

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Ken H. Liu

Date

Gut microbiome product delta-valerobetaine controls host energy metabolism

By

Ken H. Liu
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Molecular and Systems Pharmacology

Dean P. Jones, PhD
Advisor

Edward Morgan, PhD
Committee Member

Eric Ortlund, PhD
Committee Member

Thomas Ziegler, MD
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Gut microbiome product delta-valerobetaine controls host energy metabolism

By

Ken H. Liu
B.S., Georgia Institute of Technology, 2006

Advisor: Dean P. Jones, PhD

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biological and Biomedical Science: Molecular and Systems Pharmacology
2019

Abstract

Gut microbiome product delta-valerobetaine controls host energy metabolism

By Ken H. Liu

Chemical signals produced by the gut microbiome communicate with the liver to regulate systemic energy balance. The research in the dissertation concerns 1) the development and validation of chemical analysis tools suited for identifying novel chemical signals produced by the microbiome and 2) identification and characterization of microbial metabolites which impact energy metabolism in the liver. Chapter 1 serves as an overview of microbiome-host metabolic interactions and provides an experimental framework for in-depth study on this topic. Chapter 2 describes the optimization of a chemical analysis platform suitable for maximizing detection of chemicals in biological samples. Chapter 3 uses the analytical strategy described from chapter 2 to validate the quantification of approximately 200 metabolites and detection of 441 metabolites in biological samples. Chapter 4 describes the discovery of the microbiome-derived mitochondrial metabolite δ -valerobetaine (VB) and the characterization of its activity on energy metabolism in human cells and in mice. Chapter 5 contains brief concluding remarks and future avenues for exploration.

Gut microbiome product delta-valerobetaine controls host energy metabolism

By

Ken H. Liu
Bachelor of Science

Advisor: Dean P. Jones, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biological and Biomedical Science: Molecular and Systems Pharmacology
2019

Acknowledgments

There is no shortage of people here at Emory University that made this work possible. First and foremost, I have to thank my advisor Dr. Dean Jones for his mentorship, encouragement, honesty, patience, positive mindset, and for sharing his wisdom and knowledge with me. Thank you for always encouraging me to aim for the best. None of this work could have been possible without your network at Emory and beyond.

Second of all, thank you to the members of the Jones lab network past and present: thanks for all of the discussions and listening to my jokes. The contributions from Dr. Douglas Walker, Dr. Young-Mi Go, Dr. Karan Uppal, Dr. Shuzhao Li, Dr. Mary Nellis, Dr. Xin Hu, Dr. Jolyn Fernandes, Ravneet Kaur, Chunyu Ma, Bill Liang, Vilinh Tran, and Michael Orr have made the work in this dissertation possible.

Thirdly, to members of the Dr. Andrew Neish's lab and especially Bejan Saeedi: the definition of a mutualistic commensal relationship. Thank you for your support of this work and your wise words "left foot, right foot" reminding us to do things one step at a time.

Fourth, I'd like to thank my committee members Dr. Thomas Ziegler, Dr. Eric Ortlund (and lab members Sam Druzak), and Dr. Edward Morgan (and lab members Dr. Choon Lee and Grant Singer) for serving on my committee and allowing me to become an honorary member of your laboratories.

To the MSP program: Thank you for accepting me into this graduate program, making me feel part of a community and supporting my training.

To other members of the Emory community: Dr. Thota Ganesh, Dr. Samantha Yeligar, Moriah Bellissimo, Dr. Jessica Alvarez, Catherine Cioffi, and Dr. Miriam Vos: Thank you for your contributions to this dissertation.

Lastly, I have to thank my family members: my parents and sister for instilling the curiosity and desire to learn at an early age, and their unwavering support. My wife, Christina – Thank you for being. Scientific research is a lot like the journey to becoming a better husband and father: understanding complex living organisms, identifying the correct strategy for presenting and communicating difficult concepts, and the commitment and resiliency to learn new things and continuing to overcome unexpected problems. And finally, to Arthur – ask lots of questions, find your own answers, and stay curious.

Table of Contents

Chapter 1 : Introduction	1
1.1 - The epidemic of obesity and related metabolic diseases	2
1.2 - The microbiome and obesity-related metabolic disorders	5
1.3 - Chemical communication between the microbiome and host	8
1.4 - Gut-liver and microbiome-mitochondria axes	12
1.5 - Multi-omic systems biology approaches for host-microbiome interactions	16
1.6 – Hypothesis	23
1.7 – Statement of research objectives	24
Chapter 2 : Evaluating strategies for maximizing chemical detection using high-resolution metabolomics (HRM)	29
2.1 Introduction	31
2.2 Methods	35
2.3 Results	40
2.4 Discussion	57
Chapter 3 : Reference standardization for quantification and harmonization of large-scale metabolomics	67
3.1 Introduction	69
3.2 Methods	72
3.3 Results	76
3.4 Discussion	94
Chapter 4: Inhibition of mitochondrial fatty acid beta-oxidation by the microbial metabolite, δ-valerobetaine	100

4.1 Introduction	102
4.2 Methods	104
4.3 Results	112
4.4 Discussion	134
Chapter 5: Conclusions	142
5.1 - Summary	143
5.2 – Preliminary data	146
5.3 – Conclusion	150

List of figures and tables

Chapter 1

Figure 1.1 The relationship of energy balance and metabolic disorders	4
Figure 1.2 The microbiome contributes to host energy balance.....	7
Figure 1.3 Chemical communication between the microbiome and host.....	9
Figure 1.4 The liver and mitochondria as central players in energy balance	15
Figure 1.5 Systems biology paradigm of host-microbiome interactions	20

Chapter 2

Figure 2.1 High-resolution metabolomics QC.....	41
Figure 2.2 A. Total number of ions detected with percentage of ions matching at least one chemical in KEGG represented by dark portion of the bar. B. Number of chemical matches, with percentage of total database (17,554 chemicals) matched.....	43
Figure 2.3 Dual HRM platforms increase number of database matches to 72% of the KEGG chemical database. Matches obtained from left HRM platform (blue), right HRM platform (red), and both HRM methods (purple).	45
Figure 2.4 Number of ions with database matches from paired HRM platforms. Ions detected by left HRM platform (blue), right HRM platform (red), and both HRM methods (purple).	48
Figure 2.5 Comparison of ion intensity for tyrosine detection by HILIC/ESI+ and C18/ESI+. Each dot represents the normalized peak intensity for tyrosine measured in the same 30 samples across two different analytical platforms.....	49
Figure 2.6 HRM detects accurate mass matches to metabolites from a broad range of chemical classes.	51
Figure 2.7 Dual HRM using C18/ESI- (blue) with HILIC/ESI+ (red) provides in-depth coverage of human metabolic pathways using multiple adducts for chemical matching	53
Figure 2.8 Dual HRM using C18/ESI- (blue) with HILIC/ESI+ (red) provides adequate coverage of human metabolic pathways using conservative chemical matching strategy.....	54
Figure 2.9 Representative chemical detection with eight HRM platforms.....	56

Chapter 3

Figure 3.1 General workflow for metabolite identification and quantification.....	77
Figure 3.2 Summary of adducts detected and number of quantifiable metabolites.....	79

Figure 3.3 Selection of consistently detected adducts for quantification	81
Figure 3.4 Evaluation of inter-reference ion ratios of multiple ion forms shows selection of ion form used is important for reference standardization reproducibility	83
Figure 3.5 Consistency of amino acid concentrations measured in heparin NIST1950 using EDTA reference plasma Qstd3	86
Figure 3.6 Reference sample metabolite concentrations and previously reported ranges	88
Figure 3.7 Application of reference standardization to 200 study samples analyzed across a 10 batch study over a one week period.....	91
Figure 3.8 Application of reference standardization to 3,677 study samples analyzed across 17 studies over a 17 month period.	93
Appendix 1 (from Chapter 3). Table of validated metabolites (441) and estimated concentrations of metabolites in Qstd3 (214), NIST1950 (198), and Chear (205) pooled plasma reference materials.....	151
Appendix 2 (from Chapter 3). Reference for how metabolites were classified as detected, identified, or quantifiable.....	163
Chapter 4	
Figure 4.1 Summary of metabolome pathways impacted by the microbiome	113
Figure 4.2 Summary of transcription pathways impacted by the microbiome	114
Figure 4.3 Untargeted high-resolution metabolomics analysis comparing liver and liver mitochondria of 6 germ-free (GF) and 6 conventionalized (CV) mice	116
Figure 4.4 Ion dissociation spectra (MS/MS) analysis of 160.1332 <i>m/z</i>	117
Figure 4.5 Confirmation of metabolite identity for 160.1332 as delta-valerobetaine	118
Figure 4.6 Microbial activity is required to produce VB.....	120
Figure 4.7 12 hour dose response study of VB in HepG2 cells.....	122
Figure 4.8 Mitochondrial respiration after 12 hour treatment with VB in HepG2 cells.....	124
Figure 4.9 Stable isotope tracer study to examine the immediate effect of VB treatment on cellular mitochondrial fatty acid oxidation.....	126
Figure 4.10 VB dose-response study in fed mice	129
Figure 4.11 VB dose-response study in fasted mice.....	130

Figure 4.12 VB increases hepatic triglycerides after fasting and decreases ketone body formation during fasting in male and female mice	131
Figure 4.13 Clinical phenotypes associated with circulating VB	133
Appendix 3 (Chapter 4). ¹ H NMR of synthesized δ -valerobetaine in D ₂ O	228
Appendix 4 (Chapter 4). RNAseq analysis shows the influence of the microbiome on hepatic gene expression	229

Chapter 5

Figure 5.1 A simplified experimental pipeline for understanding interaction networks between the microbiome and host	144
Figure 5.2 Working model of the VB-mitochondrial interaction network	145
Figure 5.3 The metabolic fate of VB	147
Figure 5.4 The role of VB in other tissues	149

List of abbreviations

CV% – Coefficient of variation

CV – Conventionalized

DodSR – Department of Defense Serum Repository

FCCP – Trifluoromethoxy carbonylcyanide phenylhydrazone

FTMS – Fourier transform mass spectrometry

GF – Germ-free

HILIC – Hydrophilic interaction liquid chromatography

HMDB – Human metabolome database

HRM – High-resolution metabolomics

KEGG – Kyoto Encyclopedia of Genes and Genomes

LC-MS – Liquid chromatography mass spectrometry

MS – Mass spectrometry

MS/MS – Tandem mass spectrometry

NAFLD – Non-alcoholic fatty liver disease

NIST – National Institute of Standards and Technology

NMR – Nuclear magnetic resonance

VB – Valerobetaine

SCFA – Short-chain fatty acids

SEM – Standard error of the mean

RSD – Relative standard deviation

Chapter 1 : Introduction

This chapter serves as a conceptual framework for understanding human metabolic diseases as a potential consequence of dysfunctional chemical communication between the microbiome and host. Recent studies show rises in rates of obesity-related metabolic disorders are linked to changes in the intestinal microbiota as a result of changing lifestyles, dietary habits and/or use of antibiotics. Nonetheless, the molecular mechanisms underlying these observations are not well understood. In this introductory conceptual review, an integrative systems biology framework for understanding microbiome-linked mechanisms responsible for changes to human energy metabolism is presented. Because mitochondrial dysfunction can cause uncoupling of energy expenditure with energy supply, understanding how the intestinal microbiota, and specifically metabolites produced by the microbiota influence mitochondrial function and hepatic regulation of systemic energy balance may lead to new therapeutic strategies to combat this epidemic.

1.1 - The epidemic of obesity and related metabolic diseases

Chronic diseases related to obesity, insulin resistance and metabolic syndrome are a serious public health concern worldwide. These related metabolic disorders are risk factors contributing to development and complicating the treatment of several diseases including cancer, cardiovascular disease, non-alcoholic fatty liver disease (NAFLD), type 2 diabetes, stroke, depression, and early death (1). In the United States alone, over half of the population is considered overweight or obese (2) with associated annual costs for health care and loss of economic productivity estimated to be well in excess of \$100 billion dollars. Following current trends (3), over 80% of adults in the United States are predicted to be obese or overweight by the year 2030, creating a significant burden on the economy and health care system. While strategies to reverse these trends have been implemented at multiple levels (e.g. socioeconomic interventions such as improved education and access to healthy foods, or individual diet, exercise, drugs or other medical interventions), the molecular mechanisms driving this epidemic are not entirely understood.

At a fundamental level, obesity is characterized by an excessive accumulation of fat in tissues. Fats are the primary form of energy storage in the body and exist in several forms. Normally, the coordinated regulation of energy metabolism drives adaptive responses to fluctuations in energy intake, storage, and/or expenditure. This adaptive response provides metabolic flexibility and allows increases in energy intake to be coupled to increased energy expenditure to maintain energy balance. These responses are uncoupled in obesity, resulting in increased energy storage and/or decreased energy expenditure. The accumulation of fat is associated with systemic low-grade inflammation and promotes the development of insulin resistance and associated diseases (**Figure 1.1**) (4). Understanding mechanisms underlying these

increases in energy intake and storage, or decreases in energy expenditure could drive the development of novel personalized therapeutic strategies for prevention or treatment of these related disorders (5).

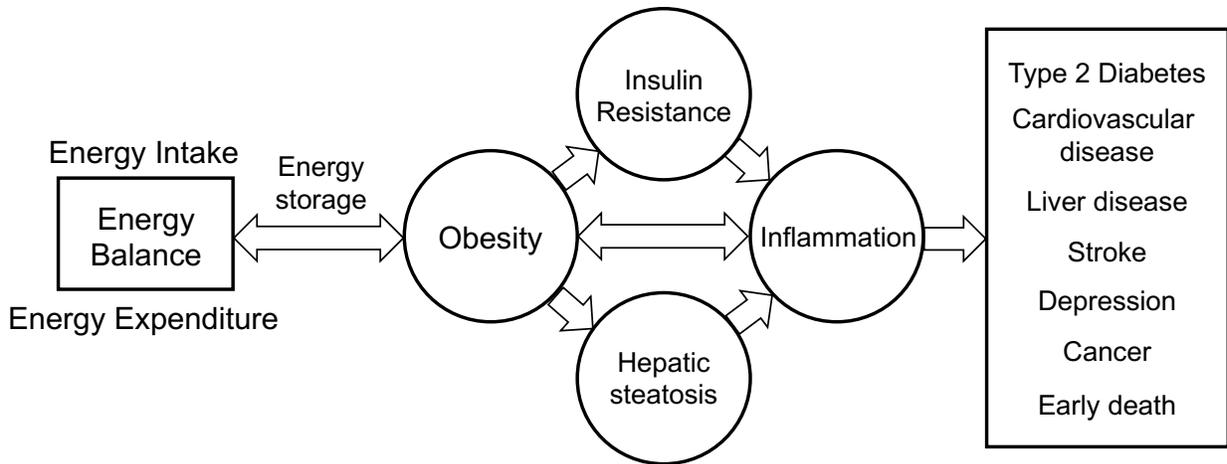


Figure 1.1 Obesity and related metabolic disorders are disorders of energy balance. Excessive accumulation of fats in tissues can exacerbate inflammatory processes and promote insulin resistance. These metabolic syndromes are risk factors for chronic diseases contributing to over \$100 billion in medical expenses and lost productivity.

