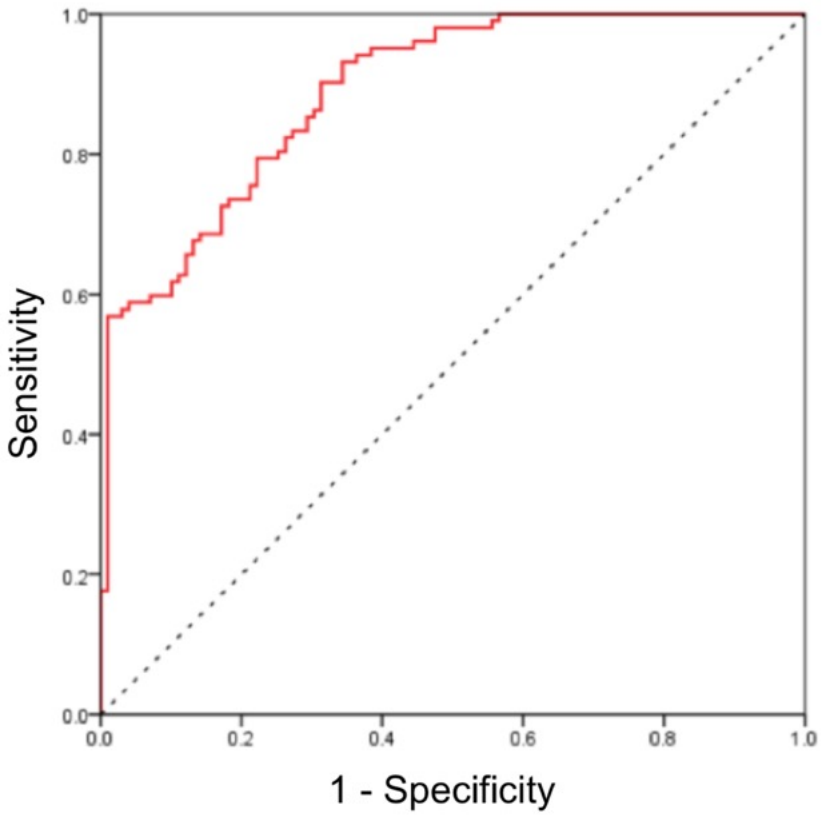


Figure 2.4 ROC curve illustrating the classification performance of biomarker panel with 6 differential metabolites and age for distinguishing between BC patients and healthy controls. [AUROC=0.89, 95% CI: 0.85-0.93, sensitivity: 0.80, specificity: 0.75]. AUC, area under curve; AUROC, area under receiver operating characteristic; BC, breast cancer; CI, confidence interval.

To evaluate the usefulness of this novel panel of 6 biomarkers in detecting early-stage breast cancer, control subjects were analyzed against only stage I and II BC patients using a combination of univariate testing, chemometric analysis, and ROC evaluation. As shown in **Table 2.3**, these 6 differential metabolites were significant at the 0.05 level when comparing stage I and II BC patients to controls, as determined by GLM univariate testing. Moreover, an orthogonal PLS-DA (OPLS-DA) model constructed using these 6 metabolites showed appreciable differences between groups (**Figure 2.5A**). Furthermore, as presented in **Figure 2.5B**, ROC analysis of the OPLS-DA model showed good classification performance (AUROC=0.87, 95% CI: 0.82-0.92). As

indicated by these results, the biomarker panel presented herein serves not only to distinguish BC patients from healthy controls but is also capable of discriminating stage I and II patients with localized disease from healthy control subjects with relatively high diagnostic accuracy, comparable to that of the all-stage cancer model.

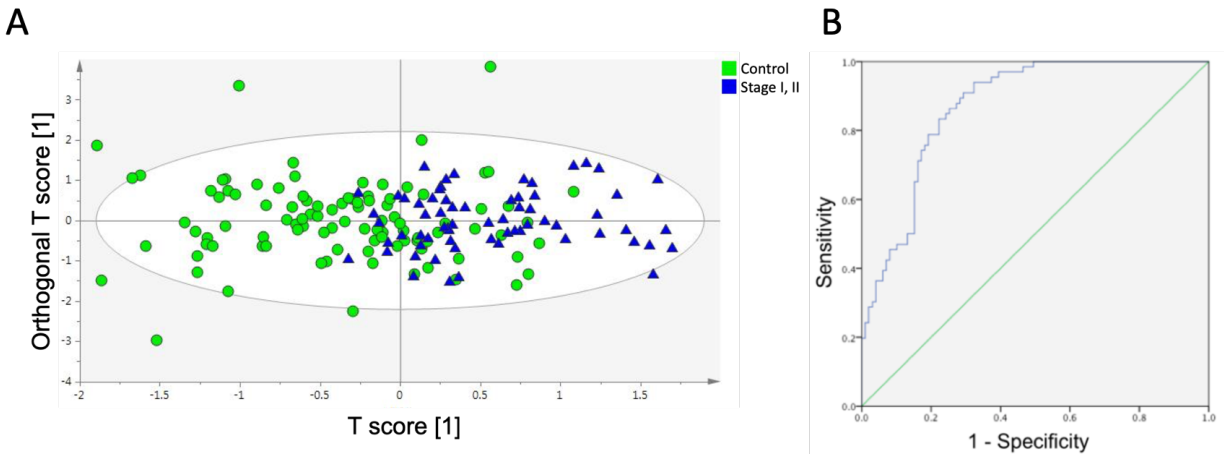


Figure 2.5 (A) Score plot of the orthogonal PLS-DA model for discrimination between stage I and II breast cancer patients (*n* = 66) and healthy controls (*n* = 99) [R^2X (cum) = 0.264, R^2Y (cum) = 0.358, Q^2 (cum) = 0.311]. Data were log_{10}-transformed, and Pareto scaled. Six differential metabolites (q < 0.05 and VIP >1) were used to construct this OPLS-DA model. (B) ROC curve illustrating the classification performance of the 6-metabolite OPLS-DA model for distinguishing between stage I and II BC patients and controls. [AUROC=0.87, 95% CI: 0.82-0.92, sensitivity: 0.86, specificity: 0.75]. AUROC, area under receiver operating characteristic; BC, breast cancer; CI, confidence interval; OPLS-DA, orthogonal partial least squares-discriminant analysis; PLS-DA, partial least squares-discriminant analysis; ROC, receiver operating characteristic.

Factor Analysis of Metabolic Data

A secondary aim of this study was to relate detected metabolites to affected pathways. To this end, metabolite data were subjected to EFA (Gorsuch, 1982). This multivariate technique was performed on a reduced correlation matrix of the 30 metabolites used for between-group

comparisons in order to determine pathways (i.e., factors) related to BC. Spectral decomposition of the experimental data matrix revealed a maximum of 4 factors (i.e., Kaiser criterion). Parallel analysis revealed only three factors accounted for more variance than random, permuted data (**Supplementary Figure 2.5**). Subsequently, 1-, 2-, and 3-factor models were extracted and rotated in conformity with oblique promax and infomax criteria, totaling 6 possible factor models. Each model was comparatively examined for percentage of total variance explained, magnitude of factor loadings, number of variables loaded onto each factor, and potential for meaningful factor interpretation and subsequent factor assignment. The 3-factor infomax model yielded the most satisfactory solution (**Table 2.5**). The findings revealed that 3 metabolites loaded significantly (>0.50) on the first factor, 3 metabolites loaded significantly on the second factor, and 3 metabolites loaded significantly on the third factor. These factors were found to be representative of the arginine/proline pathway, fatty acid biosynthesis, and tryptophan metabolism, respectively, suggesting significant alterations of these pathways in patients diagnosed with breast cancer.

Table 2.5 Reduced Factor Loading Matrix with Significantly Altered Metabolites as Variables

Variable	Factor 1 (arginine/proline metabolism)	Factor 2 (fatty acid biosynthesis)	Factor 3 (tryptophan metabolism)
Proline	0.83	0.33	0.39
Glycocyamine	0.51	0.18	0.18
Agmatine	0.75	0.03	0.37
Indole	0.12	0.08	0.91
3-Indolepropionic acid	0.01	0.23	0.59
Indole-3-acetic acid	0.39	0.49	0.71
Nonadecanoic acid	0.08	0.56	0.38
Palmitic acid	0.35	0.81	0.00
Stearic acid	0.21	0.64	0.33

Pathway Analysis of Metabolic Data

To understand the possible connection among detected plasma metabolites, we constructed metabolic pathway maps using IPA software (Krämer et al., 2014), as shown in **Figure 2.6**. MetaboAnalyst 4.0 (Chong et al., 2018) was used to perform pathway enrichment and

topology analysis (**Supplementary Figure 2.6**). It is worth mentioning that all metabolic pathways identified by our EFA as being significantly altered in BC patients were also identified by one or both of our bioinformatics analyses as being significantly altered as well; for instance, disturbances in arginine and proline metabolism as identified by EFA (**Table 2.5**) were corroborated by the results of our pathway analysis (**Figure 2.6**) and enrichment analysis (**Supplementary Figure 2.6**). As the results of our pathway, enrichment, and exploratory factor analyses are highly commensurate with each other, we can be reasonably confident that arginine/proline synthesis and degradation, fatty acid biosynthesis, and tryptophan metabolism are dysregulated in BC patients. Future studies can further target these networks for the discovery of pathway-specific biomarkers and potential therapeutic targets.

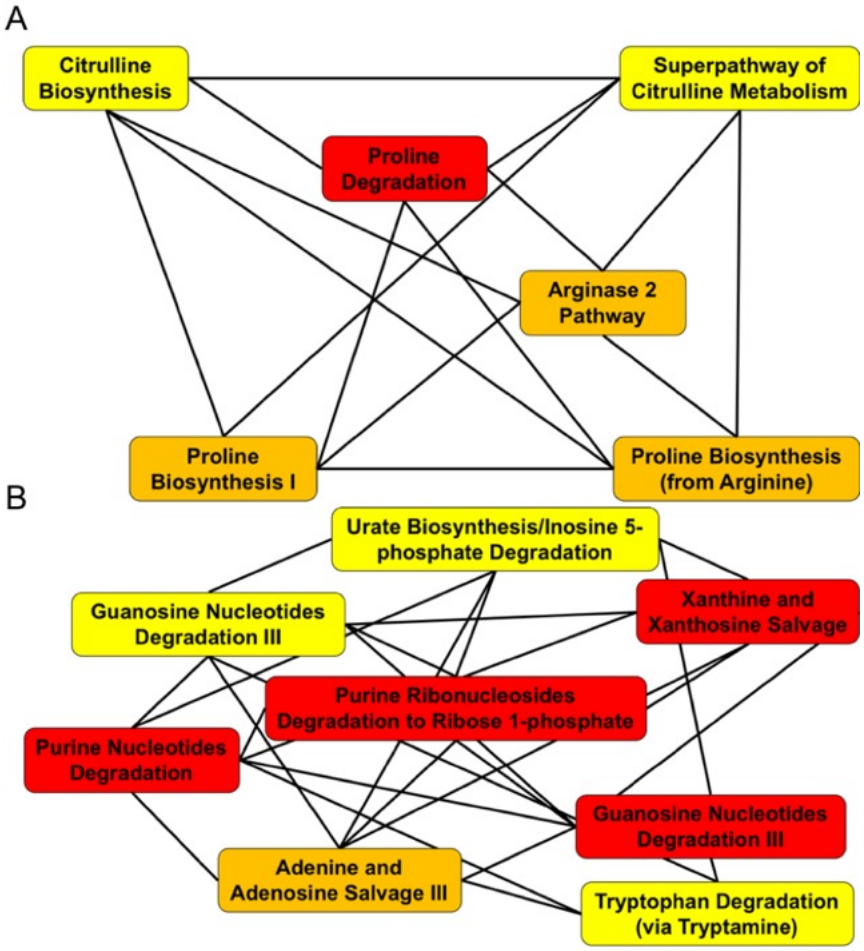


Figure 2.6 Overlapping canonical pathway constructed using identified plasma metabolites showing the interrelationship between (A) arginine/proline/citrulline metabolism, and (B) affected mechanisms of tryptophan and purine metabolism, by way of the kynurenine pathway. Yellow, orange, and red mechanisms are reflective of small, medium, and large disturbances, respectively.

A visual representation of these affected pathways is given in **Figure 2.7**. Of particular interest is the high level of agreement in results observed among the statistical analyses employed in this study. Results of our enrichment analysis, factor analysis, and pathway analysis unanimously connote arginine and proline metabolism as being altered in breast cancer patients. In this pathway, levels of agmatine and glycocyamine, both closely related to arginine, were

significantly reduced in stage III breast cancer patients; an even greater mean reduction in proline was observed in BC patients of all stages. In addition, the results of our factor and pathway analyses both showed tryptophan metabolism to be highly dysregulated in BC patients. Between cancer patients and controls, univariate testing showed indole and its product indole-3-acetate to be significantly up- and down-regulated, respectively. Finally, our results indicated fatty acid metabolism as being significantly altered between cancer patients and controls. Levels of palmitate were ubiquitously higher in BC patients of all cancer stages. The effects of breast cancer on tryptophan and fatty acid metabolism were corroborated by two separate analytic techniques, while arginine and proline metabolism was shown to be significantly affected by all three bioinformatics analyses. Our results provided strong metabolic pathway candidates for future studies looking to investigate the underlying biological mechanisms of breast cancer pathology.

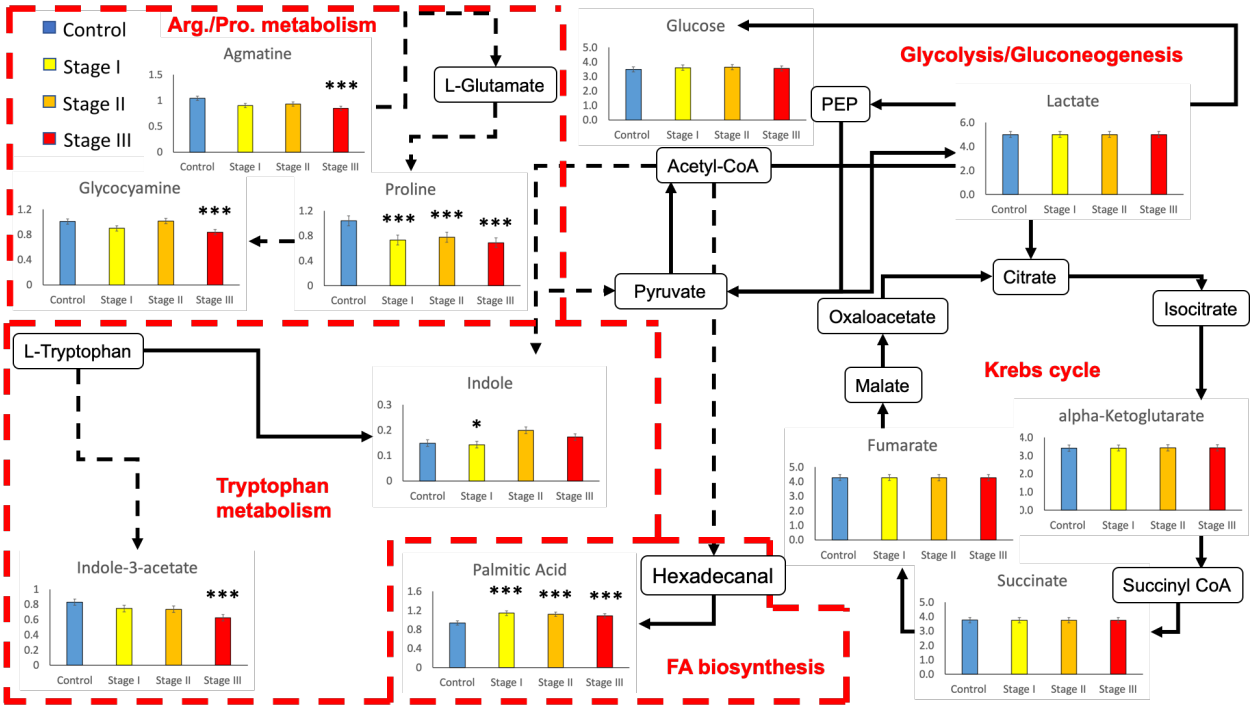


Figure 2.7 Metabolic map showing significant differences in arginine/proline, tryptophan, and fatty acid metabolism between control subjects and staged BC patients. Detected metabolites in the pathways are graphed using normalized data. Solid arrows represent direct connections between metabolites; dashed lines designate abridged connections. *, $p < 0.05$; ***, $p < 0.001$ compared to healthy controls. BC, breast cancer; FA, fatty acid; PEP, phosphoenolpyruvate.

Nonsignificant metabolites of glycolysis and the TCA cycle are also graphed alongside significantly altered metabolites in **Figure 2.7**. By and large, both lactate and alpha-ketoglutarate have been observed to be significantly increased in human breast cancer patients (Armitage & Barbas, 2014; Mishra & Ambs, 2015; D. R. Wise & Thompson, 2010), and have been attributed to the Warburg effect and glutamine addiction, respectively. However, no such increase in the levels of either metabolite was detected in our data set. This could be due to a lack of statistical power and may need to be validated with future studies

Discussion

For the last 20 years, significant interest has grown in utilizing mass spectrometry for the detection and analysis of cancer-related metabolic alterations. In doing so, these efforts have borne highly valuable diagnostic information and elucidation of probable biological mechanisms of cancer initiation and proliferation (Claudino et al., 2007; Mishra & Ambs, 2015; Patti et al., 2012). In the current study, we presented a combination of targeted metabolomics and multivariate statistics for the discovery of sensitive and specific BC metabolic biomarkers. Using this particular LC-MS/MS approach, 105 metabolites from many relevant metabolic pathways were reliably detected in both positive and negative ionization modes. Our multi-step biomarker selection, supervised model construction, and subsequent cross-validation have effectively demonstrated the robust diagnostic power of our metabolic profiling method in this study of 201 subjects.

To date, a number of studies have implemented mass spectrometry-based methods for the detection of metabolic alterations linked to breast cancer (Asiago et al., 2010; Frickenschmidt et al., 2008; Gu et al., 2011; Jain et al., 2012). Typically, these studies have effected global metabolic profiling approaches. Global analytical platforms tend to capture as many features as possible, making acquired data potentially less reliable or robust as a result. In contrast, the highly reproducible targeted LC-MS/MS method we presented in this study was observed to have a median CV value of 5.1%, producing metabolic markers previously unreported. Further enhancing our data quality and reliability, all targeted metabolites reported in this study were confirmed using pure standard compounds in lieu of applying database searches for compound annotation, as is conventionally done in global profiling approaches.

Many studies regarding biomarkers for breast cancer detection have been published (Budczies et al., 2012; Y. Chen et al., 2009; Duffy, 2006; Günther, 2015; Hilvo et al., 2011; Lee et al., 2013). For example, cancer antigen (CA) 15-3 and carcinoembryonic antigen (CEA) have

already been used clinically (Lee et al., 2013); however, they are mostly indicative of late-stage metastases and exhibit poor classification accuracy (~36-56%) for early-stage BC (Duffy, 2006). Another study analyzing breast cancer tissue samples using GC-MS discovered 13 tumor markers for discrimination between disease and normal subjects (Budczies et al., 2012) with sensitivity and specificity of roughly 80%. Notably, certain findings of the current study are echoed by previous literature. Hilvo et. al. (Hilvo et al., 2011) used LC-MS to investigate lipid metabolism in relation to breast cancer pathogenesis; they found significantly increased levels of palmitate in BC patients of all cancer stages, which was also observed in the current study. Another study found 46 urinary biomarkers for breast cancer diagnosis using LC-MS (Y. Chen et al., 2009). Interestingly, the researchers discovered significant alterations in tryptophan metabolism due mainly to lowered levels of indole and indoleacetate, a trend that was also seen in the current study through analysis of plasma samples.

The application of metabolomics technology in epidemiological and clinical studies is becoming common practice, but few multicenter metabolomics studies to increase reliability have been conducted. This study considers samples from multiple clinical locations, and we suggest a method for attenuating the confounding effect of significantly altered metabolite levels due to geographical location of sample collection. The 30 metabolites that were not found to be significantly different between controls (presumably 'housekeeping' metabolites) were selected for comparison between cancer patients and healthy subjects and for subsequent biomarker selection. Although this study proposes a solution to the problem of significant metabolite variation inherent in multicenter metabolomics studies, further studies are warranted to validate it.

In our current study, we performed univariate analysis of plasma metabolites between BC patients and healthy controls and observed significant alterations in a variety of the metabolites detected. Furthermore, significantly altered plasma metabolites with $q < 0.05$ and VIP > 1 from the enhanced PLS-DA model were selected for inclusion in the biomarker panel (**Table 2.2**).

Additionally, age was included as a clinical factor to enhance the VIP-based metabolite model. Moreover, our enrichment, pathway, and factor analyses revealed significant alterations in (1) arginine/proline degradation, (2) fatty acid biosynthesis, and (3) tryptophan metabolism.

Our results indicate proline to be significantly altered in stage I, II, and II breast cancer patients. This finding is in keeping with recent mechanistic studies that have discovered increased proline dehydrogenase (Prodh) activity to fuel proline catabolism and consequently lead to increased growth of breast cancer cells in 3D culture and *in vivo* metastasis formation (Elia et al., 2017). Consequently, recent efforts have attempted to target proline catabolism, since metastasizing cancer cells rely on by-products of proline degradation to fuel their increased energy need during the colonization of distant organs. As such, targeting proline metabolism does not affect primary cancer growth or non-transformed cells but, rather, impairs metastasis formation in unaffected organs and may lead to better prognosis.

Our results also indicate palmitate to be significantly overexpressed in BC patients at all cancer stages. As stated before, aberrant metabolism is a characteristic feature of breast cancer due to the increased energy needs of tumors. Cancer cells often induce a state of lipolysis and/or increased fatty acid synthesis in an effort to meet those needs (Wolf et al., 2006). This is known as cancer cachexia and has been estimated to account for nearly 20% of all cancer deaths. One key aspect of cancer treatment is the prevention of tumor growth which, in turn, requires a disruption in cancer cells' energy consumption. Accordingly, studies have focused on reducing levels of endogenously produced fatty acids by inhibiting the activity of enzymes responsible for fatty acid biosynthesis (Pizer et al., 1996).

The by-products of tryptophan metabolism, indole and indole-3-acetate were observed to be significantly altered in stage I and stage III BC subjects, respectively. One line of research has focused on the propagation of rapid-growing 'progressive' cancers due to a failure of the immune system to maintain control over budding tumors. As a result, there has been increased interest in

cancer's ability to escape the human immune response. Recent advances have shown the consumption of tryptophan to be critical in the escape mechanisms of tumors (Prendergast, 2011). Mechanistically, cancers have been shown to upregulate the liver enzyme tryptophan dioxygenase, thereby driving tryptophan consumption to produce kynurenine, an endogenous ligand for the aryl hydrocarbon receptor which mediates invasive tumor growth. This affected pathway allows tumors to overcome the human immune response and is, potentially, reflected in the results of our factor and pathway analyses.

There are some limitations in this study. This study considers samples from multiple clinical locations which, although critically important, is underused in metabolomics study design. However, knowledge of sample demographics is a limitation. Although statistically important clinical features such as age were known, we failed to collect other demographic information such as body mass index or smoking and drinking status, which may be confounding factors for the associations between metabolite levels and breast cancer status.

Although the analytical platform was optimized for the detection of over 400 metabolites, 30 potential markers were included in further data analysis when accounting for reliability and non-significance between control samples collected from two clinical sites. Results of our representation analysis indicate that the metabolic profile generated by these 30 metabolites is reflective of 27 metabolic pathways (see **Supplementary Figure 2.3**). Likewise, previous studies have used 20 compounds to infer tryptophan-induced pathogenesis of breast cancer (Cao et al., 2015), and 25 metabolites for discrimination between lung cancer patients and age-matched controls (X. Zhang et al., 2016). Also, we performed pathway-based analyses to observe higher-order effects due to breast cancer (S. Huang et al., 2016). It should also be noted that there is a great need for breast cancer metabolomics studies utilizing multi-center designs (Günther, 2015); our study aims to address that need while also offering a simple way to account for possible metabolic variations in samples taken from different clinical locations.

Furthermore, more research is needed to study metabolism related to molecular subtypes of breast cancer. In this study, no significant difference in metabolites was observed between ER/PR$^+$, HER2$^+$ *vs.* ER/PR$^+$, HER2$^-$, and triple negative *vs.* non-triple negative patients. However, 15 metabolites had $p < 0.05$ when comparing ER, PR, and HER2 positive and negative patients, as well as staged BC patients, although none of these 15 metabolites remained significant after FDR-correction. Future studies should also examine larger cohorts from multiple locations to further validate the altered metabolites and metabolic pathways related to BC pathogenesis discovered in our study (Jackson, 2003; Noether, 1987). To the best of our knowledge, this is the first study using an LC-MS/MS metabolomics approach to analyze plasma samples from two clinical sites for breast cancer diagnosis. Our predictive model demonstrates relatively good performance (89%) for detection of stage I and II breast cancer with specificity of 75% when sensitivity is 80%. This study provides a strong rationale for the development of larger multi-site projects to validate the findings across population groups and further advance the development of accurate tools for clinical risk prediction of breast cancer.

Conclusions

This study is part of a growing body of literature in which an LC-MS/MS targeted plasma metabolic profiling approach has been applied for the comparison of BC patients and healthy controls. To the best of our knowledge, this is the first targeted approach for breast cancer diagnosis to consider samples from multiple locations. Our results demonstrate a panel of 18 metabolites with FDR q-value < 0.05 and 6 metabolites with both $q < 0.05$ and VIP > 1. The 6 differential metabolites were also shown to be effective for the detection of early stage, localized disease (stage I and II). Application of bioinformatic methods showed underlying disturbances in metabolic pathways related to tumor growth, metastasis, and immune escape mechanisms.

Accounting for age, this metabolic profiling method can potentially provide a novel disease biomarker panel for breast cancer.

CHAPTER 3

COCCIDIOIDOMYCOSIS DETECTION USING TARGETED PLASMA AND URINE METABOLIC PROFILING

(Published in *Journal of Proteome Research*, 2019, 18: 2791—2802)

Abstract

Coccidioidomycosis, also known as Valley fever (VF), is a potentially lethal fungal infection that results in more than 200 deaths per year in the United States. Despite the important role of metabolic processes in the molecular pathogenesis of VF, robust metabolic markers to enable effective screening, rapid diagnosis, accurate surveillance, and therapeutic monitoring of VF are still lacking. In this study, we present a targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS)-based metabolic profiling approach for identifying metabolic marker candidates that could enable rapid, highly sensitive and specific VF detection. Using this targeted approach, 207 plasma metabolites and 231 urinary metabolites from many metabolic pathways of potential biological significance were reliably detected and monitored in 147 samples taken from two groups of subjects (48 VF patients and 99 non-VF controls). The results of our univariate significance testing and multivariate model development informed the construction of a 3-metabolite panel of potential plasma biomarkers and a 9-metabolite panel of potential urinary biomarkers. Receiver operating characteristic (ROC) curves generated based on orthogonal partial least squares-discriminant analysis (OPLS-DA) models showed excellent classification performance, with 94.4% sensitivity and 97.6% specificity for plasma metabolites. Urine metabolites were less accurate, demonstrating 89.7% sensitivity and 88.1% specificity. Enrichment, pathway, and network analyses revealed significant disturbances in glycine and serine metabolism, in both plasma and urine samples. To the best of our knowledge, this is the first study aiming to discover novel metabolite markers of VF, which could achieve accurate

diagnosis within 24 hrs. The results expand basic knowledge of the metabolome related to VF and potentially reveal pathways or markers that could be targeted therapeutically. This study also provides a promising basis for the development of larger multi-site projects to validate our findings across population groups and further advance the development of better clinical care for VF patients.

Introduction

Coccidioidomycosis, also known as Valley fever (VF), is a respiratory infection caused by inhalation of airborne fungal spores of *Coccidioides immitis* or *Coccidioides posadasii* (Fisher et al., 2007). These category C fungal pathogens are endemic to desert climates with mild winters and arid summers such as those in the southwestern United States, including California, Arizona, New Mexico, and Texas, and parts of northern Mexico and South America. From 1990-2008, 3089 deaths in the United States were attributed to coccidioidomycosis, or roughly 200 per year (J. Y. Huang et al., 2012). In states where VF is endemic, overall incidence is estimated to be 42.6 cases per every 100,000 persons per year (Benedict, 2013). Between 1998 and 2016, Arizona accounted for 51-79% of all reported cases of VF in the United States (Adams et al., 2017). In highly-endemic areas such as the Phoenix and Tucson metropolitan areas of Arizona, VF is estimated to account for 15-30% of all community-acquired pneumonias (CAP) (Valdivia et al., 2006), and evidence seems to suggest that diagnoses are under-reported due to low testing rates (Chang et al., 2008). Therefore, VF is a common threat to human health especially in endemic areas.

Currently, VF is difficult to diagnose due to presentation of vague symptoms that often mimic viral or bacterial pneumonias or even lung cancer (Petrini et al., 2003). In fact, the majority of people who are exposed to this virulent, dimorphic fungus never seek medical care (Smith et al., 1948), and approximately 40% of people who contract the fungal infection present with flu-like

symptoms such as fatigue, cough, fever, shortness of breath, headaches, night sweats, muscle or joint pain, and rash (Tsang et al., 2010), which can persist for weeks to months (Thompson, 2011). Approximately 5-10% of infected people will develop serious, often chronic, lung diseases (Thompson, 2011), while roughly 1% of the patients develop disseminated coccidioidomycosis as the infection spreads from the lungs to other parts of the body resulting in nodules, ulcers, skin lesions, and possible meningitis (Crum et al., 2004; Galgiani et al., 2005; Thompson, 2011).

The current mainstay diagnostics for VF are serologic testing methods, which mainly include enzyme immunoassay (EIA), complement fixation (CF), and immunodiffusion (ID). However, no single serological test offers both excellent sensitivity and specificity. Additionally, approximately 10% of immunocompetent patients and 30% of immunosuppressed patients fail to produce an adequate immunological response to the VF infection, especially in the acute phase of disease (Blair et al., 2006). Accurate diagnosis therefore relies on a combination of clinical presentation, serology, radiography, histology and culture (Ampel, 2010). These diagnostic methods are either costly, time-consuming (often greater than 2 weeks in the case of fungal culture), invasive, or indeterminable. Therefore, a fast, cost-effective, highly sensitive and specific method for the detection of VF is critically needed.

Broadly, fungal infections can induce wide and extensive alterations in metabolism (Bills & Gloer, 2016; Brown et al., 2014; Eisenreich et al., 2015; Ene et al., 2014) which provides a promising approach to detect fungal infections, such as VF. For example, fungal pathogens must assimilate local nutrients to establish an infection in their mammalian host, and metabolic flexibility is generally essential for fungal pathogenicity. Since metabolites are sensitive to subtle differences and changes in pathological status, metabolomics, the comprehensive study of small molecular-weight metabolites and their dynamic changes in biological systems (Patti et al., 2012), provides advanced methods to identify changing metabolite levels, and has resulted in the rapid discovery of disease biomarkers during the past decade (Bowers et al., 2014; G. A. N. Gowda &

Raftery, 2013; Gu et al., 2012; Jasbi, Wang, et al., 2019; Kaysen et al., 2015; Yin & Xu, 2017). Mass spectrometry (MS)-based metabolic profiling has proven to be a promising tool for analyzing metabolic alterations due to various diseases and, therefore, can provide sensitive and valuable diagnostic information (Ahn et al., 2017; Jasbi, Wang, et al., 2019; Madsen et al., 2010), pathogenesis identification (Ene et al., 2014; Poddighe et al., 2017), and potential therapeutic targets for clinical treatments (Monnerat et al., 2018) and disease monitoring (Zhu et al., 2015). Indeed, previous studies have used MS-based methods in conjunction with chemometric analyses to develop metabolic biomarker panels for the accurate diagnosis of fungal infections such as mucormycosis (Dadwal & Kontoyiannis, 2018) and aspergillosis (Savelieff & Pappalardo, 2017), as well as detection of various mycotoxins produced as a result of host-fungal interactions (Culibrk et al., 2016; Rubert et al., 2017). In this study, we present the first targeted plasma and urine liquid chromatography-tandem mass spectrometry (LC-MS/MS) profiling approach for the rapid and accurate detection of VF.

Methods

Reagents

Acetonitrile (ACN), methanol (MeOH), ammonium acetate (NH_4OAc), and acetic acid (AcOH), all LC-MS grade, were purchased from Fisher Scientific (Pittsburgh, PA). Ammonium hydroxide (NH_4OH) was bought from Sigma-Aldrich (Saint Louis, MO). DI water was provided in-house by a Water Purification System from EMD Millipore (Billerica, MA). Phosphate buffered saline (PBS) was bought from GE Healthcare Life Sciences (Logan, UT). Standard compounds corresponding to the measured metabolites were purchased from Sigma-Aldrich (Saint Louis, MO) and Fisher Scientific (Pittsburgh, PA).

Sample Collection and VF Diagnostic Criteria

The samples were collected under a previously approved IRB protocol with waived consent. Urine and plasma specimens were acquired from excess clinical specimens collected for routine standard of care at Mayo Clinic Arizona. De-identified aliquots were provided to the Arizona Metabolomics Laboratory (College of Health Solutions, Arizona State University) for processing. Samples had been frozen at –80°C until analysis. VF status was determined based on clinical evaluation using the Mayo Clinic Arizona multi-factorial criteria for the diagnosis of Coccidioidomycosis (Grys et al., 2018). This diagnostic rubric includes the evaluation of patient symptoms, radiography, serology, histology, and culture results, as no single assay or measure can currently be used alone to diagnose VF.

Sample Preparation

Frozen plasma and urine samples were first thawed overnight under 4°C. Afterward, 50 µL of each plasma sample was placed in a 2 mL Eppendorf vial while 100 µL of each urine sample was placed in a separate vial. For both plasma and urine, the initial step for protein precipitation and metabolite extraction was performed by adding 500 µL MeOH and 50 µL internal standard solution (containing 1,810.5 µM $^{13}C_3$-lactate and 142 µM $^{13}C_5$-glutamic acid). The mixture was then vortexed for 10 s and stored under –20°C for 30 min, followed by centrifugation at 14,000 RPM for 10 min at 4°C. The supernatants (450 µL and 500 µL for plasma and urine, respectively) were collected into new Eppendorf vials and dried using a CentriVap Concentrator (Labconco, Fort Scott, KS). The dried samples were reconstituted in 150 µL of 40% PBS/60% ACN and centrifuged again at 14,000 RPM at 4°C for 10 min. After that, 100 µL of supernatant was collected from each sample into an LC autosampler vial for subsequent analysis. Two pooled samples, which were a mixture of all plasma and urine samples respectively, were used as the internal quality-control (QC) samples and injected once every 10 experimental samples.

LC-MS/MS

The targeted LC-MS/MS method used here was modeled after that developed and used in a growing number of studies (Buas et al., 2017; Carroll et al., 2015; Gu et al., 2016, 2015; Li et al., 2018; Zhu et al., 2014). Briefly, all LC-MS/MS experiments were performed on an Agilent 1290 UPLC-6490 QQQ-MS system (Santa Clara, CA). Each sample was injected twice, 10 µL for analysis using negative ionization mode and 4 µL for analysis using positive ionization mode. Both chromatographic separations were performed in hydrophilic interaction chromatography (HILIC) mode on a Waters XBridge BEH Amide column (150 x 2.1 mm, 2.5 µm particle size, Waters Corporation, Milford, MA). The flow rate was 0.3 mL/min, auto-sampler temperature was kept at 4°C, and the column compartment was set to 40°C. The mobile phase was composed of Solvents A (10 mM ammonium acetate, 10 mM ammonium hydroxide in 95% H_2O/5% ACN) and B (10 mM ammonium acetate, 10 mM ammonium hydroxide in 95% ACN/5% H_2O). After an initial 1 min isocratic elution of 90% B, the percentage of Solvent B decreased to 40% at t = 11 min. The composition of Solvent B was maintained at 40% for 4 min (t = 15 min), after which the percentage of B gradually went back to 90%, to prepare for the next injection.

The mass spectrometer was equipped with an electrospray ionization (ESI) source. Targeted data acquisition was performed in multiple-reaction-monitoring (MRM) mode. We monitored 118 and 160 MRM transitions in negative and positive mode, respectively (278 transitions in total). The whole LC-MS system was controlled by Agilent MassHunter Workstation software (Santa Clara, CA). The extracted MRM peaks were integrated using Agilent MassHunter Quantitative Data Analysis software (Santa Clara, CA).

Data Analysis

Univariate testing was performed using SPSS 22.0 (SPSS Inc., Chicago, IL). Multivariate statistical analyses were performed using open-source R software and SIMCA-P (Umetrics,

Umeå, Sweden). The data were $\log_{10}$-transformed prior to model construction. Pathway analysis and integrating enrichment analysis were performed and visualized using the online MetaboAnalyst software (Chong et al., 2018).

Results

Metabolic profiles

A total of 48 VF patients and 99 non-VF controls were included in the study. Of the VF samples, 18 were plasma and 30 were urine. Of the non-VF control samples, 41 were from plasma and 58 from urine. Paired plasma and urine samples were only obtained from one patient, although on different days, given this was not a coordinated collection. Roughly half of all VF patients were taking antifungal medication at the time of sample collection. **Table 3.1** shows the clinical and demographic characteristics of patients included in the study. There was no statistically significant difference in plasma or urine metabolites between VF patients on antifungal medication and those VF patients not taking antifungals, as calculated by a Mann-Whitney *U* test (all $q > 0.05$).

Table 3.1 Clinical and demographic characteristics of study participants.

		Total *n (%)*	**Gender**		**Age**		**Clinical Course of Disease**			**Serology** *(EIA, CF or ID)*		**Antifungals?**	
			Male	Female	< 65	≥ 65	Acute	Chronic	Disseminated	Positive	Negative	Yes	No
VF (+)	Urine	30 (62.5)	19 (63.3)	11 (36.7)	15 (50)	15 (50)	17 (56.7)	10 (33.3)	3 (10)	24 (80)	6 (20)	25 (83.3)	5 (16.7)
	Plasma	18 (37.5)	11 (61.1)	7 (38.9)	10 (55.6)	8 (44.4)	10 (55.5)	3 (16.7)	5 (27.8)	16 (88.9)	2 (11.1)	5 (27.8)	13 (72.2)
VF (-)	Urine	58 (59)	34 (58.6)	24 (41.4)	31 (53.4)	27 (46.6)	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*
	Plasma	41 (41)	22 (53.7)	19 (46.3)	29 (70.7)	12 (29.3)	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*	*N/A*

In the current study, we used a large-scale, targeted LC-MS/MS approach for reliable and comprehensive VF plasma and urine metabolic profiling.(Jasbi, Wang, et al., 2019) Using this metabolic profiling system, targeted analysis of 278 MRM transitions was achieved for metabolites spanning over 20 different chemical classes (such as amino acids, carboxylic acids, pyridines, etc.) from more than 35 metabolic pathways (e.g., TCA cycle, amino acid metabolism, glycolysis, purine and pyrimidine metabolism, urea cycle, etc.) in both positive and negative ionization modes. In total, we found that 207 plasma metabolites and 231 urine metabolites were reliably detected with relative abundances > 1,000 in more than 80% of all samples. After normalization by averaged values from QC injection data, relative levels of the 207 plasma metabolites had a median coefficient of variation (CV) value of 11.91% (range: 0.46%-13.01%) with ~70% of metabolites having CV < 15%, while the 231 reliably detected urine metabolites had a median CV value of 11.37% (range: 0.02%-12.00%) with ~85% of metabolites having CV < 15% (**Supplementary Figure 3.1**).