1.2 - The microbiome and obesity-related metabolic disorders

Emerging evidence links dietary interactions with the intestinal microbiota to the development of human metabolic disorders (6-11). The human gut microbiome is comprised of over one hundred trillion bacteria, fungi, viruses, protozoa, and other microorganisms that inhabit the small and large intestines. Collectively, the metabolic potential of the microbiota exceeds the metabolic capacity of human cells (12; 13). Thus, the activity of these microorganisms resemble a hidden organ within our intestines (14), as they perform functions related to host nutrient metabolism, xenobiotic and drug metabolism, maintenance of the mucosal barrier, immunomodulation, and protection against pathogens (15; 16).

While the normal composition and function of the typical healthy microbiota is still not completely defined, obesity and poor diet quality is associated with a reduction in microbial diversity, and altered representation of bacterial genes (17-19). For example, obesity has been associated with increased carbohydrate-utilizing gram-positive *Firmicutes* (e.g. *Lactobacilli*, *Ruminococci*, *Clostridia*) and decreased short-chain fatty acid-producing (SCFA) gram-negative *Bacteroidetes* (e.g. *Bifidobacter*, *Bacteroides*) (19). Other studies have shown obesity is linked to decreased *Methanobrevibacter smithii* (17) and increased *Lactobacillus* (20; 21). Dietary changes elicit rapid changes to the composition of the intestinal microbiome. Turnbaugh et al. showed that mouse gut microbiota changed in composition within 24 hours after a shift to a high-fat, high-sugar “Western” diet from a low-fat plant-polysaccharide-based diet (22). In humans, Wu et. al showed that shifting from a high-fat/low-fiber diet to a low-fat/high-fiber diet elicited responses to the enteric phenotype, with carbohydrate-rich diets increasing abundances of *Prevotella* and animal fat-rich diets increasing abundances of *Bacteroidetes* (23). While strain-specific effects within the same bacterial families or species can result in variable directionality

of associations and differences in individual bacterial families being identified from study to study, the majority of published work shows changes to the microbiome are associated with changes to metabolic health. Whether these changes drive disease or are adaptive responses to disease is currently not well understood.

Intriguingly, germ-free (GF) mice, which are born and raised without microbiota, are protected from diet-induced obesity and gain weight less rapidly than their colonized counterparts (24; 25). This difference in weight gain occurred despite colonized mice eating 30% less calories than GF animals and was associated with increased expression of genes encoding enzymes for the digestion of otherwise indigestible polysaccharides (in colonized animals), leading to the conclusion that microbial enzymes increase energy extraction from diet (24). Studies from the same group also showed that an induction of hepatic lipogenesis and increased fat storage were associated with weight gain in colonized animals (25). Another study showed that microbiota depletion by use of antibiotics in conventional high-fat diet-fed and genetically obese mice decreased adiposity and improved insulin sensitivity and glucose tolerance (26). Additionally, transplantation of microbiota from obese mice or humans into GF-mice caused recipient mice to develop obesity and insulin resistance (27; 28). These studies demonstrate 1) the presence of the microbiome increases energy harvest and storage in the host and 2) the heritability of host traits associated with obese or lean microbiota upon transplantation. Taken together, these studies show the intestinal microbiota are an internal environmental or endobiotic contributor to host energy balance, adiposity and fat metabolism (**Figure 1.2**) (24). However, mechanisms explaining how the presence of the microbiome or specific changes to microbial composition alters host energy balance are not completely described.

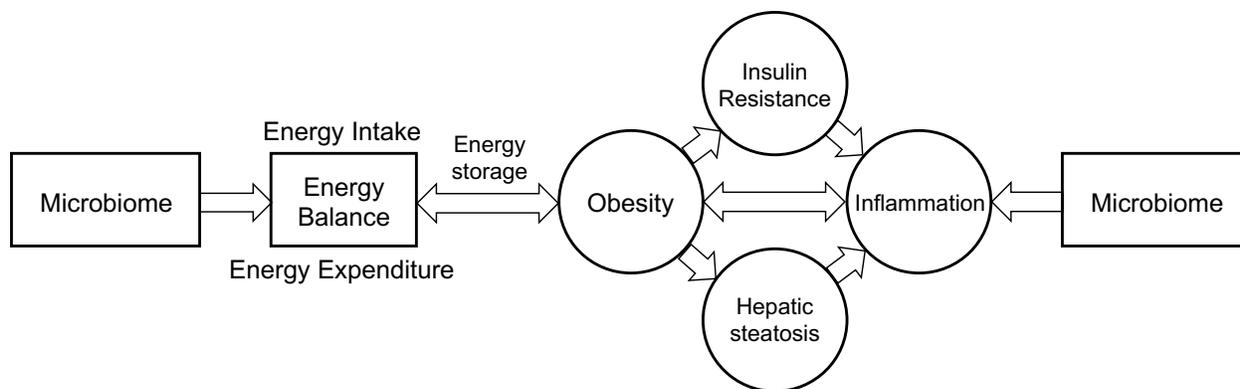


Figure 1.2 The microbiome contributes to host energy balance by promoting energy harvest from the diet and transcriptional regulation of energy metabolism in the host. Furthermore, environmental factors can increase gut barrier permeability allowing microbial products to drive a systemic low-grade inflammation contributing to metabolic dysfunction.

1.3 - Chemical communication between the microbiome and host

Recent studies have focused on elucidating the role of microbiome-derived small molecule metabolites and their specific contributions to host energy metabolism and physiology (**Figure 1.3**). Owing to their diverse genetic potential, intestinal microbes possess remarkable metabolic capacity for digestion and metabolism of a wide range of dietary macromolecules or xenobiotics, and synthesis of vitamins, hormones, and other metabolites that engage with host metabolism as metabolic intermediates, coenzymes, and signaling molecules (29; 30). Most metabolites produced by commensal gut bacteria are considered beneficial and are used by microbial communities and/or host cells in the local environment of the intestinal lumen. Nonetheless, local interactions of microbiome-derived metabolites with enterocytes (intestinal epithelia), immune cells (T-cells, dendritic cells, mast cells, macrophages, etc), and the enteric neuroendocrine system (enterochromaffin cells, vagal nerve) can elicit systemic effects on host metabolism. For example, short-chain fatty acids are used as energy substrates by host cells (31) and serve as signaling molecules for host membrane receptors, such as the G-protein coupled receptor 41/43, to regulate cellular metabolic functions (12). SCFAs also regulate the production of gut hormones Glucagon-like peptide-1 (GLP-1), Peptide YY (PYY), and ghrelin, which suppress appetite and reduce energy intake (32-35). In addition to SCFAs, bile acids, organic acids, aromatic organic acids, amino acid derivatives, indoles, methylamines, sulfur-containing metabolites, branched-chain fatty acids, lipopolysaccharides, hydrogen sulfide, ethanol, N-acylamides, and conjugated linoleic acids are other classes of metabolites associated with the activity of the intestinal microbiota with documented roles as chemical signals exchanged between microbiome and host (12; 29; 36).

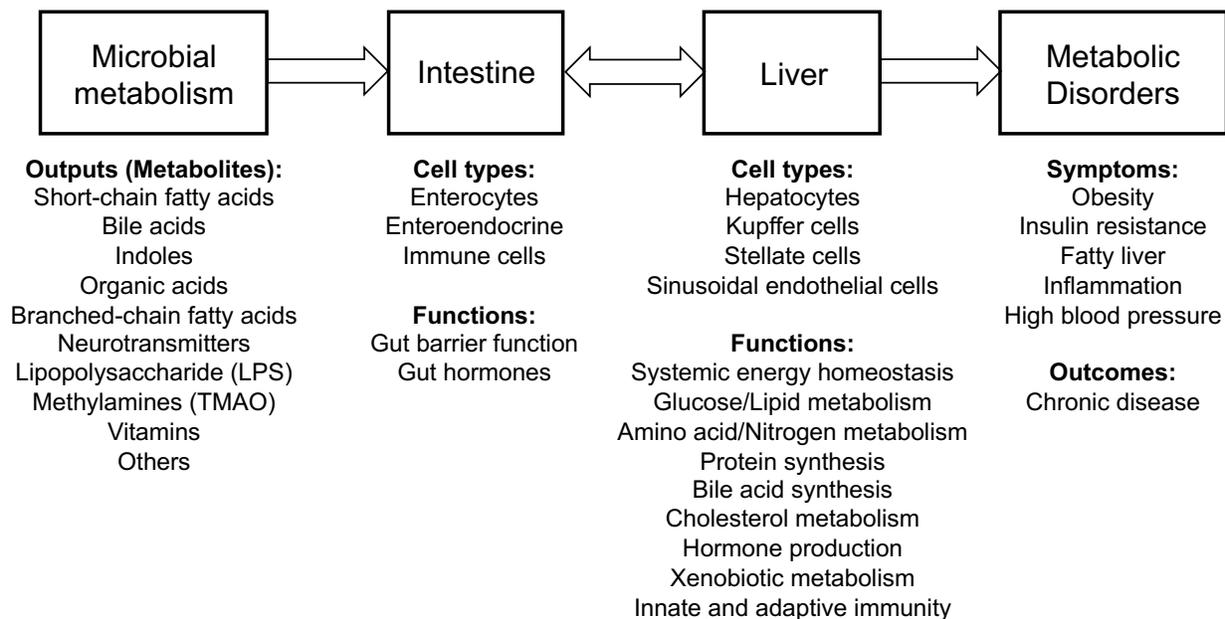


Figure 1.3 Chemical communication between the microbiome and host. Microbial enzymes convert dietary or host-derived chemicals to additional metabolites that interact with host targets in different organ systems. The specific interactions between gut-derived metabolites and specific cell populations or cellular compartments in the liver has not yet been explored. These interactions can elicit changes to cellular functions and host energy metabolism.

Approximately 10% of circulating mammalian metabolites are linked to the presence of the microbiome (37), showing the presence of microbial signals in extraintestinal compartments and tissues. As blood metabolome profiles are reflective of gut microbiome diversity (38), several microbiome-associated metabolites have been identified as potential biomarkers or risk factors for obesity-related diseases. For example, high-fat diets increase circulating bacterial lipopolysaccharides (LPS), activating CD14 and toll-like receptor-4, leading to a systemic low-grade inflammation associated with obesity, type 2 diabetes, and non-alcoholic fatty liver disease (NAFLD) (39; 40). Dysregulation of aromatic and branched-chain amino acid (BCAA) metabolism was associated with hepatic steatosis and other metabolic syndrome-related symptoms in several human studies (41-44). In cell and mouse studies, phenylacetic acid, a microbial product derived from aromatic amino acids, altered BCAA metabolism and induced hepatic steatosis (41). Other studies have shown cardiovascular and fatty liver diseases are associated with microbial metabolism of carnitine, choline and related metabolites, leading to increases in circulating concentrations of trimethylamine oxide (TMAO) (45-48). Trimethylamine (TMA) is produced by microbial enzymes, absorbed and converted to TMAO by hepatic flavin monooxygenase 3 (FMO3). TMAO is linked to pro-thrombotic, pro-inflammatory, and pro-atherogenic mechanisms in mammals (48; 49). These studies show specific interactions of host and microbiome metabolomes are associated with disease processes and have led to development of clinical biomarkers and therapeutic strategies targeting host-microbiome signaling processes.

Overall, these studies and others have demonstrated that the human metabolism is a union of host and microbial players. While several circulating microbiome-associated metabolites

associated with disease processes have previously been identified, a recent review estimates over 95% of metabolites from microbiome metabolomics studies have not yet been characterized (50). Identifying bioactive microbiome-derived metabolites, and tracing the origins of microbial metabolites from the gut, to absorption into the portal circulation, and distribution to target tissues with important metabolic functions (e.g. the liver) and cellular compartments responsible for energy metabolism (e.g. the mitochondria) will allow us to better understand the role of individual microbial metabolites in human metabolism.

1.4 - Gut-liver and microbiome-mitochondria axes

Aberrant hepatic regulation of systemic energy balance and mitochondrial dysfunction are commonly linked to obesity-related metabolic disorders. At an organismal level, the liver plays a central role in governing systemic energy balance and also acting as a metabolic and immunologic buffer between absorbed gut-derived metabolites and the systemic circulation. Similarly, at a cellular level, mitochondria govern the majority of cellular metabolism and bioenergetics thereby acting as an interface between environmental inputs and cellular functions (51). Evolutionary evidence indicates that eukaryotic mitochondria may be descended from an ancient prokaryotic endosymbiont (52), suggestive of possible conserved signaling pathways between bacteria and host mitochondria (36; 51; 53-56).

Microbial signaling influences transcriptional regulation of host energy metabolism in the liver. Bile acids, for example, participate in an exchange between the gut and liver known as the enterohepatic circulation and represent one well-known example of bidirectional gut-liver signaling. Hepatocytes synthesize primary bile acids (cholate and chenodeoxycholate in humans/cholate and muricholate in mice) from cholesterol, which are conjugated with taurine in mice/rats or glycine in humans (taurocholate/taurochenodeoxycholate, glycocholate or glycochenodeoxycholate), and secreted into the biliary tract. Upon reaching the small intestines, bile acids aid in the digestion of dietary fats, cholesterol and fat-soluble vitamins. While the majority of bile acids are actively reabsorbed and returned to the liver via the portal vein, a small percentage of host bile acids undergo microbial modifications (deconjugations, dehydrogenations, dehydroxylations) to produce secondary bile acids (deoxycholate/lithocholate), which are returned to the liver through the portal circulation (57; 58). Microbial metabolism of bile acids activate the Farnesyl-X-receptor (FXR) in the intestinal

epithelia in mice, releasing fibroblast growth factor 15/19 (FGF15/19) to the liver and decreases *de novo* bile acid biosynthesis through inhibition of Cyp7a1 (59). FXR activation affects glucose, cholesterol, and lipid metabolism and appears to play a role in the development of obesity and NAFLD (60; 61). Other transcriptional regulators shown to respond to the microbiota include cholesterol-response element-binding protein (ChREBP) (24), sterol-response element-binding protein (SREBP-1) (24), adenosine monophosphate dependent kinase (AMPK) (25), peroxisome proliferator-activated receptor gamma-alpha (PGC-1 α) (62), peroxisome proliferator-activator receptor (PPARs) (62-64), retinoid X receptor (RXR) (65), sirtuins (SIRT) (66), aryl hydrocarbon receptor (AhR) (67-69), mammalian target of rapamycin (mTOR) (70), nuclear receptor 5A (NR5A) (71), and nuclear factor erythroid-2 related factor (Nrf2) (72; 73). Activation of these and other transcriptional regulators impacts cellular metabolism, mitochondrial morphology and function.

Beyond the previously described interactions of SCFAs in the intestinal epithelia, the role of SCFAs on mitochondrial metabolism has been extensively studied (74-78). Most SCFAs are cleared from the portal circulation by the liver prior to entry into the systemic circulation. In the liver, acetate is converted to acetyl-CoA (79), which can be used for energy production or biosynthesis of cholesterol, long-chain fatty acids, and glutamine/glutamate. In both the gut and liver, propionate is converted to succinyl-CoA and then glucose (75; 80). Butyrate is oxidized in liver mitochondria and improves mitochondrial respiration and fatty acid oxidation in mice (56; 81). These and other bacteria-derived metabolites have been shown to directly target mitochondria in the intestinal epithelia and other tissues (31; 82; 83). Taken together, these examples serve to illustrate the networks of shared metabolites connecting the microbiota to host energy metabolism through transcriptional and mitochondrial channels (**Figure 1.4**). As our

understanding of how microbial metabolism impacts mitochondrial metabolism in the liver is not yet complete, elucidating these complex interactions with emerging integrated systems biology approaches may identify novel therapeutic strategies for treatment of metabolic disorders.

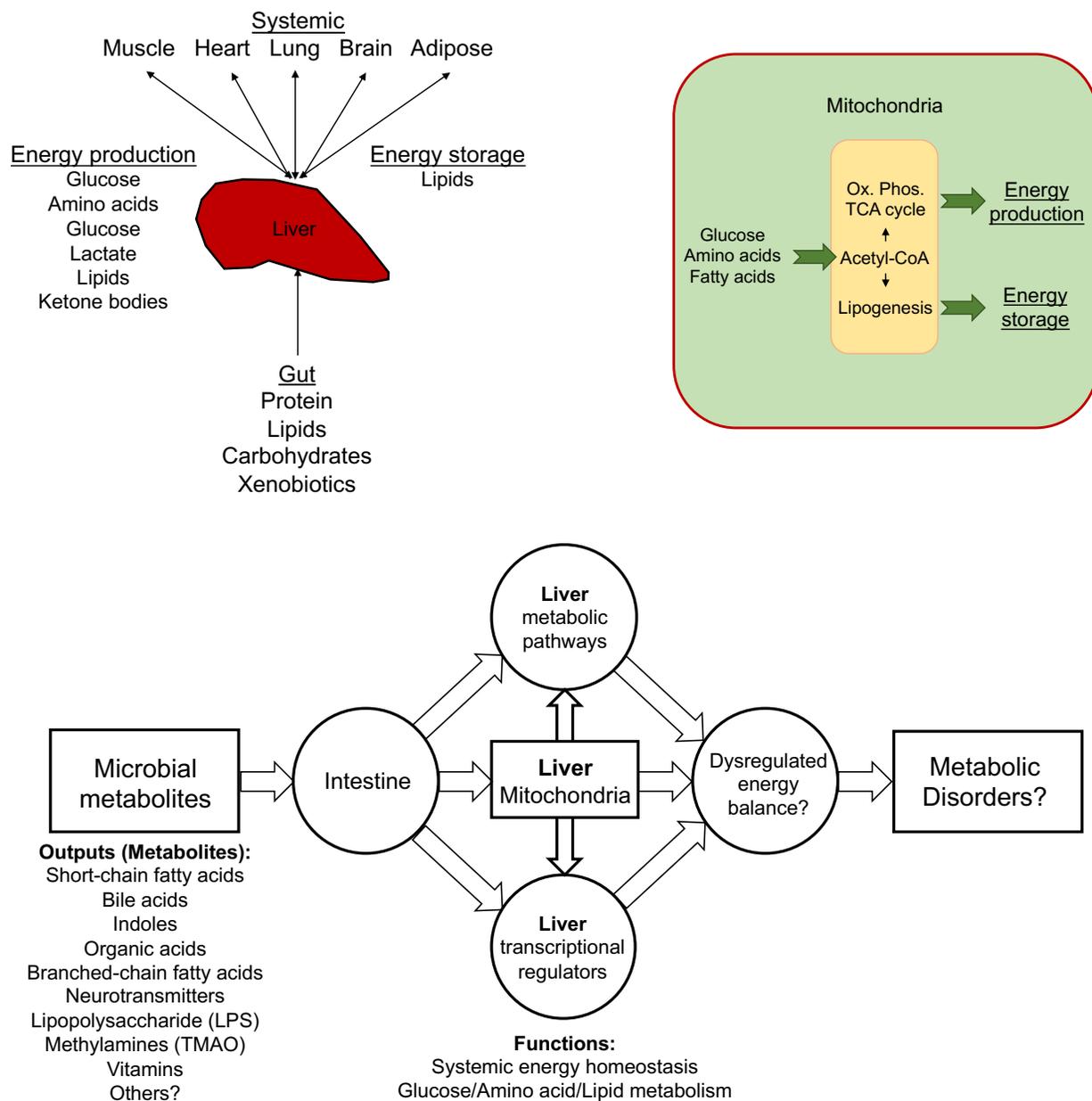


Figure 1.4 *Top*: Energy metabolism at the systemic level regulated by the liver and at the cellular level by mitochondria. *Bottom*: Because the liver is central for balancing gut-derived input with systemic energy needs, liver mitochondria are a central interface between microbial metabolites and systemic energy homeostasis. The interaction of microbiome-derived mitochondrial metabolites with cellular energy metabolism is not well-characterized.

1.5 - Multi-omic systems biology approaches for host-microbiome interactions

Technological advances have allowed the study of intricate multi-faceted host-microbiome networks related to human health, and especially metabolic diseases. In particular, advances in next-generation sequencing, mass spectrometry, and bioinformatics (84-91) have expanded the systems biology toolkit to address these critical points:

- 1. What microbial communities (16s rRNA sequencing (92-94)) are present that are associated with human metabolic health?*
- 2. What functional capabilities (provided by shotgun metagenomics), activities (metatranscriptomics), and outputs (metaproteomics, metametabolomics) does the microbiome provide to the host?*
- 3. What global host transcriptomic, proteomic, and metabolomic responses are linked to the functional outputs from the microbiome?*
- 4. What mechanisms or pathways (hypothesis-driven/or data-driven experiments in vitro/in vivo/etc.) account for cause and effect relationships through which functional outputs from the microbiome elicit host responses linked to human health or disease?*

Complementing the use of metagenomics and metatranscriptomics, which cover points 1 and 2 described above, metabolomics provides information on the functional output of the microbiome and associated host responses covering points 2-4 (95; 96). Metabolomics is focused on measuring the complete set of small molecule metabolites (i.e. metabolome) present in biological samples using nuclear magnetic resonance (NMR) or mass spectrometry. These metabolites include sugars, lipids, amino acids, nucleic acids, and other secondary metabolites which provide both the fuel and the basic materials requiring for building and sustaining life. In the systems biology paradigm, metabolites are considered an endpoint of the central dogma where

deoxyribonucleic acids (DNA) leads to ribonucleic acids (RNA), RNA leads to proteins, and proteins lead to metabolites. Metabolites are often referred to as “closest to the phenotype,” and changes to metabolite concentrations in tissues often reflect upstream changes to pathways related to or driving disease processes.

Human (and metazoan) metabolomes are comprised of contributions from host and microbial metabolism (downstream of their respective genomes, transcriptomes, and proteomes) on endogenous and exogenous (environmental, diet, xenobiotic) chemicals. Recent studies show that fecal metabolomics provide a functional readout of the intestinal microbiota (97; 98) and that blood metabolomics can predict diversity of the intestinal microbiota (38). For example, increased circulating hippuric acid has been identified as a biomarker of gut microbial diversity (99; 100). Several other metabolites have been identified as markers of abnormal intestinal microbiota activity (i.e. dysbiosis) including benzoic acid (29; 101), phenylacetate (29; 102; 103), phenylpropionate (38; 100), (ortho, meta, para)-cresol (29), para-hydroxybenzoic acid (29), para-hydroxyphenylpropionate (38), indoxyl sulfate (104), D-lactate (105), and D-arabinitol (106). Thus, measurements of metabolites associated with the activity of bacteria, fungi, or other microorganisms in distinct host compartments can be used to estimate the contribution of the microbiota to metabolism in that tissue. Nevertheless, the majority of mass spectral features detected in microbiome-metabolomics datasets are not identified, hindering efforts to characterize their pharmacologic and pharmacokinetic properties (50).

The host transcriptional response to the microbiome can be measured by the use of microarray or RNA-sequencing (86)-based transcriptomics. In contrast to the metabolome, which is typically considered closest to the phenotype, the transcriptome, or the complete set of RNA molecules present in a sample, is typically considered closest to the genotype. RNA

molecules exist in several forms (rRNA, tRNA, mRNA, lncRNA, siRNA, shRNA), and serve multiple roles within cells, including catalysis, regulation of gene expression, and protein synthesis. Global gene expression is governed by the activity of transcriptional regulators (nuclear hormone receptors (NHRs), coactivators and corepressors for NHRs, kinases, GPCRs, and others) that respond to biological information about the cellular environment and state. Microbial small molecules, alterations to cellular state, changes in cell cycle, aging, circadian rhythms, and post-translational modifications can represent signals sensed by transcriptional regulators to alter gene expression. Databases containing tissue and organism-specific gene expression profiles linked to disease, drugs/toxicants use, biological pathways, cellular components, metabolic processes, or transcription factors facilitate functional annotation of transcriptomics data and can be used for mechanistic inference of genomic targets impacted by the microbiota.

Multi-omic study designs are becoming increasingly popular by offering a holistic perspective on the interaction of molecular networks linking the microbiome to host physiology (**Figure 1.5**). For example, these approaches have been employed to understand how the microbiota are linked to host metabolism (7; 62; 64; 83; 99; 107; 108), circadian rhythms, and diurnal variations (109-113). While these top-down studies have enhanced our ability to understand these relationships, our knowledge of microbial functions, metabolites, and pathways is far from complete. The majority of genes identified in metagenomic studies are not matched in functional databases or do not map to known microbial genes (114). Taken together with the observation that most features in microbiome-metabolomics datasets are similarly unidentified, this suggests that most functional relationships between the microbiome and host remain to be elucidated. Thus, use of computational databases for functional annotation of -omics data

(which rely on prior knowledge) will miss poorly characterized pathways, genes or metabolites that could have important biological significance. Ultimately, if the goal is for these insights to develop into novel therapeutic strategies, these hidden relationships need to be characterized and evaluated as druggable targets.

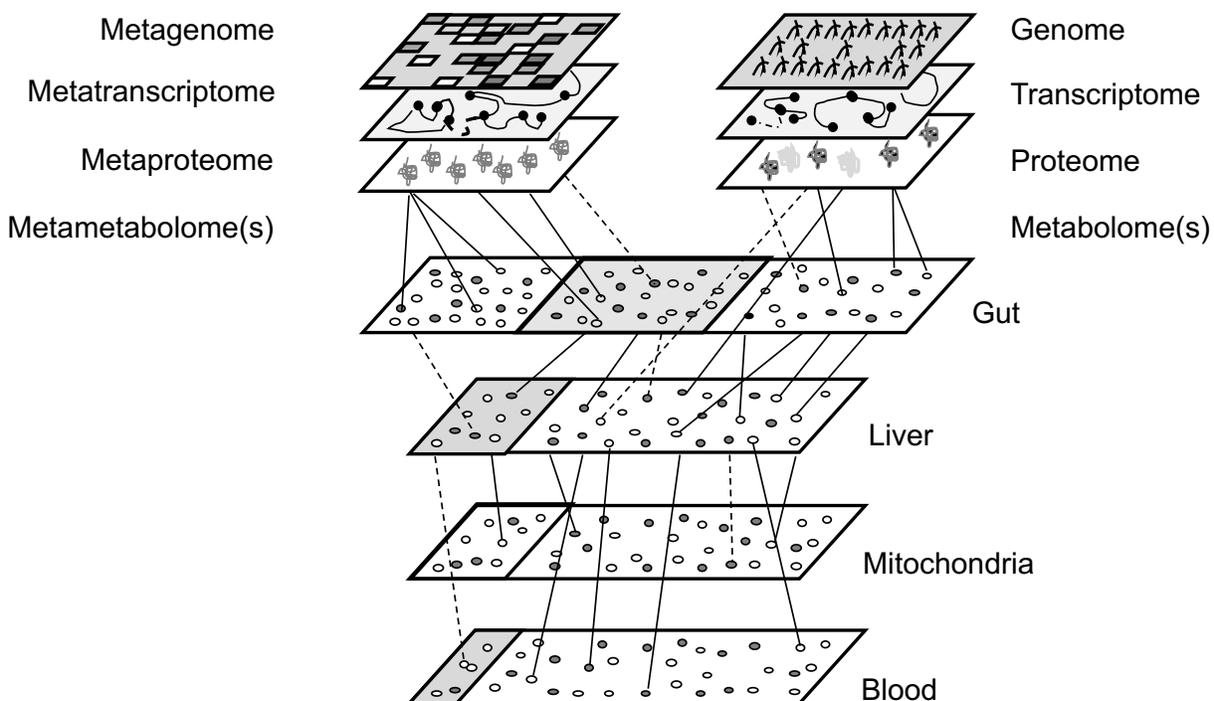


Figure 1.5 Systems biology paradigm of host-microbiome interactions. While 10% of the circulating mammalian metabolome is microbial in origin (shaded gray), little is known about the microbial contribution to the mitochondrial metabolome (unshaded). Many features from metagenomics and metabolomics datasets are poorly annotated and require detailed investigations to elucidate biological pathways and functions linked to individual microbial metabolites. Open circles represent uncharacterized metabolites. Colored circles represent characterized metabolites. Lines connecting dots between different compartments reflect either transport or metabolic processes.