Statistical analyses

Of the 207 reliably detected plasma metabolites, 106 showed statistical significance between VF patients and non-VF counterparts, as determined by a Wilcoxon rank rum test (see **Supplementary Table 3.1** for the complete list of significant plasma metabolites, their associated *p*- and *q*-values, and directional changes). Of the 231 reliably detected urine metabolites, 20 metabolites showed statistical significance between VF patients and non-VF controls (**Supplementary Table 3.2**). Volcano plots of the tested plasma and urine metabolites showing significance and fold change values are presented in **Figure 3.1**.

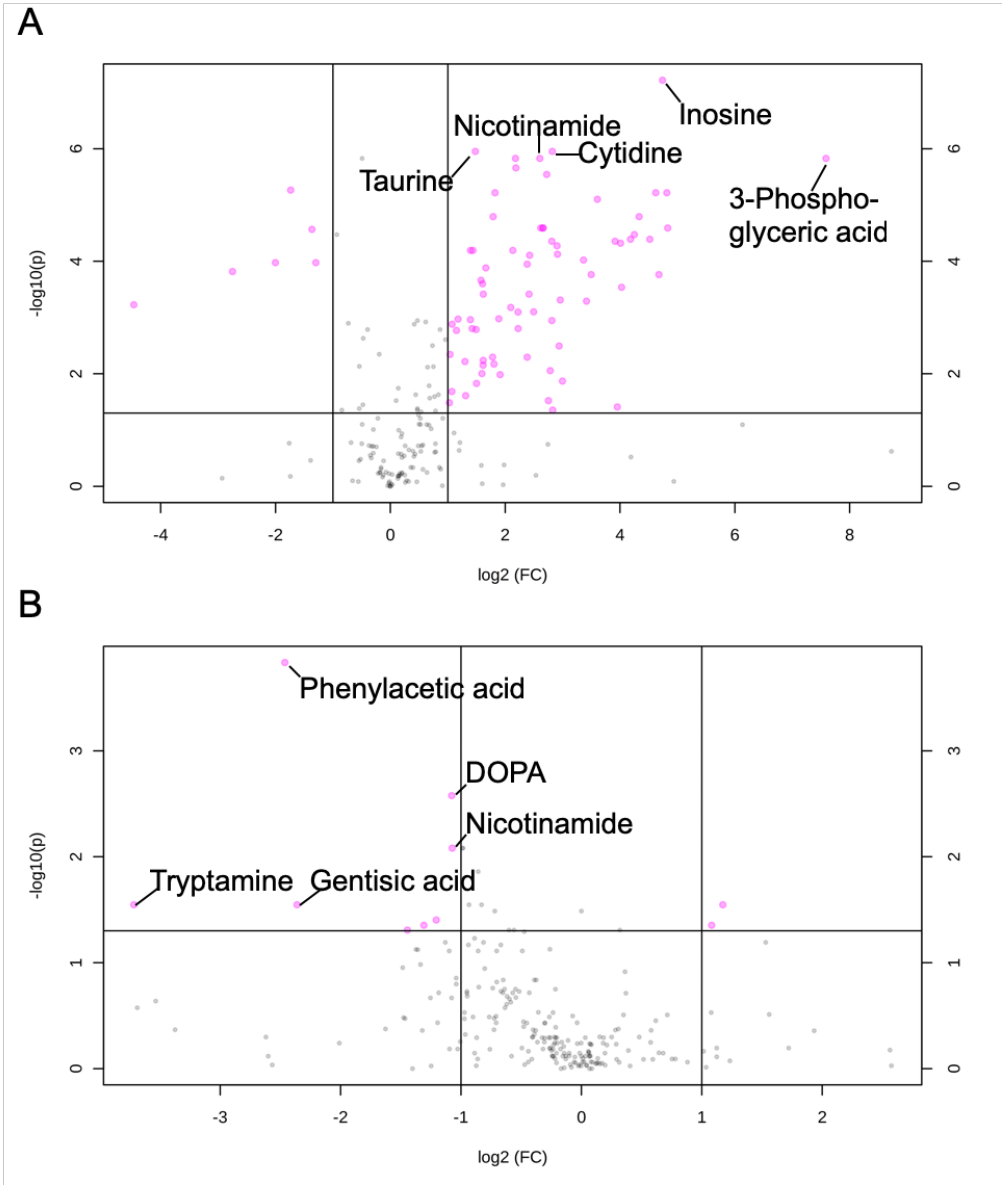


Figure 3.1 Volcano plot of (A) 207 plasma metabolites, and (B) 231 urine metabolites comparing Valley fever/non-Valley fever controls. Top 5 metabolites are labeled. Fold change (FC) threshold: 2.0; false discovery rate (FDR)-adjusted *p*-value threshold: 0.05. Unequal group variance was assumed, non-parametric test was used.

To further explore potential biomarkers for discrimination between VF patients and non-VF controls, levels of the 106 significant plasma metabolites and 20 significant urine metabolites were selected to establish initial partial least squares-discriminant analysis (PLS-DA) models. As

can be seen in **Supplementary Figure 3.2**, a separation trend was observed in the initial PLS-DA score plots. The plasma PLS-DA model (**Supplementary Figure 3.2A, B**) showed superior predictive and explanatory capacity to the PLS-DA model constructed from significant urine metabolites (**Supplementary Figure 3.2C, D**) as validated by permutation testing with 200 iterations [Plasma: R^2X (cum) = 0.973, R^2Y (cum) = 0.862, Q^2 (cum) = 0.789; Urine: R^2X (cum) = 0.847, R^2Y (cum) = 0.627, Q^2 (cum) = 0.501]. Variable importance in projection (VIP) scores were obtained from the initial PLS-DA models. As shown in **Figure 3.2**, 3 plasma metabolites were observed to have VIPs > 2, and 9 urine metabolites had VIPs > 1. In an effort to develop thrifty models that rely on the fewest number of predictors while accounting for as much variance as possible, enhanced orthogonal partial least squares-discriminant analysis (OPLS-DA) models were constructed using the 3 plasma metabolites that were both significant (q < 0.001) and had VIPs > 2 and the 9 urine metabolites that were significant (q < 0.05) and had VIPs > 1, respectively. As shown in **Figure 3.3**, separation was clearly observed in both the plasma and urine OPLS-DA models, with the plasma metabolite model again outperforming the urinary model [Plasma: R^2X (cum) = 0.668, R^2Y (cum) = 0.739, Q^2 (cum) = 0.723; Urine: R^2X (cum) = 0.302, R^2Y (cum) = 0.416, Q^2 (cum) = 0.389]. Significance information and fold change values for the final panel of 3 plasma and 9 urine metabolites can be found in **Table 3.2**; box plots of the plasma biomarker panel are provided in **Figure 3.4** while urine biomarkers are plotted in **Figure 3.5.**

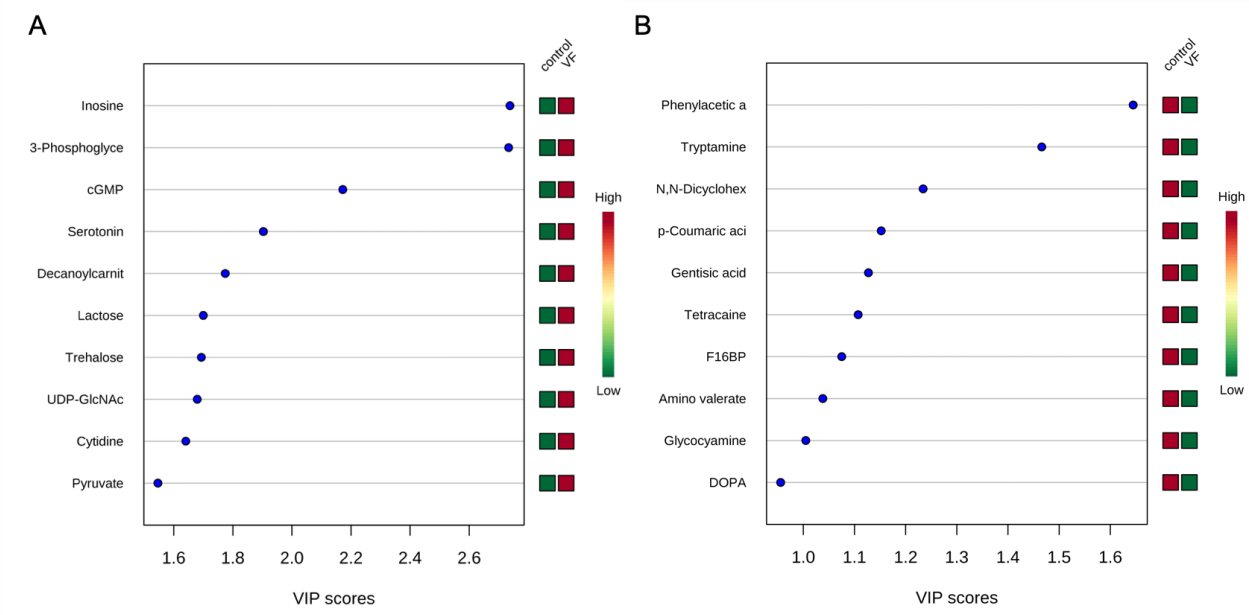


Figure 3.2 Variable Importance in Projection (VIP) scores of plasma and urine partial least squares-discriminant analysis (PLS-DA) models constructed using 106 and 20 metabolites, respectively, for discrimination between Valley fever patients and non-Valley fever controls. Top 10 important contributors to model projection are shown. (A) Three plasma metabolites were observed to have VIPs > 2. (B) Nine urine metabolites were observed to have VIPs > 1.

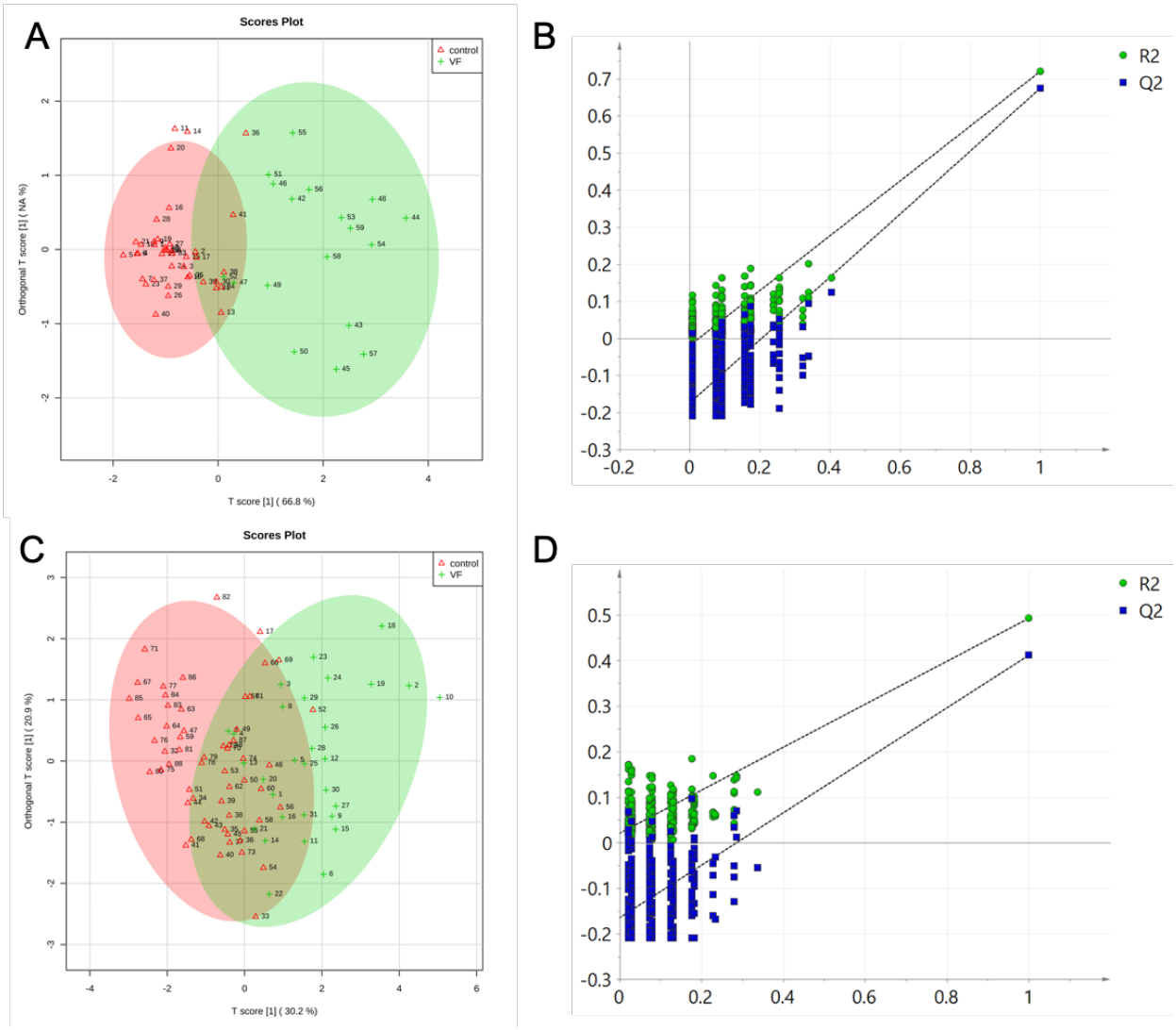


Figure 3.3 Orthogonal partial least squares-discriminant analysis (OPLS-DA) performed on log_{10}-transformed plasma and urine metabolite data: (A) score plot of 3 significant and important plasma metabolites accounting for 66.8% of variance, (B) statistical validation of plasma OPLS-DA model [R^2X (cum) = 0.668, R^2Y (cum) = 0.739, Q^2 (cum) = 0.723] by permutation testing (n = 200), (C) score plot of 9 significant and important urinary metabolites accounting for 51.1% of variance, (D) statistical validation of urinary OPLS-DA model [R^2X (cum) = 0.302, R^2Y (cum) = 0.416, Q^2 (cum) = 0.389] by permutation testing (n = 200).

Table 3.2 Significance and fold change details for final panel of plasma and urine markers.

Metabolite	Plasma		Urine	
	FDR *q*	Fold Change	FDR *q*	Fold Change
Inosine	< 0.0001	26.724		
3-Phosphoglyceric acid	< 0.0001	192.44		
cGMP	0.0003	16.312		
Phenylacetic acid			1.4E-4	0.181
Amino valerate			0.0082	0.380
Glycocyamine			0.0082	0.379
Tryptamine			0.0284	0.076
Gentisic acid			0.0284	0.194
p-Coumaric acid			0.0325	0.191
N,N'-Dicyclohexylurea			0.0395	0.434
F16BP			0.0443	0.404
Tetracaine			0.0491	0.367

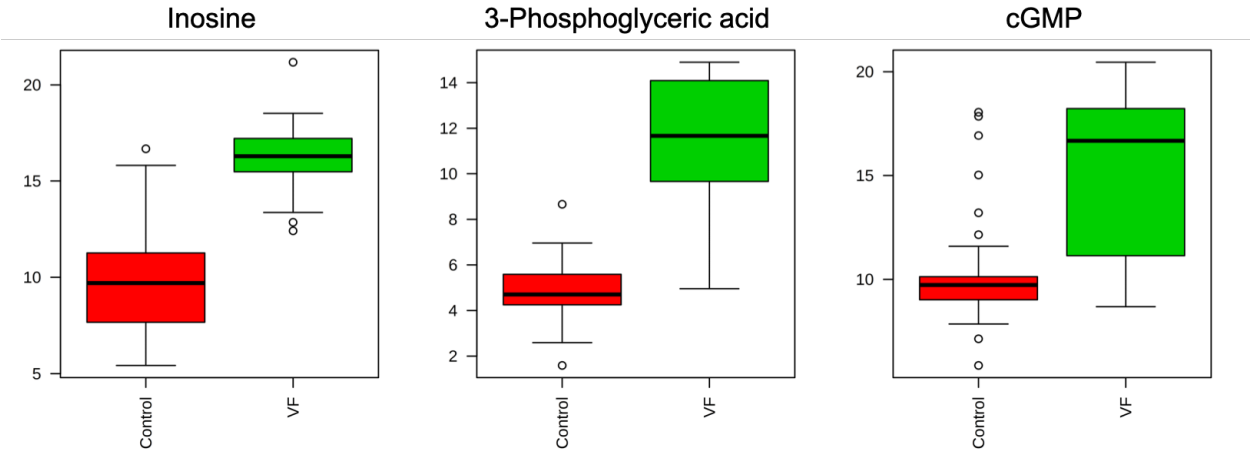


Figure 3.4 Box plots of candidate plasma markers (all $q < 0.001$ and VIP > 2) for Valley fever (VF) detection. Data were $\log_{10}$ normalized.

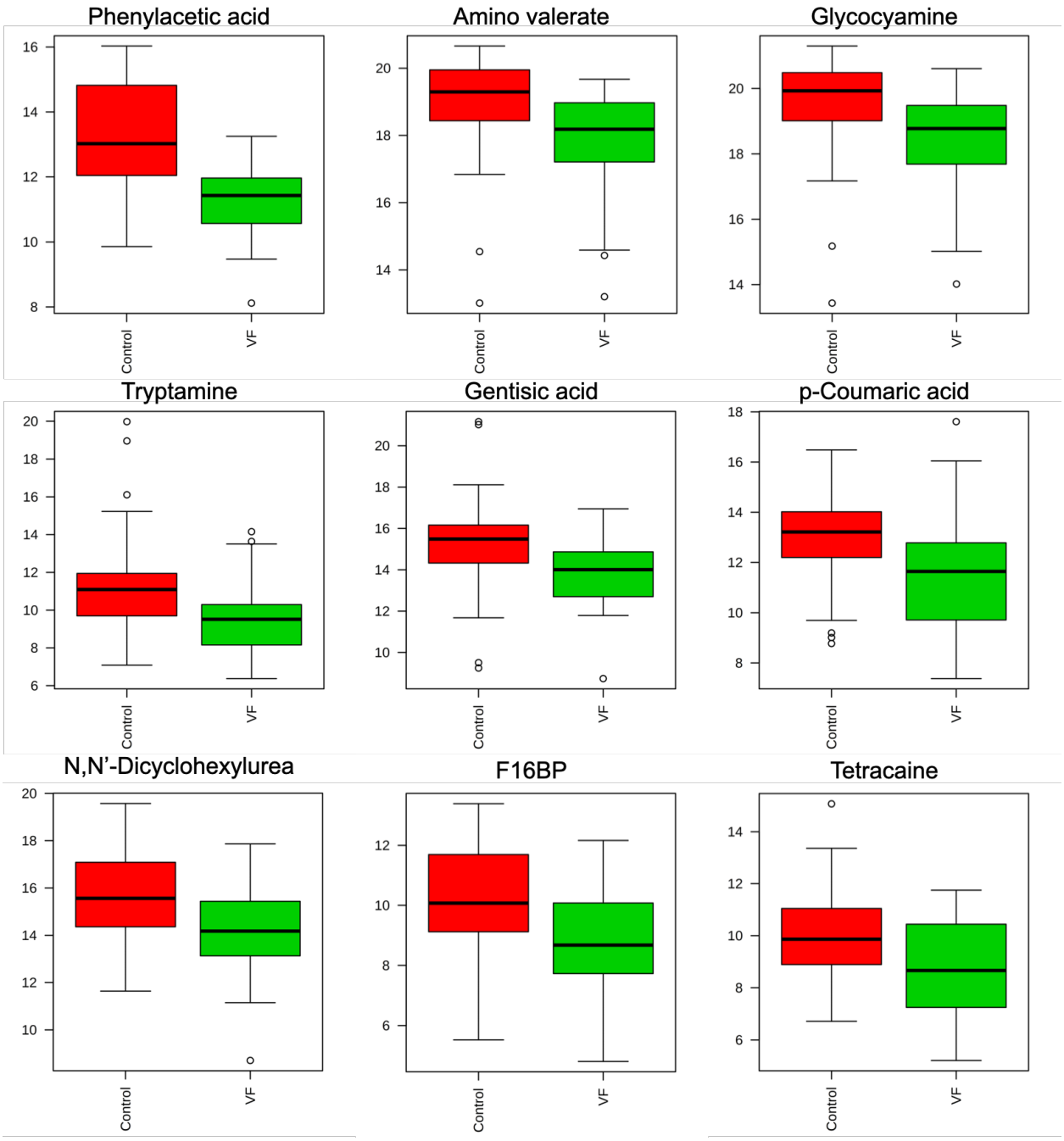


Figure 3.5 Box plots of candidate urine markers (all $q < 0.05$ and variable importance in projection (VIP) > 1) for Valley fever (VF) detection. Data were $\log_{10}$ normalized.

Receiver operating characteristic (ROC) analysis was performed to determine the classification performance of the enhanced plasma and urine OPLS-DA models. As evidenced by the ROC curve shown in **Figure 3.6A**, the OPLS-DA model constructed using only 3 significant

and important plasma metabolites demonstrated near-perfect classification accuracy (AUC = 0.995), excellent sensitivity (0.994) and specificity (0.976). The OPLS-DA model constructed using the 9 significant and important urine metabolites also showed excellent overall accuracy (AUC = 0.929), high sensitivity (0.897) and good specificity (0.881) (**Figure 3.6B**).

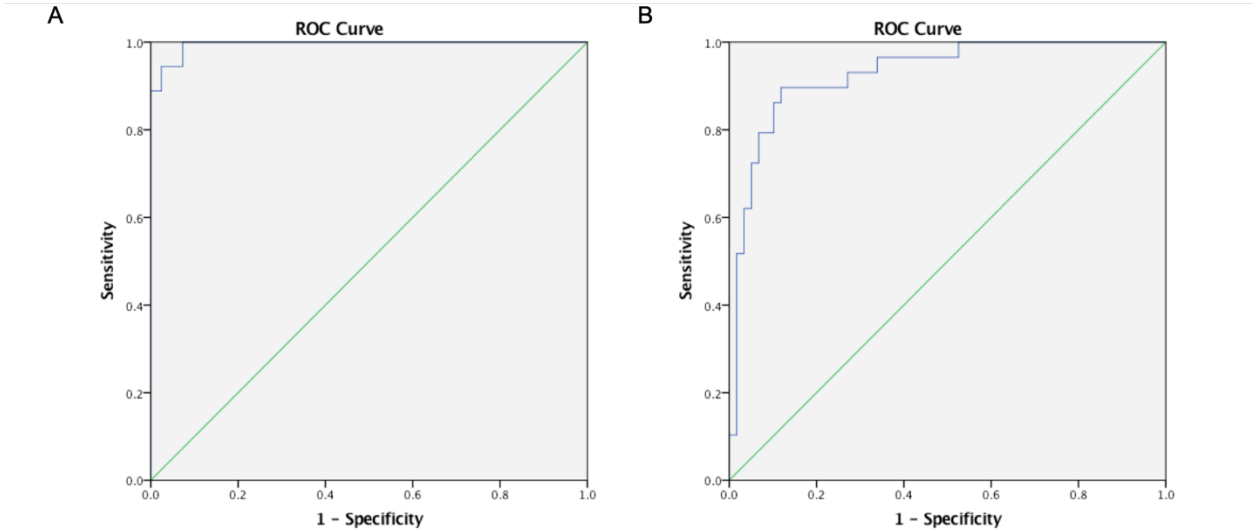


Figure 6. ROC analysis of (A) 3-metabolite plasma OPLS-DA model [AUC = 0.995 (95% CI: 0.983-1.00), sensitivity = 0.944 when specificity = 0.976] and (B) 9-metabolite urine OPLS-DA model [AUC = 0.929 (95% CI: 0.873-0.985), sensitivity = 0.897 when specificity = 0.881].

To analyze the discriminatory ability of the plasma and urine biomarker panels in accurately detecting VF patients of varying clinical course and seropositivity, univariate significance testing was performed on the plasma metabolite panel (inosine, 3-phosphoglyceric acid, cGMP) and urinary metabolite panel (phenylacetic acid, tryptamine, N,N-dicyclohexylurea, p-coumaric acid, gentisic acid, tetracaine, fructose-1,6-bisphosphate, amino valerate, glycocyamine) to monitor any potential changes in their levels attributable to clinical course or serology status. Notably, our results showed no significant change in these 12 marker candidates between VF patients with acute, chronic, or disseminated disease or VF patients with positive, negative, or indeterminant serology results (all $p > 0.10$).

Pathway Analyses of Metabolic Data

Enrichment analysis was conducted using KEGG database searches and metabolite intensities for both plasma and urine data. Enrichment analysis of 207 reliably detected plasma metabolites showed significant (*p*) disturbances in alanine metabolism (0.005) and amino sugar metabolism (0.014). Enrichment analysis of 231 reliably detected urine metabolites was also conducted. Although non-significant, results indicated a high magnitude of fold enrichment (VF/control) in phenylacetate metabolism (0.167) and alanine metabolism (0.315). Enriched pathways as determined by analysis of all reliably detected plasma and urine metabolites are shown as separated motifs in **Figure 3.7**.

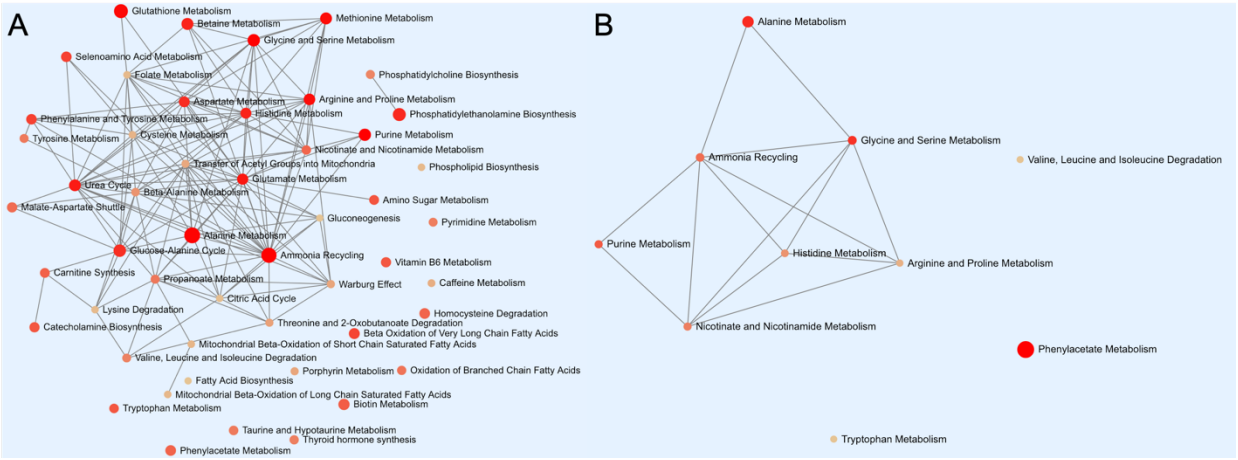


Figure 3.7 Results of enrichment analysis using (A) 207 reliably detected plasma metabolites and (B) 231 reliably detected urinary metabolites.

Pathway analysis was also performed in order to determine significantly affected pathways in VF patients. Although no pathway was observed to have large impact coefficients (> 0.50), 4 pathways were shown to be significantly affected in both plasma and urine samples (**Figure 3.8**). Agreement between plasma and urinary analyses revealed significant disturbances in nicotinate and nicotinamide metabolism and ammonia recycling.

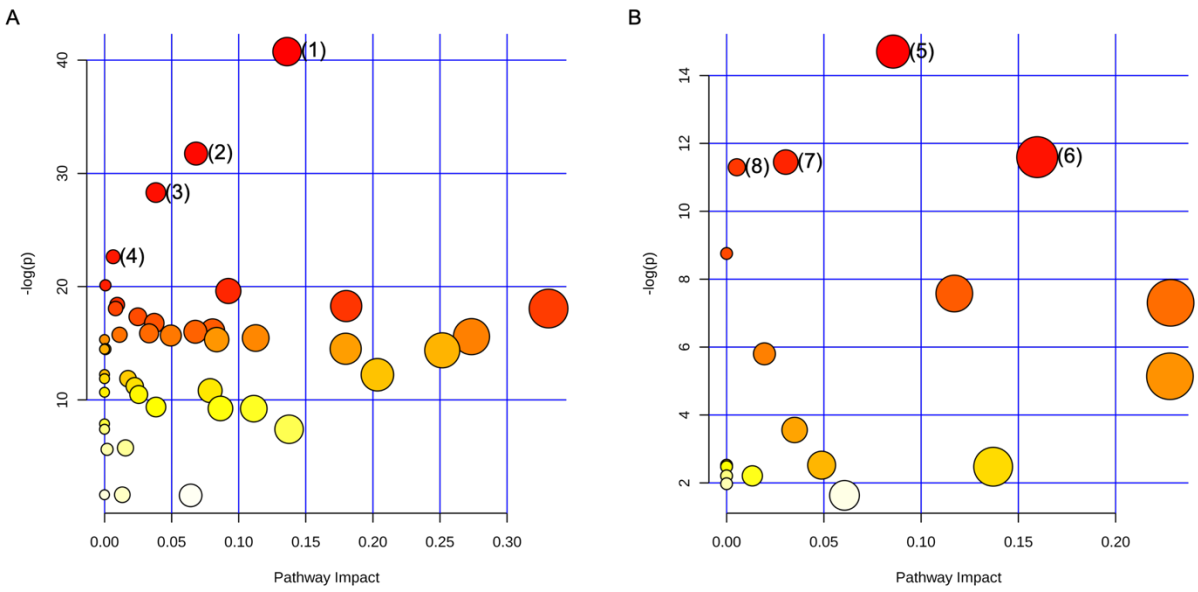


Figure 3.8 The metabolome view of pathway analysis comparing Valley fever (VF) patients and controls using (A) plasma samples [(1) glycine and serine metabolism, (2) purine metabolism, (3) nicotinate and nicotinamide metabolism, (4) ammonia recycling]; and (B) urine samples [(5) nicotinate and nicotinamide metabolism, (6) ammonia recycling, (7) phenylalanine metabolism, and (8) arginine and proline metabolism]. Data were $\log_{10}$-transformed prior to analysis.

Discussion

Roughly 10% of patients who contract VF will develop serious chronic diseases or potentially fatal disseminated diseases. Diagnosis of VF remains difficult as currently available diagnostic techniques are inaccurate, nonspecific, and time-consuming. As a result, various diagnostic techniques must be used in conjunction for detection of VF, increasing cost and time to diagnosis. Significant metabolic alterations have previously been shown in response to various fungal infections and have demonstrated potential for use as diagnostic biomarkers (Culibrk et al., 2016; Dadwal & Kontoyiannis, 2018; Rubert

et al., 2017; Savelieff & Pappalardo, 2017). For the last two decades, significant innovations in mass spectrometry-based metabolic profiling and analysis of disease-related alterations have been made and, in doing so, these efforts have borne highly sensitive and valuable diagnostic information (Ellis & Goodacre, 2006; Emwas et al., 2013; Jasbi, Wang, et al., 2019; D. Wang et al., 2019; A. Zhang et al., 2012). In the current study, we explored a combination of targeted metabolic profiling and multivariate statistical analysis for the discovery of sensitive and specific metabolite biomarkers for relatively rapid VF detection. We have used this particular method to detect 207 plasma metabolites and 231 urine metabolites from many relevant metabolic pathways. Our multi-step biomarker selection, model construction, and cross validation have demonstrated the robust diagnostic power of this metabolic profiling method in this study of 147 subjects.

Although a number of studies have performed mass spectrometry-based proteomic and transcriptomic analysis for detection of biological alterations in response to coccidioidomycosis infection (Albuquerque et al., 2014; Berber et al., 2001; Lewis et al., 2015; Mitchell et al., 2018; Whiston et al., 2012), no study has, to date, applied metabolomics for the accurate detection of coccidioidomycosis in humans (Sharpton et al., 2009). The targeted LC-MS/MS metabolite profiling approach presented in this study determined 3 significantly altered plasma metabolites with FDR $q < 0.001$ and VIP > 2 and 9 urine metabolites with FDR $q < 0.05$ and VIP > 1, which informed the construction of enhanced OPLS-DA models for the diagnosis of VF. The combination of these 3 plasma metabolites had a diagnostic sensitivity and specificity of 94.4% and 97.6%, respectively, with an AUC of 0.995. Additionally, our urine panel of 9 metabolites provided a diagnostic sensitivity of 89.7% and specificity of 88.1%, with an AUC of 0.929. Although this urine panel was less accurate than the plasma panel, it represents an increase in diagnostic accuracy over currently available urinary antigen tests, which only report sensitivity of

around 70% (Durkin et al., 2008). While one of our future directions is to combine the plasma and urine biomarkers into a single statistical model, the realization of such a model was not possible in the current study given that paired plasma and urine samples were only collected from one patient, and on different days. Therefore, we have focused on the development of thrifty, independent models given the available data. Moreover, the current time-to-diagnosis ranges from a few days using skin-testing methods (Johnson et al., 2012; Wack et al., 2015), which are contingent on an immune response and cannot differentiate current and past infections, to an unacceptably long period of 2 weeks in the case of laboratory culture. In contrast, our metabolomics approach has great potential to achieve accurate diagnosis of VF within 24 hours. Therefore, our results suggest that metabolomics methods can provide a notable improvement to VF diagnosis over currently available serological diagnostics.

Importantly, no significant differences in levels of the 12 marker candidates were observed between VF patients with acute, chronic, or disseminated disease or VF patients with positive, negative, or indeterminant serology results, indicating that both the plasma and urine metabolite panels achieved accurate diagnoses of VF irrespective of clinical course or serological status. Acutely ill patients were detected, indicating that these metabolic markers are present early in infection. Likewise, patients who were chronically ill or had disseminated diseases were also detected, suggesting that these markers persist and are present in extrapulmonary diseases. Although candidate markers presented in this study are capable of accurate VF diagnosis irrespective of stage, and therefore fulfill a critical need in current diagnostic testing, an ideal biomarker panel capable of identifying disease course would further aid clinical decisions. In order to reach this level of analysis, future metabolic biomarker panels should be designed with special attention paid to characterizing differential, stage-dependent metabolites. In addition, perhaps the most

interesting outcome was the correct identification of seronegative patients. These patients were clinically determined to have VF by the Mayo Clinic Arizona diagnostic rubric but were negative by at least one serological test. Frequently, these seronegative or serologically indeterminant patients are symptomatic and have positive radiography and/or histology tests, further complicating the process of differential diagnosis. These cases represent a difficult to diagnose subset of patients and, therefore, an assay that can correctly assess these patients is of critical need. Results of our urinary and plasma biomarker panels demonstrate the significant potential of our LC/MS-MS method to accurately classify this subset of patients.

Although previously unreported in association with Valley fever, several of the plasma and urinary biomarkers identified in our diagnostic assay have been shown to be critical in the initiation and propagation of related fungal diseases. Chitty and colleagues (2017) showed production of inosine via adenylosuccinate lyase to be essential for DNA and RNA synthesis as well as energy production of *Cryptococcus neoformans* in a murine model (Chitty et al., 2017). In another murine model, Alves de Castro *et. al.* (2018) demonstrated cGMP to be a vital component of Sch9, a serine/threonine kinase responsible for target of rapamycin (TOR) signaling, essential for virulence of *Aspergillus fumigatus* (Alves de Castro et al., 2016). In the current study, levels of inosine were found to be increased more than 26-fold in VF patients whereas levels of cGMP exhibited a 16-fold increase between VF patients and controls. Furthermore, phenylacetic acid has been shown to severely limit the proliferative capacity of *Rhizoctonia solani* in plants (Bartz et al., 2013). Similarly, p-coumaric acid was recently demonstrated to significantly inhibit the growth of *Colletotrichum spp. in vitro* (Roy et al., 2018). Levels of both metabolites were decreased by nearly 90% in VF patients as compared to controls. Additionally, fructose-1,6-bisphosphate (F16BP) has been shown to prevent mortality from active *Candida*

albicans bloodstream infection in mice (Santos et al., 2012). Interestingly, levels of F16BP were reduced by almost 60% in VF patients as compared to controls. Further investigation of these candidate markers and their role in *Coccidioides* virulence and survival in a host environment is warranted.

While results of our plasma and urine enrichment analyses are commensurate with each other, they have not been, to the best of our knowledge, previously reported in the literature in regard to VF. Enrichment analysis of plasma and urine data indicated significant and high-impact changes, respectively, in alanine metabolism. Additionally, analysis of plasma data indicated significant enrichment in amino sugar metabolism, while analysis of urine metabolites revealed a more than two-fold reduction in phenylacetate metabolism in response to active VF infection. Therefore, the significant differences and high magnitude effects observed herein provide valuable target pathways for future experimental studies.

The current understanding of the primary and secondary metabolites produced by *Coccidioides sp*. is limited, but metabolism is known to be a key factor in pathogenesis (Culibrk et al., 2016; Dadwal & Kontoyiannis, 2018; Rubert et al., 2017; Savelieff & Pappalardo, 2017). Sharpton *et. al.* used genomic sequencing evidence to suggest that coccidioidal genetic diversion away from its closest genetic relative, *Uncinocarpus reesii* (a non-pathogen), is at least partially the result of acquiring and adapting genes involved in metabolism, membrane biology, and mycotoxin production (Sharpton et al., 2009). They hypothesized that these changes led to metabolic and morphological phenotypes that enabled survival within a living host, ultimately resulting in disease. Interestingly, they found that the subtilisin N domain-containing gene family of serine proteases were significantly increased in fungi of the order *Onygenales*. Serine proteases have not only been implicated in the pathogenicity of *Aspergillus fumigatus* (Monod et al., 2002), but

also more recently a serine/threonine protein kinase and a serine/threonine phosphatase suggested to be involved in *C. posadasii* virulence (Narra et al., 2016). An avirulent *cps1* knockout strain of *C. posadasii* had 7-fold less serine/threonine protein kinase and 3-fold less serine/threonine phosphatase transcription by RNA-seq than its virulent wild-type parent. It is therefore not surprising that glycine and serine metabolism was found to be significantly disturbed in plasma and urine from VF patients in this study.

Both plasma and urinary analyses in this study revealed significant increases in nicotinamide metabolism and ammonia recycling. Increased nicotinamide may be a general host response to fungal infection, as it is involved in innate immune cell function and has previously been shown to decrease enzyme activity in *Candida* and *Trichophyton spp.* infections (Belenky et al., 2007; Ciebiada-Adamiec et al., 2010). Nicotinamide is currently being investigated for an antifungal therapeutic strategy, as it was shown to cause a loss of cell viability and reduce virulence in a mouse model of *C. albicans* infection (Wurtele et al., 2010). On the other hand, increased ammonia recycling could be a coccidioidal mechanism to evade host immune defenses and maximize nitrogen utilization to support the high demand of amino acid synthesis of rapidly dividing cells during infection, similar to that in cancer cells (Spinelli et al., 2017). *Coccidioides spp.* grow in alkaline soil conditions, and they convert the pH of the pulmonary microenvironment to alkaline concentrations during infection by releasing ammonia and urease when spherules rupture during the parasitic cycle (H. Z. Wise et al., 2013). Ammonia and urease elicit a non-protective innate host inflammatory response which contributes significantly to pathogenesis by inducing host cell damage without clearing the fungus (Mirbod-Donovan et al., 2006). Mice infected with a *C. posadasii* strain with double mutant knockout with deleted urease and ureidoglycolate hydrolase, an enzyme upstream of ammonia

synthesis from allantoate, were significantly less virulent and showed increased survival and better granuloma formation than the wild-type controls.