To address molecular understanding of gut-host energy regulation and to characterize unidentified relationships, two general experimental approaches can be taken: 1) Target-driven discovery aims to identify functional output (metabolites, transcripts, proteins) from the microbiota that specifically interacts with a known target in the host or 2) Metabolite, gene product, or protein-driven discovery aims to identify a novel functional output from the microbiota followed by efforts to elucidate cellular targets and processes that are changed by modifying the amount of ligand, gene product or protein in the organism. This dissertation uses contemporary high-resolution mass spectrometry-based metabolomics and bioinformatics methods to obtain both target-driven and systems level understanding of mechanisms governing gut-host energy regulation.

Conclusion

Given the rising rates of human metabolic diseases, it is of critical importance to understand the relationship of the microbiome with human energy metabolism. Recent advancements in mass spectrometry-based metabolomics and bioinformatics tools for integration of multi-omic data provide approaches to better characterize the interactions between small molecules produced by the microbiome and host energy metabolism. The knowledge gained from these studies can be used to drive the development of strategies for treatment of human metabolic disorders.

1.6 – Hypothesis

Diet and the intestinal microbiome contribute to obesity-related metabolic disorders, often initiated by and exacerbated by excessive fat accumulation in tissues. At present, the majority of microbiome-derived metabolites have not yet been studied in terms of their distribution and activity in host tissues. While an estimated 10% of the circulating mammalian metabolome is associated with the activity of the microbiota (37), the contribution of the microbiota to the liver and liver mitochondria metabolomes (115-117) has not yet been completely characterized. Because germ-free mice are protected from diet induced obesity, we hypothesize that uncharacterized microbiome-derived metabolites impact hepatic mitochondrial function and hepatic regulation of systemic energy metabolism to increase energy storage or decrease energy expenditure.

1.7 – Statement of research objectives

The purpose of this dissertation research was to obtain an understanding of intestinal microbiome product(s) which control host energy metabolism. The dissertation chapters described below are designed to test the above stated hypothesis. Chapter 2 is a published manuscript in Journal of Occupational and Environmental Medicine that describes my work where I systematically evaluated several analytical methodologies to determine a metabolomics platform suitable for maximizing detection of known and unidentified metabolites. Chapter 3 is a submitted manuscript which describes my analysis of over 700 individual analytical standards using the optimized methodology to establish a list of known metabolites for harmonization and quantification of metabolomics data, which provides the basis for identifying and quantifying novel microbial metabolites in study samples. Chapter 4 is a manuscript which applies the foundational work established in chapters 2 and 3 to identify δ -valerobetaine (VB), a novel mitochondrial metabolite produced by the microbiome, and to characterize the potential contribution of VB to obesity, fatty liver, and human metabolic disorders.

References

1. O'Neill S, O'Driscoll L. 2015. Metabolic syndrome: a closer look at the growing epidemic and its associated pathologies. *Obes Rev* 16:1-12
2. Hruby A, Hu FB. 2015. The Epidemiology of Obesity: A Big Picture. *Pharmacoeconomics* 33:673-89
3. Mozumdar A, Liguori G. 2011. Persistent increase of prevalence of metabolic syndrome among U.S. adults: NHANES III to NHANES 1999-2006. *Diabetes Care* 34:216-9
4. Dandona P, Aljada A, Falck E, Bandyopadhyay A, Bandyopadhyay A. Inflammation: the link between insulin resistance, obesity and diabetes.
5. Flores M, Glusman G, Brogaard K, Price ND, Hood L. 2013. P4 medicine: how systems medicine will transform the healthcare sector and society. *Personalized Medicine* 10:565-76
6. Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. Dysbiosis of the gut microbiota in disease.
7. Dumas MB, R. H.; Toye, A.; Cloarec, O.; Blancher, C.; Rothwell, A.; Fearnside, J.; Tatoud, R.; Blanc, V.; Lindon, J. C.; Mitchell, S. C.; Holmes, E.; McCarthy, M. I.; Scott, J.; Gauguier, D., Nicholson, J. K. 2006. Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *PNAS* 103:12511-6
8. Holmes E, Li JV, Marchesi JR, Nicholson JK. 2012. Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metab* 16:559-64
9. Sonnenburg JL, Backhed F. Diet-microbiota interactions as moderators of human metabolism.
10. Clemente JC, Ursell LK, Parfrey LW, Knight R. 2012. The impact of the gut microbiota on human health: an integrative view. *Cell* 148:1258-70
11. Palau-Rodriguez M, Tulipani S, Isabel Queipo-Ortuno M, Urpi-Sarda M, Tinahones FJ, Andres-Lacueva C. 2015. Metabolomic insights into the intricate gut microbial-host interaction in the development of obesity and type 2 diabetes. *Front Microbiol* 6:1151
12. Cani PD, Van Hul M, Lefort C, Depommier C, Rastelli M, Everard A. 2019. Microbial regulation of organismal energy homeostasis. *Nature Metabolism* 1:34-46
13. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, et al. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464:59-65
14. O'Hara AM, Shanahan F. 2006. The gut flora as a forgotten organ. *EMBO Rep* 7:688-93
15. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Nageshwar Reddy D. 2015. Role of the normal gut microbiota. *World J Gastroenterol* 21:8787-803
16. Valdes AM, Walter J, Segal E, Spector TD. 2018. Role of the gut microbiota in nutrition and health. *BMJ* 361:k2179
17. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, et al. 2009. A core gut microbiome in obese and lean twins. *Nature* 457:480-4
18. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, et al. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505:559-63
19. Ley RE, Bäckhed F, Turnbaugh PJ, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology. *PNAS* 102:11070-5
20. Million M, Angelakis E, Paul M, Armougom F, Raoult D. Comparative meta-analysis of the effect of *Lactobacillus* species on weight gain in humans and animals.

21. Ignacio A, Fernandes MR, Rodrigues VA, Groppo FC, Cardoso AL, et al. 2016. Correlation between body mass index and faecal microbiota from children. *Clin Microbiol Infect* 22:258 e1-8
22. Turnbaugh PJR, V. K.; Faith, J. J.; Rey, F. E.; Knight, R.; Gordon, J. I. 2009. The Effect of Diet on the Human Gut Microbiome: A Metagenomic Analysis in Humanized Gnotobiotic Mice. *Science Translational Medicine* 1:10
23. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, et al. 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334:105-8
24. Bäckhed FD, H.; Wang, T.; Hooper, L. V.; Koh, G. Y.; Nagy, A.; Semenkovich, C. F.; Gordon, J. I. 2004. The gut microbiota as an environmental factor that regulates fat storage. *PNAS* 101:15718-23
25. Bäckhed FM, J. K.; Semenkovich, C. F.; Gordon, J. I. 2007. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *PNAS* 104:979-84
26. Suarez-Zamorano N, Fabbiano S, Chevalier C, Stojanovic O, Colin DJ, et al. Microbiota depletion promotes browning of white adipose tissue and reduces obesity.
27. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444:1027-31
28. Turnbaugh PJ, Backhed F, Fulton L, Gordon JI. 2008. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* 3:213-23
29. Holmes E, Li JV, Athanasiou T, Ashrafian H, Nicholson JK. 2011. Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. *Trends Microbiol* 19:349-59
30. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. 2016. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14:273-87
31. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, et al. 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* 13:517-26
32. Delzenne NM, Cani Pd Fau - Daubioul C, Daubioul C Fau - Neyrinck AM, Neyrinck AM. Impact of inulin and oligofructose on gastrointestinal peptides.
33. Cani PD, Hoste S, Guiot Y, Delzenne NM. 2007. Dietary non-digestible carbohydrates promote L-cell differentiation in the proximal colon of rats. *British Journal of Nutrition* 98:32-7
34. Cani PD, Bibiloni R, Knauf C, Waget A, Neyrinck AM, et al. 2008. Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes* 57:1470-81
35. Cani PD, Lecourt E, Dewulf EM, Sohet FM, Pachikian BD, et al. 2009. Gut microbiota fermentation of prebiotics increases satietogenic and incretin gut peptide production with consequences for appetite sensation and glucose response after a meal. *The American Journal of Clinical Nutrition* 90:1236-43
36. Clark A, Mach N. 2017. The Crosstalk between the Gut Microbiota and Mitochondria during Exercise. *Front Physiol* 8:319

37. Wikoff WRA, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G. 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *PNAS* 106:3698-703
38. Wilmanski T, Rappaport N, Earls JC, Magis AT, Manor O, et al. 2019. Blood metabolome predicts gut microbiome α -diversity in humans. *Nature Biotechnology*
39. Cani PD, Amar J, Iglesias MA, Poggi M, Knauf C, et al. 2007. Metabolic Endotoxemia Initiates Obesity and Insulin Resistance. *Diabetes* 56:1761-72
40. Cani PD, Bibiloni R, Knauf C, Waget A, Neyrinck AM, et al. 2008. Changes in Gut Microbiota Control Metabolic Endotoxemia-Induced Inflammation in High-Fat Diet-Induced Obesity and Diabetes in Mice. *Diabetes* 57:1470-81
41. Hoyles L, Fernandez-Real JM, Federici M, Serino M, Abbott J, et al. 2018. Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nat Med* 24:1070-80
42. Lustgarten MS, Price LL, Chale A, Fielding RA. 2014. Metabolites related to gut bacterial metabolism, peroxisome proliferator-activated receptor- α activation, and insulin sensitivity are associated with physical function in functionally-limited older adults. *Aging Cell* 13:918-25
43. Peng KY, Watt MJ, Rensen S, Greve JW, Huynh K, et al. 2018. Mitochondrial dysfunction-related lipid changes occur in nonalcoholic fatty liver disease progression. *J Lipid Res* 59:1977-86
44. Jin R, Banton S, Tran VT, Konomi JV, Li S, et al. Amino Acid Metabolism is Altered in Adolescents with Nonalcoholic Fatty Liver Disease-An Untargeted, High Resolution Metabolomics Study.
45. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, et al. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472:57-63
46. Corbin KD, Zeisel SH. 2012. Choline metabolism provides novel insights into nonalcoholic fatty liver disease and its progression. *Curr Opin Gastroenterol* 28:159-65
47. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, et al. 2013. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 19:576-85
48. Koeth RA, Levison BS, Culley MK, Buffa JA, Wang Z, et al. 2014. gamma-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab* 20:799-812
49. Zhu W, Gregory JC, Org E, Buffa JA, Gupta N, et al. 2016. Gut Microbial Metabolite TMAO Enhances Platelet Hyperreactivity and Thrombosis Risk. *Cell* 165:111-24
50. Peisl BYL, Schymanski EL, Wilmes P. 2018. Dark matter in host-microbiome metabolomics: Tackling the unknowns-A review. *Anal Chim Acta* 1037:13-27
51. Go YM, Fernandes J, Hu X, Uppal K, Jones DP. Mitochondrial network responses in oxidative physiology and disease.
52. Sagan L. 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14:225-236
53. Pallen MJ. Time to recognise that mitochondria are bacteria?
54. Searcy DG. Metabolic integration during the evolutionary origin of mitochondria.
55. Pallen MJ. 2011. Time to recognise that mitochondria are bacteria? *Trends Microbiol* 19:58-64

56. Mollica MP, Mattace Raso G, Cavaliere G, Trinchese G, De Filippo C, et al. 2017. Butyrate Regulates Liver Mitochondrial Function, Efficiency, and Dynamics in Insulin-Resistant Obese Mice. *Diabetes* 66:1405-18
57. Chiang JY. Bile acid metabolism and signaling.
58. Wahlstrom A, Sayin SI, Marschall HU, Backhed F. Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism.
59. Sayin SI, Wahlstrom A Fau - Felin J, Felin J Fau - Jantti S, Jantti S Fau - Marschall H-U, Marschall Hu Fau - Bamberg K, et al. Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist.
60. Parséus A, Sommer N, Sommer F, Caesar R, Molinaro A, et al. 2017. Microbiota-induced obesity requires farnesoid X receptor. *Gut* 66:429-37
61. Gonzalez FJ, Jiang C, Patterson AD. 2016. An Intestinal Microbiota-Farnesoid X Receptor Axis Modulates Metabolic Disease. *Gastroenterology* 151:845-59
62. Kindt A, Liebisch G, Clavel T, Haller D, Hormannsperger G, et al. 2018. The gut microbiota promotes hepatic fatty acid desaturation and elongation in mice. *Nat Commun* 9:3760
63. Manoharan I, Suryawanshi A, Hong Y, Ranganathan P, Shanmugam A, et al. 2016. Homeostatic PPARalpha Signaling Limits Inflammatory Responses to Commensal Microbiota in the Intestine. *J Immunol* 196:4739-49
64. Wang R, Li H, Yang X, Xue X, Deng L, et al. 2018. Genetically Obese Human Gut Microbiota Induces Liver Steatosis in Germ-Free Mice Fed on Normal Diet. *Front Microbiol* 9:1602
65. Swann JR, Want EJ, Geier FM, Spagou K, Wilson ID, et al. 2011. Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. *Proc Natl Acad Sci U S A* 108 Suppl 1:4523-30
66. Wellman AS, Metukuri MR, Kazgan N, Xu X, Xu Q, et al. Intestinal Epithelial Sirtuin 1 Regulates Intestinal Inflammation During Aging in Mice by Altering the Intestinal Microbiota.
67. Jin U-H, Lee S-O, Sridharan G, Lee K, Davidson LA, et al. 2014. Microbiome-Derived Tryptophan Metabolites and Their Aryl Hydrocarbon Receptor-Dependent Agonist and Antagonist Activities. *Molecular Pharmacology* 85:777-88
68. Korecka A, Dona A, Lahiri S, Tett AJ, Al-Asmakh M, et al. 2016. Bidirectional communication between the Aryl hydrocarbon Receptor (AhR) and the microbiome tunes host metabolism. *Npj Biofilms And Microbiomes* 2:16014
69. Krishnan S, Ding Y, Saedi N, Choi M, Sridharan GV, et al. 2018. Gut Microbiota-Derived Tryptophan Metabolites Modulate Inflammatory Response in Hepatocytes and Macrophages. *Cell Rep* 23:1099-111
70. Koh A, Molinaro A, Ståhlman M, Khan MT, Schmidt C, et al. 2018. Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* 175:947-61.e17
71. Lin CJ, Wang MC. 2017. Microbial metabolites regulate host lipid metabolism through NR5A-Hedgehog signalling. *Nat Cell Biol* 19:550-7
72. Saedi BJ, Robinson BS, Owens J, Liu K, Eboka R, et al. 2019. Regulation of the Hepatic Antioxidant Response by the Probiotic *Lactobacillus rhamnosus* GG. *The FASEB Journal* 33:369.5-.5

73. Li B, Evivie Se Fau - Lu J, Lu J Fau - Jiao Y, Jiao Y Fau - Wang C, Wang C Fau - Li Z, et al. Lactobacillus helveticus KLDS1.8701 alleviates d-galactose-induced aging by regulating Nrf-2 and gut microbiota in mice.
74. Jacob AD, Elkins N Fau - Reiss OK, Reiss Ok Fau - Chan L, Chan L Fau - Shapiro JI, Shapiro JI. Effects of acetate on energy metabolism and function in the isolated perfused rat heart.
75. Perry RJ, Borders CB, Cline GW, Zhang XM, Alves TC, et al. 2016. Propionate Increases Hepatic Pyruvate Cycling and Anaplerosis and Alters Mitochondrial Metabolism. *J Biol Chem* 291:12161-70
76. De Vadder F, Kovatcheva-Datchary P, Goncalves D, Vinera J, Zitoun C, et al. 2014. Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* 156:84-96
77. Frye RE, Rose S, Chacko J, Wynne R, Bennuri SC, et al. 2016. Modulation of mitochondrial function by the microbiome metabolite propionic acid in autism and control cell lines. *Transl Psychiatry* 6:e927-e
78. Ji X, Zhou F, Zhang Y, Deng R, Xu W, et al. 2019. Butyrate stimulates hepatic gluconeogenesis in mouse primary hepatocytes. *Exp Ther Med* 17:1677-87
79. Smyth DH. 1947. The Rate and Site of Acetate Metabolism in the Body. *J. Physiol.* 105:299-315
80. De Vadder F, Kovatcheva-Datchary P, Zitoun C, Duchampt A, Backhed F, Mithieux G. 2016. Microbiota-Produced Succinate Improves Glucose Homeostasis via Intestinal Gluconeogenesis. *Cell Metab* 24:151-7
81. Leloir LFM, J. M. 1944. Butyrate Oxidation by Liver Enzymes. *J. Biol. Chem* 153:53-60
82. Tao S, Niu L, Cai L, Geng Y, Hua C, et al. N-(3-oxododecanoyl)-l-homoserine lactone modulates mitochondrial function and suppresses proliferation in intestinal goblet cells.
83. Chella Krishnan K, Kurt Z, Barrere-Cain R, Sabir S, Das A, et al. 2018. Integration of Multi-omics Data from Mouse Diversity Panel Highlights Mitochondrial Dysfunction in Non-alcoholic Fatty Liver Disease. *Cell Systems* 6:103-15.e7
84. Franzosa EA, Huang K, Meadow JF, Gevers D, Lemon KP, et al. Identifying personal microbiomes using metagenomic codes.
85. Franzosa EA, Hsu T, Sirota-Madi A, Shafquat A, Abu-Ali G, et al. Sequencing and beyond: integrating molecular 'omics' for microbial community profiling.
86. Wang ZG, M.; Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 10:57-63
87. Hamady M, Knight R. 2009. Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res* 19:1141-52
88. Greenblum S, Turnbaugh PJ, Borenstein E. 2012. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc Natl Acad Sci U S A* 109:594-9
89. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. 2013. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. *Metabolomics* 9:S132-S43
90. Uppal KS, Q. A.; Strobel, F. H.; Pittard, W. S.; Gernert, K. M.; Yu, T.; Jones, D. P. 2013. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics* 14

91. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, et al. 2013. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol* 9:e1003123
92. Woo PC, Lau Sk Fau - Teng JLL, Teng JI Fau - Tse H, Tse H Fau - Yuen KY, Yuen KY. Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories.
93. Sunagawa S, Mende Dr Fau - Zeller G, Zeller G Fau - Izquierdo-Carrasco F, Izquierdo-Carrasco F Fau - Berger SA, Berger Sa Fau - Kultima JR, et al. Metagenomic species profiling using universal phylogenetic marker genes.
94. Janda JM, Abbott SL. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 45:2761-4
95. Uppal K, Walker DI, Liu K, Li S, Go YM, Jones DP. 2016. Computational Metabolomics: A Framework for the Million Metabolome. *Chem Res Toxicol* 29:1956-75
96. Walker D, Go Y-M, Liu K, Pennell K, Jones D. 2016. Population Screening for Biological and Environmental Properties of the Human Metabolic Phenotype. 167-211. Number of 167-211 pp.
97. Zierer J, Jackson MA, Kastenmuller G, Mangino M, Long T, et al. 2018. The fecal metabolome as a functional readout of the gut microbiome. *Nat Genet* 50:790-5
98. Zhao Y, Wu J, Li JV, Zhou NY, Tang H, Wang Y. 2013. Gut microbiota composition modifies fecal metabolic profiles in mice. *J Proteome Res* 12:2987-99
99. Claus SP, Tsang TM, Wang Y, Cloarec O, Skordi E, et al. 2008. Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. *Mol Syst Biol* 4:219
100. Pallister T, Jackson MA, Martin TC, Zierer J, Jennings A, et al. 2017. Hippurate as a metabolomic marker of gut microbiome diversity: Modulation by diet and relationship to metabolic syndrome. *Sci Rep* 7:13670
101. Williams HRT, Cox IJ, Walker DG, Cobbold JFL, Taylor-Robinson SD, et al. 2010. Differences in gut microbial metabolism are responsible for reduced hippurate synthesis in Crohn's disease. *BMC Gastroenterology* 10:108
102. Teufel R, Mascaraque V, Ismail W, Voss M, Perera J, et al. 2010. Bacterial phenylalanine and phenylacetate catabolic pathway revealed. *Proc Natl Acad Sci U S A* 107:14390-5
103. Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites.
104. Weber D, Oefner PJ, Hiergeist A, Koestler J, Gessner A, et al. 2015. Low urinary indoxyl sulfate levels early after transplantation reflect a disrupted microbiome and are associated with poor outcome. *Blood* 126:1723
105. Bulik-Sullivan EC, Roy S, Elliott RJ, Kassam Z, Lichtman SN, et al. 2018. Intestinal Microbial and Metabolic Alterations Following Successful Fecal Microbiota Transplant for D-Lactic Acidosis. *Journal of Pediatric Gastroenterology and Nutrition* 67
106. Christensson B, Sigmundsdottir G Fau - Larsson L, Larsson L. D-arabinitol--a marker for invasive candidiasis.
107. Martin FP, Dumas ME, Wang Y, Legido-Quigley C, Yap IK, et al. 2007. A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. *Mol Syst Biol* 3:112

108. Murakami M, Tognini P, Liu Y, Eckel-Mahan KL, Baldi P, Sassone-Corsi P. 2016. Gut microbiota directs PPARgamma-driven reprogramming of the liver circadian clock by nutritional challenge. *EMBO Rep* 17:1292-303
109. Deaver JA, Eum SY, Toborek M. 2018. Circadian Disruption Changes Gut Microbiome Taxa and Functional Gene Composition. *Front Microbiol* 9:737
110. Montagner A, Korecka A, Polizzi A, Lippi Y, Blum Y, et al. 2016. Hepatic circadian clock oscillators and nuclear receptors integrate microbiome-derived signals. *Sci Rep* 6:20127
111. Hooper LV, Wang Y, Kuang Z. 2018. Metabolism, the microbiome, and the circadian clock. *The FASEB Journal* 32:101.2-2
112. Thaïss CA, Levy M, Korem T, Dohnalova L, Shapiro H, et al. 2016. Microbiota Diurnal Rhythmicity Programs Host Transcriptome Oscillations. *Cell* 167:1495-510 e12
113. Weger BD, Gobet C, Yeung J, Martin E, Jimenez S, et al. 2018. The Mouse Microbiome Is Required for Sex-Specific Diurnal Rhythms of Gene Expression and Metabolism. *Cell Metab*
114. Thomas AM, Segata N. 2019. Multiple levels of the unknown in microbiome research. *BMC Biology* 17:48
115. Go YM, Uppal K Fau - Walker DI, Walker Di Fau - Tran V, Tran V Fau - Dury L, Dury L Fau - Strobel FH, et al. Mitochondrial metabolomics using high-resolution Fourier-transform mass spectrometry.
116. Roede JR, Park Y, Li S, Strobel FH, Jones DP. 2012. Detailed mitochondrial phenotyping by high resolution metabolomics. *PLoS One* 7:e33020
117. Go YM, Fernandes J, Hu X, Uppal K, Jones DP. 2018. Mitochondrial network responses in oxidative physiology and disease. *Free Radic Biol Med* 116:31-40

Chapter 2 : Evaluating strategies for maximizing chemical detection using high-resolution metabolomics (HRM)

This chapter (previously published in the Journal of Occupational and Environmental Medicine, doi: 10.1097/JOM.0000000000000773) establishes a chemical detection platform which maximizes the ability to detect known and unidentified metabolites, which serves as a foundation for the work described in chapter 3 and 4.

High-Resolution Metabolomics Assessment of Military Personnel: Evaluating Analytical Strategies for Chemical Detection

Ken H. Liu, Douglas I. Walker, Karan Uppal, ViLinh Tran, Patricia Rohrbeck, Timothy M. Mallon, Dean P. Jones

Abstract

Objective: The aim of this study was to maximize detection of serum metabolites with high-resolution metabolomics (HRM).

Methods: Department of Defense Serum Repository (DoDSR) samples were analyzed using ultra-high resolution mass spectrometry with three complementary chromatographic phases and four ionization modes. Chemical coverage was evaluated by number of ions detected and accurate mass matches to a human metabolomics database.

Results: Individual HRM platforms provided accurate mass matches for up to 58% of the KEGG metabolite database. Combining two analytical methods increased matches to 72%, and included metabolites in most major human metabolic pathways and chemical classes. Detection and feature quality varied by analytical configuration.

Conclusions: Dual chromatography HRM with positive and negative electrospray ionization provides an effective generalized method for metabolic assessment of military personnel.

2.1 Introduction

Military personnel can be exposed to many environmental chemicals, and analytical methods which target specific chemicals or small groups of chemicals are inadequate to address the problem (1). High-resolution metabolomics (HRM) provides an affordable, high-throughput platform capable of advanced clinical chemistry measurements, environmental chemical surveillance and bioeffect monitoring suitable for precision medicine. HRM relies on a combination of chromatography coupled to ultra-high resolution mass spectrometry and advanced computational approaches for spectral feature alignment, peak integration, and feature extraction (2-6). With this workflow, HRM is capable of reproducibly measuring greater than 10,000 unique spectral features (defined by a characteristic mass-to-charge ratio (m/z), retention time, and intensity) using small volumes of biological specimens. The resulting chemical profile, often referred to as the metabolic phenotype, includes a broad range of chemical classes and metabolic pathways (2-4; 6-8). A number of studies have used HRM-based approaches to identify metabolic changes associated with a variety of disease, clinical or exposure settings, including Parkinson disease (9), pulmonary tuberculosis (10), HIV-1 infection (11), age-related macular degeneration (12), lung transplantation (13), alcohol abuse (14), and cadmium exposure (15). Furthermore, application of untargeted HRM approaches to population screening and clinical use provides improved capabilities for biomarker discovery and identifying unknown chemical exposures, with simultaneous measurements of metabolic network and pathway associations (16-18) for enhanced understanding of human health and disease.

Estimates of the metabolome range from 2,000 to >100,000 metabolites from endogenous sources (e.g., lipids, carbohydrates, nucleotides, amino acids, metabolic intermediates, signaling molecules, small peptides) and exogenous dietary and environmental sources (1; 5; 19; 20). The

ability to detect the very large number of chemicals is complicated by the very broad range in abundance of chemicals in biologic systems, which span at least eight orders of magnitude in human plasma (21). Online chemical databases such as Metlin (over 240,000 entries), the Kyoto Encyclopedia of Genes and Genomes (KEGG) (over 17,000 entries), or the Human Metabolome Database (HMDB) (over 40,000 entries) facilitate efforts to obtain unequivocal chemical identification of spectral features obtained from HRM analysis, but do not provide such identification alone. Despite this, databases provide useful resources to evaluate different analytical platforms, even without absolute identification. For instance, if one assumes that different platforms generate the same frequency of incorrect matches, then comparison of results provides an estimate of which platform gives better coverage. Similarly, the fraction of ions measured that do not have a match in the databases provides an estimate of the completeness of the databases. While imprecise, these approaches provide ways to compare performance of different platforms.

Chemicals present in the metabolome span a wide range of physicochemical properties (mass, polarity, abundance, lipophilicity, pKa), and numerous analytical strategies are available to analyze this complex mixture, including Nuclear Magnetic Resonance (NMR) or Fourier Transform Infrared Spectroscopy (FTIR) based methods, or mass spectrometric methods (MS) coupled to gas (GC), liquid chromatography (LC) or direct injection (DI) (6; 22-34). NMR and FTIR-based methods provide fast quantification of high abundance metabolites with minimal sample processing. Due to increased sensitivity enabling detection of low abundance chemicals, MS-based metabolomics provides improved capability for in-depth metabolic profiling.