These proof-of-concept diagnostic metabolites are encouraging; however, there are a number of limitations that need to be addressed in future studies. Firstly, none of the non-VF plasma samples and only 2 of the non-VF urine samples tested (1 bacterial pneumonia of an unknown species and 1 fungal pneumonia due to *Candida sp*. with *Microascus sp*.) had non-VF CAP. Although there were no false positive reactions in these 2 non-VF CAP patients, a greater number of this type of patients should be evaluated, as any diagnostic for VF should discern VF from other pneumonial etiologies including bacterial, viral, and other fungal pneumonia-causing genera like *Histoplasma, Blastomyces, Aspergillus*, and *Cryptococcal spp*. Secondly, we assessed a relatively homogenous population of patients residing in Arizona. The most prevalent *Coccidioides sp.* in Arizona is *C. posadasii*, whereas in California *C. immitis* is more prevalent. Additional VF positive samples should be collected from California and South America and evaluated for possible metabolite variation between strains and patients of different regions. Additionally, further studies need to be performed to discriminate whether these significantly produced metabolites are generated by the host or pathogen. Although this determination is not essential for diagnostic purposes, metabolite biomarkers produced by the host are more likely to be variable in heterogeneic patient groups with varying comorbidities, whereas metabolites produced by the pathogen should be more consistently present. If particular metabolites could be attributed to the pathogen, novel inhibitors of metabolic precursors could be investigated as therapeutic targets. Future studies that evaluate metabolite production and consumption from *in vitro Coccidioides spp.* fungal cultures or mouse infection studies can help determine pathogen specific metabolites and metabolic pathways for targeted treatment.

This study is part of a growing body of literature in which an MS-based method has been utilized for disease biomarker discovery and accurate diagnosis (Ahn et al., 2017; Bowers et al., 2014; Buas et al., 2017; G. A. N. Gowda & Raftery, 2013; Kaysen et al., 2015; D. Wang et al., 2019; Yin & Xu, 2017; A. Zhang et al., 2012). To the best of our knowledge, this is the first metabolomics approach for coccidioidomycosis diagnosis. Additionally, in lieu of applying database searches for compound annotation in global profiling, we tested all targeted metabolites reported in this study with pure standard compounds, allowing for improved relative quantification and analytical precision. Results of our fold change analysis, significance testing, as well as enrichment and pathway analyses indicate metabolites and pathways previously shown to be crucial for immune response inhibition, reproduction, dissemination, and pathogenic severity of fungal diseases, broadly defined. Likewise, the metabolites and associated metabolic pathways identified in this study may inform the development of new antifungal treatments for Valley fever.

Conclusions

In this study, we performed comparisons of plasma and urine metabolites from VF patients and non-VF controls using a targeted LC-MS/MS metabolic profiling approach (Jasbi, Wang, et al., 2019) and observed significant alterations in a variety of the metabolites detected. Our results demonstrate the utility of a panel of 3 plasma metabolites with FDR q-values < 0.001 and VIPs > 2 and 9 urine metabolites with q-values < 0.05 and VIPs > 1 for the rapid and accurate diagnosis of VF. These differential metabolites were used to construct predictive classification models that showed high sensitivity, specificity, and overall performance for VF patients of all clinical courses, including seronegative patients. Application of bioinformatic methods expanded basic

knowledge of the metabolome related to VF and showed ubiquitous disturbances in glycine and serine metabolism that could be targeted therapeutically in future studies. As evidenced by our results, this metabolic profiling method can potentially serve as a novel approach for rapid and routine VF diagnosis, with significant advantages to current diagnostic methods. In addition, this study provides a strong basis for larger multi-site projects to validate our findings across different population groups and further advances the development of improved clinical care for VF patients.

CHAPTER 4

METABOLIC PROFILING OF NEOCORTICAL TISSUE DISCRIMINATES ALZHEIMER'S DISEASE FROM MILD COGNITIVE IMPARIMENT, HIGH PATHOLOGY CONTROLS, AND NORMAL CONTROLS

(Published in *Journal of Proteome Research*, 2021, 20: 4303—4317)

Abstract

Alzheimer's disease (AD) is the most common cause of dementia, accounting for an estimated 60 to 80% of cases, and is the sixth-leading cause of death in the United States. While considerable advancements have been made in the clinical care of AD, it remains a complicated disorder that can be difficult to identify definitively in its earliest stages. Recently, mass spectrometry (MS)-based metabolomics has shown significant potential for elucidation of disease mechanisms and identification of therapeutic targets as well diagnostic and prognostic markers that may be useful in resolving some of the difficulties affecting clinical AD studies, such as effective stratification. In this study, complementary gas chromatography- and liquid chromatography-MS platforms were used to detect and monitor 2,080 metabolites and features in 48 post-mortem tissue samples harvested from the superior frontal gyrus of male and female subjects. Samples were taken from four groups: 12 normal control (NC) patients, 12 cognitively normal subjects characterized as high pathology controls (HPC), 12 subjects with non-specific mild cognitive impairment (MCI), and 12 subjects with AD. Multivariate statistics informed the construction and cross-validation ($p < 0.01$) of partial least squares-discriminant analysis (PLS-DA) models defined by a 9-metabolite panel of disease markers (lauric acid, stearic acid, myristic acid, palmitic acid, palmitoleic acid, and four unidentified mass spectral features). Receiver operating characteristic analysis showed high predictive accuracy of

the resulting PLS-DA models for discrimination of NC (97%), HPC (92%), MCI (~96%), and AD (~96%) groups. Pathway analysis revealed significant disturbances in lysine degradation, fatty acid metabolism, and the degradation of branched-chain amino acids. Network analysis showed significant enrichment of 11 enzymes, predominantly within the mitochondria. The results expand basic knowledge of the metabolome related to AD and reveal pathways that can be targeted therapeutically. This study also provides a promising basis for the development of larger multi-site projects to validate these candidate markers in readily available biospecimens such as blood to enable the effective screening, rapid diagnosis, accurate surveillance, and therapeutic monitoring of AD. All raw mass spectrometry data have been deposited to MassIVE (dataset identifier MSV000087165).

Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder marked primarily by cognitive decline and dementia (Costa et al., 2019; Qiu et al., 2007), in addition to the accumulation of extracellular amyloid β (Aβ) plaques and intracellular neurofibrillary tau tangles (Belbin et al., 2011; Hölscher, 2005). AD is the most common cause of adult dementia, accounting for 60 to 80% of cases worldwide (Yi et al., 2017), and is the sixth leading cause of death in the United States (Alzheimer's Association Report, 2020). Currently, AD affects more than 5.8 million Americans (Snowden et al., 2017; Yi et al., 2017), with prevalence expected to triple by 2050 (Hendrix et al., 2016). In the United States, total payments in 2020 for health care, long-term care, and hospice services are estimated to be $305 billion (Alzheimer's Association Report, 2020). Consequently, AD represents a significant threat to human health and exerts a substantial financial and societal impact.

Considerable advancements have been made in the ability to accurately diagnose AD, largely owing to the development of positron emission tomography (PET) scans for detection of plaques and tangles, as well as cerebral spinal fluid (CSF) and plasma tests for AD-associated biomarkers (Hane et al., 2017). However, treatment of patients and appropriate evaluation of the outcomes of clinical studies remain complicated by the many unknowns involved in AD. Individuals with mild cognitive impairment (MCI) may be in the early stages of AD or may be affected by an unrelated disease process, confounding studies of early interventional treatments. In addition, some individuals with intermediate to high levels of AD-associated pathology (plaques and tangles) do not display cognitive deficits (Hyman et al., 2012), indicating at least a partial disconnect between those specific pathologies and cognitive function. These resilient individuals, known as high pathology controls (HPC), can sustain AD-consistent pathology such as high amyloid loads, synaptic and neuronal demise, demyelination, and atrophy, while simultaneously remaining cognitively intact (Hyman et al., 2012; Maarouf et al., 2011). Although critical to the understanding of pathology- and cognition-specific metabolic alterations underlying AD (Beach et al., 2015; Hyman et al., 2012; Maarouf et al., 2011), this group remains understudied. It is also common for AD to be found in conjunction with other pathologies such as synucleinopathy, TDP-43, or microinfarcts (Nelson et al., 2010). Therefore, the ability to more effectively stratify and subgroup individuals would likely produce clearer and more actionable results. Furthermore, provisional diagnosis relies on a combination of mental status testing, neuropsychological tests, interviews with friends and family, laboratory tests, and various brain imaging techniques such as magnetic resonance imaging (MRI), computerized tomography (CT), and PET (Kang et al., 2016; McKhann et al., 1984; Seyfried et al., 2017). These conventional diagnostic methods show low specificity against other dementias (70%) and only moderate sensitivity (80%) (Iverson et

al., 2010). Additionally, these criteria are unable to capture early brain pathology that may predate symptoms by as much as 30 years (Sancesario & Bernardini, 2018); without timely diagnosis, patients are less likely to access appropriate treatment options that may slow disease progression (R González-Domínguez et al., 2017; Roberts & Knopman, 2013). Consequently, there is a critical need for highly sensitive and specific markers of AD that may enable early disease detection as well as identification of potential drug targets, improved prognosis, and monitoring of therapeutic response.

A growing body of evidence suggests that perturbations in various metabolic pathways play a significant role in AD (An et al., 2018; Han et al., 2011; Varma et al., 2018; G. Wang et al., 2014; Zetterberg & Burnham, 2019). Most notably, the mitochondrial cascade hypothesis states that widespread mitochondrial metabolic dysfunction is a strong characteristic of AD and plays a role in the accumulation of Aβ plaques (Hane et al., 2017; G. Wang et al., 2014; J. M. Wilkins & Trushina, 2018). Furthermore, studies have also shown long- and short-chain fatty acids to play an important role in AD pathology, exerting both protective and pathogenic effects (Peña-Bautista et al., 2019; Tynkkynen et al., 2018; Varma et al., 2018; G. Wang et al., 2014; J. M. Wilkins & Trushina, 2018; Yi et al., 2017). Alterations in glycerophospholipid (Klavins et al., 2015) and phosphatidylcholine metabolism (Costa et al., 2019; Klavins et al., 2015; Tynkkynen et al., 2018; J. M. Wilkins & Trushina, 2018; Yi et al., 2017) have also been strongly linked to AD in previous studies utilizing metabolomics-based approaches. More recently, an emerging role of the gut microbiome in AD has been identified (Mahmoudian-Dehkordi et al., 2019), and metabolic profiling studies have demonstrated the potential for bile acids in discriminating subclinical forms of AD (Baloni et al., 2020; Pan et al., 2017). Indeed, with advanced deterioration of cortical vasculature as seen in late-stage AD, it is possible for bile acids to be found in the central nervous system (Graham et al., 2018; Griffiths et al.,

2019; Nho et al., 2019). Metabolomics, the scientific study of metabolic composition and pathways present in biological systems (Zhanghan Chen et al., 2019; Chong et al., 2018; Havelund et al., 2017; Ismail et al., 2019; Jové et al., 2014; Patti et al., 2012; J. M. Wilkins & Trushina, 2018), has facilitated the accurate characterization of various metabolomes for advances in disease classification, drug therapy, and biomarker discovery. More specifically, mass spectrometry (MS) is an analytical approach in metabolomics that allows for the accurate detection and quantification of metabolites in biological samples (Beach et al., 2015; Gu et al., 2013; Kind et al., 2009). Various forms of untargeted and targeted MS-based metabolomic assays have been implemented to profile biochemical processes in AD pathology. Methods mainly include direct infusion-MS (R. González-Domínguez et al., 2014), ultrahigh resolution-MS enabled by liquid chromatography (LC)-Orbitrap (Lin et al., 2013), and gas chromatography (GC)-time-of-flight-MS (G. Wang et al., 2014), which have been used in many previous studies for therapeutic drug targeting, disease characterization, and potential diagnostic biomarkers.

The current study employs hyphenated MS-based assays that combine both targeted and untargeted metabolomics approaches to detect aqueous metabolites, lipids, fatty acids, and bile acids, in addition to profiling unidentified features. A total of 2,080 metabolites/features were detected in 48 samples of superior frontal gyrus tissue taken from four groups of patients: normal control (NC), HPC, MCI, and AD. Multivariate significance testing and model estimation constructed cross-validated partial least squares-discriminant analysis (PLS-DA) models confined to a highly predictive 9-metabolite panel of potential biomarkers capable of distinguishing each clinical group with high sensitivity and specificity. In conjunction with pathway and enrichment analyses, the current study corroborates findings of previous literature and adds to basic knowledge of the metabolome related to neurodegenerative decline as well as the behavioral symptoms

of AD-induced dementia; this study offers a large-scale analysis of molecular alterations associated with AD pathogenesis and progression, potentially supporting future drug development and prevention efforts. Importantly, this study provides clinically relevant candidate biomarkers capable of accurate post-mortem classification which may, eventually, prove useful to *in vivo* diagnosis and disease monitoring.

Methods

Reagents

Acetonitrile (ACN), methanol (MeOH), ammonium acetate (NH_4OAc), acetic acid (AcOH), and isopropanol (IPA), all LC-MS grade, were purchased from Fisher Scientific (Pittsburgh, PA). Ammonium hydroxide (NH_4OH), methyl tert-butyl ether (MTBE), *O*-methylhydroxylamine hydrochloride (MeOX), and N-Methyl-N-(tert-butyldimethylsilyl) trifluoroacetamide (MTBSTFA) were bought from Sigma-Aldrich (Saint Louis, MO). High performance LC grade chloroform ($CHCl_3$) was obtained from VWR (Radnor, PA). Deionized water was provided in-house by a water purification system from EMD Millipore (Billerica, MA). Phosphate buffered saline (PBS) was bought from GE Healthcare Life Sciences (Logan, UT). Standard compounds corresponding to measured aqueous metabolites/features were purchased from Sigma-Aldrich and Fisher Scientific. Lipid standards were purchased from Fisher Scientific, Sigma-Aldrich, and Avanti Polar Lipids (Alabaster, AL).

Clinical Samples

Frozen tissue from the superior frontal gyrus of male and female subjects was obtained from the Arizona Study of Aging and Neurodegenerative Disorders/Brain and Body Donation Program at the Banner Sun Health Research Institute (BSHRI) in Sun City,

Arizona (Beach et al., 2015). Samples were collected under a previously approved institutional review board (IRB) protocol with broad consent for usage of biospecimens (WIRB Protocol #20120821). All research protocols were conducted in accordance with the principles expressed in the Declaration of Helsinki. Subjects were divided into four groups by clinical status (n = 12 for all groups) based on assessment of post-mortem brain pathology and cognitive status before death. These groups were: normal control (NC) subjects with criteria not met for AD neuropathology, cognitively normal subjects with intermediate AD pathology characterized as high pathology controls (HPC), subjects with non-specific mild cognitive impairment (MCI) and intermediate AD neuropathology, and subjects with dementia and high neuropathology with criteria met for AD. For definition of groups by AD pathology, the National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease were used (Hyman et al., 2012). Relevant clinical characteristics were provided by BSHRI for each subject such as age, sex, APOE genotype, post-mortem interval (PMI), Mini-Mental State Examination (MMSE), and Braak score. Additionally, measures of brain pathology including cerebral amyloid angiopathy (CAA), amyloid plaques and neurofibrillary tangles were taken from either a frontal area of the brain or a compilation of sampling from various brain regions.

Targeted LC-MS/MS Aqueous Profiling

For tissue lysates, 400 mg pieces of frozen superior frontal gyrus were hand homogenized in 400 µL of ice-cold sterile PBS containing a protease/phosphatase inhibitor cocktail (Halt, Thermo Scientific). Three samples for which less tissue was available were homogenized in equal ratios (weight to volume) of the PBS/inhibitor solution. Homogenized samples were sonicated on ice in a biosafety cabinet at a 40%

amplitude for a total time of 1 min, with alternating on/off sequences of 15 sec. Samples were then centrifuged for 30 min at 14,000 RPM at 4°C. The supernatants and pellets were stored separately at -80°C until analysis.

Prior to LC-MS/MS targeted measurement, frozen tissue supernatant samples were first thawed overnight under 4°C. Afterward, 50 µL of each sample were placed in a 2 mL Eppendorf vial. The initial step for protein precipitation and metabolite extraction was performed by adding 500 µL MeOH and 50 µL internal standard solution (containing 1,810.5 µM $^{13}C_3$-lactate and 142 µM $^{13}C_5$-glutamic acid). The mixture was then vortexed for 10 s and stored at -20°C for 30 min, followed by centrifugation at 14,000 RPM for 10 min at 4°C. The supernatants (450 µL) were collected into new Eppendorf vials and dried using a CentriVap Concentrator. The dried samples were reconstituted in 150 µL of 40% PBS/60% ACN and centrifuged again at 14,000 RPM at 4°C for 10 min. Afterward, 100 µL of supernatant was collected from each sample into an LC autosampler vial for subsequent analysis. A pooled sample, which was a mixture of all experimental samples, was used as the quality control (QC) sample and injected once every 10 experimental samples.

The targeted LC-MS/MS method used here was modeled after that developed and used in a growing number of studies (Buas et al., 2017; Carroll et al., 2015; Gu et al., 2016, 2015; Li et al., 2018; Zhu et al., 2014). Briefly, all LC-MS/MS experiments were performed on an Agilent 1290 UPLC-6490 QQQ-MS system. Each supernatant sample was injected twice, 10 µL for analysis using negative ionization mode and 4 µL for analysis using positive ionization mode. Both chromatographic separations were performed in hydrophilic interaction chromatography mode on a Waters XBridge BEH Amide column (150 x 2.1 mm, 2.5 µm particle size, Waters Corporation, Milford, MA). The flow rate was 0.3 mL/min, auto-sampler temperature was kept at 4°C, and the column compartment was

set to 40°C. The mobile phase was composed of Solvents A (10 mM NH_4OAc, 10 mM NH_4OH in 95% H_2O/5% ACN) and B (10 mM NH_4OAc, 10 mM NH_4OH in 95% ACN/5% H_2O). After an initial 1 min isocratic elution of 90% B, the percentage of Solvent B decreased to 40% at t = 11 min. The composition of Solvent B was maintained at 40% for 4 min (t = 15 min), after which the percentage of B gradually went back to 90%, to prepare for the next injection. The mass spectrometer was equipped with an electrospray ionization (ESI) source. Targeted data acquisition was performed in multiple-reaction-monitoring (MRM) mode. For targeted data acquisition, we monitored 118 and 160 MRM transitions in negative and positive mode, respectively (278 transitions in total). The whole LC-MS system was controlled by Agilent MassHunter Workstation software. The extracted MRM peaks were integrated using Agilent MassHunter Quantitative Data Analysis software.

Targeted LC-MS/MS Lipidomics

Tissue samples were thawed under 4°C. Then, 200 µL 10x diluted PBS and 80 µL of MeOH containing 50 µM PC (17:0, 17:0) and PG (17:0, 17:0) internal standards were added to 20 mg of each thawed sample. A ½ spoonful of stainless-steel micro beads was added to each sample, which was subsequently homogenized for 20 sec. Afterward, 400 µL of MTBE was added to each sample (MTBE:MeOH:H_2O = 10:2:5, v/v/v) and vortexed for 30 sec followed by sonication in ice bath for 20 min. Lastly, samples were centrifuged at 14,000 RPM to separate phases. The upper MTBE layer (300 µL) was extracted, transferred to new 1.5 mL Eppendorf tubes, dried in a Vacufuge Plus Evaporator (Hamburg, Germany), and then reconstituted with 100 µL 1:1 $CHCl_3$/MeOH. Each sample (80 µL) was then transferred to a LC−MS vial for LC-MS/MS targeted lipidomics analysis, while the remaining 20 µL was pooled to create a QC sample.

For lipidomic profiling, all mass spectrometry experiments were done on an Agilent 1290 LC-6490 QQQ-MS (Santa Clara, CA), and 4 µL was injected for positive ionization, whereas 6 µL was used in negative ion mode injections. Both modes used reverse phase chromatography with a Waters XSelect HSS T3 column (150 × 2.1 mm, 2.5 µm particle size; Waters Corporation, Milford, MA). The flow rate through the column was maintained at 0.3 mL/min. The mobile phase Solvent A was composed of 10 mM NH_4OAc in 60% H2O/ 40% ACN. Solvent B consisted of 10 mM NH_4OAc in 90% IPA/10% ACN. An isocratic elution was used with 50% Solvent B for 3 min before its percentage was gradually increased to 100% over the next 12 min. Following 10 min of continued 100% Solvent B, at t = 25 min, the percent of B was decreased gradually back to 50% to prepare for the next sample injection. The set of lipids covered in our LC-MS/MS lipidomics assay were the same as those in our previous study (Eghlimi et al., 2020), and 357 lipids were selected from various lipid classes including fatty acids, glycolipids, glycerophospholipids, sphingolipids, etc. (Eghlimi et al., 2020). Lipid standards were used to test the MRM and retention time (RT) for each individual lipid.

Untargeted GC-MS Aqueous Profiling

The aqueous bottom layer (180 µL) from the MTBE extraction described above was collected into a new Eppendorf tube for derivatization prior to untargeted metabolic profiling with GC-MS. The collected bottom layer was dried under vacuum at 37ºC for 4 h using a CentriVap Concentrator (Labconco, Fort Scott, KS). The residues were first derivatized with 40 µL of 20 mg/mL MeOX solution in pyridine under 60ºC for 90 min. Next, 60 µL of MTBSTFA containing d_{27}-mysristic acid were added, and the mixture was incubated at 60ºC for 30 min. The samples were then vortexed for 30 sec, followed by

centrifugation at 14,000 RPM for 10 min. Finally, 70 µL of supernatant were collected from each sample into new glass vials for GC-MS analysis.

GC-MS conditions used here were mainly adopted from previous studies (Gu et al., 2013; Kind et al., 2009). Briefly, GC-MS experiments were performed on an Agilent 7820A GC-5977B MSD system (Santa Clara, CA) by injecting 1 µL of prepared samples. Helium was used as the carrier gas with a constant flow rate of 1.2 mL/min. The separation of metabolites was achieved using an Agilent HP-5ms capillary column (30 m x 250 µm x 0.25 µm). The column temperature was maintained at 60°C for 1 min, increased at a rate of 10°C/min to 325°C, and then held at this temperature for 10 min. Mass spectral signals were recorded at an m/z range of 50-600. Data extraction was performed using Agilent MassHunter Profinder software. A batch recursive feature extraction algorithm for small molecules was used, and peaks were filtered so that only peaks with absolute height ≥ 1,000 counts were included. An RT tolerance of 0.10 min was established, and extraction was limited to the largest 1,000 compound groups. Results were filtered if the overall identification score was less than 75.

Long Chain Fatty Acids (LCFAs)

Weighed 20 mg samples were added to separate Eppendorf tubes and prepared using the same protocol as that outlined for LC-MS/MS lipidomics. Derivatization was performed using the same protocol as that outlined for GC-MS untargeted profiling. For analysis of LCFAs, 60 µL of supernatant was transferred to a glass vial for GC-MS analysis, while 20 µL was pooled from each sample for QC analysis. GC-MS method was the same as that for GC-MS untargeted profiling.

Short Chain Fatty Acids (SCFAs)

Frozen tissue samples were first thawed overnight under 4°C. Afterward, 20 mg of each sample was homogenized with 5 µL hexanoic acid-3,3,3 (internal standard), 15 µL sodium hydroxide (NaOH [0.5 M]), and 500 µL methanol (MeOH). Following storage at -20°C for 20 min and centrifugation at 14,000 RPM for 10 min, 450 µL of supernatant were collected and sample pH was adjusted to 10 by adding 30 µL of NaOH:H_2O (1:4, v:v). Samples were then dried, and they were measured using the same protocol as that outlined for GC-MS untargeted profiling (Gu et al., 2021).

Bile Acids

Sample preparation techniques used here are well established and described in the previous literature (Dempsey et al., 2019; Ginos et al., 2018; Gutierrez et al., 2019; Li et al., 2018; Scoville et al., 2019). Briefly, 50 mg of each tissue sample were homogenized with methanol (500 µL) and then vortexed for 10 sec. Samples were stored at −20°C for 20 min, followed by sonication in an ice bath for 10 min and then centrifugation at 14,000 RPM for 15 min at 4°C. Supernatants (450 µL) were vacuum dried and then reconstituted in 100 µL MeOH/H_2O (1:1, v/v). Each prepared sample (2 µL) was injected into the LC-MS system (Agilent 1290 UPLC-6490 QQQ-MS) for analysis using negative ionization mode. The mobile phase was composed of 5 mM NH_4OAc in H_2O with 0.1% AcOH (A) and ACN with 0.1% AcOH (B). After a 1 min of isocratic elution of 75% Solvent A, the content percentage decreased to 5% A at t = 15 min. The composition of Solvent A was then maintained at 5% for 10 min, followed by an increase to 75% at t = 25 min. The MS parameters were the same as those reported for targeted LC-MS/MS aqueous profiling, except that 55 bile acids were included in the detection panel (Gutierrez et al., 2019). Samples were spiked with mixtures of standard compounds to validate bile acid identities.

Data Analysis

Following peak integration, metabolites were filtered for reliability and only those with QC coefficient of variation (CV) < 20% and relative abundance of 1,000 in > 80% of samples were retained for analysis. Data were then normalized by tissue weight and lysate volume as appropriate. The data were $\log_{10}$-transformed and Pareto scaled prior to model construction. Univariate testing was performed using SPSS 22.0 (SPSS Inc., Chicago, IL). Multivariate statistical analyses were performed using open-source R software. Pathway and integrating enzyme enrichment analysis were performed and visualized using MetaboAnalyst v4.0 (Chong et al., 2018).

Results

Clinical Characteristics

A total of 2,080 metabolites and mass spectral features were reliably detected after signal and QC filtering. Of these, 728 were identified with retention times and/or fragment mass spectra using chemical standards, while 1,352 were unidentified *m/z* values. All raw mass spectrometry data have been deposited to MassIVE (dataset identifier MSV000087165). A total of 48 subjects were included in this study: NC (*n* = 12), HPC (*n* = 12), MCI (*n* = 12), and AD (*n* = 12). **Figure 4.1** shows a graphical schema of the analytical workflow. **Table 4.1** shows the clinical information of subjects, while **Table 4.2** shows the neuropathological characteristics of study subjects. Subjects between all experimental groups were age- and sex-matched such that no statistically significant difference was observed between groups ($p > 0.05$). Principal component analysis (PCA) conducted with all reliably detected metabolites (i.e., filtered metabolites) between all groups and QC samples was performed, and 95% confidence intervals were evaluated for potential outliers. QC samples were highly clustered, suggesting good system

performance. However, the initial PCA revealed one outlier (HPC subject) which, upon confirming extensive non-ignorable missingness, was removed from subsequent analyses (**Supplementary Figure 4.1**).

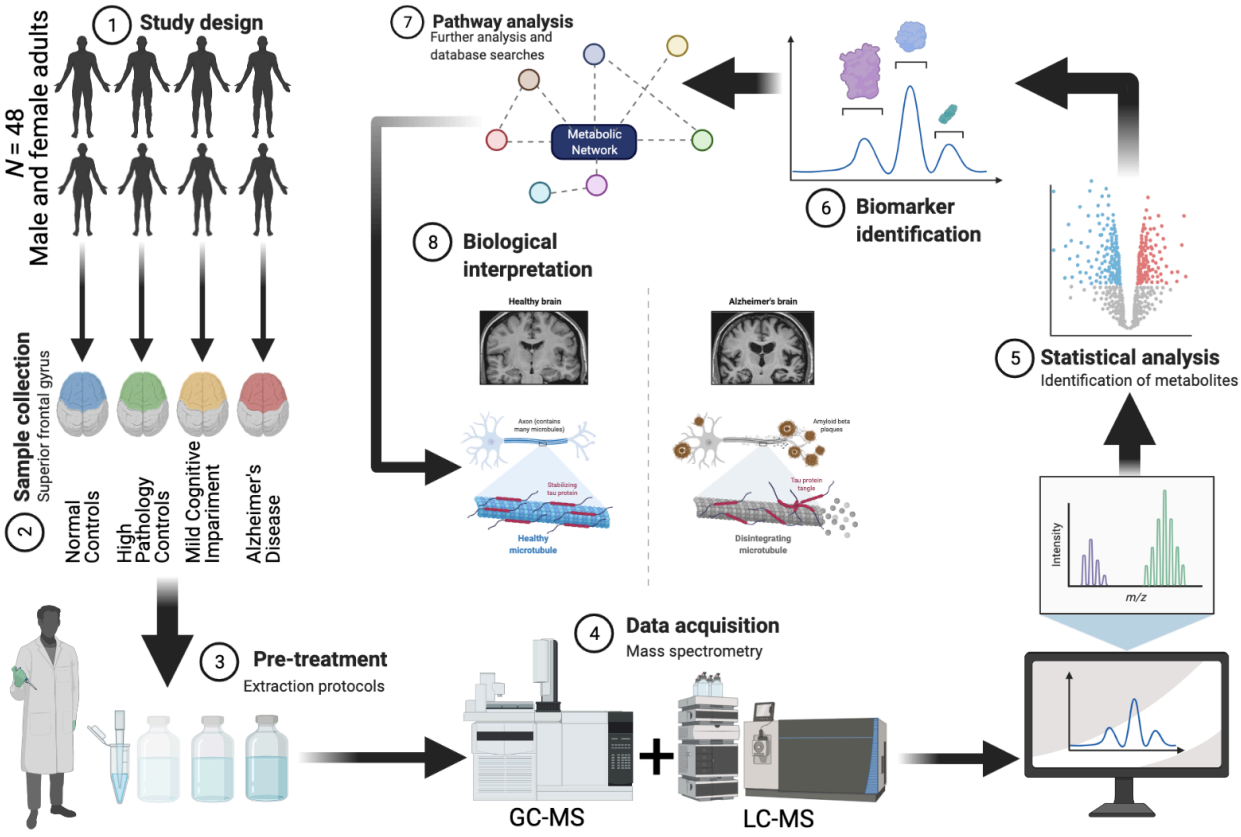


Figure 4.1 Overview of the analytical workflow of the current study. Created with BioRender.com.

Table 4.1 Clinical information of study subjects.

	Non-Demented Normal Controls (*n*=12)	High Pathology Controls (*n*=12)	Mild Cognitive Impairment (*n*=12)	Alzheimer's Disease (*n*=12)
Age in years, mean (SD)	79.7 (12.2)	90.3 (5.1)	88.1 (8.6)	81.3 (8.3)
Sex of subjects (male/female)	7/5	6/6	8/4	7/5
Post-mortem interval mean, hours (SD)	2.93 (1.0)	2.84 (0.9)	3.06 (0.8)	3.2 (0.5)
Mini-mental state examination score mean (SD)	28.8 (0.8)*	27.6 (1.4)	26.0 (3.4)**	10.3 (7.6)
APOE alleles: number of subjects with each genotype 2/3, 2/4, 3/3, 3/4	0,0,8,4	2,0,7,3	0,1,9,2	0,0,6,6

*Scores not available for two subjects.
**Score not available for one subject.

Table 4.2 Neuropathological characteristics of study subjects.

	Non-Demented Normal Controls (*n*=12)	High Pathology Controls (*n*=12)	Mild Cognitive Impairment (*n*=12)	Alzheimer's Disease (*n*=12)
Frontal plaque and tangle scores (number of subjects scoring none, sparse, moderate, or frequent)				
Plaque score (based on CERAD)	10,2,0,0	0,0,0,12	1,1,1,9	0,0,0,12
Tangle score	11,0,0,0†	9,3,0,0	5,6,1,0	0,2,1,9
Braak staging (number of subjects stage 0 to stage VI)	0,5,3,4,0,0	0,0,1,2,9,0	0,0,1,1,9,1,0	0,0,0,0,6,6

†Score not available for one subject.

Case (MCI, AD) vs Control (NC, HPC)

To assess broad differences in metabolic profiles, groups were collapsed among case (MCI and AD) and control (NC and HPC). Initial t-testing between case and control revealed two highly significant and predictive metabolites with $p < 0.001$ and univariate area under curve (AUC) > 0.90: lauric acid and myristic acid. Box plots of these metabolites are given in **Supplementary Figure 4.2**. In addition, a partial least squares-discriminant analysis (PLS-DA) model was constructed using levels of lauric and myristic acid, and receiver operating characteristic (ROC) analysis was conducted using model-

implied values to assess performance. As shown in **Figure 4.2**, the resulting PLS-DA scores plot showed appreciable separation between collapsed case and control groups, and ROC analysis by 100-fold leave-one-out cross validation (LOOCV) showed an overall accuracy of 95%, more than either metabolite individually.

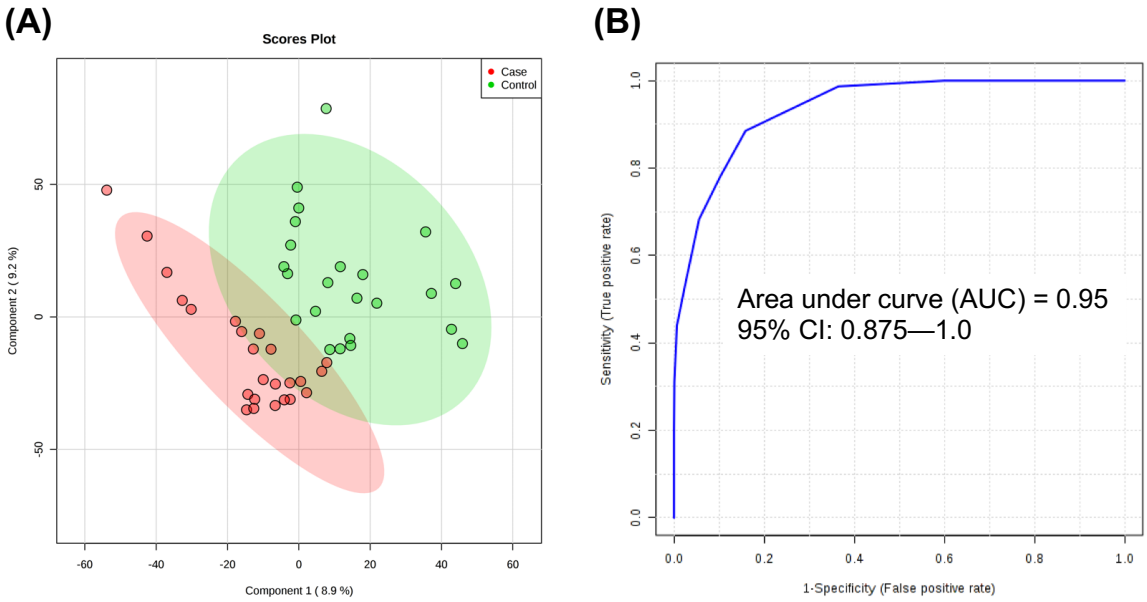


Figure 4.2 PLS-DA and ROC analysis of case (MCI and AD) and control (NC and HPC) constructed using levels of lauric acid and myristic acid: (A) Scores plot of PLS-DA model (R^2X = 0.593, R^2Y = 0.814, R^2Q = 0.701; 10-fold cross validated Q^2 = -0.183) and (B) ROC analysis by 100-fold leave-one-out cross validation (LOOCV) of model-implied values showing AUC = 0.95. PLS-DA, partial least squares-discriminant analysis; ROC, receiver operating characteristic.

NC/HPC vs Other Groups

To analyze differences among groups individually, we first compared metabolic profiles of NC vs HPC/MCI/AD, and HPC vs NC/MCI/AD. Multivariate analysis of variance (MANOVA) testing and associated post-hoc comparisons were performed to identify

significant metabolites while PLS-DA was used to operationalize those metabolites for classification. Well-established risk factors (age, sex, and APOE status) (Liu & Zhou, 2013; Riedel et al., 2016) were included in the MANOVA as covariates and all significance values adjusted accordingly. As outlined in **Table 4.3**, four metabolites (lauric acid, myristic acid, stearic acid, palmitic acid) showed significant main effects, as evidenced by LSD-controlled $p < 0.05$, and were mostly predictive of NC and HPC groups, irrespective of biological risk. Testing for group by age, group by sex, and group by allele interactions revealed no significant effects (p values = 0.077—0.991). Further analysis of a full-factorial GLM showed no significant main effect of age, sex, or APOE allele nor any significant interactions among them (all $p > 0.05$). Normalized box plots of these metabolites between all groups are shown in **Figure 4.3**. A PLS-DA model was constructed using levels of these four significant metabolites, and internal validation was performed using a 100-iteration permutation test. More than 98% of total variance between the four study groups was explained by the first two components, and permutation testing revealed the model to be statistically sound (observed $p < 0.001$). To assess the predictive performance of this unified biomarker panel, the resulting PLS-DA model was subjected to ROC analysis with 100-fold LOOCV. ROC curves for each comparison are provided in **Supplementary Figure 4.3**. Evaluation of model accuracy showed high classification performance for discrimination of NC samples (96.6%, **Supplementary Figure 4.3A**) and HPC samples (91.7%, **Supplementary Figure 4.3B**). We further used a random forest (RF) classifier for group predictions (**Supplementary Figure 4.4**) but found the RF models exhibited poor generalization (OOB error = 0.511) and, therefore, suboptimal predictive performance of case and control (AUC = 0.917), as compared to PLS classification (AUC = 0.95). One potential explanation for this may be the small sample size (48 subjects).

Table 4.3 Significant between-group metabolites as determined by MANOVA testing, focusing on NC and HPC groups.

Metabolite	*F*-value	*p*-value	LSD *p*	Significant post-hoc comparisons
Lauric acid	23.360	3.98E-9	8.29E-6	HPC - AD; NC - AD; HPC - MCI; NC - HPC; NC - MCI
Myristic acid	16.727	3.92E-7	2.49E-4	HPC - AD; NC - AD; HPC - MCI; HPC - NC; NC - MCI
Stearic acid	9.911	4.34E-5	0.036	HPC - AD; NC - AD; HPC - MCI; HPC - NC; NC - MCI
Palmitic acid	9.133	8.56E-5	0.036	HPC - AD; HPC - MCI; HPC - NC; NC - MCI

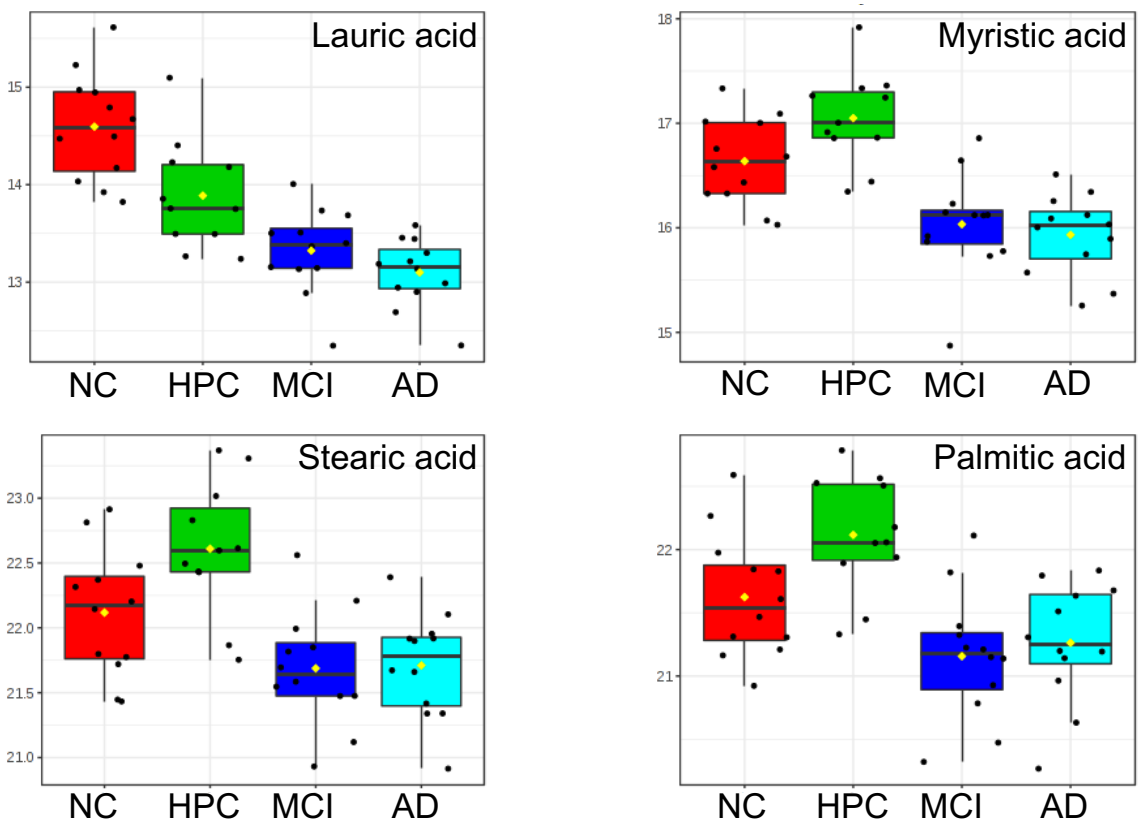


Figure 4.3 Relative abundances of four metabolites found to be significant between groups (LSD $p < 0.05$) by multivariate analysis of variance (MANOVA) testing. Data were $\log_{10}$-transformed and Pareto scaled prior to plotting. NC, normal control; HPC, high pathology control; MCI, mild cognitive impairment; AD, Alzheimer's disease.

MCI vs Other Groups

To increase model performance for discrimination of the MCI subgroup (*see* **Supplementary Figure 4.3C**), other groups were compared sequentially for identification of significant/predictive metabolites and model construction. Comparison of MCI and NC groups revealed lauric acid to be both highly significant ($p < 0.001$) and predictive (AUC = 0.993) (**Supplementary Figure 4.5**). Comparison of MCI and HPC groups revealed four metabolites (myristic acid, palmitic acid, stearic acid, palmitoleic acid) to have AUC > 0.90 and FDR $q < 0.05$ (**Figure 4.4**), while comparison of MCI and AD groups revealed four unidentified features (from untargeted GC-MS) with AUC > 0.80 and $q < 0.05$ (**Figure 4.4**). Candidate metabolites for the classification of MCI and HPC samples were used to construct an independent PLS-DA model, while candidate markers for MCI discrimination from AD samples were ported to construct a separate PLS-DA model. ROC analysis showed a predictive accuracy of 96.6% for the identification of MCI samples from high pathology controls (**Figure 4.5A**). Meanwhile, ROC analysis of the PLS-DA model constructed using levels of four unidentified features with an average AUC ~ 0.811 showed an appreciable improvement in accuracy relative to each univariate AUC; classification accuracy of the PLS-DA model exhibited an AUC = 0.917 for discrimination of MCI and AD groups (**Figure 4.5B**).

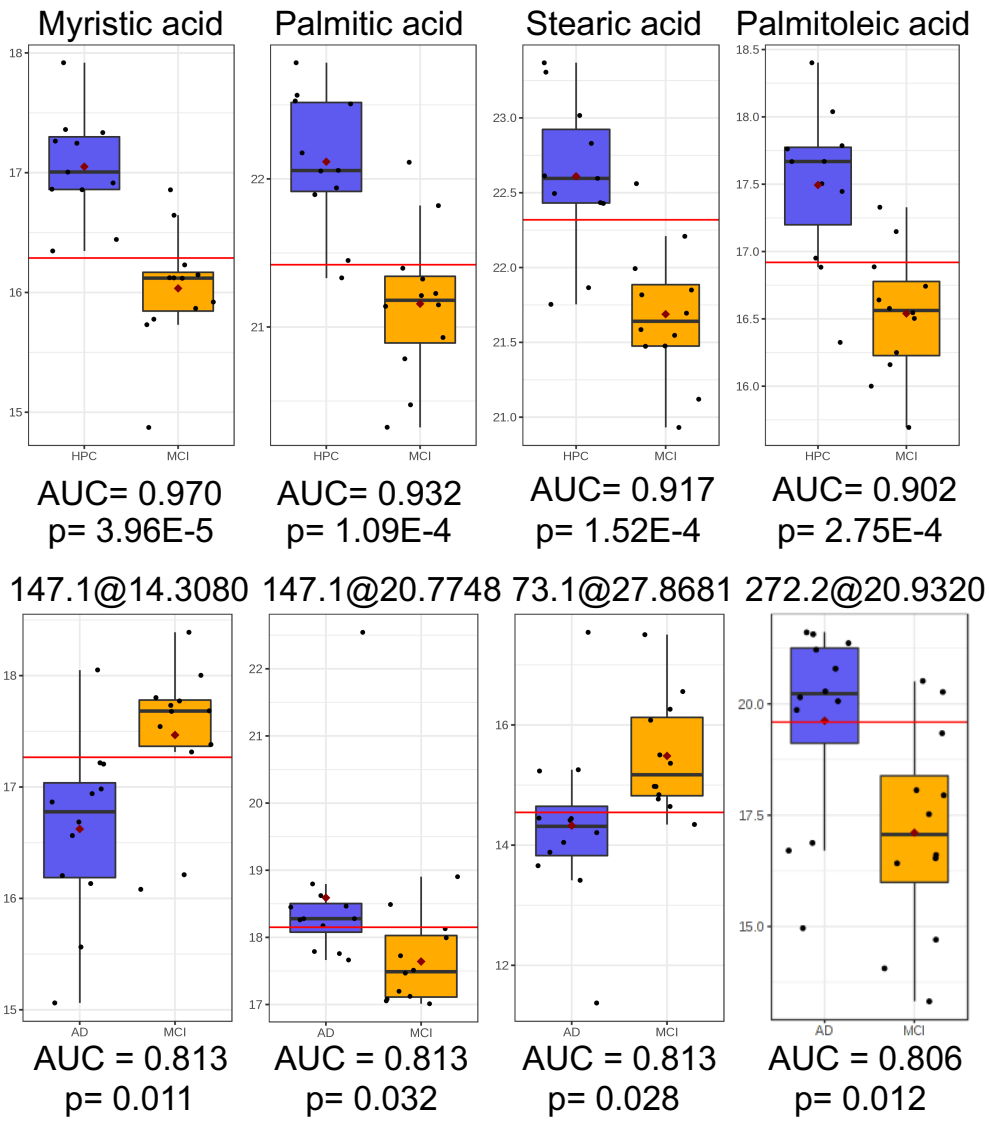


Figure 4.4 *Top row:* Relative abundances of myristic acid, palmitic acid, stearic acid, and palmitoleic acid with high predictive accuracy (AUC > 0.90) and significance (FDR q < 0.05) in univariate ROC analysis and t-testing between HPC and MCI groups. *Bottom row:* Relative abundances of four unidentified features from untargeted GC-MS analysis with good predictive accuracy (AUC > 0.80) and significance (FDR q < 0.05) in univariate ROC analysis and t-testing between MCI and AD groups. MCI, mild cognitive impairment; AD, Alzheimer's disease.

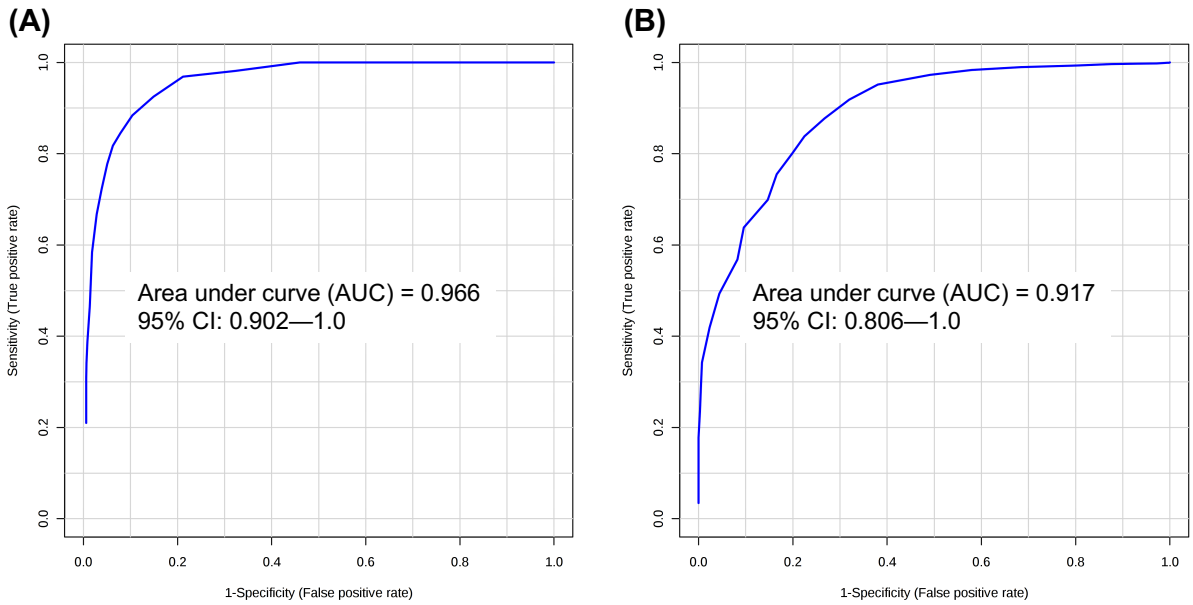


Figure 4.5 ROC analysis of classification performance by 100-fold LOOCV: (A) The PLS-DA model constructed using levels of myristic acid, palmitic acid, stearic acid, and palmitoleic acid (observed p = 0.005) for classification of MCI and HPC samples (AUC = 0.966), and (B) The PLS-DA model constructed using levels of four significant unidentified features (observed p = 0.022) for classification of MCI and AD samples (AUC = 0.917).

AD vs Other Groups

Relevant groups were also compared to AD samples for enhanced identification of disease. Univariate ROC analysis and independent samples t-testing of NC and AD samples showed lauric acid to be highly significant ($p < 0.001$) and predictive (AUC > 0.99) (*see* **Supplementary Figure 4.6**). For classification of AD samples from high pathology controls, t-testing revealed the unified biomarker panel in **Figure 4.3** as being significantly altered between groups ($p < 0.01$) and to have high predictive potential (AUC > 0.90). Direct comparisons between AD and HPC groups for these metabolites are visualized as box plots in **Supplementary Figure 4.7**. For discrimination of these groups, an additional

PLS-DA model was constructed using the aforementioned candidate markers, and ROC analysis revealed high predictive accuracy (94.8%) for discrimination of AD and HPC samples (**Figure 4.6**).

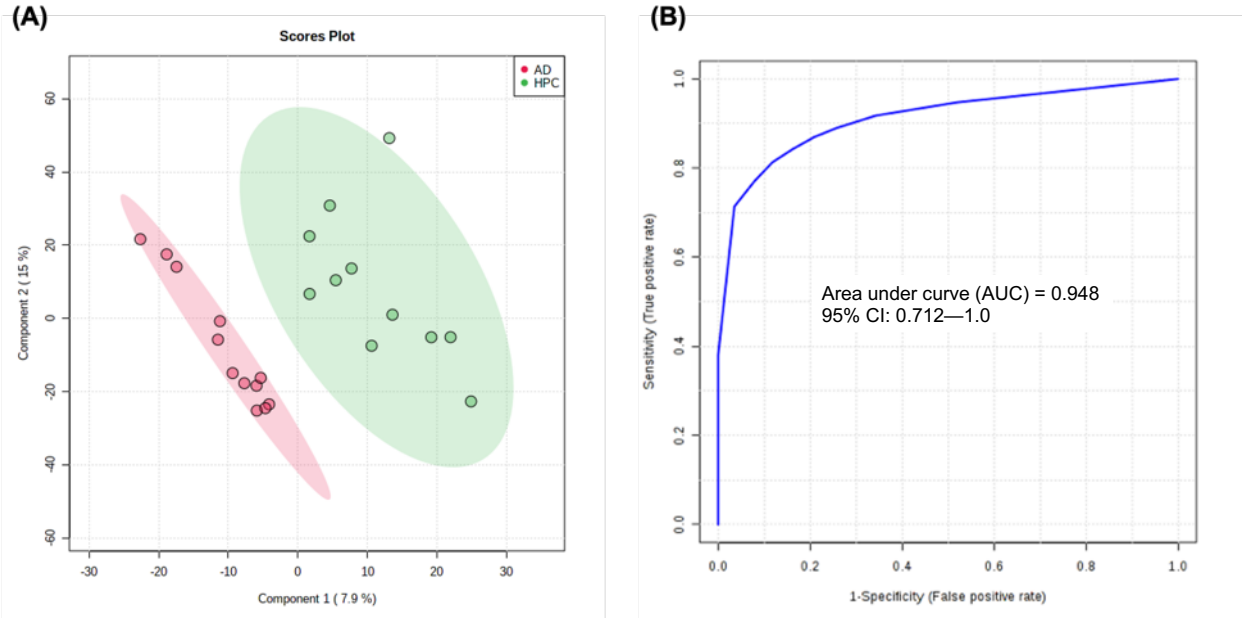


Figure 4.6 (A) PLS-DA model constructed using levels of myristic acid, lauric acid, palmitic acid, and stearic acid for classification of AD and HPC groups (observed p = 0.007), (B) ROC analysis of PLS-DA model by 100-fold LOOCV showing AUC = 0.948. HPC, high pathology control; AD, Alzheimer's disease.

Correlation Analysis of Candidate Markers and Clinical/neuropathological Characteristics

To assess relevant associations between the set of candidate markers and measures of brain pathology and disease progression, a correlation analysis was performed, and measures of association strength and significance were evaluated. A visualization of association strength between correlation variables is given in **Figure 4.7**. Full details regarding magnitude of association (r) and significance of association (p) can be found in **Supplementary Table 4.1**. In total, 4 associations had $r > 0.5$ or < -0.5 and p

< 0.05. Lauric acid showed strong, significant associations with frontal plaque (r = -0.598, p < 0.001), total plaque (r = -0.579, p < 0.001), total tangle (r = -0.507, p < 0.001), and Braak score (r = -0.539, p < 0.001). A PLS-DA model was articulated using this set of four neuropathological characteristics and significant between-group metabolites (lauric acid, myristic acid, stearic acid, and palmitic acid). With the inclusion of these clinical markers, a clear separation of AD, MCI, and HPC groups from normal controls was observed (*see* **Supplementary Figure 4.8**). Furthermore, correlation analysis between age, sex, APOE allele, and all 2,080 reliably detected metabolites/features was also performed; no association was observed to be both strongly correlated (r > |0.5|) and statistically significant (p > 0.05).

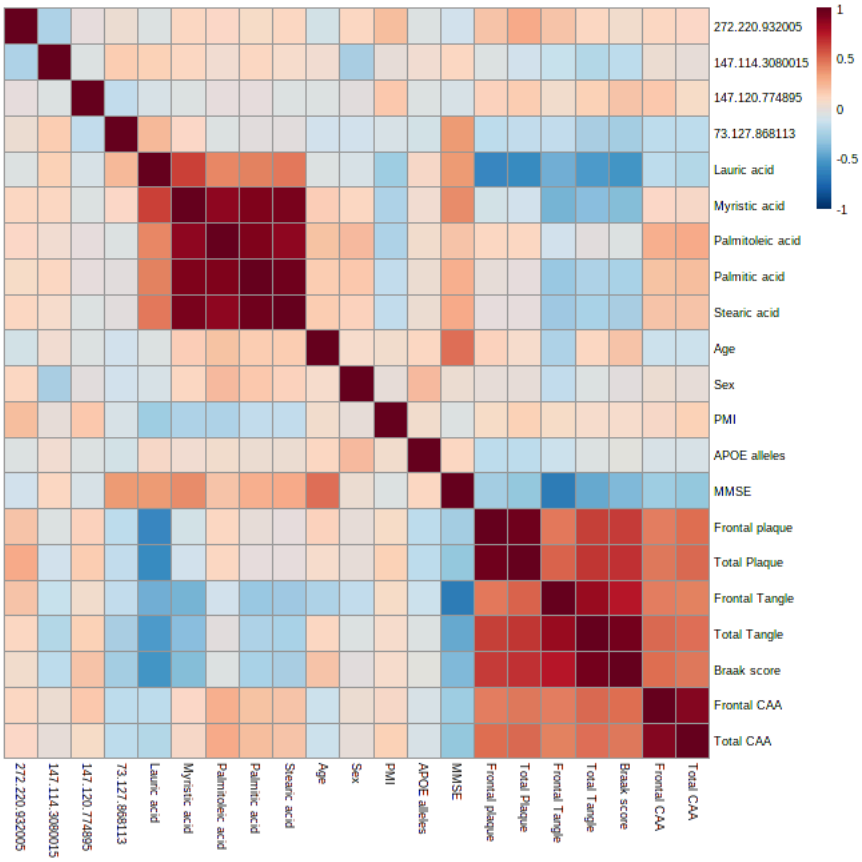


Figure 4.7 Correlation coefficients among the panel of candidate markers and clinical characteristics. CAA, cerebral amyloid angiopathy; PMI, postmortem interval; MMSE, Mini Mental State Examination.

Pathway and Enzyme Enrichment Analyses of Metabolic Data

Subjects were grouped as case (MCI and AD) and control (NC and HPC) for analysis of significantly impacted pathways in response to Alzheimer's progression. Pathway and enzyme analysis was conducted using KEGG database searches and metabolite intensities (**Figure 4.8**). Pathways were mapped to the human metabolome and only identified metabolites confirmed with authentic standards (i.e., based on retention time and MS^2 fragmentation for LC-MS/MS and retention index for GC-MS) were included in the analysis. Three pathways were observed to have large impact coefficients (>0.5):

(1) linoleic acid metabolism, (2) alanine, aspartate, and glutamate metabolism, and (3) arginine and proline metabolism. Importantly, three pathways were found to be significantly affected ($p < 0.05$) as a result of increased AD pathogenesis. Namely, those were lysine degradation, fatty acid metabolism, and valine, leucine, and isoleucine degradation.

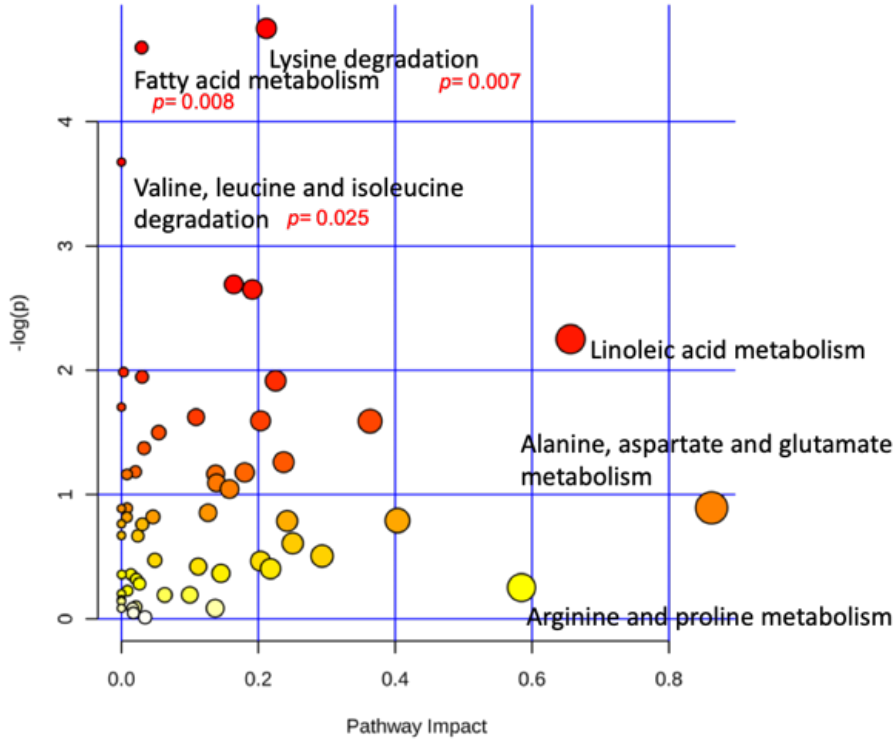


Figure 4.8 Metabolome view of pathway analysis conducted using levels of all reliably detected metabolites showing significantly altered pathways ($p < 0.05$) and those with high impact (> 0.50).

Subjects were dichotomously grouped as case/control, and enrichment analysis was conducted using a library containing 912 metabolic sets that are predicted to be changed in the case of dysfunctional enzymes using a genome-scale network model of human metabolism (**Supplementary Figure 4.9**). Eleven enzymes were found to be

significantly enriched ($p < 0.05$). Notably, seven of those were mitochondrial enzymes indicated against the background of mitochondrial pathways. Full results of the enzyme enrichment analysis are displayed in **Supplementary Table 4.2**.

Discussion

For the last two decades, significant innovations in MS-based metabolic profiling and analysis of disease-related alterations have been made and, in doing so, these efforts have borne highly sensitive and valuable diagnostic information (Andrisic et al., 2018; G. A. N. Gowda & Raftery, 2013; Lawal et al., 2017). In the current study, we explored a combination of targeted and untargeted metabolic profiling in addition to advanced multivariate statistical analysis for the discovery of sensitive and specific metabolite biomarkers for rapid AD classification post-mortem. To capture the diversity of metabolites involved in AD pathobiology, we have used this method to detect 2,080 metabolites of the superior frontal gyrus from many biologically relevant metabolic pathways. Our multi-step biomarker selection, model construction, and cross validation have demonstrated the robust diagnostic power of this metabolic profiling method in this study of 48 NC, HPC, MCI and AD subjects. Additionally, we have applied complementary LC/GC-MS approaches for enhanced monitoring of the metabolome related to AD and, cumulatively, our results show clinically relevant disturbances in energy metabolism and substrate utilization.

The metabolite profiling approach presented in this study determined 5 fatty acids capable of discriminating AD patients from NC and HPC samples with an average AUC of 97%. Recent metabolomics studies have also shown perturbations in fatty acid metabolism across differing Alzheimer pathologies. It was found that the dysregulation of sphingolipids and glycerophospholipids, long-chain fatty acids, and unsaturated fatty acids

have been associated with AD (Snowden et al., 2017; Varma et al., 2018; G. Wang et al., 2014). Similarly, significant disturbances in fatty acid metabolism were also observed in the current study (p = 0.008). More specific to our results, a recent serum profiling approach demonstrated the high predictive accuracy of palmitoleic acid, myristic acid, linoleic acid, and palmitic acid in differentiating central cognitive impairment in AD (J. Wang et al., 2020). Notably, these metabolites were also flagged as candidate markers in our profiling of neocortical tissue. Medium-chain fatty acids like lauric acid, which is found in high levels in coconut oil, have been proposed as possible nutritional therapies for the treatment of cognitive decline (Chatterjee et al., 2020; de la Rubia Ortí et al., 2018; Nafar et al., 2017), and a significant difference in lauric acid was also observed in the AD and NC groups in this study. The markedly reduced levels of these fatty acids observed in our AD subjects could be linked to the impaired glucose metabolism that is well-documented in AD patients (Burns et al., 2013; Zhichun Chen & Zhong, 2013; Szablewski, 2016). Declines in the levels of the identified fatty acids in conjunction with decreased glucose metabolism might suggest that β-oxidation of fatty acids, which is generally low in the brain, is being upregulated to support the energy needs of the brain in AD patients. Supplementation of the fatty acids that can be rapidly metabolized might help support the energy needs of the brain, potentially ameliorating symptoms. This might account for the data cited in the above referenced reports which suggest that the addition of lauric acid to the diet (via coconut oil), may improve some symptoms in AD patients. Lauric acid is known to cross the blood-brain barrier (Spector, 1988), and dietary lauric acid might therefore be accessible as an energy source for the brain (Fernando et al., 2015). The results of the current study warrant further investigation of the therapeutic potential of lauric acid for the treatment and prevention of AD.

Previous studies have shown evidence for brain glucose dysregulation in AD as characterized by higher brain tissue glucose concentration, reduced glycolytic flux, and lower GLUT3 expression as a function of increasing AD pathogenesis (An et al., 2018; Butterfield & Halliwell, 2019; Zhichun Chen & Zhong, 2013). Interestingly, the literature has shown involvement of the Warburg effect in non-tumor disease processes (Zhe Chen et al., 2017) and, in the context of AD, loss of brain aerobic glycolysis as a function of normal human aging is associated with increased tau deposition in preclinical AD (Goyal et al., 2017; Vlassenko et al., 2018). In addition, previous results have shown impaired hypothalamic insulin signaling to be associated with elevated BCAA levels in a mouse model of AD (Ruiz et al., 2016), while defects in BCAA metabolism have in-turn been shown to drive primary AD neuropathology (Ruiz et al., 2016). A prospective cohort study of over 22,000 participants found significant associations between circulating BCAAs and risk of incident dementia and AD (Tynkkynen et al., 2018). It has been shown that defects in BCAA metabolism, and subsequent accumulation, can lead to the phosphorylation of tau proteins and the incidence of AD (H. Li et al., 2018). Other studies have found post-translational modifications to the stabilizing tau proteins, which were induced by lysine residues. It has been proposed that these modifications may play an integral role in the pathobiology of tau protein (Kontaxi et al., 2017). Our pathway analysis also revealed similar results with a significant degradation of lysine ($p = 0.007$) and BCAAs ($p = 0.025$), potentially signifying the underlying pathophysiology of AD. Given the recent failure of numerous billion-dollar clinical trials targeting traditionally hypothesized AD mechanisms such as reduced acetylcholine, Aβ plaques/neurofibrillary tangles, and tau protein (Albensi, 2019), our enzyme and pathway enrichment results further corroborate previous evidence of widespread mitochondrial dysfunction concomitant with Aβ pathology and AD progression (Albensi, 2019; Swerdlow, 2018; H. M. Wilkins & Swerdlow, 2015), providing

compelling evidence for mitochondrial bioenergetics as a novel therapeutic target for preventing/slowing the onset/progression of AD.

Overall, our findings led to an integrated hypothesis describing the pathophysiology of AD in **Figure 4.9** and are conceptualized with respect to the widespread mitochondrial dysfunction observed in our results. In **Figure 4.9**, the darker red areas (to the right) are more increased with AD pathology, and greater enrichment is observed in those pathways. As can be seen, with increased AD pathogenicity (darker red areas), significant metabolic reprogramming is observed. Specifically, a decrease in aerobic glycolysis (lighter red areas) is followed by a shift toward degradation of BCAA for energy production, mostly associated with HPC and MCI subgroups (darker red areas). With even greater disease progression, further metabolic reprogramming is observed; fatty acids are progressively utilized for generation of ATP via increased β-oxidation activity and generation of $FADH_2$ and NADH for oxidative phosphorylation in the electron transport chain (darker red areas). Preference for fatty acid substrates was most pronounced in the MCI and AD subgroups.

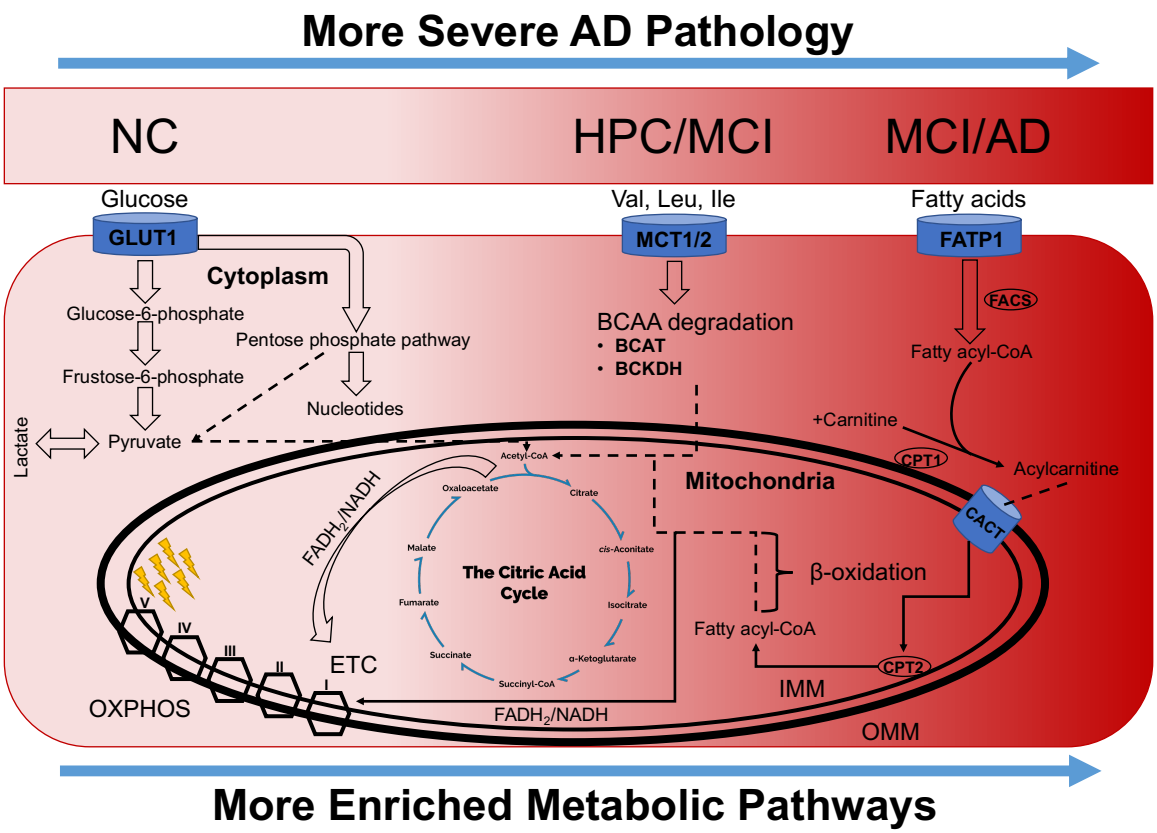


Figure 4.9 Conceptual schema articulating observed changes in substrate utilization and energy production as a function of increasing AD pathogenesis. The darker red areas (to the right) are more increased with AD pathology, and greater enrichment is observed in those pathways. Results show reduced aerobic glycolysis and increased degradation of BCAA associated with HPC and MCI groups as compared to NC. Meanwhile, a preference for fatty acid substrates is seen in MCI and AD groups, with further reductions in aerobic glycolysis as compared to NC. ETC, electron transport chain; IMM, inner mitochondrial membrane; OMM, outer mitochondrial membrane; OXPHOS, oxidative phosphorylation.

Additionally, we evaluated levels of 4 unidentified features with $p < 0.05$ and FC > 2, which informed the construction of independent PLS-DA models for enhanced classification of AD from MCI samples. The combination of these 4 features had a diagnostic sensitivity and specificity of 84.1% and 86.3%, respectively (AUC= 0.917).

Although accurate tests for AD pathology with high severity are currently available (i.e., PET amyloid and tau, CSF amyloid and tau, plasma tau), diagnostic tests useful for intermediate (MCI) and low (HPC) pathology levels are still lacking. In this study, diagnosis of HPC and MCI subgroups was achieved with more than 90% overall AUC. In addition to enabling mass screening, realization of these findings in plasma or CSF may inform clinical trial selection via improved study stratification.

Strengths and Limitations

A major strength of the study lies in the well-characterized BSHRI cohort (Beach et al., 2015) with measures of cognitive status and neuropathological examination at death. Furthermore, inclusion of traditionally understudied HPC and MCI groups allowed for the metabolic characterization of asymptomatic individuals with AD-consistent pathology and non-AD individuals with cognitive decline, respectively. Cumulatively, our panel of candidate markers shows potential for classification of individuals with early brain pathology and other dementias. Although this approach is not suited for *in vivo* diagnosis of AD given the impracticality of brain biopsy for living individuals, the putative metabolite markers/metabolic pathways and associated models reported herein serve as a strong proof-of-principle for their use as therapeutic targets. Furthermore, if validated in readily available biospecimens with minimally invasive sample collection (i.e., from blood draw), this novel panel of candidate markers may enable AD diagnosis in living patients and subsequently enhanced treatment options. Additionally, we applied six distinct metabolomics assays encompassing complementary GC and LC techniques to ensure maximal coverage of the brain metabolome and were able to monitor more than 2,000 metabolites and features. Given the known benefit of complementary MS platforms for elucidation of AD pathology (Fiehn, 2016; R González-Domínguez et al., 2017; Raúl

González-Domínguez et al., 2018), our large-scale multi-platform metabolomics approach utilizing both targeted and untargeted profiling enables comprehensive pathway and enzyme analysis, a key strength of this study to previous literature.

The main limitation of this study is the relatively small sample size. Moreover, our samples were taken cross-sectionally and therefore cannot infer longitudinal changes in metabolite information over time. Also, samples were only taken from a single brain region; inferences to other AD-associated brain structures are unknown. Nevertheless, conventional power was achieved for all biomarker analyses ($\beta < 0.2$), and models were internally validated ($p < 0.01$). It should also be noted that, while the current study infers many proposed alterations in enzymes and pathways using metabolite-level data, the results cannot determine whether the purported markers and associated pathways are drivers of AD pathology or are themselves effects of other latent pathologies; future studies are warranted to investigate this relationship mechanistically. Additionally, of the 48 subjects considered in this study, there were none with APOE 4/4 genotype. There was only one subject with APOE 2/4 genotype and 15 subjects with APOE 3/4 genotype. As APOE 4/4 individuals have the highest risk for AD, it would be useful for future studies to consider subjects with this genotype and monitor metabolites in this group. Our results merit further investigation in a larger sample with serial cognitive assessments taken during life as well as tissue samples collected from distinct brain regions both resistant and vulnerable to AD pathology to monitor possible differential changes between tissue types.

Conclusions

This study is part of a growing body of literature in which MS-based metabolomics methods have been utilized for disease biomarker discovery and accurate classification

(Buas et al., 2017; Chong et al., 2018; Gu et al., 2016; R. Li et al., 2018). We performed comparisons of brain tissue metabolites from AD patients, MCI samples, as well as high pathology and normal controls using both targeted LC-MS/MS metabolic profiling and an untargeted GC-MS approach (Klavins et al., 2015). Our results demonstrate significant alterations in a variety of the metabolites, mainly fatty acids, which are characteristic of different groups. Furthermore, we evaluated the performance of 4 unidentified metabolic features and, through multivariate model construction, achieved an overall classification performance of >90% for comparison of AD and MCI patients, which has the potential to fulfill critical clinical needs (Sun et al., 2020). Application of bioinformatic methods expanded basic knowledge of the metabolome related to AD and showed decreased glycolytic function with increased degradation of BCAAs and β-oxidation of fatty acids associated with increased AD pathogenicity. Results of our fold change analysis, significance testing, and pathway analysis indicate metabolites and pathways previously shown to be crucial in immune response inhibition (Godyń et al., 2016) and increased AD severity (Blennow et al., 2006). Likewise, the metabolites and associated metabolic pathways and enzymes identified in this study may inform the development of new therapeutic treatments for AD. In addition, this study provides a strong basis for larger multi-site projects to validate our findings across different population groups and further advances the development of improved clinical care for AD patients.

CHAPTER 5

CONCLUSIONS

A number of issues impede the translation of metabolomics studies. Broadly defined, these issues are conceptually related to study design, pre-analytic considerations, data analysis and interpretation, and the regulatory and scientific backdrop of translational efforts.

Regarding study design, a few key issues deserve special consideration. First, sample sizes in metabolomics studies are commonly below that needed to reach conventional power (Bujak et al., 2015; Considine, 2019). Although such a lack of satisfactory power would not cause us to doubt our confidence in observed metabolite differences, it fundamentally undermines the degree to which study results may be generalizable to the larger population from which the sample was drawn. Relatedly, validation studies are often missing, and findings are sometimes poorly corroborated or even contradictory (Xia et al., 2013). Critically, many metabolomics studies are clinically uninformed and, as such, are inherently unsuited to overcoming the ambiguities faced by physicians during differential diagnosis (G. A. N. Gowda & Raftery, 2013). Furthermore, some metabolomics studies suffer from a lack of clinical parallelization and thus cannot prove improved efficacy relative to current standard of care (Odom & Sutton, 2021). Moreover, most metabolomics studies are cross-sectional and thus inappropriately designed for monitoring long-term disease course, which is a primary research interest for many diseases posing the greatest threats to human health (G. A. N. Gowda & Raftery, 2013). Finally, many diagnostic metabolomics studies include improper biological samples necessitating collection protocols that are either too invasive, time consuming, or expensive to facilitate routine diagnostic testing and disease monitoring (Lichtenberg et al., 2021). Therefore, whenever possible, researchers should endeavor to collect

biospecimens that can be easily scaled to facilitate expanded testing via end-user sample kits, such as with dried blood spot (DBS) and exhaled breath condensate testing.

There are also pre-analytic aspects of the research that are often not considered at the outset of studies that later preclude the translation of metabolomics results into formalized LDTs. For instance, although it is well known that variations in sample storage, freeze/thaw cycling, temperature and preservative use may impact metabolite profiles and lead to erroneous results (Raúl González-Domínguez et al., 2020; X. Wang et al., 2019), there are no standardized sample handling protocols. Additionally, there are no standard operating protocols (SOPs) for preparation of various biospecimens which may lead to incongruous findings (Dias & Koal, 2016). Furthermore, given the cost and limited availability of many internal standards, only approximately 10% of metabolomics studies are truly quantitative (Pinu et al., 2019), thereby preventing translational efforts such as validation between laboratories and instruments.

Following data collection, post-analytical differences in data processing and results reporting further confuse a clear translational pathway to LDT formalization (Dias & Koal, 2016). Differences in mass spectral peak integration, sample filtering, missing value imputation, data normalization (including transformation and scaling), as well as methods for outlier detection and reconciliation of technical batch effects vary widely in metabolomics and are far too often poorly reported or not reported at all (van den Berg et al., 2006). Of significant detriment to translational efforts is the lack of a central database for indexing metabolite markers (Dias & Koal, 2016). Although metabolomics has greatly benefited from the creation and maintenance of general metabolite databases such as the Human Metabolome Database (HMDB) (Wishart, Guo, et al., 2022) and METLIN (Guijas et al., 2018), as well as fit-for-purpose databases such as the Natural Products Magnetic Resonance Database (NP-MRD) (Wishart, Sayeeda, et al., 2022), the curation of

biological and statistical characteristics of metabolite markers has enjoyed considerably less attention from the scientific community. Without an extensively indexed database of metabolite biomarkers, efforts to consolidate findings to establish simple, non-invasive, rapid, accurate, and cost-effective laboratory testing remain significantly obstructed (Considine, 2019).