Recent advances in MS technology, such as very high mass resolution (>60,000 resolving power), mass accuracy (<5 ppm), and increased scan speed of ultra-high resolution mass

spectrometers, decreases the requirements for separation of chemicals prior to detection (35-37).

Nonetheless, the use of liquid chromatography (LC) separation prior to MS improves chemical coverage, sensitivity, and quantification, especially for complex biofluid analysis (6).

Commonly used LC strategies include reverse phase (C_{18}) chromatography, hydrophilic interaction liquid chromatography (HILIC) and anion exchange (AE) chromatography. A variety of ion sources are also available for ionization of analytes prior to introduction into the mass spectrometer, including electrospray ionization (ESI) (38; 39) and atmospheric pressure chemical ionization (APCI) (30). Complementary ionization and separation approaches for metabolomics analyses have been applied for metabolic phenotyping studies of human populations. Dunn et. al (20), Want et. al (40), Rabinowitz et. al (29), Patti et. al (41), Psychogios et. al (42) and others have made considerable progress demonstrating the utility of different instrumentation. However, direct comparison of different chromatographic/ionization platforms for HRM is not available.

To identify the optimal analytical strategy for HRM profiling of serum obtained from military personnel, we analyzed a set of thirty non-identified serum samples obtained from the United States Department of Defense Serum Repository (DoDSR) with different combinations of HILIC, C_{18} , or AE chromatography and ionization strategies [positive (+) and negative (-) ESI and APCI]. We compare the total number of reproducible ions detected, defined as m/z features (accurate mass mass/charge with associated retention time and ion intensity) detected with each configuration and matches to known chemicals in the Kyoto Encyclopedia of Genes and Genomes Human Metabolite database (KEGG) (43; 44). This strategy allowed us to perform an estimate of metabolic coverage of known and unknown chemicals based on the total number of ions detected and KEGG matches from common adducts. We followed this estimation by

comparing detection of representative metabolites with each HRM strategy. This systematic evaluation will provide guidance on an optimal analytical configuration that could be used for future metabolic phenotyping studies in military personnel.

2.2 Methods

Samples

Both Emory University and the USUHS IRBs reviewed and approved the research protocol as non-human subjects population health surveillance. Thirty unidentified serum samples (i.e., unknown source or date of collection) were obtained from the Department of Defense Serum Repository (DoDSR) for analysis. The repository consists of approximately 50 million serum samples originally collected for mandatory armed forces personnel HIV testing (45). Samples were collected according to standard DoDSR protocols (45) at military treatment facilities, shipped at 4-8°C and placed in long-term storage at -30°C. Prior to analysis, specimens were thawed, mixed on a vortex mixer, and 500 µL was aliquoted into separate micro centrifuge tubes, refrozen and shipped on dry ice to Emory University according to standard DoDSR protocols (45).

Sample Preparation

An internal standard mixture consisting of 8 stable isotope internal standards was prepared in LCMS grade acetonitrile (Fluka Analytical). This mixture broadly represents different classes of small molecules for high-resolution metabolomics analysis. These chemicals included [¹³C₆]-D-glucose, [¹⁵N]-indole, [¹³C₅]-L-glutamic acid, [¹⁵N]-L-tyrosine, [trimethyl-¹³C₃]-caffeine, [3,3-¹³C₂]-cystine, [¹⁵N,¹³C₅]-L-methionine, and [¹³C₅, ¹⁵N₂]-L-glutamine. All internal standards were obtained from Cambridge Isotope Laboratories and >98% pure, according to the manufacturer's certificate of analyses. Accurate masses for the "M+H", "M+Na", or "M-H" adducts for each internal standard compound were used to verify the presence of standard in each sample. Each sample was prepared according to Soltow et al. (6) by adding 130 µL of acetonitrile containing the internal standard mixture to 65 µL of serum.

Following mixing and incubation on ice for 30 min, precipitated proteins were pelleted with 10 min centrifugation at $16,100 \times g$ at 4°C . The resulting supernatant was transferred into autosampler vials and maintained at 4°C for the duration of the analysis (<24 hours). Pooled plasma reference samples (Qstd) and the NIST SRM 1950 certified plasma reference standard (46) were prepared and included at the beginning and end of every analytical run for post hoc quantification and standardization.

Liquid chromatography

Chromatographic separation was performed on a Dionex Ultimate 3000 UHPLC with a dual column compartment for column switching. This setup allowed an analytical separation to be performed on one column, while a second column was washed and conditioned prior to the next injection. For each set of analyses, a single chromatographic retention mechanism was employed, with the ionization polarity alternating between injections. Mobile phase A consisted of 2% formic acid (Sigma-Aldrich, analytical grade 27001) in LCMS grade water (Fluka Analytical Chromasolv LCMS grade). Mobile phase B consisted of LCMS grade acetonitrile, and mobile phase C consisted of LCMS grade water. Following each analytical injection, each column was washed and reconditioned at the starting mobile phase conditions for 20 min. The flow rate for all analytical separations was $350 \mu\text{L}/\text{min}$, injection volume was $10 \mu\text{L}$ and all samples were analyzed in triplicate to ensure analytical reproducibility. Previous analyses with a 10 min gradient showed the need for an extended wash period, as some late-eluting chemicals were not completely resolved within the 10 min runtime. As a result, each analytical separation was performed for 20 minutes.

Reverse Phase (C_{18}) chromatography: Higgins C_{18} 100×2.1 mm (TS-1021-C185) columns were used for reversed phase separation. For C_{18} /APCI or ESI+ analysis, the initial mobile phase

conditions consisted of 5% A, 15% B, and 80% C for four min. This was followed by a 10 min linear gradient to 5% A and 95% B, which was then held for 6 min. For C₁₈/APCI or ESI-analysis, the initial conditions were 80% A, 15% B, and 5% C for 4 min, increased to 95% B and 5% C for 10 min, which was then held for 6 min.

Anion Exchange (AE) chromatography: Hamilton PRP-X110 100 × 2.1 mm (79743) columns were used for AE chromatography. For AE/ESI+ and ESI- analysis, the initial mobile phase conditions consisted of 5% A, 50% B, and 45% C held for 2 min, increased to 50% A, and 50% B for min and held for 13 min.

Hydrophilic Interaction Liquid Chromatography (HILIC): Supelco Ascentis Express HILIC 100 × 2.1 mm mm (53939-U) columns were used for HILIC chromatography. For HILIC/ESI+ analysis, the initial mobile phase conditions consisted of 8% A, 90% B, and 2% C, held for 4 min, increased to 50% A, 45% B, and 5% C for 10 min, and held for 6 min. For HILIC/ESI-analysis, the initial mobile phase conditions consisted of 98% B, and 2% C, held for 4 min, increased to 5% A, 45% B, and 50% C for 10 min and held for 6 min.

Ultra-high resolution mass spectrometry

Mass spectral detection was performed with a Thermo Scientific Q-Exactive HF mass spectrometer in continuous full scan mode at 70,000 resolution (scan range 85-1,275 m/z for all analyses other than AE, AE scan range was 100-1,500 m/z). This difference in mass range resulted in approximately 200 fewer features detected by omission of 85-100 m/z , and 500 more features detected due to inclusion of 1,275 to 1,500 m/z . Ion source conditions were optimized for both negative and positive ionization through systematic variation of different operational parameters to maximize the signal intensity of a representative chemical mixture infused into the source at appropriate mobile phase flow rate and composition. The automatic gain control

(AGC) target was set at 10^6 with a maximum ion injection time at 200 ms. Positive mode conditions were: spray voltage, 4500 V; capillary temperature, 275°C. Negative mode conditions were: spray voltage 3200 V; capillary temperature, 320°C. For both modes, sheath and auxiliary gas flow rates were maintained at 45 and 5 (arbitrary units), respectively. The S-Lens RF level was set at 69 for both negative and positive mode.

Data extraction and analysis

Instrument .RAW files were converted to .CDF format and mass spectral features were extracted and aligned using apLCMS (2; 4) with modifications by xMSanalyzer (3). The apLCMS software includes baseline subtraction, noise filtering (based upon a feature being present in at least 10% of spectra), retention time alignment (30 second maximum drift allowed) (2). xMSanalyzer was used with default parameter settings, and all samples met quality control criteria for mass alignment of internal standards, total feature detection and reproducibility of replicates. To further ensure analytical reproducibility and minimize measurement variability (assessed by ion technical replicate CV), a feature must be detected on at least two out of three technical replicates and features with greater than median 50% CV for technical replicates were removed from subsequent analyses. Due to the desire to identify low abundance chemicals, which may not be present in every individual, we did not select features based on a threshold fractional detection. To estimate the number of chemicals detected in a single analysis, we performed tentative metabolite annotation by using the *feat.batch.annotation.KEGG* function in xMSanalyzer using an m/z search tolerance of ± 10 ppm and “M+H”, “M+Na”, “M+K”, “M-H₂O+H”, “M+ACN+H”, “M+ACN+Na”, “M+2Na-H” adducts for positive mode; “M-H”, “M+Cl”, “M-H₂O-H”, and “M+Na-2H” adducts for negative mode. The 10 ppm window is based on previous studies showing that even though most ions are within 5 ppm mass accuracy,

the apLCMS data extraction algorithm can result in greater differences between observed and exact mass (2; 4; 6). We also annotated metabolites with a conservative approach by only searching for “M+H” ions in positive mode or “M-H” ions in negative mode. In both cases, we minimized redundancies by eliminating duplicate KEGG Compound IDs. We mapped unique chemical matches onto metabolic pathways using the KEGG mapper tool accessed at http://www.genome.jp/kegg/tool/map_pathway2.html. The online *BioVenn* tool (<http://www.cmbi.ru.nl/cdd/biovenn/>) was used to compare unique and overlapping chemical matches and ions detected by different HRM strategies (47).

2.3 Results

Characteristics of HRM data

Figure 2.1 contains histograms of feature triplicate median coefficient of variation (CV), distribution of m/z as a function of retention time (RT), and ion detection as a function of RT for each mode of chromatography and ionization. All platforms except $C_{18}/APCI+$ had median CV < 30%, equivalent to Standard Error of the Mean (SEM) < 17.4%, indicating large numbers of ions are detected with sufficient reproducibility for health evaluation (**Figure 2.1, left panels**). The distribution of m/z as a function of RT provides a way to visualize elution profiles. These results (**Figure 2.1, middle panels**) show that metabolites are differentially retained depending on the choice of AE, C_{18} , or HILIC chromatography and that ion elution profiles are similar for positive and negative ionization modes. Histograms of ions detected as a function of RT (**Figure 2.1, right panels**) show that maximal ion detection occurs during the initial wash-through volume, likely due to the salt content of plasma. For both C_{18} and HILIC, subsequent peaks were also consistent with the known mixture of hydrophobic and hydrophilic metabolites in plasma.

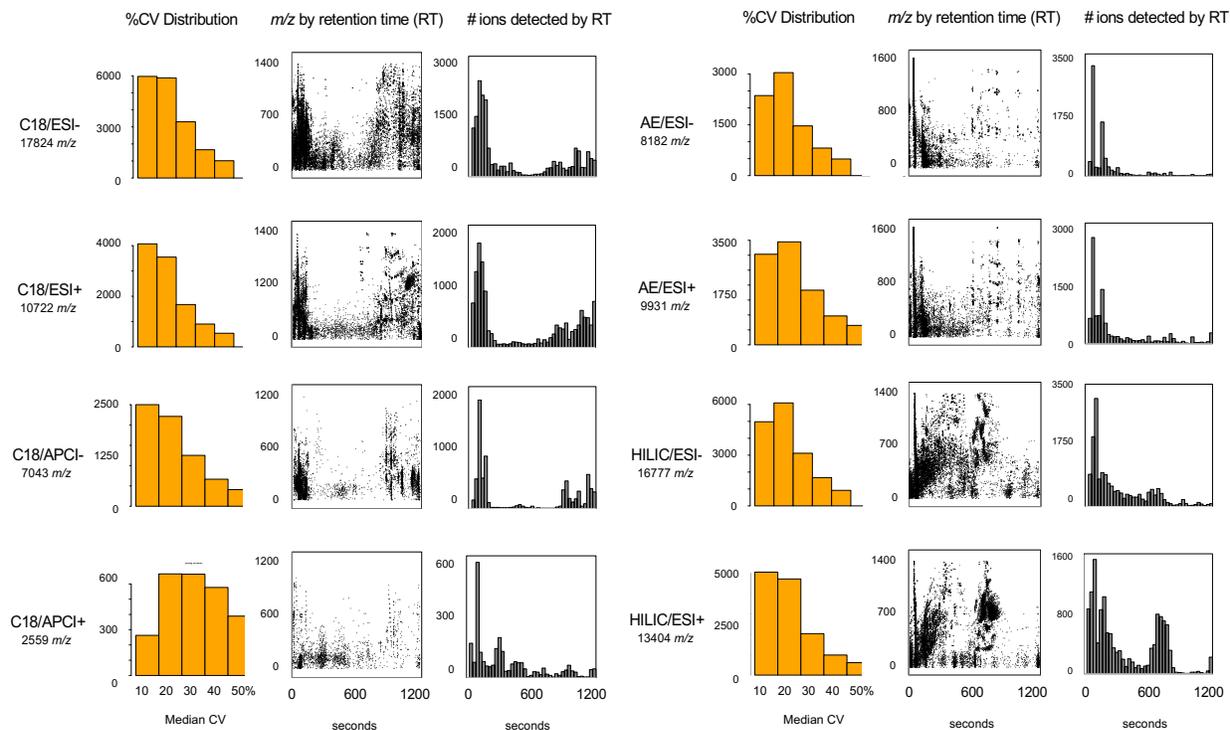


Figure 2.1 High-resolution metabolomics QC showing most ions are reproducibly detected with median CV% less than 17.4%.

Total number of m/z features detected and KEGG matches for individual HRM platform

The total number of ions detected for each analysis platform varied from 17,824 for C₁₈/ESI- to 2,559 for C₁₈/APCI+ (**Figure 2.2a**). Analytical configurations using ESI detected more ions than configurations using APCI. Negative ionization resulted in a greater number of detected ions than the comparable analyses with positive ionization, except for AE: HILIC/ESI- (16,777) versus HILIC/ESI+ (13,404); C₁₈/ESI- (17,824) versus C₁₈/ESI+ (10,722); C₁₈/APCI- (7,043) versus C₁₈/APCI+ (2,559); AE/ESI+ (9,931) versus AE/ESI- (8,182). To estimate the number of ions which possibly represent known metabolites, we searched the KEGG database for chemical matches within 10 ppm mass accuracy corresponding to the adduct forms described above. The results showed that only the configurations with relatively lower number of detected ions had more than 50% database matches, suggesting that many detected ions could be derived from uncharacterized metabolites.