Finally, many ambiguities in the current regulatory and scientific landscape of laboratory testing may be to blame for the paucity of translated results. Regarding intellectual property of naturally occurring molecules, uncertainty in patenting rights disincentivizes costly validation efforts critically needed for regulatory approval (Lichtenberg et al., 2021), resulting in a cycle of redundant discovery with meager returns on investment. In addition to these issues in the regulatory sphere, metabolomics must contend with the cultural dilemmas common to all sciences. Chief among these is a skewed scientific incentive that leverages professional advancement for continuous publication, thereby emphasizing discovery with little consideration or praise for researchers engaged in the time-consuming work of translational science.

Nevertheless, these impediments to translational throughput may be largely ameliorated with a conscious detailing of potential solutions. These suggestions are graphically outlined in **Figure 5.1** with respect to the regulatory and scientific topography of translational metabolomics, improvements in study design, and pre-/post-analytical considerations to increase the generalizability and clinical utility of results.

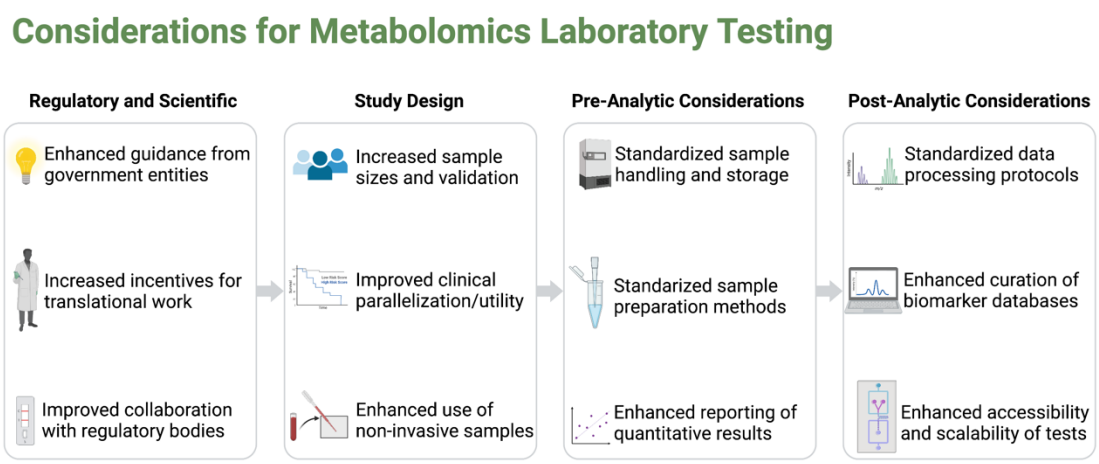


Figure 5.1 Considerations for metabolomics laboratory testing and suggestions for improving translational output. Created with BioRender.com.

Translation of metabolomics findings to consumer facing LDTs would be greatly incentivized by enhanced guidance from government entities regarding the allocation of intellectual property rights, the patentability of metabolites or their relationship to disease, and requirements of CLIA certification for academic laboratories. Top-down, a general paradigm shift in scientific philosophy is required; for instance, journals can do much to encourage high quality translational work by insisting on validation of purported biomarkers as prerequisite to publication. Furthermore, regulatory obstacles and uncertainties can be overcome by close collaboration with regulatory agencies from project conception to review and approval and even post-market surveillance efforts such as external proficiency testing.

Changes to study design can also lessen the translational burden. Although many well-powered metabolomics studies have been conducted, most biomarker studies are under-powered (Bujak et al., 2015; Considine, 2019) and thus suffer from a high probability of type II error. As such, there is increased risk of overfitting and decreased potential for generalizable results. Although not always possible, large well-characterized

cohorts are necessary to power the confident detection of many thousands of metabolites. When not feasible, mining of publicly available datasets combined with *in silico* methods may be a convenient means of increasing sample size and power (Ghosh et al., 2020). Similarly, samples are often drawn from a single study site within a narrow collection window and results may therefore be highly variable across geographic and chronological distance. Consequently, larger, multi-site longitudinal studies are needed to significantly enhance translational success (G. A. N. Gowda & Raftery, 2013). Additionally, metabolomics studies of disease biomarkers are often either not bench-marked against current standards of care or are ill suited for differential diagnosis (G. A. N. Gowda & Raftery, 2013; Odom & Sutton, 2021). Therefore, studies should be designed with special emphasis paid to clinical parallelization and inclusion of any clinically relevant disease groups. Finally, many biomarker papers in the metabolomics space profile inappropriate biospecimen types that are either too invasive, costly, or time-consuming to allow routine testing and monitoring (Lichtenberg et al., 2021; Pinu et al., 2019). Therefore, future studies should investigate disease signatures in conducive sample types such as DBS or exhaled breath condensate.

To mitigate the degree of confounding variance introduced into metabolomics experiments, standardization of sample storage, handling, and preparation protocols is critically needed (Dias & Koal, 2016). A mutually agreed-upon set of operating procedures promulgated by a network of working groups representative of various metabolomics societies would systematically reduce inconsistencies in replication between laboratories. Furthermore, research should attempt to validate previous disease-associated metabolites whenever possible, rather than pursuing novel biomarker discovery efforts (Griffiths et al., 2010; Xia et al., 2013); as such, there is a greater need for targeted metabolite detection, whereas untargeted profiling should be used in exploratory

investigations where no metabolites have previously been associated with the variable of interest. To facilitate this gradual accrual of translational research, authors should endeavor to report structurally identified metabolites whenever possible rather than unknown or tentatively unidentified mass spectral features. Furthermore, absolute quantitation of metabolite concentrations is essential to translation of results between studies.

There are also post-analytical considerations which, if made appropriately, can facilitate discovery of public value. At a fundamental level, standard data practices related to peak processing or normalization procedures should be formally enumerated or, at the very least, be clearly described in sufficient detail to allow replication (van den Berg et al., 2006). Although a basic expectation of all scientific reporting, these crucial details are often incompletely reported or entirely missing from metabolomic biomarker studies. Furthermore, in addition to uploading all raw mass spectral data and relevant subject metadata to publicly available depositories, researchers should detail biomarker results in an appropriate database such as MarkerDB (Wishart et al., 2021). Notably, researchers should index key details that may inform future biomarker validation efforts such as confidence intervals, predictive algorithms, diagnostic sensitivity and specificity, permutation and model fit statistics, measures of effect size, and any relevant statistical corrections. Lastly, translational research should always be undertaken with an eye toward the end user, whether that be a physician or the consumer. As such, researchers in the field should endeavor to reduce the analytical and statistical complexity of metabolomics testing (Pinu et al., 2019). To realize this goal, researchers should strive to produce LDTs on accessible, scalable platforms with automated data processing and reporting such as mini-MS, lab-on-chip, and point-of-care testing formats.

Indeed, metabolomics as a science has facilitated critical advances over the last two decades (Patti et al., 2012). Unfortunately, various factors have historically hindered the translational throughput of metabolomics results (G. A. N. Gowda & Raftery, 2013; Lichtenberg et al., 2021). It is the expressed hope of this author that a thoughtful consideration of these factors and their potential solutions will accelerate the translation of candidate biomarker panels to validated LDTs with high clinical utility.

REFERENCES

Adams, D. A., Thomas, K. R., Jajosky, R. A., Foster, L., Baroi, G., Sharp, P., Onweh, D. H., Schley, A. W., & Anderson, W. J. (2017). *Summary of Notifiable Infectious Diseases and Conditions — United States, 2015*. https://www.cdc.gov/mmwr/mmwr_nd/index.html

Ahn, J., Kim, J., Hwang, J., Song, J., Kim, K., & Cha, H. S. (2017). Urinary Metabolomic Profiling to Identify Potential Biomarkers for the Diagnosis of Behcet's Disease by Gas Chromatography/Time-of-Flight−Mass Spectrometry. *International Journal of Molecular Sciences*, *18*(11), 2309. https://doi.org/10.3390/ijms18112309

Albensi, B. C. (2019). Dysfunction of mitochondria: Implications for Alzheimer's disease. *International Review of Neurobiology*, *145*, 13–27. https://doi.org/10.1016/bs.irn.2019.03.001

Albuquerque, P., Paes, H. C., Tavares, A. H., Fernandes, L., Bocca, A. L., Silva-Pereira, I., Sueli Soares Felipe, M., & Moraes Nicola, A. (2014). Transcriptomics of the Host–Pathogen Interaction in Paracoccidioidomycosis. In *Transcriptomics in Health and Disease* (pp. 265–287). Springer International Publishing. https://doi.org/10.1007/978-3-319-11985-4_14

Alves de Castro, P., dos Reis, T. F., Dolan, S. K., Oliveira Manfiolli, A., Brown, N. A., Jones, G. W., Doyle, S., Riaño-Pachón, D. M., Squina, F. M., Caldana, C., Singh, A., Del Poeta, M., Hagiwara, D., Silva-Rocha, R., & Goldman, G. H. (2016). The *Aspergillus fumigatus* SchA ^{SCH9} kinase modulates SakA ^{HOG1} MAP kinase activity and it is essential for virulence. *Molecular Microbiology*, *102*(4), 642–671. https://doi.org/10.1111/mmi.13484

Alzheimer's Association Report. (2020). 2020 Alzheimer's disease facts and figures. *Alzheimer's and Dementia*, *16*(3), 391–460. https://doi.org/10.1002/alz.12068

Ampel, N. M. (2010). The diagnosis of coccidioidomycosis. *F1000 Medicine Reports*, *2*, 2. https://doi.org/10.3410/M2-2

An, Y., Varma, V. R., Varma, S., Casanova, R., Dammer, E., Pletnikova, O., Chia, C. W., Egan, J. M., Ferrucci, L., Troncoso, J., Levey, A. I., Lah, J., Seyfried, N. T., Legido-Quigley, C., O'Brien, R., & Thambisetty, M. (2018). Evidence for brain glucose dysregulation in Alzheimer's disease. *Alzheimer's and Dementia*, *14*(3), 318–329. https://doi.org/10.1016/j.jalz.2017.09.011

Andrisic, L., Dudzik, D., Barbas, C., Milkovic, L., Grune, T., & Zarkovic, N. (2018). Short overview on metabolomics approach to study pathophysiology of oxidative stress in cancer. *Redox Biology*, *14*, 47–58. https://doi.org/10.1016/j.redox.2017.08.009

Armitage, E. G., & Barbas, C. (2014). Metabolomics in cancer biomarker discovery: Current trends and future perspectives. *Journal of Pharmaceutical and Biomedical Analysis*, *87*, 1–11. https://doi.org/10.1016/j.jpba.2013.08.041

Asiago, V. M., Alvarado, L. Z., Shanaiah, N., Gowda, G. A. N., Owusu-Sarfo, K., Ballas,

R. A., & Raftery, D. (2010). Early Detection of Recurrent Breast Cancer Using Metabolite Profiling. *Cancer Research*, *70*(21), 8309–8318. https://doi.org/10.1158/0008-5472.CAN-10-1319

Bain, J. R., Stevens, R. D., Wenner, B. R., Ilkayeva, O., Muoio, D. M., & Newgard, C. B. (2009). Metabolomics Applied to Diabetes Research: Moving From Information to Knowledge. *Diabetes*, *58*(11), 2429–2443. https://doi.org/10.2337/db09-0580

Baloni, P., Funk, C. C., Yan, J., Yurkovich, J. T., Kueider-Paisley, A., Nho, K., Heinken, A., Jia, W., Mahmoudian-Dehkordi, S., Louie, G., Saykin, A. J., Arnold, M., Kastenmüller, G., Griffiths, W. J., Thiele, I., Kaddurah-Daouk, R., Doraiswamy, P. M., Blach, C., Moseley, A., … Price, N. D. (2020). Metabolic Network Analysis Reveals Altered Bile Acid Synthesis and Metabolism in Alzheimer's Disease. *Cell Reports Medicine*, *1*(8), 100138. https://doi.org/10.1016/j.xcrm.2020.100138

Bartz, F. E., Glassbrook, N. J., Danehower, D. A., & Cubeta, M. A. (2013). Modulation of the phenylacetic acid metabolic complex by quinic acid alters the disease-causing activity of Rhizoctonia solani on tomato. *Phytochemistry*, *89*, 47–52. https://doi.org/10.1016/j.phytochem.2012.09.018

Beach, T. G., Adler, C. H., Sue, L. I., Serrano, G., Shill, H. A., Walker, D. G., Lue, L., Roher, A. E., Dugger, B. N., Maarouf, C., Birdsill, A. C., Intorcia, A., Saxon-Labelle, M., Pullen, J., Scroggins, A., Filon, J., Scott, S., Hoffman, B., Garcia, A., … Sabbagh, M. N. (2015). Arizona Study of Aging and Neurodegenerative Disorders and Brain and Body Donation Program. *Neuropathology*, *35*(4), 354–389. https://doi.org/10.1111/neup.12189

Beckonert, O., Keun, H. C., Ebbels, T. M. D., Bundy, J., Holmes, E., Lindon, J. C., & Nicholson, J. K. (2007). Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols*, *2*(11), 2692–2703. https://doi.org/10.1038/nprot.2007.376

Beger, R. D. (2013). A review of applications of metabolomics in cancer. *Metabolites*, *3*(3), 552–574. https://doi.org/10.3390/metabo3030552

Belbin, O., Brown, K., Shi, H., Medway, C., Abraham, R., Passmore, P., Mann, D., Smith, A. D., Holmes, C., McGuinness, B., Craig, D., Warden, D., Heun, R., Kölsch, H., Love, S., Kalsheker, N., Williams, J., Owen, M. J., Carrasquillo, M., … Kehoe, P. G. (2011). A multi-center study of ACE and the risk of late-onset Alzheimer's disease. *Journal of Alzheimer's Disease*, *24*(3), 587–597. https://doi.org/10.3233/JAD-2011-101914

Belenky, P., Bogan, K. L., & Brenner, C. (2007). NAD+ metabolism in health and disease. *Trends in Biochemical Sciences*, *32*(1), 12–19. https://doi.org/10.1016/j.tibs.2006.11.006

Benedict, K. (2013). Increase in Reported Coccidioidomycosis — United States, 1998–2011. In *Morbidity and Mortality Weekly Report*. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6212a1.htm?s_cid=mm6212a1_e

Berber, A., Gómez-Santos, R., Fanghänel, G., & Sánchez-Reyes, L. (2001). Anthropometric indexes in the prediction of type 2 diabetes mellitus, hypertension and dyslipidaemia in a Mexican population. *International Journal of Obesity*, *25*(12), 1794–1799. https://doi.org/10.1038/sj.ijo.0801827

Bills, G. F., & Gloer, J. B. (2016). Biologically Active Secondary Metabolites from the Fungi. In *The Fungal Kingdom* (Vol. 4, Issue 6, pp. 1087–1119). American Society of Microbiology. https://doi.org/10.1128/microbiolspec.FUNK-0009-2016

Blair, J. E., Coakley, B., Santelli, A. C., Hentz, J. G., & Wengenack, N. L. (2006). Serologic testing for symptomatic coccidioidomycosis in immunocompetent and immunosuppressed hosts. *Mycopathologia*, *162*(5), 317–324. https://doi.org/10.1007/s11046-006-0062-5

Blennow, K., de Leon, M. J., & Zetterberg, H. (2006). Alzheimer's disease. *Lancet*, *368*(9533), 387–403. https://doi.org/10.1016/S0140-6736(06)69113-7

Bowers, J., Hughes, E., Skill, N., Maluccio, M., & Raftery, D. (2014). Detection of hepatocellular carcinoma in hepatitis C patients: Biomarker discovery by LC–MS. *Journal of Chromatography B*, *966*, 154–162. https://doi.org/10.1016/j.jchromb.2014.02.043

Brown, A. J. P., Brown, G. D., Netea, M. G., & Gow, N. A. R. (2014). Metabolism impacts upon Candida immunogenicity and pathogenicity at multiple levels. *Trends in Microbiology*, *22*(11), 614–622. https://doi.org/10.1016/j.tim.2014.07.001

Buas, M. F., Gu, H., Djukovic, D., Zhu, J., Onstad, L., Reid, B. J., Raftery, D., & Vaughan, T. L. (2017). Candidate serum metabolite biomarkers for differentiating gastroesophageal reflux disease, Barrett's esophagus, and high-grade dysplasia/esophageal adenocarcinoma. *Metabolomics*, *13*(3), 23. https://doi.org/10.1007/s11306-016-1154-y

Budczies, J., Denkert, C., Müller, B. M., Brockmöller, S. F., Klauschen, F., Györffy, B., Dietel, M., Richter-Ehrenstein, C., Marten, U., Salek, R. M., Griffin, J. L., Hilvo, M., Orešič, M., Wohlgemuth, G., & Fiehn, O. (2012). Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue – a GC-TOFMS based metabolomics study. *BMC Genomics*, *13*(1), 334. https://doi.org/10.1186/1471-2164-13-334

Bujak, R., Struck-Lewicka, W., Markuszewski, M. J., & Kaliszan, R. (2015). Metabolomics for laboratory diagnostics. *Journal of Pharmaceutical and Biomedical Analysis*, *113*, 108–120. https://doi.org/10.1016/J.JPBA.2014.12.017

Burns, C. M., Chen, K., Kaszniak, A. W., Lee, W., Alexander, G. E., Bandy, D., Fleisher, A. S., Caselli, R. J., & Reiman, E. M. (2013). Higher serum glucose levels are associated with cerebral hypometabolism in Alzheimer regions. *Neurology*, *80*(17), 1557–1564. https://doi.org/10.1212/WNL.0b013e31828f17de

Butterfield, D. A., & Halliwell, B. (2019). Oxidative stress, dysfunctional glucose metabolism and Alzheimer disease. *Nature Reviews Neuroscience*, *20*(3), 148–160.

https://doi.org/10.1038/s41583-019-0132-6

Cao, Z.-G., Qin, X.-B., Liu, F.-F., & Zhou, L.-L. (2015). Tryptophan-induced pathogenesis of breast cancer. *Afri Health Sci*, *15*(3), 982–987. https://doi.org/10.4314/ahs.v15i3.36

Carroll, P. A., Diolaiti, D., McFerrin, L., Gu, H., Djukovic, D., Du, J., Cheng, P. F., Anderson, S., Ulrich, M., Hurley, J. B., Raftery, D., Ayer, D. E., & Eisenman, R. N. (2015). Deregulated Myc requires MondoA/Mlx for metabolic reprogramming and tumorigenesis. *Cancer Cell*, *27*(2), 271–285. https://doi.org/10.1016/j.ccell.2014.11.024

Chang, D. C., Anderson, S., Wannemuehler, K., Engelthaler, D. M., Erhart, L., Sunenshine, R. H., Burwell, L. A., & Park, B. J. (2008). Testing for coccidioidomycosis among patients with community-acquired pneumonia. *Emerging Infectious Diseases*, *14*(7), 1053–1059. https://doi.org/10.3201/eid1407.070832

Chatterjee, P., Fernando, M., Fernando, B., Dias, C. B., Shah, T., Silva, R., Williams, S., Pedrini, S., Hillebrandt, H., Goozee, K., Barin, E., Sohrabi, H. R., Garg, M., Cunnane, S., & Martins, R. N. (2020). Potential of coconut oil and medium chain triglycerides in the prevention and treatment of Alzheimer's disease. *Mechanisms of Ageing and Development*, *186*, 111209. https://doi.org/10.1016/j.mad.2020.111209

Chen, Y., Zhang, R., Song, Y., He, J., Sun, J., Bai, J., An, Z., Dong, L., Zhan, Q., & Abliz, Z. (2009). RRLC-MS/MS-based metabonomics combined with in-depth analysis of metabolic correlation network: finding potential biomarkers for breast cancer. *The Analyst*, *134*(10), 2003–2011. https://doi.org/10.1039/b907243h

Chen, Zhanghan, Li, Z., Li, H., & Jiang, Y. (2019). Metabolomics: A promising diagnostic and therapeutic implement for breast cancer. *OncoTargets and Therapy*, *12*, 6797–6811. https://doi.org/10.2147/OTT.S215628

Chen, Zhe, Liu, M., Li, L., & Chen, L. (2017). Involvement of the Warburg effect in non-tumor diseases processes. *Journal of Cellular Physiology*, *233*(4), 2839–2849. https://doi.org/10.1002/jcp.25998

Chen, Zhichun, & Zhong, C. (2013). Decoding Alzheimer's disease from perturbed cerebral glucose metabolism: Implications for diagnostic and therapeutic strategies. *Progress in Neurobiology*, *108*, 21–43. https://doi.org/10.1016/j.pneurobio.2013.06.004

Chitty, J. L., Blake, K. L., Blundell, R. D., Koh, Y. Q. A. E., Thompson, M., Robertson, A. A. B., Butler, M. S., Cooper, M. A., Kappler, U., Williams, S. J., Kobe, B., & Fraser, J. A. (2017). *Cryptococcus neoformans* ADS lyase is an enzyme essential for virulence whose crystal structure reveals features exploitable in antifungal drug design. *Journal of Biological Chemistry*, *292*(28), 11829–11839. https://doi.org/10.1074/jbc.M117.787994

Chong, J., Soufan, O., Li, C., Caraus, I., Li, S., Bourque, G., Wishart, D. S., & Xia, J. (2018). MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics

analysis. *Nucleic Acids Research*, *46*(W1), W486–W494. https://doi.org/10.1093/nar/gky310

Ciebiada-Adamiec, A., Małafiej, E., & Ciebiada, I. (2010). Inhibitory effect of nicotinamide on enzymatic activity of selected fungal strains causing skin infection. *Mycoses*, *53*(3), 204–207. https://doi.org/10.1111/j.1439-0507.2009.01696.x

Claudino, W. M., Quattrone, A., Biganzoli, L., Pestrin, M., Bertini, I., & Di Leo, A. (2007). Metabolomics: Available Results, Current Research Projects in Breast Cancer, and Future Applications. *Journal of Clinical Oncology*, *25*(19), 2840–2846. https://doi.org/10.1200/JCO.2006.09.7550

Considine, E. C. (2019). The Search for Clinically Useful Biomarkers of Complex Disease: A Data Analysis Perspective. *Metabolites*, *9*(7), 126. https://doi.org/10.3390/METABO9070126

Costa, A. C., Joaquim, H. P. G., Forlenza, O. V., Gattaz, W. F., & Talib, L. L. (2019). Three plasma metabolites in elderly patients differentiate mild cognitive impairment and Alzheimer's disease: a pilot study. *European Archives of Psychiatry and Clinical Neuroscience*, *270*(4), 483–488. https://doi.org/10.1007/s00406-019-01034-9

Crum, N. F., Lederman, E. R., Stafford, C. M., Parrish, J. S., & Wallace, M. R. (2004). Coccidioidomycosis: A Descriptive Survey of a Reemerging Disease. Clinical Characteristics and Current Controversies. *Medicine*, *83*(3), 149–175. https://doi.org/10.1097/01.md.0000126762.91040.fd

Culibrk, L., Croft, C. A., & Tebbutt, S. J. (2016). Systems Biology Approaches for Host–Fungal Interactions: An Expanding Multi-Omics Frontier. *OMICS: A Journal of Integrative Biology*, *20*(3), 127–138. https://doi.org/10.1089/omi.2015.0185

Dadwal, S. S., & Kontoyiannis, D. P. (2018). Recent advances in the molecular diagnosis of mucormycosis. *Expert Review of Molecular Diagnostics*, *18*(10), 845–854. https://doi.org/10.1080/14737159.2018.1522250

de la Rubia Ortí, J. E., García-Pardo, M. P., Drehmer, E., Sancho Cantus, D., Julián Rochina, M., Aguilar, M. A., & Hu Yang, I. (2018). Improvement of Main Cognitive Functions in Patients with Alzheimer's Disease after Treatment with Coconut Oil Enriched Mediterranean Diet: A Pilot Study. *Journal of Alzheimer's Disease*, *65*(2), 577–587. https://doi.org/10.3233/JAD-180184

Dempsey, J. L., Wang, D., Siginir, G., Fei, Q., Raftery, D., Gu, H., & Yue Cui, J. (2019). Pharmacological Activation of PXR and CAR Downregulates Distinct Bile Acid-Metabolizing Intestinal Bacteria and Alters Bile Acid Homeostasis. *Toxicological Sciences*, *168*(1), 40–60. https://doi.org/10.1093/toxsci/kfy271

DeSantis, C. E., Ma, J., Sauer, A. G., Newman, L. A., & Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. *CA: A Cancer Journal for Clinicians*, *67*(6), 439–448. https://doi.org/10.3322/caac.21412

Dias, D. A., & Koal, T. (2016). Progress in Metabolomics Standardisation and its

Significance in Future Clinical Laboratory Medicine. *The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine*, *27*(4), 343. /pmc/articles/PMC5282916/

Duffy, M. J. (2006). Serum Tumor Markers in Breast Cancer: Are They of Clinical Value? *Clinical Chemistry*, *52*(3), 345–351. https://doi.org/10.1373/clinchem.2005.059832

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A., Haselden, J. N., Nicholls, A. W., Wilson, I. D., Kell, D. B., Goodacre, R., & Human Serum Metabolome (HUSERMET) Consortium. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*, *6*(7), 1060–1083. https://doi.org/10.1038/nprot.2011.335

Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., & Griffin, J. L. (2011). Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews*, *40*(1), 387–426. https://doi.org/10.1039/b906712b

Durkin, M., Connolly, P., Kuberski, T., Myers, R., Kubak, B. M., Bruckner, D., Pegues, D., & Wheat, L. J. (2008). Diagnosis of Coccidioidomycosis with Use of the *Coccidioides* Antigen Enzyme Immunoassay. *Clinical Infectious Diseases*, *47*(8), e69–e73. https://doi.org/10.1086/592073

Eghlimi, R., Shi, X., Hrovat, J., Xi, B., & Gu, H. (2020). Triple Negative Breast Cancer Detection Using LC-MS/MS Lipidomic Profiling. *Journal of Proteome Research*, *19*(6), 2367–2378. https://doi.org/10.1021/acs.jproteome.0c00038

Eisenreich, W., Heesemann, J., Rudel, T., & Goebel, W. (2015). Metabolic Adaptations of Intracellullar Bacterial Pathogens and their Mammalian Host Cells during Infection ("Pathometabolism"). *Microbiology Spectrum*, *3*(3). https://doi.org/10.1128/microbiolspec.MBP-0002-2014

Elia, I., Broekaert, D., Christen, S., Boon, R., Radaelli, E., Orth, M. F., Verfaillie, C., Grünewald, T. G. P., & Fendt, S.-M. (2017). Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nature Communications*, *8*, 15267. https://doi.org/10.1038/ncomms15267

Ellis, D. I., & Goodacre, R. (2006). Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *The Analyst*, *131*(8), 875. https://doi.org/10.1039/b602376m

Emwas, A.-H. M., Salek, R. M., Griffin, J. L., & Merzaban, J. (2013). NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations. *Metabolomics*, *9*(5), 1048–1072. https://doi.org/10.1007/s11306-013-0524-y

Ene, I. V., Brunke, S., Brown, A. J. P., & Hube, B. (2014). Metabolism in Fungal Pathogenesis. *Cold Spring Harbor Perspectives in Medicine*, *4*(12), a019695–a019695. https://doi.org/10.1101/cshperspect.a019695

Fan, T. W. M., & Lane, A. N. (2011). NMR-based stable isotope resolved metabolomics in systems biochemistry. *Journal of Biomolecular NMR*, *49*(3–4), 267–280. https://doi.org/10.1007/s10858-011-9484-6

Fan, T. W. M., & Lane, A. N. (2012). The Handbook of Metabolomics. In R. M. Higashi (Ed.), *Handbook of Metabolomics* (Vol. 17). Springer. https://doi.org/10.1007/978-1-61779-618-0

Fan, Y., Zhou, X., Xia, T. S., Chen, Z., Li, J., Liu, Q., Alolga, R. N., Chen, Y., Lai, M. De, Li, P., Zhu, W., & Qi, L. W. (2016). Human plasma metabolomics for identifying differential metabolites and predicting molecular subtypes of breast cancer. *Oncotarget*, *7*(9), 9925–9938. https://doi.org/10.18632/oncotarget.7155

Fernando, W. M. A. D. B., Martins, I. J., Goozee, K. G., Brennan, C. S., Jayasena, V., & Martins, R. N. (2015). The role of dietary coconut for the prevention and treatment of Alzheimer's disease: Potential mechanisms of action. *British Journal of Nutrition*, *114*(1), 1–14. https://doi.org/10.1017/S0007114515001452

Fernie, A. R., Trethewey, R. N., Krotzky, A. J., & Willmitzer, L. (2004). Metabolite profiling: from diagnostics to systems biology. *Nature Reviews Molecular Cell Biology*, *5*(9), 763–769. https://doi.org/10.1038/nrm1451

Fiehn, O. (2002). Metabolomics--the link between genotypes and phenotypes. *Plant Molecular Biology*, *48*(1–2), 155–171.

Fiehn, O. (2016). Metabolomics by gas chromatography-mass spectrometry: Combined targeted and untargeted profiling. *Current Protocols in Molecular Biology*, *114*, 30.4.32. https://doi.org/10.1002/0471142727.mb3004s114

Fisher, F. S., Bultman, M. W., Johnson, S. M., Pappagianis, D., & Zaborsky, E. (2007). Coccidioides niches and habitat parameters in the southwestern United States: A matter of scale. *Annals of the New York Academy of Sciences*, *1111*(1), 47–72. https://doi.org/10.1196/annals.1406.031

Fong, M. Y., Zhou, W., Liu, L., Alontaga, A. Y., Chandra, M., Ashby, J., Chow, A., O'Connor, S. T. F., Li, S., Chin, A. R., Somlo, G., Palomares, M., Li, Z., Tremblay, J. R., Tsuyada, A., Sun, G., Reid, M. A., Wu, X., Swiderski, P., … Wang, S. E. (2015). Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. *Nature Cell Biology*, *17*(2), 183–194. https://doi.org/10.1038/ncb3094

Frickenschmidt, A., Fröhlich, H., Bullinger, D., Zell, A., Laufer, S., Gleiter, C. H., Liebich, H., & Kammerer, B. (2008). Metabonomics in cancer diagnosis: mass spectrometry-based profiling of urinary nucleosides from breast cancer patients. *Biomarkers*, *13*(4), 435–449. https://doi.org/10.1080/13547500802012858

Galgiani, J. N., Ampel, N. M., Blair, J. E., Catanzaro, A., Johnson, R. H., Stevens, D. A., & Williams, P. L. (2005). Coccidioidomycosis. *Clinical Infectious Diseases*, *41*(9), 1217–1223. https://doi.org/10.1086/496991

Ghosh, T., Zhang, W., Ghosh, D., & Kechris, K. (2020). Predictive Modeling for Metabolomics Data. *Methods in Molecular Biology*, *2104*, 313–336. https://doi.org/10.1007/978-1-0716-0239-3_16

Ginos, B. N. R., Navarro, S. L., Schwarz, Y., Gu, H., Wang, D., Randolph, T. W., Shojaie, A., Hullar, M. A. J., Lampe, P. D., Kratz, M., Neuhouser, M. L., Raftery, D., & Lampe, J. W. (2018). Circulating bile acids in healthy adults respond differently to a dietary pattern characterized by whole grains, legumes and fruits and vegetables compared to a diet high in refined grains and added sugars: A randomized, controlled, crossover feeding stud. *Metabolism*, *83*, 197–204. https://doi.org/10.1016/J.METABOL.2018.02.006

Glinton, K. E., & Elsea, S. H. (2019). Untargeted Metabolomics for Autism Spectrum Disorders: Current Status and Future Directions. *Frontiers in Psychiatry*, *10*(eCollection). https://doi.org/10.3389/FPSYT.2019.00647

Godyń, J., Jończyk, J., Panek, D., & Malawska, B. (2016). Therapeutic strategies for Alzheimer's disease in clinical trials. *Pharmacological Reports*, *68*(1), 127–138. https://doi.org/10.1016/j.pharep.2015.07.006

González-Domínguez, R., García-Barrera, T., & Gómez-Ariza, J. L. (2014). Using direct infusion mass spectrometry for serum metabolomics in Alzheimer's disease. *Analytical and Bioanalytical Chemistry*, *406*(28), 7137–7148. https://doi.org/10.1007/s00216-014-8102-3

González-Domínguez, R, Sayago, A., & Fernández-Recamales, Á. (2017). Metabolomics in Alzheimer's disease: The need of complementary analytical platforms for the identification of biomarkers to unravel the underlying pathology. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *1071*, 75–92. https://doi.org/10.1016/j.jchromb.2017.02.008

González-Domínguez, Raúl, González-Domínguez, Á., Sayago, A., & Fernández-Recamales, Á. (2018). Mass spectrometry-based metabolomic multiplatform for Alzheimer's disease research. In *Methods in Molecular Biology* (Vol. 1750, pp. 125–137). Humana Press Inc. https://doi.org/10.1007/978-1-4939-7704-8_8

González-Domínguez, Raúl, González-Domínguez, Á., Sayago, A., & Fernández-Recamales, Á. (2020). Recommendations and best practices for standardizing the pre-analytical processing of blood and urine samples in metabolomics. *Metabolites*, *10*(6), 1–18. https://doi.org/10.3390/metabo10060229

Gorsuch, R. L. (1982). *Factor Analysis* (2nd ed.). LEA.

Gowda, G. A. N., & Raftery, D. (2013). Biomarker Discovery and Translation in Metabolomics. *Current Metabolomics*, *1*(3), 227–240. https://doi.org/10.2174/2213235X113019990005

Gowda, G. N., Zhang, S., Gu, H., Asiago, V., Shanaiah, N., & Raftery, D. (2008). Metabolomics-based methods for early disease diagnostics. *Expert Review of Molecular Diagnostics*, *8*(5), 617–633. https://doi.org/10.1586/14737159.8.5.617

Goyal, M. S., Vlassenko, A. G., Blazey, T. M., Su, Y., Couture, L. E., Durbin, T. J., Bateman, R. J., Benzinger, T. L. S., Morris, J. C., & Raichle, M. E. (2017). Loss of Brain Aerobic Glycolysis in Normal Human Aging. *Cell Metabolism*, *26*(2), 353-360.e3. https://doi.org/10.1016/j.cmet.2017.07.010

Graham, S. F., Rey, N. L., Ugur, Z., Yilmaz, A., Sherman, E., Maddens, M., Bahado-Singh, R. O., Becker, K., Schulz, E., Meyerdirk, L. K., Steiner, J. A., Ma, J., & Brundin, P. (2018). Metabolomic profiling of bile acids in an experimental model of prodromal parkinson's disease. *Metabolites*, *8*(4), 71. https://doi.org/10.3390/metabo8040071

Griffin, J. L., Atherton, H., Shockcor, J., & Atzori, L. (2011). Metabolomics as a tool for cardiac research. *Nature Reviews Cardiology*, *8*(11), 630–643. https://doi.org/10.1038/nrcardio.2011.138

Griffiths, W. J., Abdel-Khalik, J., Yutuc, E., Roman, G., Warner, M., Gustafsson, J. Å., & Wang, Y. (2019). Concentrations of bile acid precursors in cerebrospinal fluid of Alzheimer's disease patients. *Free Radical Biology and Medicine*, *134*, 42–52. https://doi.org/10.1016/j.freeradbiomed.2018.12.020

Griffiths, W. J., Koal, T., Wang, Y., Kohl, M., Enot, D. P., & Deigner, H. P. (2010). Targeted metabolomics for biomarker discovery. *Angewandte Chemie (International Ed. in English)*, *49*(32), 5426–5445. https://doi.org/10.1002/ANIE.200905579

Grys, T., Brighton, A., Chang, Y., Liesman, R., Bolster, L., & Blair, J. (2018). Comparison of two FDA-cleared EIA assays for the detection of Coccidioides antibodies against a composite clinical standard. *Medical Mycology*, *in press*.

Gu, H., Carroll, P. A., Du, J., Zhu, J., Neto, F. C., Eisenman, R. N., & Raftery, D. (2016). Quantitative Method to Investigate the Balance between Metabolism and Proteome Biomass: Starting from Glycine. *Angewandte Chemie International Edition*, *55*(50), 15646–15650. https://doi.org/10.1002/anie.201609236

Gu, H., Gowda, G. A. N., Neto, F. C., Opp, M. R., & Raftery, D. (2013). RAMSY: Ratio Analysis of Mass Spectrometry to Improve Compound Identification. *Analytical Chemistry*, *85*(22), 10771–10779. https://doi.org/10.1021/ac4019268

Gu, H., Gowda, G. N., & Raftery, D. (2012). Metabolic profiling: are we en route to better diagnostic tests for cancer? *Future Oncology*, *8*(10), 1207–1210. https://doi.org/10.2217/fon.12.113

Gu, H., Jasbi, P., Patterson, J., & Yan, J. (2021). Enhanced Detection of Short-Chain Fatty Acids Using Gas Chromatography Mass Spectrometry. *Current Protocols*, *1*(6), e177. https://doi.org/10.1002/cpz1.177

Gu, H., Pan, Z., Xi, B., Asiago, V., Musselman, B., & Raftery, D. (2011). Principal component directed partial least squares analysis for combining nuclear magnetic resonance and mass spectrometry data in metabolomics: Application to the detection of breast cancer. *Analytica Chimica Acta*, *686*(1–2), 57–63. https://doi.org/10.1016/j.aca.2010.11.040

Gu, H., Zhang, P., Zhu, J., & Raftery, D. (2015). Globally Optimized Targeted Mass Spectrometry: Reliable Metabolomics Analysis with Broad Coverage. *Analytical Chemistry*, *87*(24), 12355–12362. https://doi.org/10.1021/acs.analchem.5b03812

Guijas, C., Montenegro-Burke, J. R., Domingo-Almenara, X., Palermo, A., Warth, B., Hermann, G., Koellensperger, G., Huan, T., Uritboonthai, W., Aisporna, A. E., Wolan, D. W., Spilker, M. E., Benton, H. P., & Siuzdak, G. (2018). METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Analytical Chemistry*, *90*(5), 3156–3164. https://doi.org/10.1021/acs.analchem.7b04424

Günther, U. L. (2015). Metabolomics Biomarkers for Breast Cancer. *Pathobiology*, *82*(3–4), 153–165. https://doi.org/10.1159/000430844

Guo, M., Li, X., Zhang, L., Liu, D., Du, W., Yin, D., Lyu, N., Zhao, G., Guo, C., & Tang, D. (2017). Accurate quantification of 5-Methylcytosine, 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxylcytosine in genomic DNA from breast cancer by chemical derivatization coupled with ultra performance liquid chromatography-electrospray quadrupole time . *Oncotarget*, *8*(53), 91248–91257. https://doi.org/10.18632/oncotarget.20093

Gutierrez, D., Weinstock, A., Antharam, V. C., Gu, H., Jasbi, P., Shi, X., Dirks, B., Krajmalnik-Brown, R., Maldonado, J., Guinan, J., & Thangamani, S. (2019). Antibiotic-induced gut metabolome and microbiome alterations increase the susceptibility to Candida albicans colonization in the gastrointestinal tract. *FEMS Microbiology Ecology*, *96*(1), fiz187. https://doi.org/10.1093/femsec/fiz187

Hakimi, A. A., Reznik, E., Lee, C.-H., Creighton, C. J., Brannon, A. R., Luna, A., Aksoy, B. A., Liu, E. M., Shen, R., Lee, W., Chen, Y., Stirdivant, S. M., Russo, P., Chen, Y. B., Tickoo, S. K., Reuter, V. E., Cheng, E. H., Sander, C., & Hsieh, J. J. (2016). An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. *Cancer Cell*, *29*(1), 104–116. https://doi.org/10.1016/j.ccell.2015.12.004

Halama, A., Möller, G., & Adamski, J. (2011). Metabolic Signatures in Apoptotic Human Cancer Cell Lines. *OMICS: A Journal of Integrative Biology*, *15*(5), 325–335. https://doi.org/10.1089/omi.2010.0121

Halama, A., Riesen, N., Möller, G., Hrabě de Angelis, M., & Adamski, J. (2013). Identification of biomarkers for apoptosis in cancer cell lines using metabolomics: tools for individualized medicine. *Journal of Internal Medicine*, *274*(5), 425–439. https://doi.org/10.1111/joim.12117

Han, X., Rozen, S., Boyle, S. H., Hellegers, C., Cheng, H., Burke, J. R., Welsh-Bohmer, K. A., Doraiswamy, P. M., & Kaddurah-Daouk, R. (2011). Metabolomics in early Alzheimer's disease: Identification of altered plasma sphingolipidome using shotgun lipidomics. *PLoS ONE*, *6*(7), e21643. https://doi.org/10.1371/journal.pone.0021643

Hane, F. T., Robinson, M., Lee, B. Y., Bai, O., Leonenko, Z., & Albert, M. S. (2017). Recent Progress in Alzheimer's Disease Research, Part 3: Diagnosis and Treatment. *Journal of Alzheimer's Disease*, *57*(3), 645–665. https://doi.org/10.3233/JAD-160907

Havelund, J. F., Heegaard, N. H. H., Færgeman, N. J. K., & Gramsbergen, J. B. (2017). Biomarker research in parkinson's disease using metabolite profiling. *Metabolites*, *7*(3), 42. https://doi.org/10.3390/metabo7030042

Hayes, J., Peruzzi, P. P., & Lawler, S. (2014). MicroRNAs in cancer: biomarkers, functions and therapy. *Trends in Molecular Medicine*, *20*(8), 460–469. https://doi.org/10.1016/j.molmed.2014.06.005

Hendrix, J. A., Bateman, R. J., Brashear, H. R., Duggan, C., Carrillo, M. C., Bain, L. J., DeMattos, R., Katz, R. G., Ostrowitzki, S., Siemers, E., Sperling, R., & Vitolo, O. V. (2016). Challenges, solutions, and recommendations for Alzheimer's disease combination therapy. *Alzheimer's and Dementia*, *12*(5), 623–630. https://doi.org/10.1016/j.jalz.2016.02.007

Hilvo, M., Denkert, C., Lehtinen, L., Muller, B., Brockmoller, S., Seppanen-Laakso, T., Budczies, J., Bucher, E., Yetukuri, L., Castillo, S., Berg, E., Nygren, H., Sysi-Aho, M., Griffin, J. L., Fiehn, O., Loibl, S., Richter-Ehrenstein, C., Radke, C., Hyotylainen, T., … Oresic, M. (2011). Novel Theranostic Opportunities Offered by Characterization of Altered Membrane Lipid Metabolism in Breast Cancer Progression. *Cancer Research*, *71*(9), 3236–3245. https://doi.org/10.1158/0008-5472.CAN-10-3894

Hölscher, C. (2005). Development of beta-amyloid-induced neurodegeneration in Alzheimer's disease and novel neuroprotective strategies. *Reviews in the Neurosciences*, *16*(3), 181–212. https://doi.org/10.1515/REVNEURO.2005.16.3.181

Horvath, A. R. (2013). From Evidence to Best Practice in Laboratory Medicine. *The Clinical Biochemist Reviews*, *34*(2), 60. /pmc/articles/PMC3799219/

Huang, J. Y., Bristow, B., Shafir, S., & Sorvillo, F. (2012). Coccidioidomycosis-associated Deaths, United States, 1990–2008. *Emerging Infectious Diseases*, *18*(11), 1723–1728. https://doi.org/10.3201/eid1811.120752

Huang, S., Chong, N., Lewis, N. E., Jia, W., Xie, G., & Garmire, L. X. (2016). Novel personalized pathway-based metabolomics models reveal key metabolic pathways for breast cancer diagnosis. *Genome Medicine*, *8*(1), 34. https://doi.org/10.1186/s13073-016-0289-9

Hyman, B. T., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Carrillo, M. C., Dickson, D. W., Duyckaerts, C., Frosch, M. P., Masliah, E., Mirra, S. S., Nelson, P. T., Schneider, J. A., Thal, D. R., Thies, B., Trojanowski, J. Q., Vinters, H. V., & Montine, T. J. (2012). National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease. *Alzheimer's and Dementia*, *8*(1), 1–13. https://doi.org/10.1016/j.jalz.2011.10.007

Ingram, D., Sanders, K., Kolybaba, M., & Lopez, D. (1997). Case-control study of phyto-oestrogens and breast cancer. *The Lancet*, *350*(9083), 990–994. https://doi.org/10.1016/S0140-6736(97)01339-1

Ismail, I. T., Showalter, M. R., & Fiehn, O. (2019). Inborn errors of metabolism in the era of untargeted metabolomics and lipidomics. *Metabolites*, *9*(10), 242.

https://doi.org/10.3390/metabo9100242

Iverson, D. J., Gronseth, G. S., Reger, M. A., Classen, S., Dubinsky, R. M., & Rizzo, M. (2010). Practice parameter update: Evaluation and management of driving risk in dementia: Report of the quality standards subcommittee of the American academy of neurology. *Neurology*, *74*(16), 1316–1324. https://doi.org/10.1212/WNL.0b013e3181da3b0f

Jackson, D. L. (2003). Revisiting Sample Size and Number of Parameter Estimates: Some Support for the N:q Hypothesis. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 128–141. https://doi.org/10.1207/S15328007SEM1001_6

Jain, M., Nilsson, R., Sharma, S., Madhusudhan, N., Kitami, T., Souza, A. L., Kafri, R., Kirschner, M. W., Clish, C. B., & Mootha, V. K. (2012). Metabolite Profiling Identifies a Key Role for Glycine in Rapid Cancer Cell Proliferation. *Science*, *336*(6084), 1040–1044. https://doi.org/10.1126/science.1218595

Jasbi, P., Mitchell, N. M., Shi, X., Grys, T. E., Wei, Y., Liu, L., Lake, D. F., & Gu, H. (2019). Coccidioidomycosis Detection Using Targeted Plasma and Urine Metabolic Profiling. *Journal of Proteome Research*, *18*(7), 2791–2802. https://doi.org/10.1021/acs.jproteome.9b00100

Jasbi, P., Shi, X., Chu, P., Elliott, N., Hudson, H., Jones, D., Serrano, G., Chow, B., Beach, T. G., Liu, L., Jentarra, G., & Gu, H. (2021). Metabolic Profiling of Neocortical Tissue Discriminates Alzheimer's Disease from Mild Cognitive Impairment, High Pathology Controls, and Normal Controls. *Journal of Proteome Research*, *20*(9), 4303–4317. https://doi.org/10.1021/ACS.JPROTEOME.1C00290

Jasbi, P., Wang, D., Cheng, S. L., Fei, Q., Cui, J. Y., Liu, L., Wei, Y., Raftery, D., & Gu, H. (2019). Breast cancer detection using targeted plasma metabolomics. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, *1105*, 26–37. https://doi.org/10.1016/j.jchromb.2018.11.029

Jiang, H., Walter, S. D., Brown, P. E., & Chiarelli, A. M. (2016). Estimation of screening sensitivity and sojourn time from an organized screening program. *Cancer Epidemiology*, *44*, 178–185. https://doi.org/10.1016/j.canep.2016.08.021

Johnson, R., Kernerman, S. M., Sawtelle, B. G., Rastogi, S. C., Nielsen, H. S., & Ampel, N. M. (2012). A Reformulated Spherule-Derived Coccidioidin (Spherusol) to Detect Delayed-Type Hypersensitivity in Coccidioidomycosis. *Mycopathologia*, *174*(5–6), 353–358. https://doi.org/10.1007/s11046-012-9555-6

Jové, M., Collado, R., Quiles, J. L., Ramírez-Tortosa, M. C., Sol, J., Ruiz-Sanjuan, M., Fernandez, M., de la Torre Cabrera, C., Ramírez-Tortosa, C., Granados-Principal, S., Sánchez-Rovira, P., & Pamplona, R. (2017). A plasma metabolomic signature discloses human breast cancer. *Oncotarget*, *8*(12), 19522–19533. https://doi.org/10.18632/oncotarget.14521

Jové, M., Portero-Otín, M., Naudí, A., Ferrer, I., & Pamplona, R. (2014). Metabolomics of human brain aging and age-related neurodegenerative diseases. *Journal of*

Neuropathology and Experimental Neurology, *73*(7), 640–657. https://doi.org/10.1097/NEN.0000000000000091

Kang, S., Jeong, H., Baek, J. H., Lee, S. J., Han, S. H., Cho, H. J., Kim, H., Hong, H. S., Kim, Y. H., Yi, E. C., Seo, S. W., Na, D. L., Hwang, D., & Mook-Jung, I. (2016). PiB-PET Imaging-Based Serum Proteome Profiles Predict Mild Cognitive Impairment and Alzheimer's Disease. *Journal of Alzheimer's Disease*, *53*(4), 1563–1576. https://doi.org/10.3233/JAD-160025

Katz-Brull, R., & Degani, H. (1996). Kinetics of choline transport and phosphorylation in human breast cancer cells; NMR application of the zero trans method. *Anticancer Research*, *16*(3B), 1375–1380. http://www.ncbi.nlm.nih.gov/pubmed/8694504

Kaysen, G. A., Johansen, K. L., Chertow, G. M., Dalrymple, L. S., Kornak, J., Grimes, B., Dwyer, T., Chassy, A. W., & Fiehn, O. (2015). Associations of Trimethylamine N-Oxide With Nutritional and Inflammatory Biomarkers and Cardiovascular Outcomes in Patients New to Dialysis. *Journal of Renal Nutrition*, *25*(4), 351–356. https://doi.org/10.1053/j.jrn.2015.02.006

Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., & Fiehn, O. (2009). FiehnLib: Mass Spectral and Retention Index Libraries for Metabolomics Based on Quadrupole and Time-of-Flight Gas Chromatography/Mass Spectrometry. *Analytical Chemistry*, *81*(24), 10038–10048. https://doi.org/10.1021/ac9019522

Klavins, K., Koal, T., Dallmann, G., Marksteiner, J., Kemmler, G., & Humpel, C. (2015). The ratio of phosphatidylcholines to lysophosphatidylcholines in plasma differentiates healthy controls from patients with Alzheimer's disease and mild cognitive impairment. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, *1*(3), 295–302. https://doi.org/10.1016/j.dadm.2015.05.003

Kontaxi, C., Piccardo, P., & Gill, A. C. (2017). Lysine-directed post-translational modifications of tau protein in Alzheimer's disease and related tauopathies. *Frontiers in Molecular Biosciences*, *4*. https://doi.org/10.3389/fmolb.2017.00056

Krämer, A., Green, J., Pollard, J., & Tugendreich, S. (2014). Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics*, *30*(4), 523–530. https://doi.org/10.1093/bioinformatics/btt703

Lawal, O., Ahmed, W. M., Nijsen, T. M. E., Goodacre, R., & Fowler, S. J. (2017). Exhaled breath analysis: a review of 'breath-taking' methods for off-line analysis. *Metabolomics*, *13*(10), 1–16. https://doi.org/10.1007/s11306-017-1241-8

Lee, J. S., Park, S., Park, J. M., Cho, J. H., Kim, S. Il, & Park, B.-W. (2013). Elevated levels of serum tumor markers CA 15-3 and CEA are prognostic factors for diagnosis of metastatic breast cancers. *Breast Cancer Research and Treatment*, *141*(3), 477–484. https://doi.org/10.1007/s10549-013-2695-7

Lewis, E. R. G., David, V. R., Doyle, A. L., Rajabi, K., Kiefer, J. A., Pirrotte, P., & Barker, B. M. (2015). Differences in Host Innate Responses among Coccidioides Isolates in a Murine Model of Pulmonary Coccidioidomycosis. *Eukaryotic Cell*, *14*(10), 1043–

1053. https://doi.org/10.1128/EC.00122-15

Li, C. Y., Dempsey, J. L., Wang, D., Lee, S. W., Weigel, K. M., Fei, Q., Bhatt, D. K., Prasad, B., Raftery, D., Gu, H., & Cui, J. Y. (2018). PBDEs altered gut microbiome and bile acid homeostasis in male C57BL/6 mice. *Drug Metabolism and Disposition*, *46*(8), 1226–1240. https://doi.org/10.1124/dmd.118.081547

Li, H., Ye, D., Xie, W., Hua, F., Yang, Y., Wu, J., Gu, A., Ren, Y., & Mao, K. (2018). Defect of branched-chain amino acid metabolism promotes the development of Alzheimer's disease by targeting the mTOR signaling. *Bioscience Reports*, *38*(4), BSR20180127. https://doi.org/10.1042/BSR20180127

Li, R., Grimm, S. A., Mav, D., Gu, H., Djukovic, D., Shah, R., Merrick, B. A., Raftery, D., & Wade, P. A. (2018). Transcriptome and DNA Methylome Analysis in a Mouse Model of Diet-Induced Obesity Predicts Increased Risk of Colorectal Cancer. *Cell Reports*, *22*(3), 624–637. https://doi.org/10.1016/j.celrep.2017.12.071

Lichtenberg, S., Trifonova, O. P., Maslov, D. L., Balashova, E. E., & Lokhov, P. G. (2021). Metabolomic Laboratory-Developed Tests: Current Status and Perspectives. *Metabolites*, *11*(7), 423. https://doi.org/10.3390/METABO11070423

Lin, S., Liu, H., Kanawati, B., Liu, L., Dong, J., Li, M., Huang, J., Schmitt-Kopplin, P., & Cai, Z. (2013). Hippocampal metabolomics using ultrahigh-resolution mass spectrometry reveals neuroinflammation from Alzheimer's disease in CRND8 mice. *Analytical and Bioanalytical Chemistry*, *405*(15), 5105–5117. https://doi.org/10.1007/s00216-013-6825-1

Lindon, J. C., & Nicholson, J. K. (2014). The emergent role of metabolic phenotyping in dynamic patient stratification. *Expert Opinion on Drug Metabolism & Toxicology*, *10*(7), 915–919. https://doi.org/10.1517/17425255.2014.922954

Liu, D., & Zhou, X. H. (2013). ROC analysis in biomarker combination with covariate adjustment. *Academic Radiology*, *20*(7), 874–882. https://doi.org/10.1016/j.acra.2013.03.009

Lyman, G. H., Somerfield, M. R., Bosserman, L. D., Perkins, C. L., Weaver, D. L., & Giuliano, A. E. (2017). Sentinel Lymph Node Biopsy for Patients With Early-Stage Breast Cancer: American Society of Clinical Oncology Clinical Practice Guideline Update. *Journal of Clinical Oncology*, *35*(5), 561–564. https://doi.org/10.1200/JCO.2016.71.0947

Maarouf, C. L., Daugs, I. D., Kokjohn, T. A., Walker, D. G., Hunter, J. M., Kruchowsky, J. C., Woltjer, R., Kaye, J., Castaño, E. M., Sabbagh, M. N., Beach, T. G., & Roher, A. E. (2011). Alzheimer's disease and non-demented high pathology control nonagenarians: Comparing and contrasting the biochemistry of cognitively successful aging. *PLoS ONE*, *6*(11), e27291. https://doi.org/10.1371/journal.pone.0027291

Madsen, R., Lundstedt, T., & Trygg, J. (2010). Chemometrics in metabolomics—A review in human disease diagnosis. *Analytica Chimica Acta*, *659*(1–2), 23–33.

https://doi.org/10.1016/J.ACA.2009.11.042

Mahmoudian-Dehkordi, S., Arnold, M., Nho, K., Ahmad, S., Jia, W., Xie, G., Louie, G., Kueider-Paisley, A., Moseley, M. A., Thompson, J. W., St John Williams, L., Tenenbaum, J. D., Blach, C., Baillie, R., Han, X., Bhattacharyya, S., Toledo, J. B., Schafferer, S., Klein, S., … Kaddurah-Daouk, R. (2019). Altered bile acid profile associates with cognitive impairment in Alzheimer's disease—An emerging role for gut microbiome. *Alzheimer's and Dementia*, *15*(1), 76–92. https://doi.org/10.1016/j.jalz.2018.07.217

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology*, *34*(7), 939–944. https://doi.org/10.1212/wnl.34.7.939

Mirbod-Donovan, F., Schaller, R., Hung, C.-Y., Xue, J., Reichard, U., & Cole, G. T. (2006). Urease produced by Coccidioides posadasii contributes to the virulence of this respiratory pathogen. *Infection and Immunity*, *74*(1), 504–515. https://doi.org/10.1128/IAI.74.1.504-515.2006

Mishra, P., & Ambs, S. (2015). Metabolic signatures of human breast cancer. *Molecular & Cellular Oncology*, *2*(3), e992217. https://doi.org/10.4161/23723556.2014.992217

Mitchell, N. M., Sherrard, A. L., Dasari, S., Magee, D. M., Grys, T. E., & Lake, D. F. (2018). Proteogenomic Re-Annotation of Coccidioides posadasii Strain Silveira. *PROTEOMICS*, *18*(1), 1700173. https://doi.org/10.1002/pmic.201700173

Monnerat, G., Seara, F. A. C., Evaristo, J. A. M., Carneiro, G., Evaristo, G. P. C., Domont, G., Nascimento, J. H. M., Mill, J. G., Nogueira, F. C. S., & Campos de Carvalho, A. C. (2018). Aging-related compensated hypogonadism: Role of metabolomic analysis in physiopathological and therapeutic evaluation. *The Journal of Steroid Biochemistry and Molecular Biology*, *183*, 39–50. https://doi.org/10.1016/j.jsbmb.2018.05.005

Monod, M., Capoccia, S., Léchenne, B., Zaugg, C., Holdom, M., & Jousson, O. (2002). Secreted proteases from pathogenic fungi. *International Journal of Medical Microbiology*, *292*(5–6), 405–419. https://doi.org/10.1078/1438-4221-00223

More, T. H., RoyChoudhury, S., Christie, J., Taunk, K., Mane, A., Santra, M. K., Chaudhury, K., & Rapole, S. (2018). Metabolomic alterations in invasive ductal carcinoma of breast: A comprehensive metabolomic study using tissue and serum samples. *Oncotarget*, *9*(2), 2678–2696. https://doi.org/10.18632/oncotarget.23626

Nafar, F., Clarke, J. P., & Mearow, K. M. (2017). Coconut oil protects cortical neurons from amyloid beta toxicity by enhancing signaling of cell survival pathways. *Neurochemistry International*, *105*, 64–79. https://doi.org/10.1016/j.neuint.2017.01.008

Narra, H. P., Shubitz, L. F., Mandel, M. A., Trinh, H. T., Griffin, K., Buntzman, A. S., Frelinger, J. A., Galgiani, J. N., & Orbach, M. J. (2016). A *Coccidioides posadasii*

CPS1 Deletion Mutant Is Avirulent and Protects Mice from Lethal Infection. *Infection and Immunity*, *84*(10), 3007–3016. https://doi.org/10.1128/IAI.00633-16

Nelson, P. T., Abner, E. L., Schmitt, F. A., Kryscio, R. J., Jicha, G. A., Smith, C. D., Davis, D. G., Poduska, J. W., Patel, E., Mendiondo, M. S., & Markesbery, W. R. (2010). Modeling the association between 43 different clinical and pathological variables and the severity of cognitive impairment in a large autopsy cohort of elderly persons. *Brain Pathology*, *20*(1), 66–79. https://doi.org/10.1111/j.1750-3639.2008.00244.x

Nho, K., Kueider-Paisley, A., Mahmoudian-Dehkordi, S., Arnold, M., Risacher, S. L., Louie, G., Blach, C., Baillie, R., Han, X., Kastenmüller, G., Jia, W., Xie, G., Ahmad, S., Hankemeier, T., van Duijn, C. M., Trojanowski, J. Q., Shaw, L. M., Weiner, M. W., Doraiswamy, P. M., … Kaddurah-Daouk, R. (2019). Altered bile acid profile in mild cognitive impairment and Alzheimer's disease: Relationship to neuroimaging and CSF biomarkers. *Alzheimer's and Dementia*, *15*(2), 232–244. https://doi.org/10.1016/j.jalz.2018.08.012

Noether, G. E. (1987). Sample Size Determination for Some Common Nonparametric Tests. *Journal of the American Statistical Association*, *82*(398), 645–647. https://doi.org/10.1080/01621459.1987.10478478

Odom, J. D., & Sutton, V. R. (2021). Metabolomics in Clinical Practice: Improving Diagnosis and Informing Management. *Clinical Chemistry*, *67*(12), 1606–1617. https://doi.org/10.1093/CLINCHEM/HVAB184

Onitilo, A. A., Engel, J. M., Greenlee, R. T., & Mukesh, B. N. (2009). Breast Cancer Subtypes Based on ER/PR and Her2 Expression: Comparison of Clinicopathologic Features and Survival. *Clinical Medicine & Research*, *7*(1–2), 4–13. https://doi.org/10.3121/cmr.2009.825

Pan, X., Elliott, C. T., McGuinness, B., Passmore, P., Kehoe, P. G., Hölscher, C., McClean, P. L., Graham, S. F., & Green, B. D. (2017). Metabolomic profiling of bile acids in clinical and experimental samples of Alzheimer's disease. *Metabolites*, *7*(2), 28. https://doi.org/10.3390/metabo7020028

Patel, S., & Ahmed, S. (2015). Emerging field of metabolomics: big promise for cancer biomarker identification and drug discovery. *Journal of Pharmaceutical and Biomedical Analysis*, *107*, 63–74. https://doi.org/10.1016/j.jpba.2014.12.020

Patti, G. J., Yanes, O., & Siuzdak, G. (2012). Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology*, *13*(4), 263–269. https://doi.org/10.1038/nrm3314

Peña-Bautista, C., Roca, M., Hervás, D., Cuevas, A., López-Cuevas, R., Vento, M., Baquero, M., García-Blanco, A., & Cháfer-Pericás, C. (2019). Plasma metabolomics in early Alzheimer's disease patients diagnosed with amyloid biomarker. *Journal of Proteomics*, *200*, 144–152. https://doi.org/10.1016/j.jprot.2019.04.008

Petrini, B., Sköld, C. M., Bronner, U., & Elmberger, G. (2003). Coccidioidomycosis mimicking lung cancer. *Respiration; International Review of Thoracic Diseases*,

70(6), 651–654. https://doi.org/10.1159/000075215

Pinu, F. R., Goldansaz, S. A., & Jaine, J. (2019). Translational Metabolomics: Current Challenges and Future Opportunities. *Metabolites*, *9*(6), 108. https://doi.org/10.3390/METABO9060108

Pizer, E. S., Jackisch, C., Wood, F. D., Pasternack, G. R., Davidson, N. E., & Kuhajda, F. P. (1996). Inhibition of fatty acid synthesis induces programmed cell death in human breast cancer cells. *Cancer Research*, *56*(12), 2745–2747. http://www.ncbi.nlm.nih.gov/pubmed/8665507

Poddighe, S., Murgia, F., Lorefice, L., Liggi, S., Cocco, E., Marrosu, M. G., & Atzori, L. (2017). Metabolomic analysis identifies altered metabolic pathways in Multiple Sclerosis. *The International Journal of Biochemistry & Cell Biology*, *93*, 148–155. https://doi.org/10.1016/j.biocel.2017.07.004

Prendergast, G. C. (2011). Why tumours eat tryptophan. *Nature*, *478*(7368), 192–194. https://doi.org/10.1038/478192a

Qiu, C., De Ronchi, D., & Fratiglioni, L. (2007). The epidemiology of the dementias: An update. *Current Opinion in Psychiatry*, *20*(4), 380–385. https://doi.org/10.1097/YCO.0b013e32816ebc7b

Reaves, M. L., & Rabinowitz, J. D. (2011). Metabolomics in systems microbiology. *Current Opinion in Biotechnology*, *22*(1), 17–25. https://doi.org/10.1016/j.copbio.2010.10.001

Rhee, E. P., Ho, J. E., Chen, M.-H., Shen, D., Cheng, S., Larson, M. G., Ghorbani, A., Shi, X., Helenius, I. T., O'Donnell, C. J., Souza, A. L., Deik, A., Pierce, K. A., Bullock, K., Walford, G. A., Vasan, R. S., Florez, J. C., Clish, C., Yeh, J.-R. J., … Gerszten, R. E. (2013). A Genome-wide Association Study of the Human Metabolome in a Community-Based Cohort. *Cell Metabolism*, *18*(1), 130–143. https://doi.org/10.1016/j.cmet.2013.06.013

Riedel, B. C., Thompson, P. M., & Brinton, R. D. (2016). Age, APOE and sex: Triad of risk of Alzheimer's disease. *Journal of Steroid Biochemistry and Molecular Biology*, *160*, 134–147. https://doi.org/10.1016/j.jsbmb.2016.03.012

Roberts, R., & Knopman, D. S. (2013). Classification and epidemiology of MCI. *Clinics in Geriatric Medicine*, *29*(4), 753–772. https://doi.org/10.1016/j.cger.2013.07.003

Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, H. T., Abebe, M., Abebe, Z., Abejie, A. N., Abera, S. F., Abil, O. Z., Abraha, H. N., … Murray, C. J. L. (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7

Roy, S., Nuckles, E., & Archbold, D. D. (2018). Effects of Phenolic Compounds on Growth of Colletotrichum spp. In Vitro. *Current Microbiology*, *75*(5), 550–556.

https://doi.org/10.1007/s00284-017-1415-7

Rubert, J., Righetti, L., Stranska-Zachariasova, M., Dzuman, Z., Chrpova, J., Dall'Asta, C., & Hajslova, J. (2017). Untargeted metabolomics based on ultra-high-performance liquid chromatography–high-resolution mass spectrometry merged with chemometrics: A new predictable tool for an early detection of mycotoxins. *Food Chemistry*, *224*, 423–431. https://doi.org/10.1016/j.foodchem.2016.11.132

Ruiz, H. H., Chi, T., Shin, A. C., Lindtner, C., Hsieh, W., Ehrlich, M., Gandy, S., & Buettner, C. (2016). Increased susceptibility to metabolic dysregulation in a mouse model of Alzheimer's disease is associated with impaired hypothalamic insulin signaling and elevated BCAA levels. *Alzheimer's and Dementia*, *12*(8), 851–861. https://doi.org/10.1016/j.jalz.2016.01.008

Sancesario, G. M., & Bernardini, S. (2018). Alzheimer's disease in the omics era. *Clinical Biochemistry*, *59*, 9–16. https://doi.org/10.1016/j.clinbiochem.2018.06.011

Santos, R. C. V., Moresco, R. N., Peña Rico, M. A., Susperregui, A. R. G., Rosa, J. L., Bartrons, R., Ventura, F., Mário, D. N., Alves, S. H., Tatsch, E., Kober, H., de Mello, R. O., Scherer, P., Dias, H. B., & de Oliveira, J. R. (2012). Fructose-1,6-Bisphosphate Reduces the Mortality in Candida albicans Bloodstream Infection and Prevents the Septic-Induced Platelet Decrease. *Inflammation*, *35*(4), 1256–1261. https://doi.org/10.1007/s10753-012-9436-7

Savelieff, M. G., & Pappalardo, L. (2017). Novel cutting-edge metabolite-based diagnostic tools for aspergillosis. *PLOS Pathogens*, *13*(9), e1006486. https://doi.org/10.1371/journal.ppat.1006486

Scalbert, A., Brennan, L., Fiehn, O., Hankemeier, T., Kristal, B. S., van Ommen, B., Pujos-Guillot, E., Verheij, E., Wishart, D., & Wopereis, S. (2009). Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics*, *5*(4), 435–458. https://doi.org/10.1007/s11306-009-0168-0

Scoville, D. K., Li, C. Y., Wang, D., Dempsey, J. L., Raftery, D., Mani, S., Gu, H., & Cui, J. Y. (2019). Polybrominated Diphenyl Ethers and Gut Microbiome Modulate Metabolic Syndrome-Related Aqueous Metabolites in Mice. *Drug Metabolism and Disposition*, *47*(8), 928–940. https://doi.org/10.1124/dmd.119.086538

Seyfried, N. T., Dammer, E. B., Swarup, V., Nandakumar, D., Duong, D. M., Yin, L., Deng, Q., Nguyen, T., Hales, C. M., Wingo, T., Glass, J., Gearing, M., Thambisetty, M., Troncoso, J. C., Geschwind, D. H., Lah, J. J., & Levey, A. I. (2017). A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer's Disease. *Cell Systems*, *4*(1), 60-72.e4. https://doi.org/10.1016/j.cels.2016.11.006

Shah, S. H., Kraus, W. E., & Newgard, C. B. (2012). Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases form and function. *Circulation*, *126*(9), 1110–1120. https://doi.org/10.1161/CIRCULATIONAHA.111.060368

Sharfstein, J. (2015). FDA Regulation of Laboratory-Developed Diagnostic Tests: Protect the Public, Advance the Science. *JAMA*, *313*(7), 667–668. https://doi.org/10.1001/JAMA.2014.18135

Sharpton, T. J., Stajich, J. E., Rounsley, S. D., Gardner, M. J., Wortman, J. R., Jordar, V. S., Maiti, R., Kodira, C. D., Neafsey, D. E., Zeng, Q., Hung, C.-Y., McMahan, C., Muszewska, A., Grynberg, M., Mandel, M. A., Kellner, E. M., Barker, B. M., Galgiani, J. N., Orbach, M. J., … Taylor, J. W. (2009). Comparative genomic analyses of the human fungal pathogens Coccidioides and their relatives. *Genome Research*, *19*(10), 1722–1731. https://doi.org/10.1101/gr.087551.108

Shen, Y., & Zelen, M. (2001). Screening Sensitivity and Sojourn Time From Breast Cancer Early Detection Clinical Trials: Mammograms and Physical Examinations. *Journal of Clinical Oncology*, *19*(15), 3490–3499. https://doi.org/10.1200/JCO.2001.19.15.3490

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer statistics, 2018. *CA: A Cancer Journal for Clinicians*, *68*(1), 7–30. https://doi.org/10.3322/caac.21442

Sikaris, K. A. (2017). Enhancing the Clinical Value of Medical Laboratory Testing. *The Clinical Biochemist Reviews*, *38*(3), 114. /pmc/articles/PMC5759162/

Smith, C. E., Whiting, E. G., Baker, E. E., Rosenberger, H. G., Beard, R. R., & Saito, M. T. (1948). The use of coccidioidin. *American Review of Tuberculosis*, *57*, 330–360.

Smith, R. A., Andrews, K. S., Brooks, D., Fedewa, S. A., Manassaram-Baptiste, D., Saslow, D., Brawley, O. W., & Wender, R. C. (2018). Cancer screening in the United States, 2018: A review of current American Cancer Society guidelines and current issues in cancer screening. *CA: A Cancer Journal for Clinicians*, *68*(4), 297–316. https://doi.org/10.3322/caac.21446

Snowden, S. G., Ebshiana, A. A., Hye, A., An, Y., Pletnikova, O., O'Brien, R., Troncoso, J., Legido-Quigley, C., & Thambisetty, M. (2017). Association between fatty acid metabolism in the brain and Alzheimer disease neuropathology and cognitive performance: A nontargeted metabolomic study. *PLoS Medicine*, *14*(3), 1–19. https://doi.org/10.1371/journal.pmed.1002266

Spector, R. (1988). Fatty Acid Transport Through the Blood-Brain Barrier. *Journal of Neurochemistry*, *50*(2), 639–643. https://doi.org/10.1111/j.1471-4159.1988.tb02958.x

Sperber, H., Mathieu, J., Wang, Y., Ferreccio, A., Hesson, J., Xu, Z., Fischer, K. A., Devi, A., Detraux, D., Gu, H., Battle, S. L., Showalter, M., Valensisi, C., Bielas, J. H., Ericson, N. G., Margaretha, L., Robitaille, A. M., Margineantu, D., Fiehn, O., … Ruohola-Baker, H. (2015). The metabolome regulates the epigenetic landscape during naive-to-primed human embryonic stem cell transition. *Nature Cell Biology*, *17*(12), 1523–1535. https://doi.org/10.1038/ncb3264

Spinelli, J. B., Yoon, H., Ringel, A. E., Jeanfavre, S., Clish, C. B., & Haigis, M. C. (2017). Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass. *Science*, *358*(6365), 941–946.

https://doi.org/10.1126/science.aam9305

Sun, C., Gao, M., Wang, F., Yun, Y., Sun, Q., Guo, R., Yan, C., Sun, X., & Li, Y. (2020). Serum metabolomic profiling in patients with Alzheimer Disease and Amnestic Mild Cognitive Impairment by GC/MS. *Biomedical Chromatography*, *34*(9), e4875. https://doi.org/10.1002/bmc.4875

Swerdlow, R. H. (2018). Mitochondria and Mitochondrial Cascades in Alzheimer's Disease. *Journal of Alzheimer's Disease*, *62*(3), 1403–1416. https://doi.org/10.3233/JAD-170585

Szablewski, L. (2016). Glucose Transporters in Brain: In Health and in Alzheimer's Disease. *Journal of Alzheimer's Disease*, *55*(4), 1307–1320. https://doi.org/10.3233/JAD-160841

Thompson, G. (2011). Pulmonary Coccidioidomycosis. *Seminars in Respiratory and Critical Care Medicine*, *32*(06), 754–763. https://doi.org/10.1055/s-0031-1295723

Torre, L. A., Islami, F., Siegel, R. L., Ward, E. M., & Jemal, A. (2017). Global Cancer in Women: Burden and Trends. *Cancer Epidemiology Biomarkers & Prevention*, *26*(4), 444–457. https://doi.org/10.1158/1055-9965.EPI-16-0858

Tsang, C. A., Anderson, S. M., Imholte, S. B., Erhart, L. M., Chen, S., Park, B. J., Christ, C., Komatsu, K. K., Chiller, T., & Sunenshine, R. H. (2010). Enhanced surveillance of coccidioidomycosis, Arizona, USA, 2007-2008. *Emerging Infectious Diseases*, *16*(11), 1738–1744. https://doi.org/10.3201/eid1611.100475

Tynkkynen, J., Chouraki, V., van der Lee, S. J., Hernesniemi, J., Yang, Q., Li, S., Beiser, A., Larson, M. G., Sääksjärvi, K., Shipley, M. J., Singh-Manoux, A., Gerszten, R. E., Wang, T. J., Havulinna, A. S., Würtz, P., Fischer, K., Demirkan, A., Ikram, M. A., Amin, N., … Salomaa, V. (2018). Association of branched-chain amino acids and other circulating metabolites with risk of incident dementia and Alzheimer's disease: A prospective study in eight cohorts. *Alzheimer's and Dementia*, *14*(6), 723–733. https://doi.org/10.1016/j.jalz.2018.01.003

Valdivia, L., Nix, D., Wright, M., Lindberg, E., Fagan, T., Lieberman, D., Stoffer, T., Ampel, N. M., & Galgiani, J. N. (2006). Coccidioidomycosis as a common cause of community-acquired pneumonia. *Emerging Infectious Diseases*, *12*(6), 958–962. https://doi.org/10.3201/eid1206.060028

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, *7*, 142. https://doi.org/10.1186/1471-2164-7-142

Varma, V. R., Oommen, A. M., Varma, S., Casanova, R., An, Y., Andrews, R. M., O'Brien, R., Pletnikova, O., Troncoso, J. C., Toledo, J., Baillie, R., Arnold, M., Kastenmueller, G., Nho, K., Doraiswamy, P. M., Saykin, A. J., Kaddurah-Daouk, R., Legido-Quigley, C., & Thambisetty, M. (2018). Brain and blood metabolite signatures of pathology and progression in Alzheimer disease: A targeted metabolomics study. *PLoS*

Medicine, *15*(1), 1–31. https://doi.org/10.1371/journal.pmed.1002482

Vlassenko, A. G., Gordon, B. A., Goyal, M. S., Su, Y., Blazey, T. M., Durbin, T. J., Couture, L. E., Christensen, J. J., Jafri, H., Morris, J. C., Raichle, M. E., & Benzinger, T. L. S. (2018). Aerobic glycolysis and tau deposition in preclinical Alzheimer's disease. *Neurobiology of Aging*, *67*, 95–98. https://doi.org/10.1016/j.neurobiolaging.2018.03.014

Wack, E. E., Ampel, N. M., Sunenshine, R. H., & Galgiani, J. N. (2015). The Return of Delayed-Type Hypersensitivity Skin Testing for Coccidioidomycosis. *Clinical Infectious Diseases*, *61*(5), 787–791. https://doi.org/10.1093/cid/civ388

Wang, D., Cheng, S. L., Fei, Q., Gu, H., Raftery, D., Cao, B., Sun, X., Yan, J., Zhang, C., & Wang, J. (2019). Metabolic profiling identifies phospholipids as potential serum biomarkers for schizophrenia. *Psychiatry Research*, *272*, 18–29. https://doi.org/10.1016/j.psychres.2018.12.008

Wang, G., Zhou, Y., Huang, F. J., Tang, H. D., Xu, X. H., Liu, J. J., Wang, Y., Deng, Y. L., Ren, R. J., Xu, W., Ma, J. F., Zhang, Y. N., Zhao, A. H., Chen, S. Di, & Jia, W. (2014). Plasma metabolite profiles of Alzheimer's disease and mild cognitive impairment. *Journal of Proteome Research*, *13*(5), 2649–2658. https://doi.org/10.1021/pr5000895

Wang, J., Wei, R., Xie, G., Arnold, M., Kueider-Paisley, A., Louie, G., Mahmoudian-Dehkordi, S., Blach, C., Baillie, R., Han, X., De Jager, P. L., Bennett, D. A., Kaddurah-Daouk, R., & Jia, W. (2020). Peripheral serum metabolomic profiles inform central cognitive impairment. *Scientific Reports*, *10*(1), 14059. https://doi.org/10.1038/s41598-020-70703-w

Wang, X., Gu, H., Palma-Duran, S. A., Fierro, A., Jasbi, P., Shi, X., Bresette, W., & Tasevska, N. (2019). Influence of storage conditions and preservatives on metabolite fingerprints in urine. *Metabolites*, *9*(10), 203. https://doi.org/10.3390/metabo9100203

Wei, Y., Jasbi, P., Shi, X., Turner, C., Hrovat, J., Liu, L., Rabena, Y., Porter, P., & Gu, H. (2021). Early Breast Cancer Detection Using Untargeted and Targeted Metabolomics. *Journal of Proteome Research*, *20*(6), 3124–3133. https://doi.org/10.1021/ACS.JPROTEOME.1C00019

Whiston, E., Zhang Wise, H., Sharpton, T. J., Jui, G., Cole, G. T., & Taylor, J. W. (2012). Comparative transcriptomics of the saprobic and parasitic growth phases in Coccidioides spp. *PloS One*, *7*(7), e41034. https://doi.org/10.1371/journal.pone.0041034

Wilkins, H. M., & Swerdlow, R. H. (2015). Relationships Between Mitochondria and Neuroinflammation: Implications for Alzheimer's Disease. *Current Topics in Medicinal Chemistry*, *16*(8), 849–857. https://doi.org/10.2174/1568026615666150827095102

Wilkins, J. M., & Trushina, E. (2018). Application of Metabolomics in Alzheimer's Disease. *Frontiers in Neurology*, *8*, 80. https://doi.org/10.3389/fneur.2017.00719

Wilmanski, T., Rappaport, N., Earls, J. C., Magis, A. T., Manor, O., Lovejoy, J., Omenn, G. S., Hood, L., Gibbons, S. M., & Price, N. D. (2019). Blood metabolome predicts gut microbiome α-diversity in humans. *Nature Biotechnology*, *37*(10), 1217–1228. https://doi.org/10.1038/S41587-019-0233-9

Wise, D. R., & Thompson, C. B. (2010). Glutamine addiction: a new therapeutic target in cancer. *Trends in Biochemical Sciences*, *35*(8), 427–433. https://doi.org/10.1016/j.tibs.2010.05.003

Wise, H. Z., Hung, C.-Y., Whiston, E., Taylor, J. W., & Cole, G. T. (2013). Extracellular ammonia at sites of pulmonary infection with Coccidioides posadasii contributes to severity of the respiratory disease. *Microbial Pathogenesis*, *59–60*, 19–28. https://doi.org/10.1016/j.micpath.2013.04.003

Wishart, D. S., Bartok, B., Oler, E., Liang, K. Y. H., Budinski, Z., Berjanskii, M., Guo, A., Cao, X., & Wilson, M. (2021). MarkerDB: an online database of molecular biomarkers. *Nucleic Acids Research*, *49*(D1), D1259–D1267. https://doi.org/10.1093/NAR/GKAA1067

Wishart, D. S., Guo, A. C., Oler, E., Wang, F., Anjum, A., Peters, H., Dizon, R., Sayeeda, Z., Tian, S., Lee, B. L., Berjanskii, M., Mah, R., Yamamoto, M., Jovel, J., Torres-Calzada, C., Hiebert-Giesbrecht, M., Lui, V. W., Varshavi, D., Varshavi, D., … Gautam, V. (2022). HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Research*, *50*(D1), D622–D631. https://doi.org/10.1093/NAR/GKAB1062

Wishart, D. S., Sayeeda, Z., Budinski, Z., Guo, A., Lee, B. L., Berjanskii, M., Rout, M., Peters, H., Dizon, R., Mah, R., Torres-Calzada, C., Hiebert-Giesbrecht, M., Varshavi, D., Varshavi, D., Oler, E., Allen, D., Cao, X., Gautam, V., Maras, A., … Cort, J. R. (2022). NP-MRD: the Natural Products Magnetic Resonance Database. *Nucleic Acids Research*, *50*(D1), D665–D677. https://doi.org/10.1093/NAR/GKAB1052

Wolf, I., Sadetzki, S., Kanety, H., Kundel, Y., Pariente, C., Epstein, N., Oberman, B., Catane, R., Kaufman, B., & Shimon, I. (2006). Adiponectin, ghrelin, and leptin in cancer cachexia in breast and colon cancer patients. *Cancer*, *106*(4), 966–973. https://doi.org/10.1002/cncr.21690

Wurtele, H., Tsao, S., Lépine, G., Mullick, A., Tremblay, J., Drogaris, P., Lee, E.-H., Thibault, P., Verreault, A., & Raymond, M. (2010). Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy. *Nature Medicine*, *16*(7), 774–780. https://doi.org/10.1038/nm.2175

Xia, J., Broadhurst, D. I., Wilson, M., & Wishart, D. S. (2013). Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics*, *9*(2), 280–299. https://doi.org/10.1007/s11306-012-0482-9

Yi, L., Liu, W., Wang, Z., Ren, D., & Peng, W. (2017). Characterizing Alzheimer's disease through metabolomics and investigating anti-Alzheimer's disease effects of natural products. *Annals of the New York Academy of Sciences*, *1398*(1), 130–141. https://doi.org/10.1111/nyas.13385

Yin, P., & Xu, G. (2017). Metabolomics Toward Biomarker Discovery. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 1619, pp. 467–475). https://doi.org/10.1007/978-1-4939-7057-5_32

Zetterberg, H., & Burnham, S. C. (2019). Blood-based molecular biomarkers for Alzheimer's disease. *Molecular Brain*, *12*(26), 1–7. https://doi.org/10.1186/s13041-019-0448-1

Zhang, A., Sun, H., & Wang, X. (2012). Saliva Metabolomics Opens Door to Biomarker Discovery, Disease Diagnosis, and Treatment. *Applied Biochemistry and Biotechnology*, *168*(6), 1718–1727. https://doi.org/10.1007/s12010-012-9891-5

Zhang, X., Zhu, X., Wang, C., Zhang, H., & Cai, Z. (2016). Non-targeted and targeted metabolomics approaches to diagnosing lung cancer and predicting patient prognosis. *Oncotarget*, *7*(39), 63437–63448. https://doi.org/10.18632/oncotarget.11521

Zhu, J., Djukovic, D., Deng, L., Gu, H., Himmati, F., Abu Zaid, M., Chiorean, E. G., & Raftery, D. (2015). Targeted serum metabolite profiling and sequential metabolite ratio analysis for colorectal cancer progression monitoring. *Analytical and Bioanalytical Chemistry*, *407*(26), 7857–7863. https://doi.org/10.1007/s00216-015-8984-8

Zhu, J., Djukovic, D., Deng, L., Gu, H., Himmati, F., Chiorean, E. G., & Raftery, D. (2014). Colorectal Cancer Detection Using Targeted Serum Metabolic Profiling. *Journal of Proteome Research*, *13*(9), 4120–4130. https://doi.org/10.1021/pr500494u

APPENDIX A

CHAPTER 2 SUPPLEMENTARY MATERIALS

Supplementary Table 2.1 Partial list of targeted metabolites and associated major metabolic pathways.