To estimate the coverage of known metabolites provided by each configuration, we calculated the percentage of the KEGG chemical database (17,554 total chemicals) that could be matched by a single method. The results (**Figure 2.2B**) showed that the best platforms (HILIC/ESI+ and AE/ESI+) could detect up to 58% of the KEGG chemical database. In combination with the results in **Figure 2.2A**, indicating the majority of detected ions are not matched to chemicals in KEGG, these results suggest the KEGG database is incomplete in coverage of the human metabolome and that a single analytic platform may capture only up to 60% of known human metabolites.

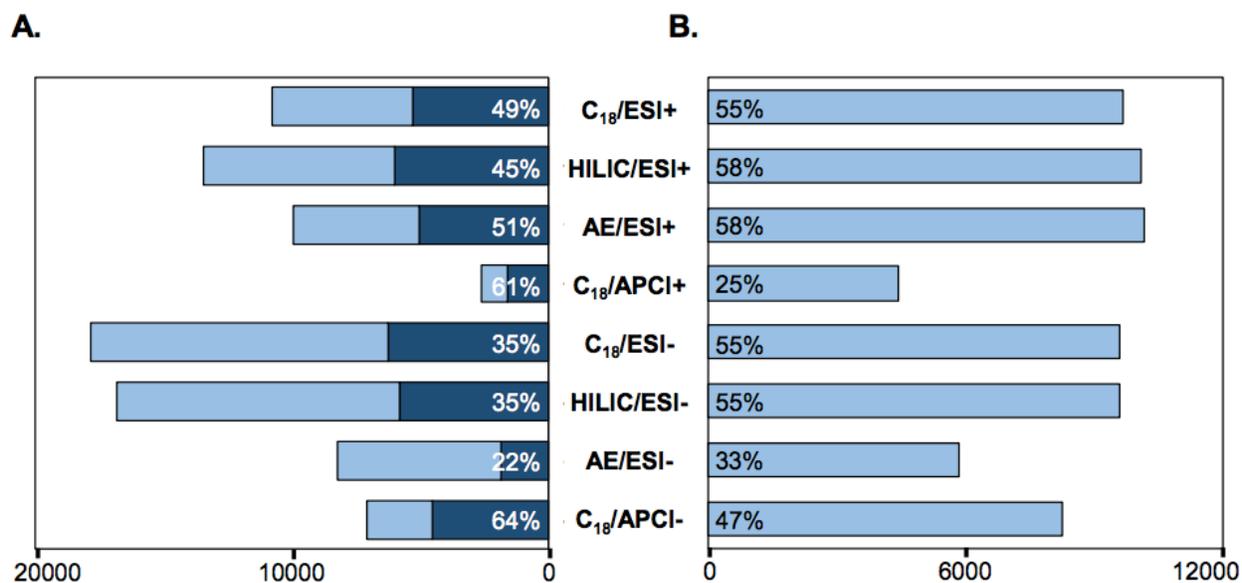


Figure 2.2 A. Total number of ions detected with percentage of ions matching at least one chemical in KEGG represented by dark portion of the bar. B. Number of chemical matches, with percentage of total database (17,554 chemicals) matched.

Dual HRM platforms increase metabolic coverage

To determine the extent that dual HRM strategies could improve the number of database matches, we examined pairwise combinations of ionizations and chromatography. Similar total ion count was obtained with any paired combination of positive and negative ESI using either HILIC or C₁₈ chromatography. To determine combinations that provided the maximum number of database matches, we examined paired LC and ionization configurations using number of non-overlapping and overlapping matches to the KEGG database. The pair consisting of HILIC/ESI+ and C₁₈/ESI- analysis had the highest number of chemical matches (12,712), which included matches to 72% of the chemicals present in the KEGG database (**Figure 2.3**). A number of dual HRM strategies achieved greater than 70% coverage of the KEGG database. For example, using HILIC chromatography with dual polarity (ESI+ and ESI-) provided 12,550 matches. Using C₁₈ chromatography with ESI+ and ESI- resulted in a total of 12,454 matches. The combination of C₁₈/ESI+ with HILIC/ESI- matched to 12,271 chemicals in KEGG. Thus, either of these dual HRM platforms appears to provide an effective way to increase the number of database matches by approximately 25% over individual HRM strategies.

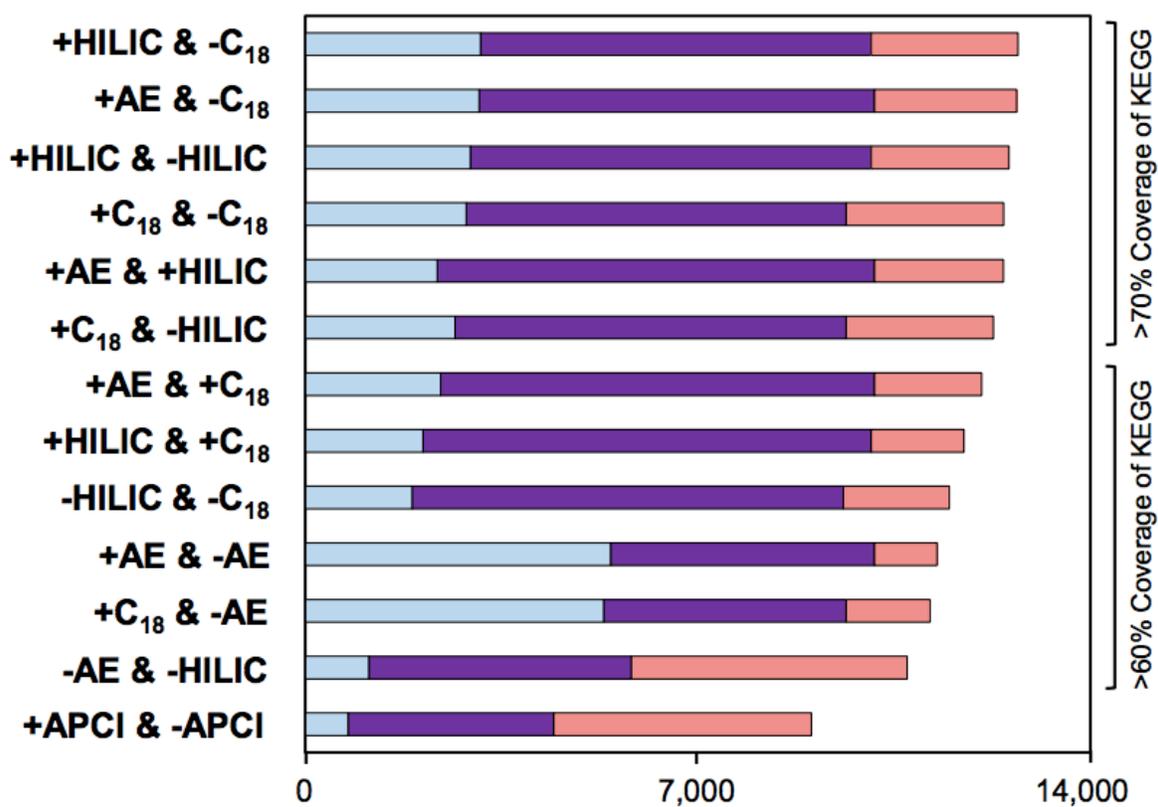


Figure 2.3 Dual HRM platforms increase number of database matches to 72% of the KEGG chemical database. Matches obtained from left HRM platform (blue), right HRM platform (red), and both HRM methods (purple).

The use of ESI/APCI or C18/HILIC/AE increases ion detection for HRM

To determine which pair of columns provided maximal number of ions with matches to the KEGG database, we compared the effect of chromatography on the detection of ions matching at least one chemical with comparisons for positive and negative ionization completed separately (**Figure 2.4**). Direct comparisons of ions obtained in negative mode with ions obtained in positive mode are not possible unless molecular ions for chemicals are identified. Therefore, this analysis is to compare the effect of chromatography on the ability to detect unique ions that could match chemicals in KEGG. The use of different columns resulted in a relative increase in non-redundant ions, i.e., likely increasing the total number of chemicals detected. C₁₈/ESI⁻ detected 3,642 ions with database matches and HILIC/ESI⁻ detected 4,578 ions with database matches, with only 975 of these matches being the same for the two platforms. Similar increases in the number of ions with at least one database match were observed when any two chromatographic strategies were used. Thus, the data show that in a dual platform analysis, the use of two different column types improves chemical coverage.

In comparison to the large overlap in the number of chemical matches obtained for HILIC/ESI⁺ and C₁₈/ESI⁺ (68% of ions have m/z within 10 ppm), only 13% of ions with database match for at least one chemical were the same for HILIC/ESI⁺ and C₁₈/ESI⁺. These results show that detected ions with chemical matches are different depending on the HRM configuration, and that employing multiple analytical approaches can increase the detection of ions with chemical matches. Furthermore, this data suggests that the chemicals detected by both platforms were high abundance, present in multiple adduct forms and potentially suitable for internal cross-validation within a dual-chromatography protocol (6). Support for this concept

was provided by comparison of the signal for tyrosine in the HILIC/ESI+ and C₁₈/ESI+ comparison (**Figure 2.5**).

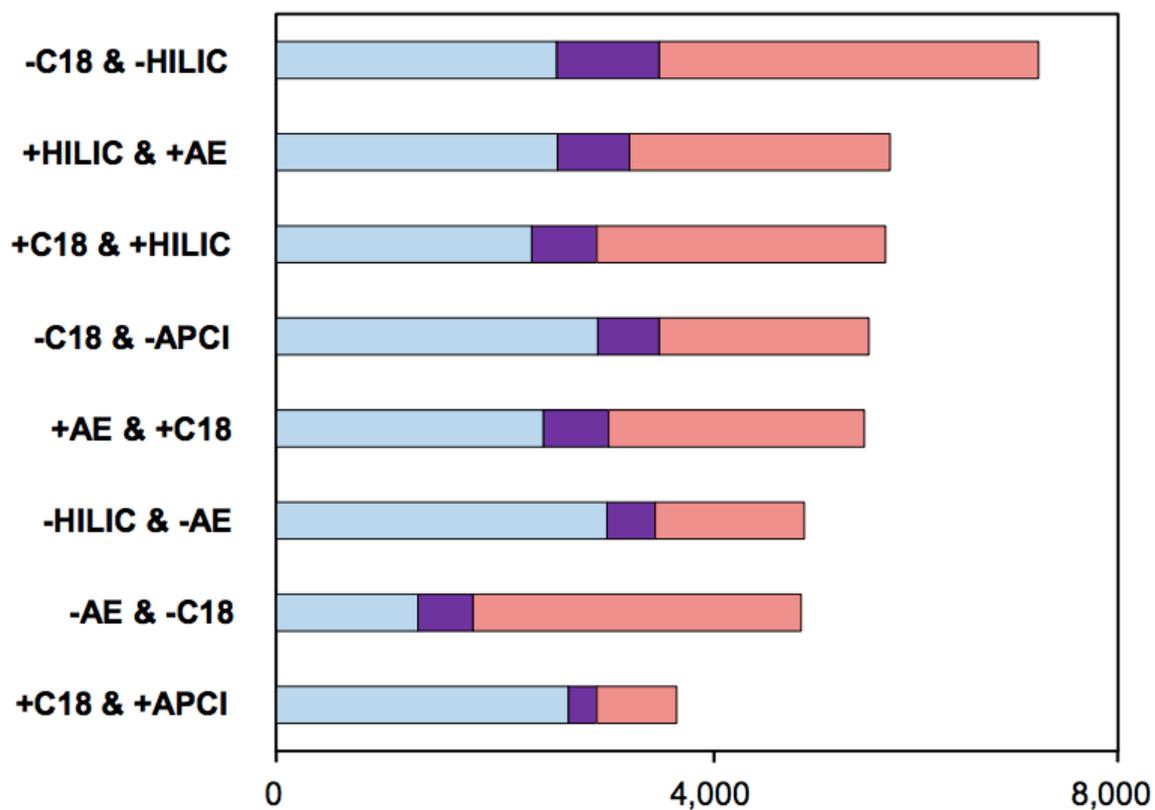


Figure 2.4 Number of ions with database matches from paired HRM platforms. Ions detected by left HRM platform (blue), right HRM platform (red), and both HRM methods (purple).