Metabolites	Major Pathway(s)
2-Aminoisobutyric acid	Amino acid/pyrimidine metabolism
3-Indolepropionic acid	Tryptophan Cycle
4-Pyridoxic acid	Vitamins/B6
5-Aminolevulinic acid	Glycine, serine and threonine metabolism/Porphyrin and chlorophyll
Acetylglycine	Amino Acid metabolism
Agmatine	Arginine and proline metabolism
Asparagine	Amino Acid
Betaine	Amino acids metabolism/Gly,Ser, Thr metabolism
Cytidine	Nucleotide/Pyrimidine metabolism
Decanoylcarnitine	Acylcarnitine
Glycocyamine	Glycine, serine and threonine metabolism/ Arginine and proline metabolism
Hypoxanthine	Nucleotide
Indole	Tryptophan Cycle
Indole-3-acetic acid	Tryptophan Cycle
N,N'-Dicyclohexylurea	
Pantothenic acid	Pantothenate and CoA biosynthesis/beta-Alanine metabolism
Proline	Amino Acid
TMAO	Gut flora metabolism / Redox
Taurine	Amino acids metabolism/Sulfur metabolism
Urocanic acid	Histidine metabolism
2,3-Dihydroxybenzoic acid	Benzoate degradation
2-Hydroxybenzoic acid	Phenylalanine metabolism
3-Indoxylsulfate	Tryptophan Cycle
Citraconic acid	Amino Acid metabolism/Val, Leu, IL
Gentisic acid	Tyrosine metabolism
Myoinositol	Inositol phosphate metabolism/Galactose metabolism
Nonadecanoic acid	Fatty acid
Palmitic acid	Fatty acid metabolism
Stearic acid	Fatty acid metabolism
Xanthine	Nucleotide
1-Methyladenosine	Nucleotide/Purine metabolism
2-aminoadipic acid	Lysine biosynthesis/7-ketocholesterol
7-ketocholesterol	Cholesterol metabolism
Acetylcarnitine	Fatty acid metabolism
Acetylglucosamine	Amino sugar and nucleotide sugar metabolism
Adenosine	Nucleotide/Purine metabolism
Adenosyl-L-homocysteine	Cysteine and methionine metabolism
Alanine	Amino Acid
Anthranilic acid	Amino Acid metabolism/Trp, Phe, Tyr
Arginine	Amino Acid
Caffeine	Caffeine metabolism
Carnitine	Amino acids metabolism/Lys
Carnosine	Histidine metabolism

Choline	Vitamins
Citrulline	Urea Cycle
Creatine	Amino acids metabolism/Arg, Gly
Creatinine	Amino acids metabolism/Arg, Gly
Cystathionine	Amino acids metabolism/cys
Cystine	Amino Acid
Dimethylarginine	Arginine metabolism
GDP	Nucleotide/Purine metabolism
Glutamic acid	Amino Acid
Glycine	Amino Acid
Hippuric acid	Phenylalanine metabolism
Histidine	Amino Acid
Homoserine	Amino acids metabolism/Thr, Met, Asp
Hydrocortisone	Steroid hormone biosynthesis
Inosine	Nucleotide/Purine metabolism
Isoleucine	Valine, leucine and isoleucine degradation
Kynurenic acid	Tryptophan Cycle/Amino Acid metabolism
Kynurenine	Tryptophan Cycle
L-Alloisoleucine	Valine, leucine and isoleucine degradation
Leucine	Amino Acid
Lysine	Amino Acid
Methionine	Amino Acid
Methylhistamine	Amino acids metabolism/His
N-Acetylethanolamine	
Nicotinamide	Nicotinate and nicotinamide metabolism
Norleucine	Amino acid
Ornithine	Urea cycle
Phenylalanine	Amino Acid
Pipecolinic acid	Lysine degradation
Pyroglutamic acid	Amino acids/Glu
Serine	Amino Acid
Threonine	Amino Acid
Tryptophan	Tryptophan Cycle
Tyrosine	Amino Acid
UDP-GlcNAc	Amino sugar and nucleotide sugar metabolism
Uridine	Nucleotide/Pyrimidine metabolism
Valine	Amino Acid
2-Pyrrolidone-5-carboxylic acid	D-Glutamine and D-glutamate metabolism
3-Methyl-2-oxovaleric acid	Amino Acid
3-Phenyllactic acid	Tropane, piperidine and pyridine alkaloid biosynthesis
4-Methyl-2-oxopentanoic acid	Valine, leucine and isoleucine biosynthesis/degradation
Adipic acid	Caprolactam degradation
Allantoin	Nucleotide Degradation
Dextrose	Glycolysis / Gluconeogenesis
Fumarate	TCA cycle
Gluconic acid	Pentose phosphate pathway
Glucose	Glycolysis/sugar
Ketoisoleucine	Isoleucine metabolism

Ketoleucine	Valine, leucine and isoleucine degradation/biosynthesis
Lactate	Glycolysis/TCA
Suberic acid	Fatty acid metabolism
Succinate	TCA Cycle
Urate	Nucleotide/Purine metabolism
alpha-Hydroxyisovaleric acid	Valine, leucine and isoleucine metabolism
alpha-KG	TCA cycle/Amino acid metabolism

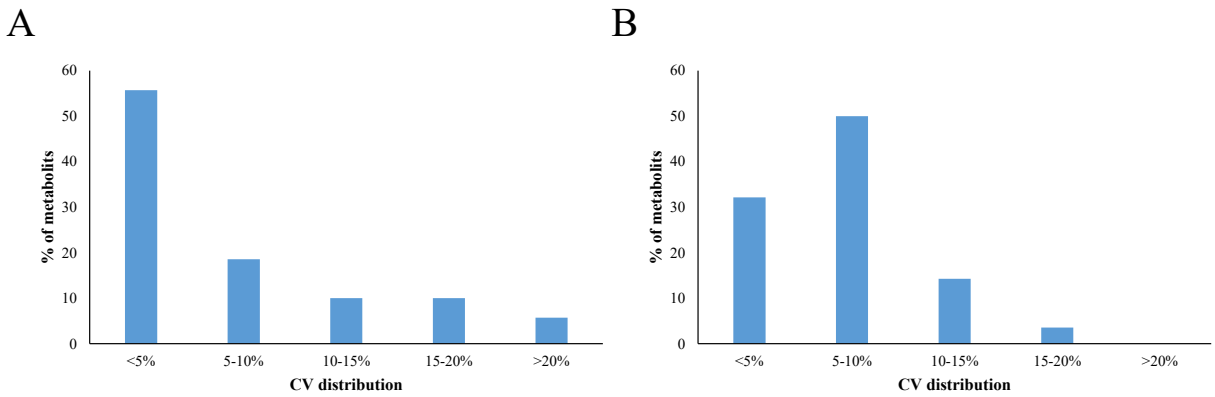
Supplementary Table 2.2 Comparative results of 98 metabolites between BC patients and healthy controls, and between two groups of control samples.

No.	Metabolite	q^a	q^b
1	Cystine	7.27E-16	2.40E-18
2	7-Ketocholesterol	9.39E-16	3.32E-20
3	Ornithine	1.94E-11	1.47E-07
4	Glutamic acid	1.22E-10	4.65E-18
5	Succinate	1.47E-10	1.85E-07
6	Acetylcarnitine	5.90E-09	1.54E-04
7	alpha-KG	8.26E-09	3.51E-22
8	Pipecolinic acid	3.60E-08	2.21E-03
9	Tyrosine	9.00E-08	2.02E-05
10	Proline	1.01E-07	5.58E-02
11	Pyroglutamic acid	1.61E-07	4.98E-03
12	Phenylalanine	3.79E-07	2.29E-08
13	Allantoin	3.96E-07	1.09E-09
14	Lactate	1.28E-06	2.80E-06
15	2-Pyrrolidone-5-carboxylic acid	3.94E-06	8.16E-05
16	Alanine	4.02E-06	8.10E-04
17	Suberic acid	5.09E-06	1.93E-10
18	Valine	5.33E-06	4.13E-04
19	L-Alloisoleucine	1.49E-05	2.46E-07
20	Isoleucine	1.54E-05	3.22E-07
21	Norleucine	1.54E-05	3.36E-07
22	Leucine	1.54E-05	2.84E-07
23	Adenosine	2.47E-05	4.81E-08
24	2-Hydroxybenzoic acid	2.49E-05	2.12E-01
25	Caffeine	2.66E-05	2.11E-04
26	Carnosine	2.72E-05	5.34E-03
27	Choline	4.86E-05	1.57E-03
28	Hippuric acid	5.19E-05	1.73E-02
29	GDP	5.21E-05	1.07E-03
30	Myoinositol	5.35E-05	8.40E-02
31	Acetylglucosamine	5.39E-05	1.22E-02
32	Creatinine	1.52E-04	1.84E-06
33	3-Phenyllactic acid	2.56E-04	4.07E-05
34	Lysine	3.17E-04	3.99E-05
35	Kynurenic acid	4.66E-04	9.78E-03
36	Homoserine	4.76E-04	8.62E-04
37	Palmitic acid	5.20E-04	5.54E-01

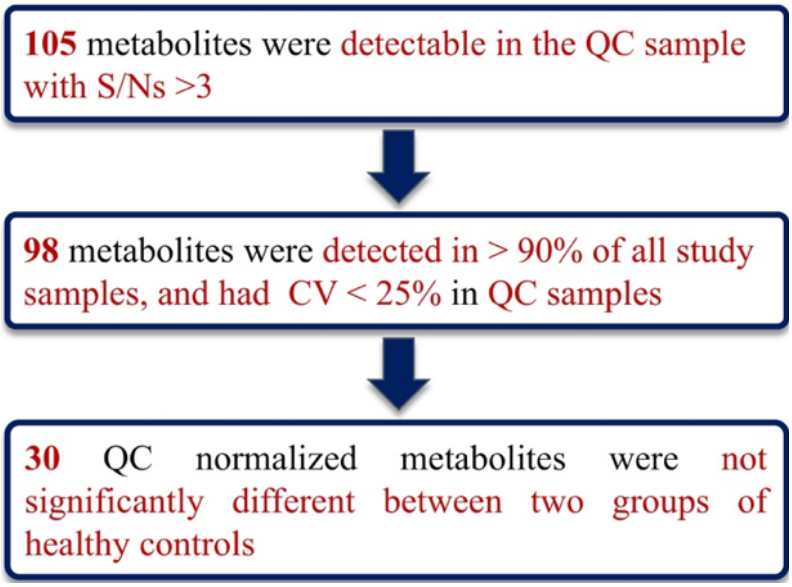
38	Threonine	5.30E-04	1.03E-03
39	Adipic acid	8.72E-04	2.56E-12
40	Hypoxanthine	8.73E-04	7.41E-01
41	2-Aminoadipic acid	8.98E-04	9.38E-03
42	Histidine	1.01E-03	4.71E-05
43	Carnitine	2.33E-03	1.14E-08
44	Glucose	2.68E-03	3.41E-02
45	Urate	2.95E-03	2.16E-03
46	4-Pyridoxic acid	2.95E-03	8.94E-01
47	Gentisic acid	3.85E-03	1.49E-01
48	Kynurenine	4.32E-03	1.17E-04
49	Cytidine	4.43E-03	6.65E-01
50	Serine	5.55E-03	1.06E-03
51	Agmatine	5.86E-03	7.80E-01
52	UDP-GlcNAc	5.92E-03	1.35E-10
53	Dextrose	7.39E-03	2.67E-02
54	Stearic acid	8.52E-03	9.85E-01
55	Fumarate	1.34E-02	1.02E-02
56	Pantothenic acid	1.44E-02	4.20E-01
57	Tryptophan	1.59E-02	1.53E-07
58	Ketoleucine	2.14E-02	4.65E-02
59	5-Aminolevulinic acid	2.24E-02	1.11E-01
60	Acetylglycine	3.01E-02	5.22E-01
61	Ketoisoleucine	3.08E-02	2.23E-02
62	Methionine	3.54E-02	3.43E-02
63	1-Methyladenosine	3.54E-02	5.32E-12
64	2,3-Dihydroxybenzoic acid	3.57E-02	4.77E-01
65	3-Methyl-2-oxovaleric acid	3.59E-02	3.46E-02
66	4-Methyl-2-oxopentanoic acid	3.81E-02	2.91E-02
67	Anthranilic acid	4.58E-02	9.93E-06
68	N,N'-Dicyclohexylurea	6.27E-02	6.18E-02
69	Creatine	6.54E-02	4.18E-02
70	2-Aminoisobutyric acid	6.57E-02	5.83E-02
71	TMAO	7.54E-02	2.61E-01
72	3-Indoxylsulfate	9.05E-02	6.88E-02
73	Nicotinamide	9.37E-02	1.83E-02
74	Decanoylcarnitine	1.49E-01	6.03E-01
75	Indole-3-acetic acid	1.54E-01	1.30E-01
76	Asparagine	1.71E-01	9.90E-01
77	Adenosyl-L-homocysteine	1.95E-01	1.11E-05
78	Citrulline	2.28E-01	1.56E-02
79	Urocanic acid	2.92E-01	1.36E-01

80	Nonadecanoic acid	3.30E-01	5.20E-02
81	N-Acetylethanolamine	3.44E-01	2.02E-03
82	Taurine	3.75E-01	7.47E-02
83	Uridine	4.01E-01	1.37E-03
84	Cystathionine	4.33E-01	1.32E-03
85	Hydrocortisone	4.91E-01	2.52E-03
86	3-Indolepropionic acid	6.39E-01	7.32E-01
87	Gluconic acid	6.54E-01	2.38E-04
88	Inosine	6.56E-01	4.08E-04
89	Dimethylarginine	6.98E-01	1.71E-04
90	Glycocyamine	7.12E-01	1.10E-01
91	Betaine	7.15E-01	1.38E-01
92	alpha-Hydroxyisovaleric acid	8.24E-01	1.19E-02
93	Citraconic acid	8.53E-01	5.84E-01
94	Arginine	9.24E-01	1.68E-02
95	Indole	9.79E-01	8.69E-02
96	Methylhistamine	9.89E-01	3.32E-03
97	Glycine	9.91E-01	2.28E-02
98	Xanthine	9.95E-01	2.69E-01

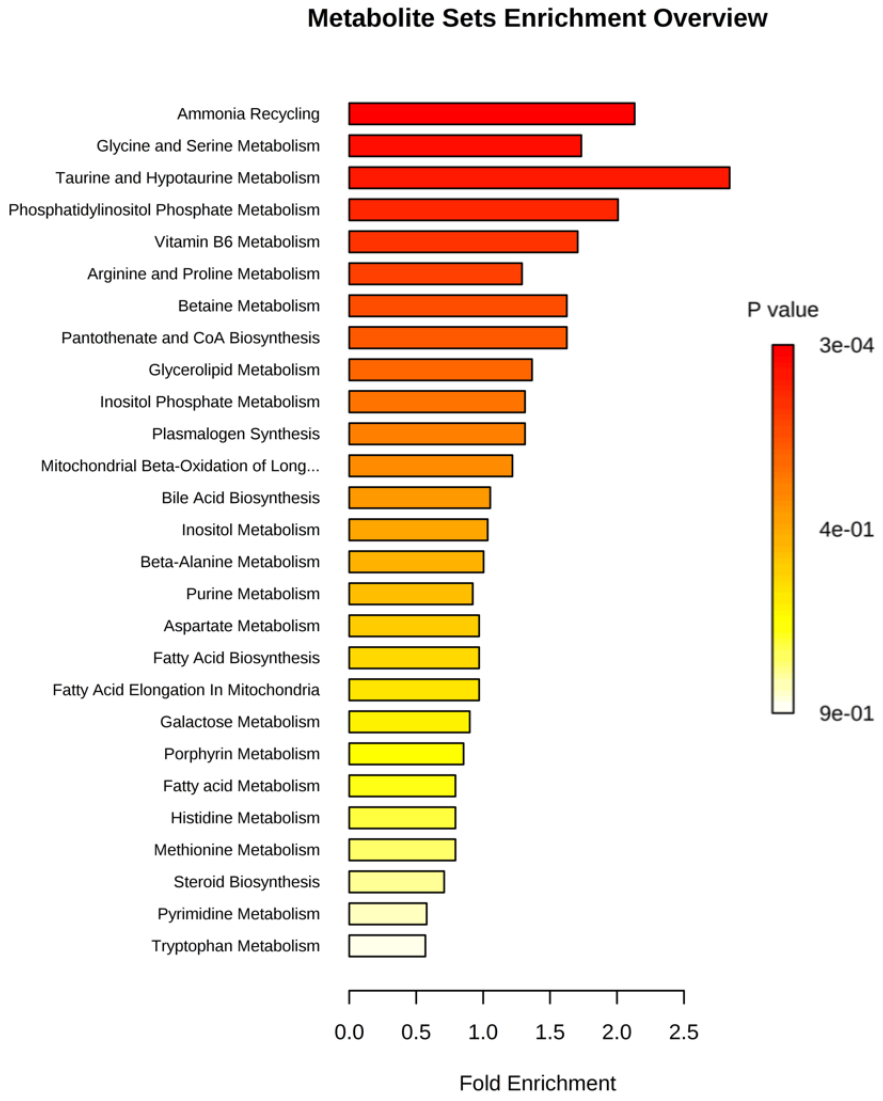
[a]*q*-values are FDR corrections for *p*-values comparing BC patients and all controls.
[b]*q*-values are FDR corrections for *p*-values comparing the two sets of controls from different centers.
q-values in red indicate <0.05.



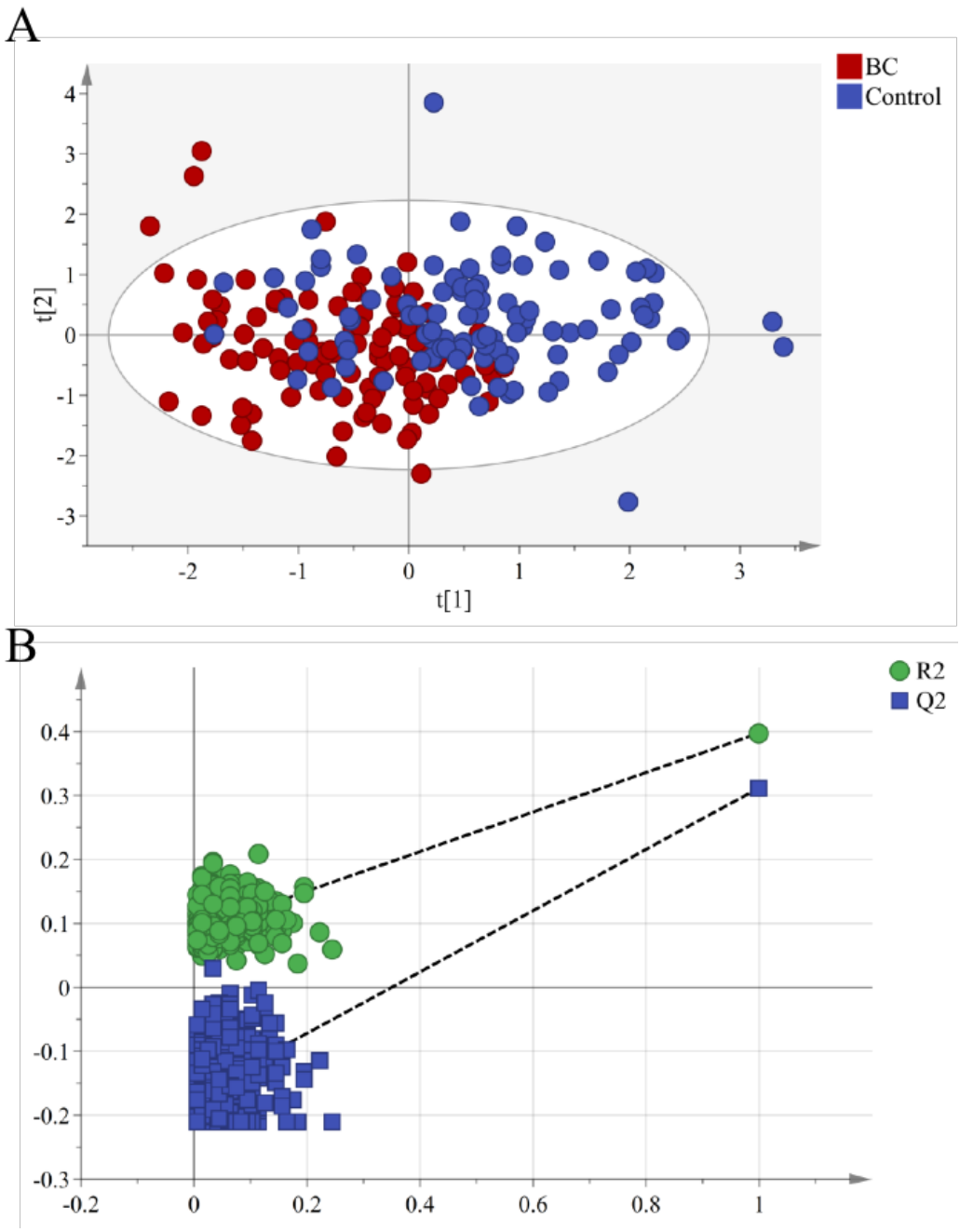
Supplementary Figure 2.1 Distribution of coefficient of variation (CV) values of all measured metabolites in this study. (A) CV distribution in positive mode detection; (B) CV distribution in negative mode detection. QC median CV: 5.1%, with ~87% of metabolites having CV < 15%.



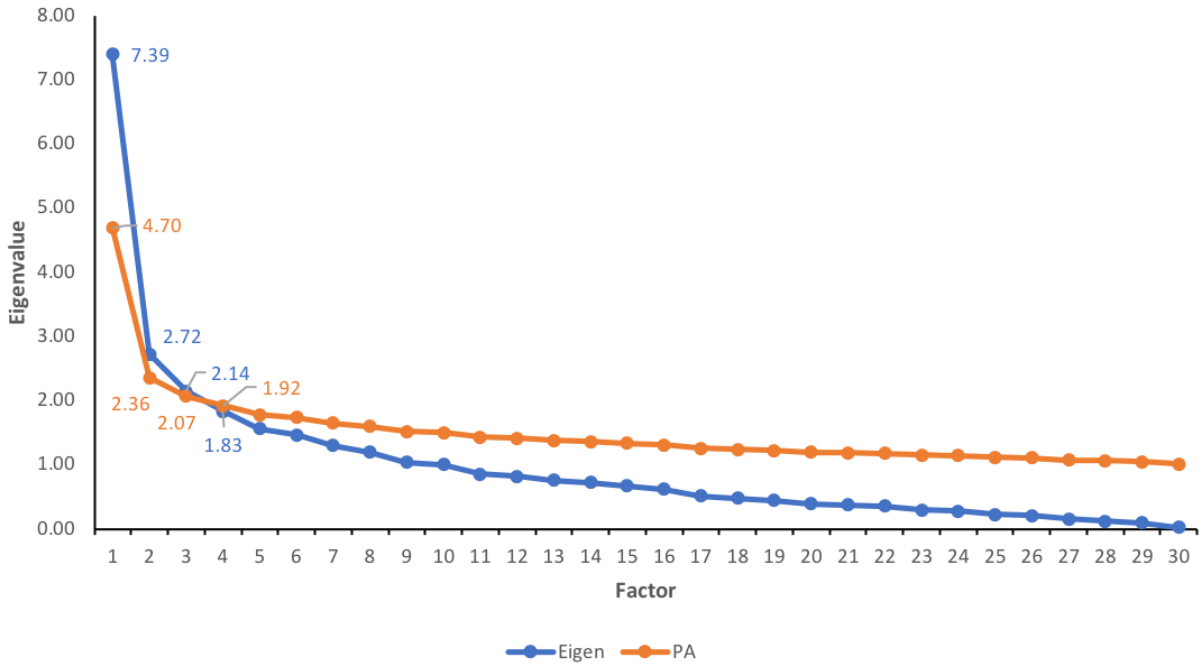
Supplementary Figure 2.2 Flow diagram depicting process of metabolite selection for biomarker analysis.



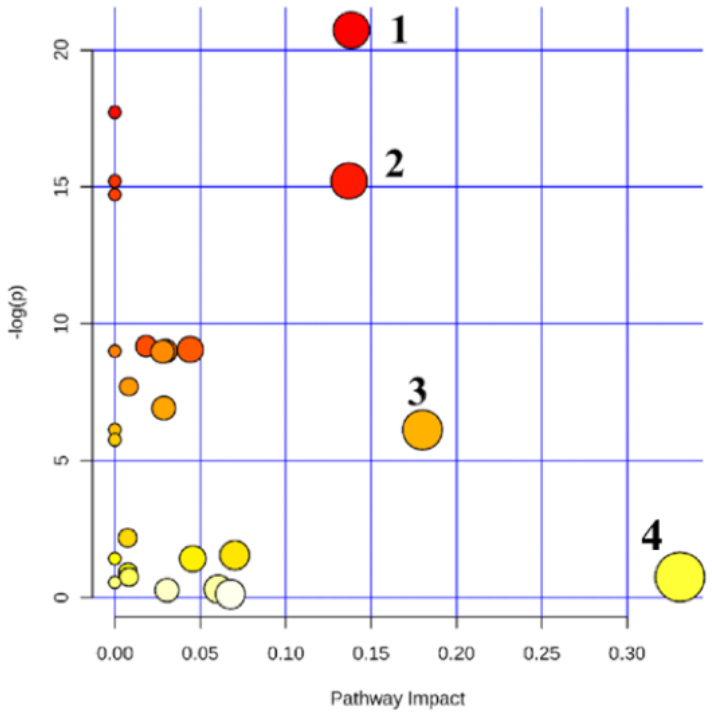
Supplementary Figure 2.3 Over-representation analysis conducted using relative abundances of 30 reliably detected metabolites used for comparison between BC patients and controls. Results indicate these 30 metabolites to be significantly reflective of 7 metabolic pathways: ammonia recycling ($p < 0.001$), glycine and serine metabolism ($p = 0.005$), taurine and hypotaurine metabolism ($p = 0.010$), phosphatidylinositol phosphate metabolism ($p = 0.032$), vitamin B6 metabolism ($p = 0.037$), arginine and proline metabolism ($p = 0.039$), and betaine metabolism ($p = 0.042$).



Supplementary Figure 2.4 PLS-DA model constructed using 30 metabolites for discrimination between breast cancer patients and healthy controls (R^2X (cum) = 0.291, R^2Y (cum) = 0.398, Q^2 (cum) = 0.312). (A) Score plot: data were log_{10}-transformed, and Pareto scaled. (B) Statistical validation of the PLS-DA model by permutation testing (n=300).



Supplementary Figure 2.5 Parallel analysis for 30 between-group metabolites. Plotted eigenvalues obtained from breast cancer data set (Eigen) were compared to random data from an equivalent matrix (PA). Only the first 3 factors were shown to have eigenvalues greater than random data and, as such, only 3 factors were retained.



Supplementary Figure 2.6 The metabolome view of pathway enrichment analysis comparing BC patients and controls. (1) Arginine and proline metabolism; (2) Inositol phosphate metabolism; (3) Pantothenate and CoA biosynthesis; (4) Fatty acid biosynthesis.

APPENDIX B

CHAPTER 3 SUPPLEMENTARY MATERIALS

Supplementary Table 3.1 Complete list of plasma metabolites found to be significant between VF patients and non-VF controls.

Metabolite	Mean (SD) of non-VF	Mean (SD) of VF	*p*-value	*q*-value[a]	VF/non-VF
Phenylpyruvic acid	767.098 (909.487)	2652.667 (1399.143)	< 0.0001	0	Down
m-Hydroxyphenylacetic acid	5413.902 (4204.522)	35754.333 (72110.464)	< 0.0001	0	Down
Phenylglyoxylic acid	201.610 (206.345)	5677.500 (13815.467)	< 0.0001	0	Down
4-Ethylbenzoic acid	3414.463 (1562.858)	1789.611 (772.786)	< 0.0001	0	Up
Aconitic acid	134894.098 (125561.923)	614132.833 (704041.673)	< 0.0001	0	Down
Glucuronic acid	10691.463 (8835.651)	67688.000 (89264.837)	< 0.0001	0	Down
3-Phosphoglyceric acid	47.000 (65.085)	9044.889 (10224.959)	< 0.0001	0	Down
Trehalose	501.366 (733.602)	12350.278 (21092.154)	< 0.0001	0	Down
Lactose	1525.366 (2124.076)	43415.333 (77212.041)	< 0.0001	0	Down
Mucic acid	1695.073 (1834.075)	32251.667 (81455.680)	< 0.0001	0	Down
UDP	444402.707 (80869.195)	315901.556 (61426.519)	< 0.0001	0	Up
Urate	20902.341 (15600.697)	8111.889 (15749.978)	< 0.0001	0	Up
Nicotinamide	80973.098 (65498.482)	491977.167 (341431.680)	< 0.0001	0	Down
Adenosine	126.951 (290.795)	2306.111 (8547.169)	< 0.0001	0	Down
Inosine	8418.439 (20891.451)	224975.056 (542154.620)	< 0.0001	0	Down
Cytidine	6217.195 (8535.782)	43921.722 (40874.749)	< 0.0001	0	Down
Serotonin	574.354 (1471.699)	13168.889 (16156.079)	< 0.0001	0	Down
Kynurenine	67817.341 (48058.141)	239900.333 (154387.537)	< 0.0001	0	Down
Acetylglucosamine	2347.659 (1571.320)	14808.556 (29497.084)	< 0.0001	0	Down
HIAA	22163.146 (16978.735)	100382.944 (67347.661)	< 0.0001	0	Down
Taurine	445508.293 (319693.230)	1242645.556 (546361.931)	< 0.0001	0	Down
Neopterin	114.610 (92.163)	2309.111 (6325.796)	< 0.0001	0	Down
N-Acetylneuraminic acid	7837.854 (6214.353)	48149.667 (65529.408)	< 0.0001	0	Down

Serine	636530.805 (465498.745)	191022.222 (122889.050)	< 0.0001	0	Up
ATP	256.110 (429.587)	3117.583 (6387.790)	< 0.0001	0	Down
4,3-Cresotic acid	543.024 (586.677)	5591.889 (11345.426)	< 0.0001	0.0001	Down
Glutaconic acid	57052.000 (51630.259)	297695.111 (505144.435)	< 0.0001	0.0001	Down
Fumarate	3904.537 (5361.854)	10277.167 (6802.435)	< 0.0001	0.0001	Down
R5P	1106.610 (1110.623)	4854.000 (4039.972)	< 0.0001	0.0001	Down
Phosphocreatine	276465.268 (175998.461)	112381.611 (53614.620)	< 0.0001	0.0001	Up
Hypoxanthine	36082.902 (59246.688)	253440.500 (375073.278)	< 0.0001	0.0001	Down
Creatinine	6806479.073 (4136343.484)	18434021.389 (20597092.113)	< 0.0001	0.0001	Down
Cytosine	774.220 (625.257)	5823.667 (9532.608)	< 0.0001	0.0001	Down
Amiloride	2137.073 (1768.553)	32232.778 (41054.770)	< 0.0001	0.0001	Down
Acetyl-L-glutamine	4870.244 (6320.621)	36424.889 (90812.647)	< 0.0001	0.0001	Down
Acetylcarnitine	15264397.5 (13906895.225)	48325691.000 (34471688.498)	< 0.0001	0.0001	Down
Aspartate	144548.439 (191390.440)	36062.722 (62627.864)	< 0.0001	0.0001	Up
UDP-GlcNAc	254.341 (456.734)	4096.278 (5193.289)	< 0.0001	0.0001	Down
GDP	576.976 (995.162)	3100.167 (2829.869)	< 0.0001	0.0001	Down
Valeric acid	208295.902 (245995.384)	30966.944 (20154.547)	< 0.0001	0.0002	Up
Sorbitol	7562.390 (5745.078)	85457.889 (170824.541)	< 0.0001	0.0002	Down
alpha-KG/Adipic acid	115687.195 (128604.182)	352055.167 (285924.094)	< 0.0001	0.0002	Down
Decanoylcarnitine	801675.390 (1206714.705)	2395290.889 (2373922.031)	< 0.0001	0.0002	Down
Kynurenic acid	1168.439 (1742.876)	29907.111 (106962.679)	< 0.0001	0.0002	Down
cGMP	17250.720 (57975.418)	281385.417 (414126.069)	< 0.0001	0.0003	Down
2-Methylglutaric acid	48112.439 (52466.834)	147715.167 (117705.465)	< 0.0001	0.0004	Down
Galactonic acid	15268.390 (13596.163)	81456.444 (176100.084)	< 0.0001	0.0004	Down
Xanthurenic acid	577.561 (1182.061)	4491.278 (10014.180)	0.0001	0.0005	Down

1-Methyladenosine	146860.110 (223884.636)	1567914.056 (5396571.157)	0.0001	0.0005	Down
PGE2	7240.659 (7896.321)	327.056 (311.300)	0.0001	0.0006	Up
Dimethylarginine	508.085 (486.202)	2174.889 (2436.668)	0.0002	0.0006	Down
3-hydroxybutyric acid	528.049 (614.555)	2467.000 (4048.681)	0.0002	0.0008	Down
Adenine	1776.171 (1480.568)	10018.111 (20253.545)	0.0002	0.0008	Down
Glyoxylic acid	1218.951 (1550.379)	4513.944 (5186.019)	0.0003	0.001	Down
Pantothenic acid	30854.073 (50997.288)	81098.944 (135081.935)	0.0003	0.0011	Down
Xanthosine	858.146 (674.404)	6030.611 (12327.632)	0.0003	0.0011	Down
Glycylproline	9499.390 (26184.445)	13175.778 (13036.061)	0.0003	0.0011	Down
Glucosamine	2389.927 (10952.810)	5410.278 (9793.262)	0.0003	0.0011	Down
4-Hydroxybenzaldehyde	2147.829 (1240.685)	3290.556 (1814.624)	0.0003	0.0012	Down
Lactate	606008.780 (614066.590)	1276734.722 (692832.013)	0.0004	0.0013	Down
dTMP	12302.317 (9428.154)	7404.556 (4089.898)	0.0004	0.0013	Up
Leucic acid	4765.829 (6196.919)	12796.389 (11703.913)	0.0005	0.0016	Down
Gluconic acid	10084.341 (10236.700)	47173.444 (107661.401)	0.0005	0.0016	Down
3-Phenyllactic acid	2753.146 (1797.324)	6114.111 (4379.643)	0.0006	0.0017	Down
Pyruvate	3138.463 (7788.109)	5717.611 (6128.526)	0.0005	0.0017	Down
Cystine	36519.659 (64288.908)	102764.111 (90411.819)	0.0005	0.0017	Down
Anthranilic acid	1743.927 (1377.783)	1256.833 (2604.669)	0.0008	0.0024	Up
6-Methyl-DL-Tryptophan	6758.220 (7758.716)	13139.889 (8471.394)	0.0008	0.0025	Down
Ethylmalonic acid	562.878 (929.142)	4316.889 (11022.810)	0.0011	0.0032	Down
Imidazole	93276.585 (24883.531)	70857.722 (17972.566)	0.0011	0.0032	Up
Methylhistamine	11390.195 (6269.223)	18930.167 (8901.474)	0.0011	0.0032	Down
4-Pyridoxic acid	288231.634 (983543.794)	251933.556 (840323.263)	0.0016	0.0045	Up
Allopurinol	1760.098 (1664.661)	3616.778 (3950.359)	0.0016	0.0045	Down

Methyl α-D-glucopyranoside	526.732 (575.170)	2747.556 (4814.640)	0.0019	0.0051	Down
3-Aminobutyric acid	23391.122 (20558.345)	80367.556 (122804.668)	0.0019	0.0051	Down
5-Hydroxytryptophan	1914.585 (1785.209)	5871.944 (6218.323)	0.0021	0.0057	Down
Isovaleric acid	48101.854 (46856.729)	118393.056 (101307.610)	0.0023	0.006	Down
D-Galacturonic acid	9832.439 (11294.956)	34362.667 (49976.537)	0.0026	0.0066	Down
Normetanephrine	4696.244 (3267.454)	14385.833 (26929.888)	0.0027	0.007	Down
Nicotinate	1080.488 (4749.259)	1816.222 (2754.319)	0.003	0.0075	Down
3-hydroxykynurenine	18325.195 (10047.178)	28709.500 (14933.541)	0.0031	0.0076	Down
Gibberellic acid	446.854 (435.723)	3078.611 (5795.638)	0.0036	0.0088	Down
2-Aminoadipic acid	13945.220 (13192.151)	42123.056 (51884.925)	0.0041	0.0099	Down
Erythrose	4974.683 (4671.148)	18720.500 (38448.715)	0.0043	0.0104	Down
Mandelic acid	2177.000 (1745.405)	17376.556 (45442.269)	0.0057	0.0135	Down
Phenylbutazone	8580.000 (15741.924)	24277.222 (87556.855)	0.0063	0.0148	Down
2-Pyrrolidinone	5364.439 (3554.457)	8560.778 (23304.948)	0.007	0.0163	Down
Phthalic acid	3431.049 (5447.462)	3774.222 (12035.429)	0.0078	0.018	Down
DOPA	4966.098 (7481.359)	10435.000 (21778.280)	0.0091	0.0207	Down
Maleic acid	10884.463 (7215.685)	14780.056 (6056.841)	0.0106	0.0235	Down
Pyroglutamic acid	906652.000 (659064.874)	1619928.611 (1308957.236)	0.0106	0.0235	Down
Xylose	4946.024 (7334.651)	12278.111 (13914.277)	0.0111	0.0245	Down
2-Aminobutyric acid	2172414.488 (1325623.343)	3700377.389 (2526935.169)	0.0117	0.0255	Down
Ribose	11512.146 (7202.911)	18516.333 (8366.034)	0.0126	0.0271	Down
Suberic acid	5126.341 (3361.662)	34565.944 (74238.831)	0.0142	0.03	Down
Cadaverine	29193.512 (17452.927)	31284.000 (65752.635)	0.0142	0.03	Down
beta-Hydroxyisovaleric acid	21565.610 (22788.207)	43948.278 (40332.407)	0.0156	0.0327	Down
Glycine	73422.415 (43287.510)	52689.944 (36160.433)	0.0172	0.0355	Up

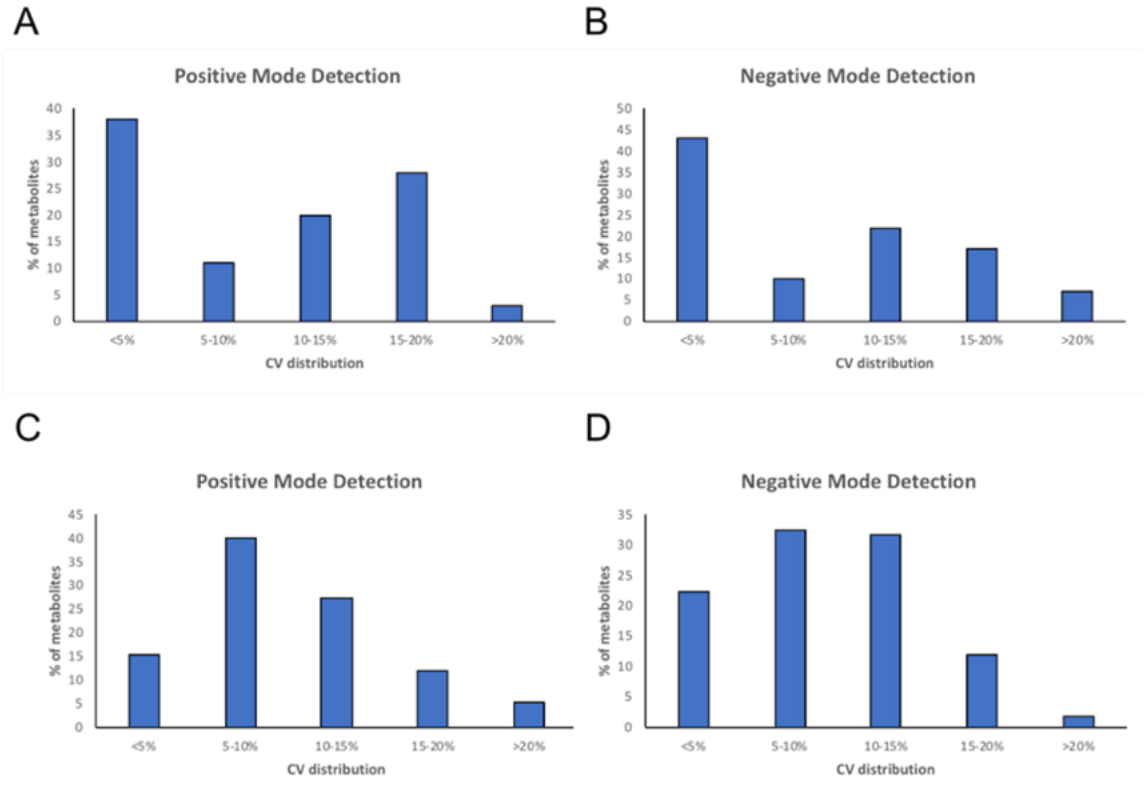
Azelaic acid	4294.366 (4658.546)	66508.111 (264243.763)	0.019	0.0389	Down
Homoserine/Threonine	1535813.927 (825122.781)	1054702.444 (530221.831)	0.0207	0.0415	Up
5-Aminolevulinic acid	989197.756 (843115.074)	1372020.833 (659096.468)	0.0207	0.0415	Down
Malonic acid	9420.439 (9521.260)	66877.833 (122333.866)	0.0226	0.0442	Down
Hydroxyproline	1052729.122 (939096.019)	1457822.056 (736465.080)	0.0226	0.0442	Down
Glutamic acid	1173856.000 (1266873.298)	653706.000 (892982.804)	0.0226	0.0442	Up
Ketoleucine	21780.976 (20912.018)	37558.278 (35324.800)	0.0237	0.0458	Down
Oxaloacetic acid	12333.049 (13049.461)	17882.667 (8632.556)	0.0247	0.0474	Down

[a]Values are FDR-corrected.

Supplementary 3.2 Urinary metabolites found to be statistically significant ($q < 0.05$) between VF patients and non-VF controls, using non-parametric t-test.

Correction[a]	Metabolite[b]	p[c]	q[d]	Fold Change[e]
Mann-Whitney	Phenylacetic acid	6.3E-7	1.4E-4	0.181
	DOPA	2.2E-5	0.0026	0.474
	Nicotinamide	1.0E-4	0.0082	0.475
	Amino valerate	1.6E-4	0.0082	0.380
	Glycocyamine	1.7E-4	0.0082	0.379
	6-Hydroxynicotinic acid	3.5E-4	0.0138	0.347
	Tryptamine	0.0011	0.0284	0.076
	Gentisic acid	0.0011	0.0284	0.194
	Glyoxylic acid	0.0012	0.0284	2.257
	Suberic acid	0.0012	0.0284	0.367
	3-Methyl-2-oxovaleric acid	0.0013	0.0284	0.339
	4-Ethylbenzoic acid	0.0017	0.0325	0.315
	p-Coumaric acid	0.0018	0.0325	0.191
	N,N-Dicyclohexylurea	0.0023	0.0395	0.434
	F16BP	0.0030	0.0443	0.404
	Dimethylarginine	0.0030	0.0443	2.117
	Tetracaine	0.0038	0.0491	0.367
	Urocanic acid	0.0040	0.0491	0.290
	Xanthosine	0.0042	0.0491	0.282
	Urate	0.0042	0.0491	0.153

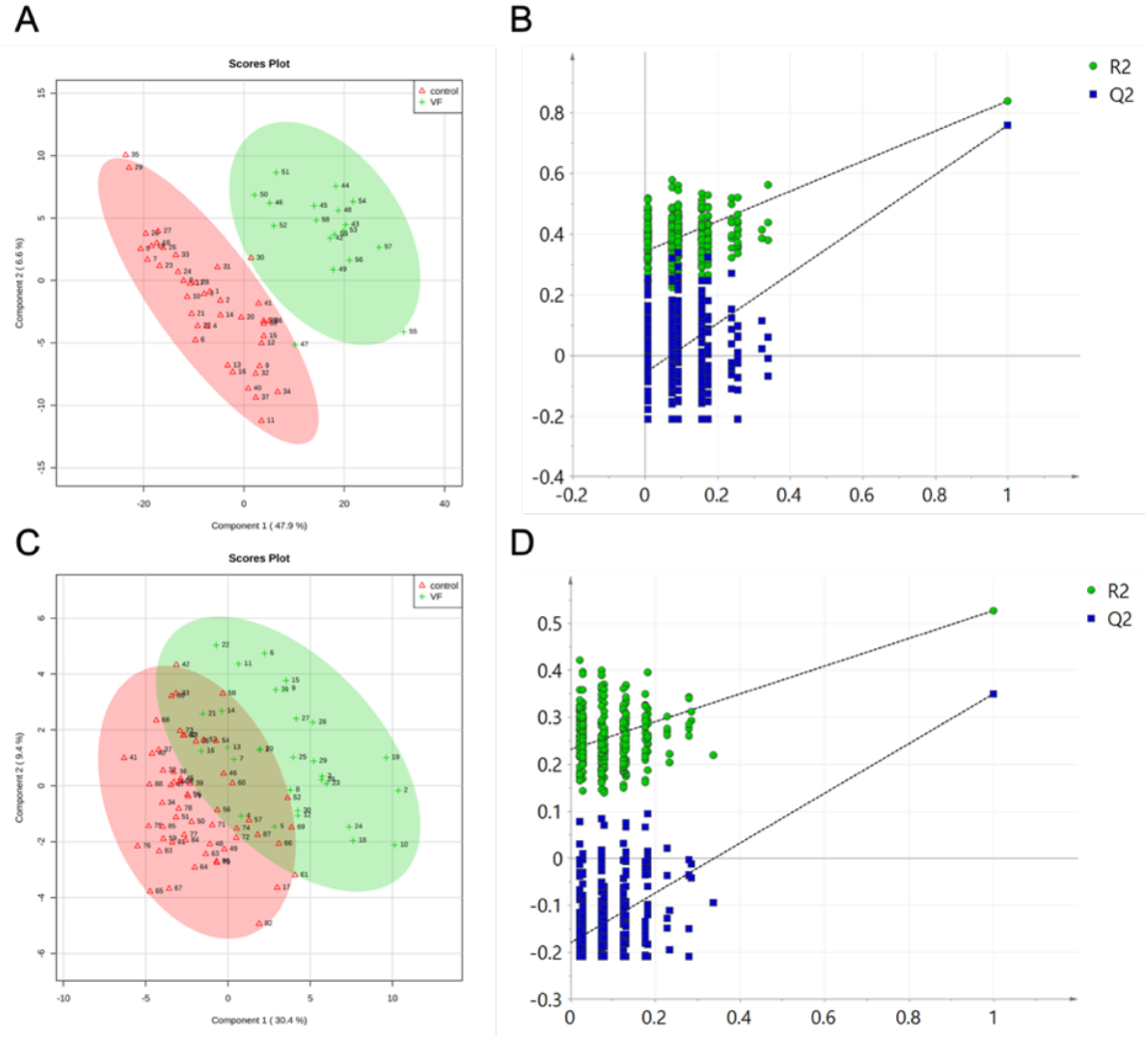
[a]Unequal group variance assumed.
[b]Data were normalized by creatinine levels.
[c]Adjusted p-value cutoff: 0.05.
[d]Non-parametric Wilcoxon rank-sum test used to adjust for multiple comparisons.
[e]Calculated as mean ratio of metabolites: VF/control.



Supplementary Figure 3.1 Distribution of coefficient of variation (CV) values of all measured plasma and urine metabolites in this study. (A) CV distribution of plasma metabolites in positive mode detection; (B) CV distribution of plasma metabolites in negative mode detection; (C) CV distribution of urine metabolites in positive mode detection; (D) CV distribution of urine metabolites in negative mode detection.

[Plasma QC CV range: 0.46%-13.01%, median CV: 11.91%, with ~70% of metabolites having CV<15%]

[Urine QC CV range: 0.02%-12.00%, median CV: 11.37%, with ~85% of metabolites having CV<15%]



Supplementary Figure 3.2 Partial least squares-discriminant analysis (PLS-DA) performed on log_{10}-transformed plasma and urine metabolite data: (A) score plot of 106 significant plasma metabolites accounting for 54.5% of variance, (B) statistical validation of plasma PLS-DA model [R^2X (cum) = 0.973, R^2Y (cum) = 0.862, Q^2 (cum) = 0.789] by permutation testing (n = 200), (C) score plot of 20 significant urinary metabolites accounting for 39.8% of variance, (D) statistical validation of urinary PLS-DA model [R^2X (cum) = 0.847, R^2Y (cum) = 0.627, Q^2 (cum) = 0.501] by permutation testing (n = 200).

APPENDIX C

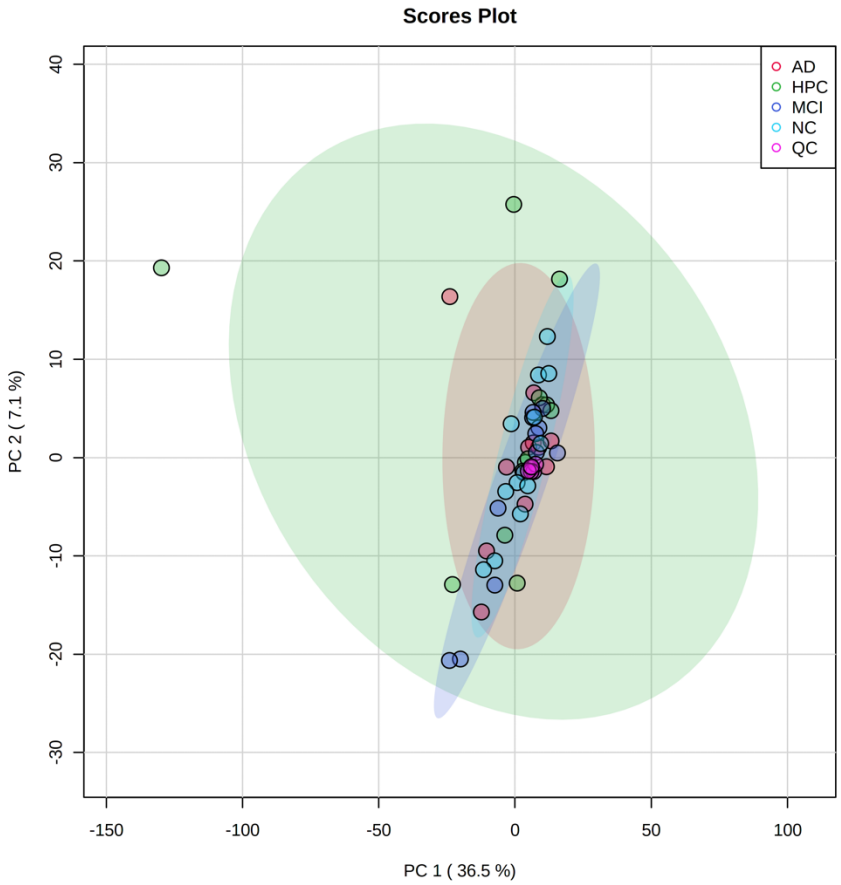
CHAPTER 4 SUPPLEMENTARY MATERIALS

Supplementary Table 4.1 Matrix of correlation analysis between candidate markers and clinical characteristics showing *r* and (*p*).

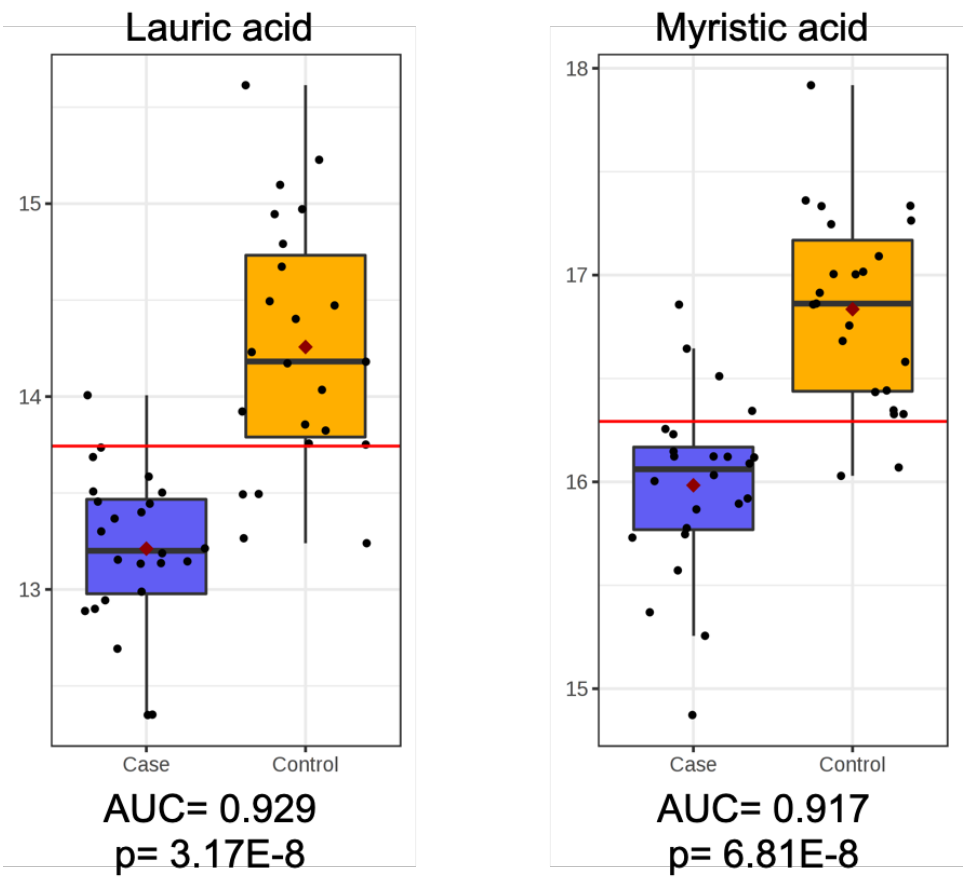
Clinical Characteristic	272.2@20.932005	147.1@14.3080015	147.1@20.774895	73.1@27.868113	Lauric acid	Myristic acid	Palmitoleic acid	Palmitic acid	Stearic acid
Age	-0.079	0.048	-0.036	-0.103	-0.043	0.150	0.210	0.168	0.168
	(0.598)	(0.748)	(0.810)	(0.490)	(0.775)	(0.313)	(0.157)	(0.259)	(0.259)
Sex	0.116	-0.246	-0.012	-0.100	-0.071	0.104	0.234	0.181	0.136
	(0.437)	(0.095)	(0.936)	(0.505)	(0.638)	(0.488)	(0.114)	(0.222)	(0.360)
PMI	0.219	0.019	0.183	-0.078	-0.269	-0.221	-0.222	-0.149	-0.154
	(0.139)	(0.897)	(0.217)	(0.604)	(0.067)	(0.135)	(0.134)	(0.318)	(0.302)
APOE alleles	-0.045	0.048	-0.047	-0.084	0.092	0.051	0.058	0.031	0.043
	(0.762)	(0.751)	(0.755)	(0.574)	(0.538)	(0.733)	(0.699)	(0.839)	(0.777)
MMSE	-0.097	0.114	-0.064	0.362	0.363	0.398	0.202	0.278	0.294
	(0.515)	(0.445)	(0.669)	(0.012)	(0.012)	(0.006)	(0.172)	(0.059)	(0.045)
Frontal plaque	0.198	-0.036	0.130	-0.160	-0.598	-0.085	0.118	0.003	-0.004
	(0.182)	(0.810)	(0.382)	(0.283)	(0.000)	(0.569)	(0.428)	(0.982)	(0.981)
Total Plaque	0.296	-0.092	0.157	-0.153	-0.579	-0.100	0.108	-0.001	-0.002
	(0.043)	(0.539)	(0.293)	(0.305)	(0.000)	(0.503)	(0.469)	(0.993)	(0.988)
Frontal Tangle	0.203	-0.129	0.070	-0.141	-0.410	-0.399	-0.109	-0.289	-0.276
	(0.172)	(0.387)	(0.639)	(0.345)	(0.004)	(0.005)	(0.467)	(0.048)	(0.061)
Total Tangle	0.111	-0.200	0.148	-0.246	-0.507	-0.350	-0.011	-0.219	-0.229
	(0.457)	(0.177)	(0.322)	(0.096)	(0.000)	(0.016)	(0.941)	(0.138)	(0.122)
Braak score	0.063	-0.157	0.198	-0.252	-0.539	-0.353	-0.046	-0.234	-0.250
	(0.672)	(0.293)	(0.182)	(0.087)	(0.000)	(0.015)	(0.757)	(0.113)	(0.091)
Frontal CAA	0.105	0.036	0.179	-0.165	-0.169	0.094	0.274	0.209	0.196
	(0.481)	(0.811)	(0.228)	(0.267)	(0.257)	(0.529)	(0.063)	(0.159)	(0.187)
Total CAA	0.100	0.005	0.083	-0.164	-0.211	0.091	0.295	0.219	0.194
	(0.502)	(0.972)	(0.579)	(0.269)	(0.155)	(0.543)	(0.044)	(0.139)	(0.191)

Supplementary Table 4.2 Full list of significantly enriched enzymes in response to Alzheimer's progression. For analysis, subjects were dichotomously grouped as case (MCI and AD) and control (NC and HPC).

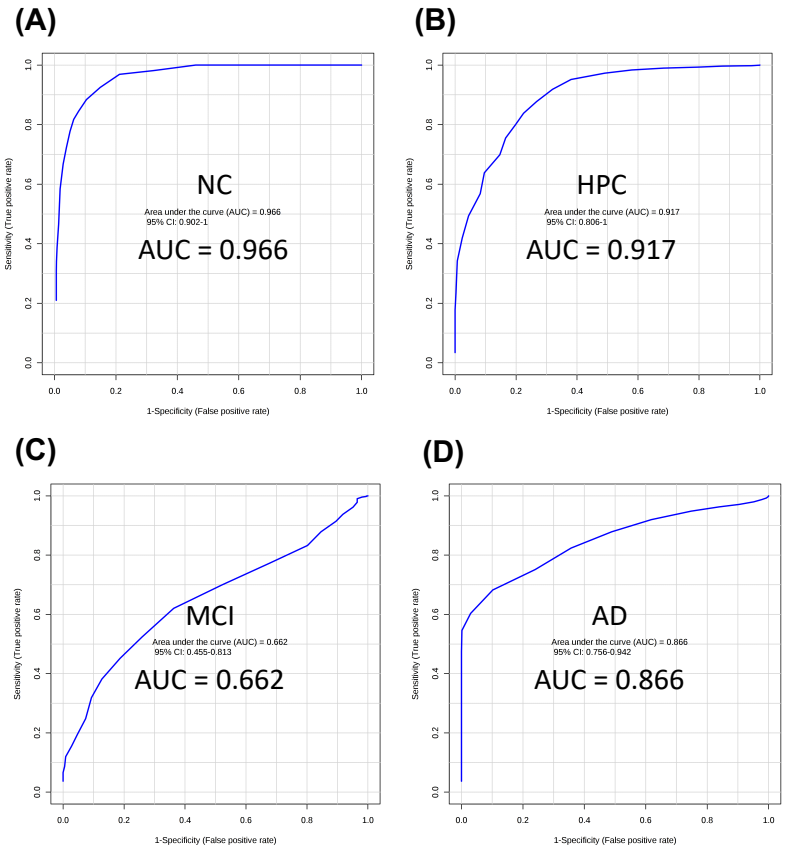
Enzyme	*p* value
Mitochondrial 2-oxovalerate dehydrogenase	0.025
3-amino-isobutyrate transport	0.025
Mitochondrial 3-amino-isobutyrate transport	0.025
3-hydroxyacyl-CoA dehydratase	0.025
Mitochondrial 3-hydroxyisobutyrate dehydrogenase	0.025
Mitochondrial 3-hydroxyisobutyryl-CoA hydrolase	0.025
Mitochondrial acyl-CoA dehydrogenase	0.025
L-3-amino-isobutanoate exchange	0.025
Mitochondrial L-3-aminoisobutyrate transaminase	0.025
Mitochondrial malonate-semialdehyde dehydrogenase	0.025
Methylmalonate-semialdehyde dehydrogenase	0.025



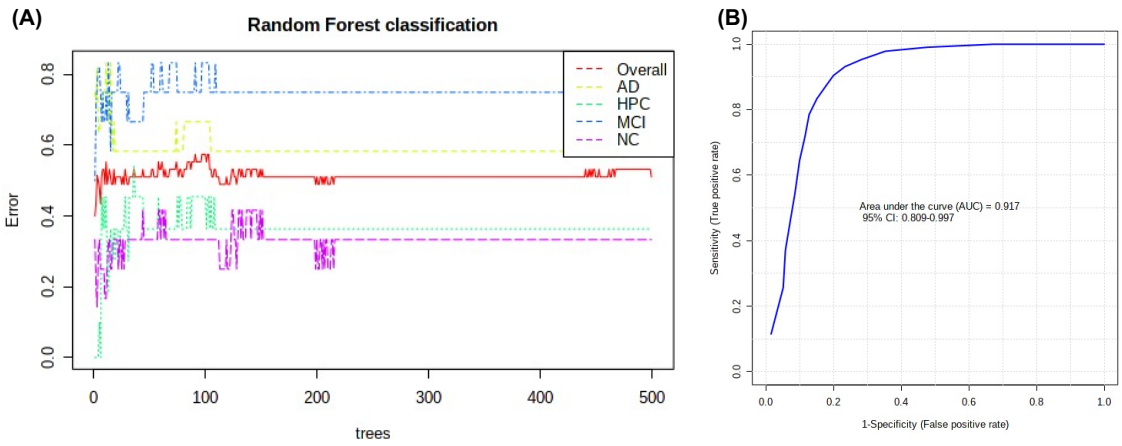
Supplementary Figure 4.1 Scores plot of principal component analysis (PCA) conducted with all reliably detected metabolites between all groups, including quality control (QC) samples; 95% confidence intervals were evaluated for potential outliers. Data for sample 13 (HPC group) was reviewed and removed from further analysis.



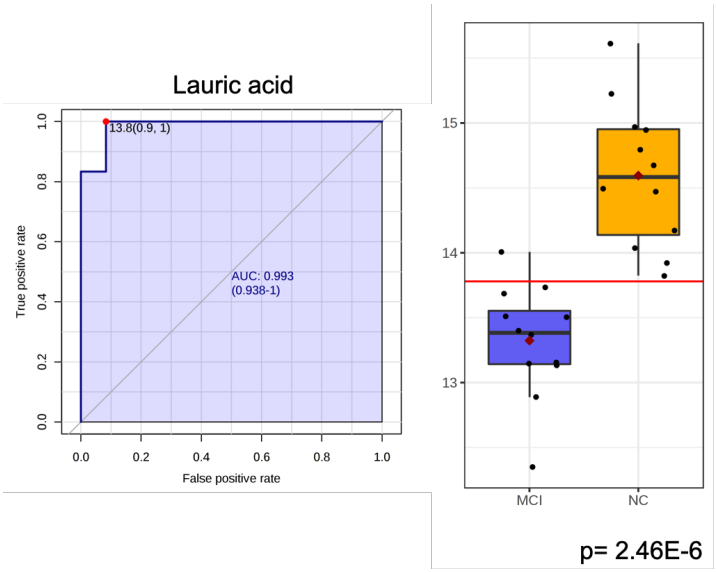
Supplementary Figure 4.2 Significant metabolites between case (MCI and AD) and control (NC and HPC) as determined by independent samples t-test. Data were $\log_{10}$-transformed and Pareto scaled prior to plotting.



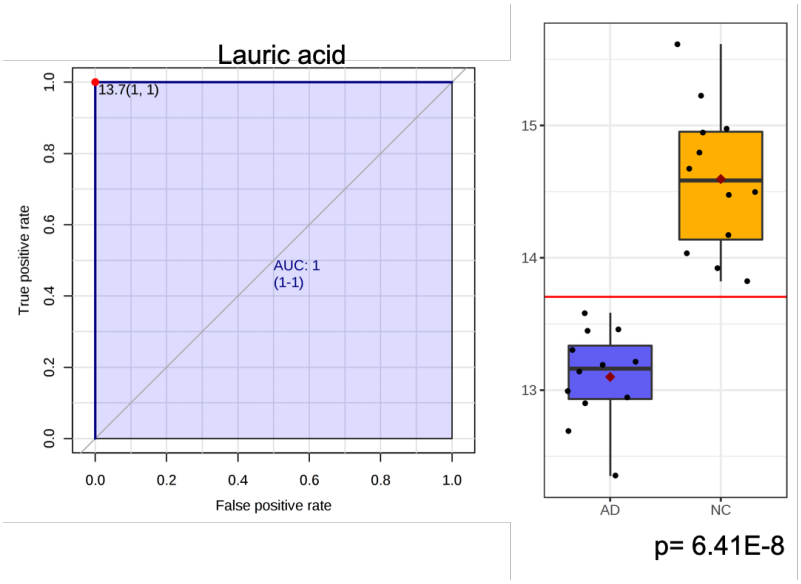
Supplementary Figure 4.3 ROC analysis of PLS-DA model constructed using levels of lauric acid, myristic acid, stearic acid, and palmitic acid (all $p < 0.05$): (A) classification of NC vs. HPC/MCI/AD, (B) classification of HPC vs. NC/MCI/AD, (C) classification of MCI vs. NC/HPC/AD, and (D) classification of AD vs. NC/HPC/MCI.



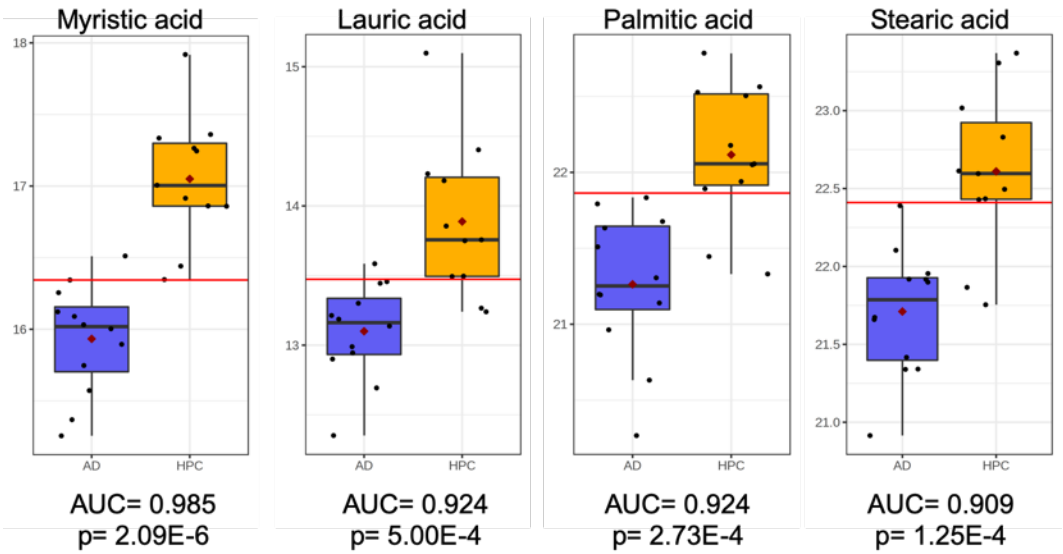
Supplementary Figure 4.4 (A) RF analysis of study groups based on levels of lauric acid, myristic acid, stearic acid, palmitic acid (OOB error = 0.511), and (B) 100-fold LOOCV ROC analysis of case (MCI/AD) and control (NC/HPC) using RF classifier (AUC = 0.917).



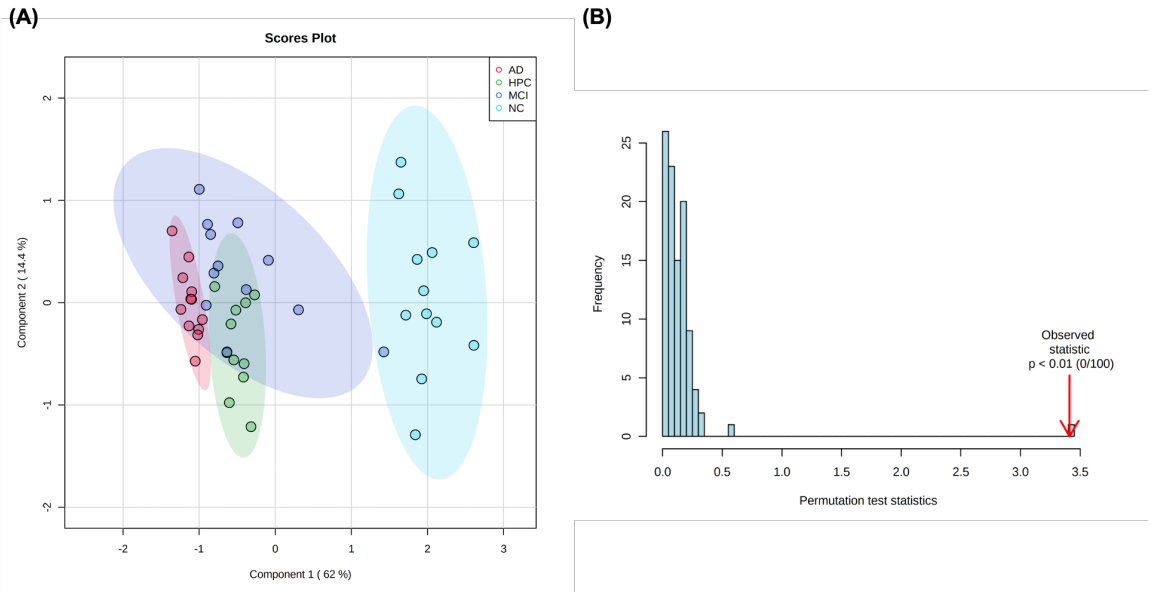
Supplementary Figure 4.5 Univariate ROC analysis and t-testing between NC and MCI groups show lauric acid to be highly predictive (AUC = 0.993) and significant (FDR q < 0.001).



Supplementary Figure 4.6 Univariate ROC analysis and t-testing between NC and AD groups show lauric acid to be highly predictive (AUC = 1.0) and significant (FDR q < 0.001).



Supplementary Figure 4.7 Univariate ROC analysis and t-testing of between HPC and AD groups show high predictive performance (AUC > 0.90) and significance (FDR q < 0.01) of myristic acid, lauric acid, palmitic acid, and stearic acid.



Supplementary Figure 4.8 (A) Scores plot of PLS-DA model constructed using significant between-group metabolites (lauric acid, myristic acid, stearic acid, palmitic acid) in conjunction with highly correlated neuropathological characteristics (frontal plaque, total plaque, total tangle, and Braak score) (R^2X = 0.710, R^2Y = 0.748, R^2Q = 0.722) and (B) Permutation testing with 100 iterations shows excellent model fit (p < 0.01).

Supplementary Figure 4.9 Network view of enzyme enrichment analysis conducted between case (MCI and AD) and control (NC and HPC). Enrichment ratios were significantly different for eleven enzymes between groups ($p < 0.05$).

BIOGRAPHICAL SKETCH

Paniz Mohajer Jasbi was born in Tehran, Iran on September 5, 1992. His secondary education was completed in 2010 at Westwood High School in Mesa, Arizona. In 2015, Paniz graduated Summa Cum Laude from Arizona State University's College of Liberal Arts and Sciences with a Bachelor of Science in Psychological Science and a minor in Philosophy. Upon graduation, he joined Dr. Elizabeth D. Phillips' conditioned food preferences laboratory as a research and lab manager. In August 2016, he entered the Graduate College at Arizona State University to pursue a Master of Science in Obesity Prevention and Management in the College of Health Solutions, where he was mentored by Dr. Elizabeth D. Phillips and Dr. Carol S. Johnston. Under the supervision of Dr. Haiwei Gu, Paniz began a PhD in Exercise and Nutritional Sciences at Arizona State University's College of Health Solutions. Since then, Paniz has authored 18 peer-reviewed journal articles, and two peer-reviewed book sections. Additionally, he has been granted a patent for his work in infectious disease. Furthermore, during his time as a doctoral student, Paniz has delivered five oral presentations from the regional to international levels and has presented 23 abstracts at scientific conferences across the United States. At Arizona State University, he has served as a teaching associate at both the graduate and undergraduate levels and has mentored numerous students in his time there. He also presently serves as an *ad hoc* reviewer of several leading journals with wide readership. Paniz has also been the recipient of the Arizona Graduate Scholar Award, the Outstanding Research Award and the Teaching Excellence Award of the Graduate College and has garnered $15,000 in research and travel funding. Upon graduation, he hopes to enter a career in biomedical diagnostics. He lives in Scottsdale, AZ with his fiancée Cassidy Turner.