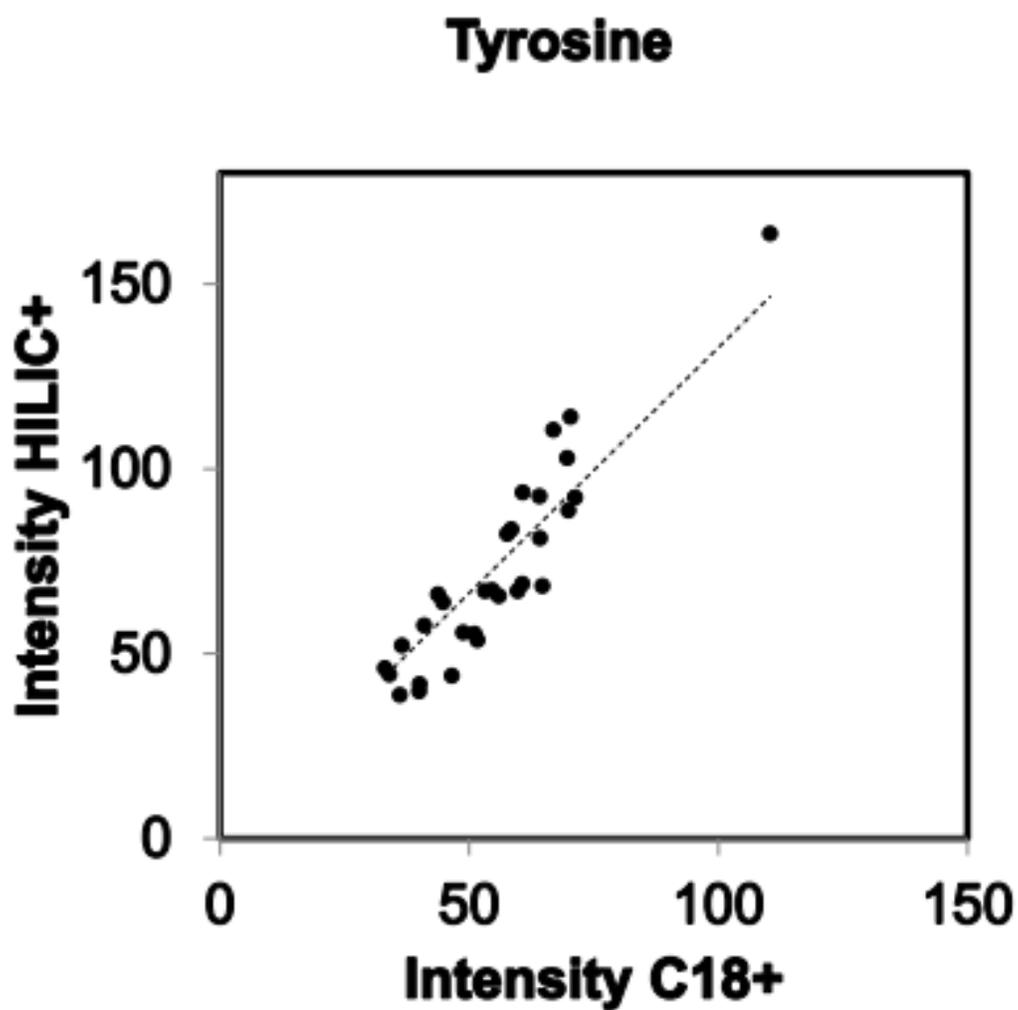


Figure 2.5 Comparison of ion intensity for tyrosine detection by HILIC/ESI+ and C18/ESI+. Each dot represents the normalized peak intensity for tyrosine measured in the same 30 samples across two different analytical platforms.

HRM detects endogenous and exogenous chemicals

Combining all eight HRM platforms results in 14,387 different database matches, providing up to 82% coverage of all metabolites in KEGG. KEGG matches were then classified based on their functional KEGG BRITE categorization (**Figure 2.6**). Approximately half of KEGG matches did not have KEGG classifications. Metabolites with roles in intermediary metabolism are termed “compounds with biological roles” and represented 3.4% of all KEGG matches while lipids represented a larger fraction, 12.3%. The remaining matches were for chemicals derived from exogenous sources (pesticides, carcinogens, pharmaceuticals, phytochemical compounds, endocrine disrupting compounds, natural toxins, and metabolites derived from natural products). These results show that diverse classes of chemicals could be detected in a single analysis, including those derived from environmental or occupational exposures.

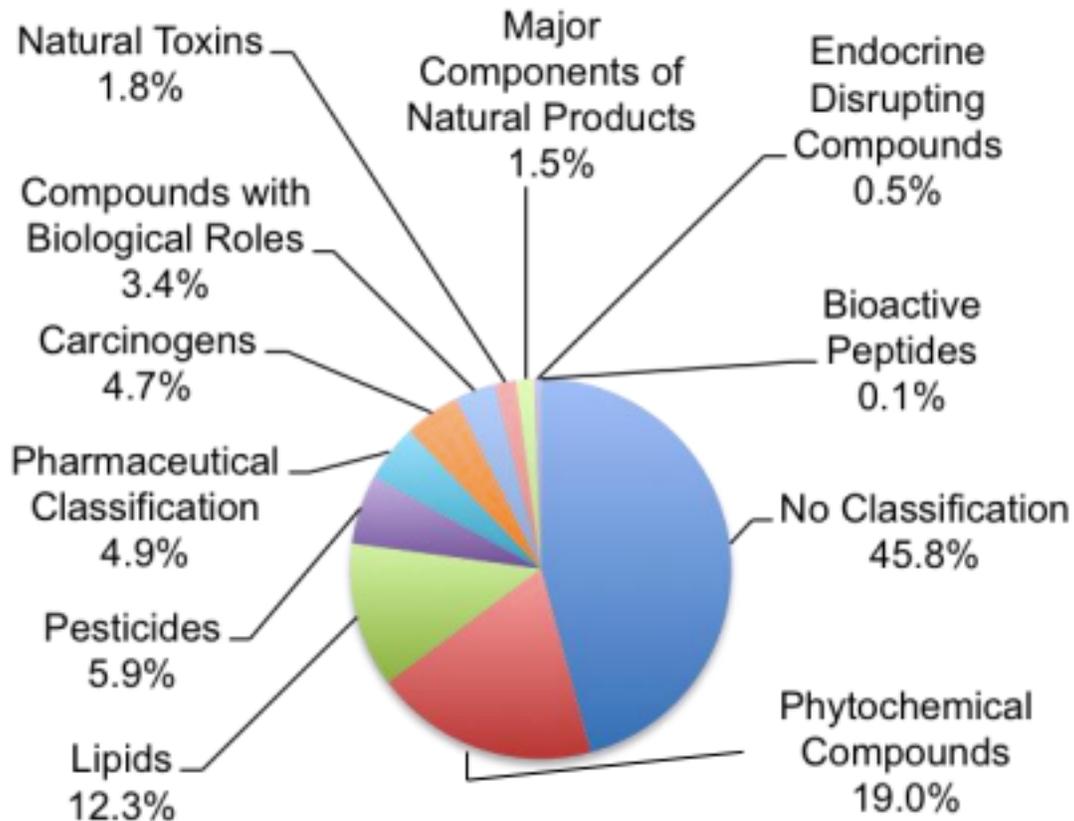


Figure 2.6 KEGG BRITE categorization of accurate mass matches (MS1) for detected features by HRM. HRM detects accurate mass matches to metabolites from a broad range of chemical classes. However, approximately half of detected metabolites are not classified in KEGG.

Pathway Coverage

To determine the coverage of metabolic pathways provided by dual chromatography, we mapped the database matches to KEGG human metabolic pathways. Combining C₁₈/ESI- with HILIC/ESI+ (**Figure 2.7**) analysis provided similar coverage (1246 database matches) of the human reference metabolic pathway in KEGG (hsa01100) as C₁₈/ESI- with C₁₈/ESI+ and HILIC/ESI- with HILIC/ESI+ (data not shown). In previous targeted MS/MS confirmation of database matches to ions detected by comparable methods (3; 6; 48), we found that 60% to 80% of these matches are correct identifications. Thus, the results show that a dual chromatography approach can provide a general assessment of metabolism, but additional confirmation is needed for conclusions concerning specific pathway effects in study populations.

We performed a comparable annotation using only ions matching [M+H] (HILIC/ESI+) or [M-H] (C₁₈/ESI-) adducts to determine the utility of this simpler annotation strategy. Results showed 757 matches to chemicals on hsa01100 were obtained, with less overlapped detection between platforms (**Figure 2.8**). Despite the lesser number of matches, this simpler strategy provided similar coverage of metabolic pathways. Although this more conservative strategy includes some incorrect matches, the analysis emphasizes that combined platforms provides substantial metabolic coverage.

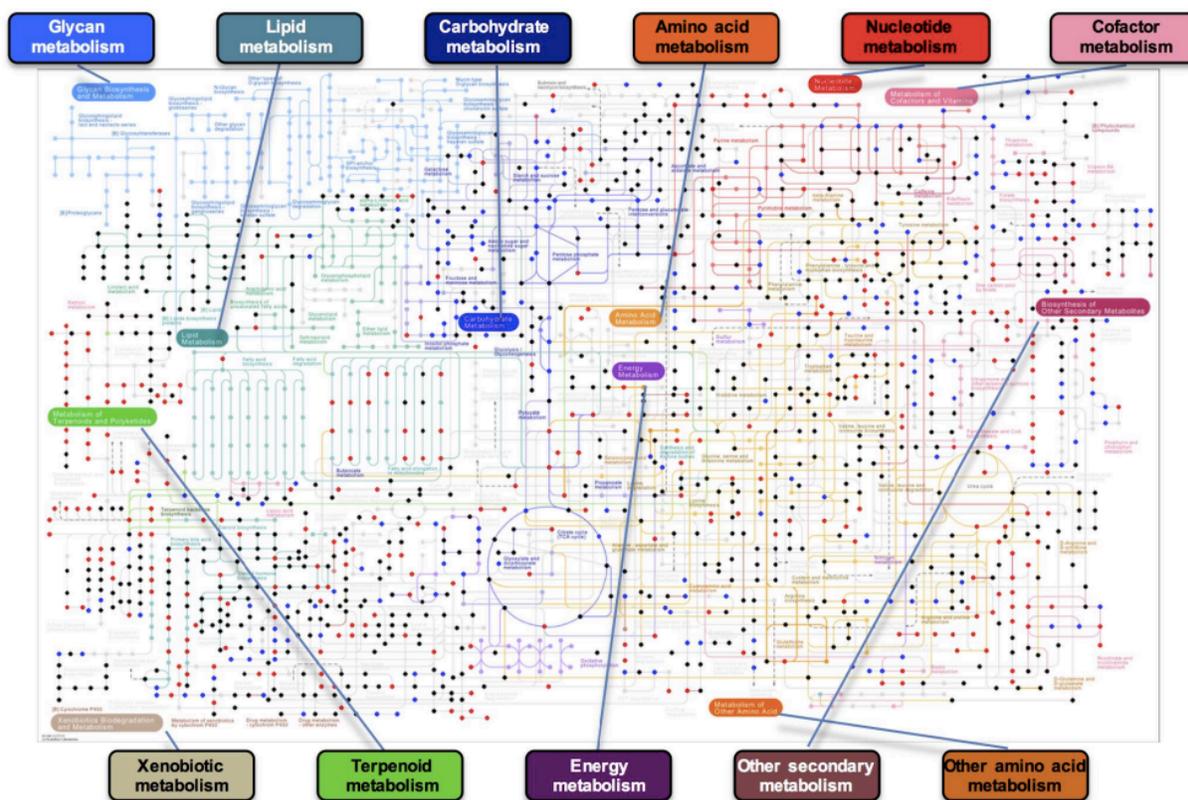


Figure 2.7 Dual HRM using C18/ESI- (blue) with HILIC/ESI+ (red) provides in-depth coverage of human metabolic pathways using multiple adducts for chemical matching. Black dots indicate metabolites matched with both HRM methods. 1246 metabolites are matched with this strategy.

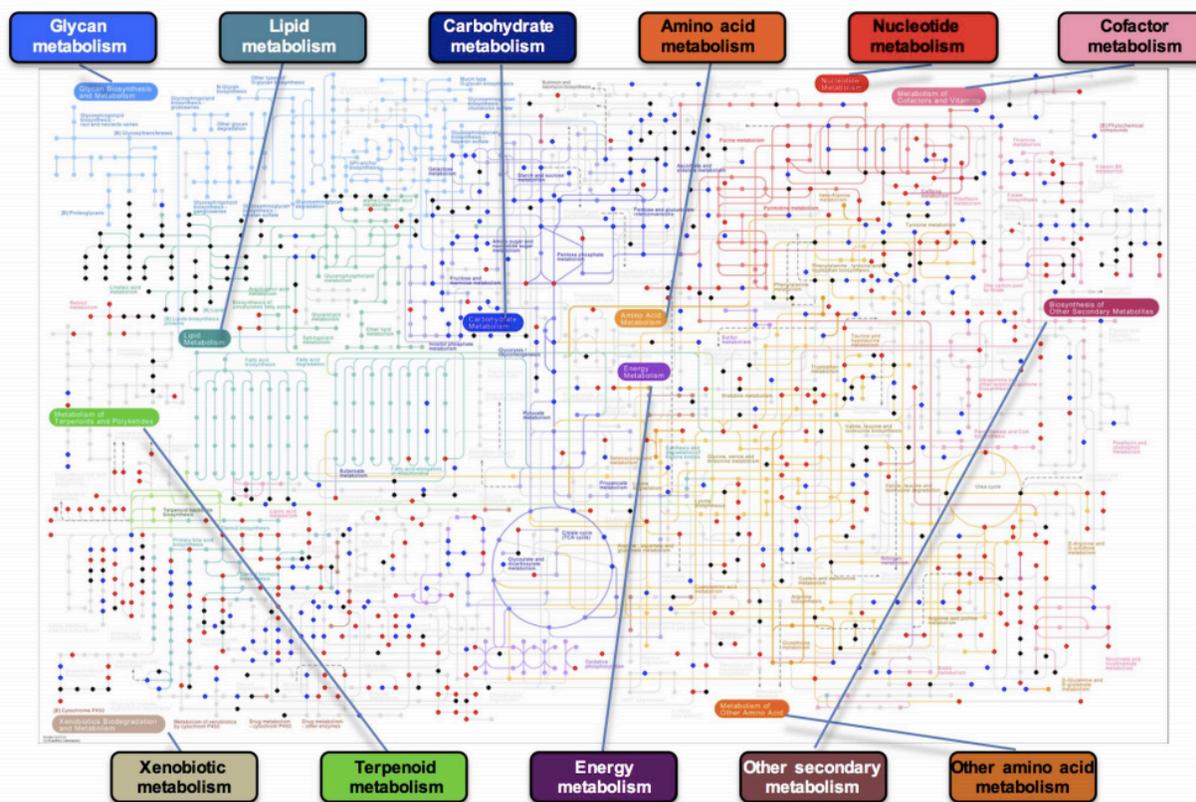


Figure 2.8 Dual HRM using C18/ESI- (blue) with HILIC/ESI+ (red) provides adequate coverage of human metabolic pathways using conservative chemical matching strategy. Black dots indicate metabolites matched with both HRM methods. 757 metabolites are matched.

Comparison of Targeted Chemical Detection

The information provided by the analyses above suggests that the combination of different polarities with different analytical columns provides optimal chemical coverage. To test this specifically, we assessed each HRM configuration for its ability to detect a list of metabolites with confirmed identities (**Figure 2.9**). [3,3-¹³C₂]-cystine, an internal standard, was detected in seven of the eight platforms. In contrast, cotinine, was only detected with positive ionization (177.1026 *m/z*) and fatty acids were better detected with negative ionization. The results show that inclusion of both polarities within a dual chromatography analysis improves the coverage of known metabolites.

Figure 8

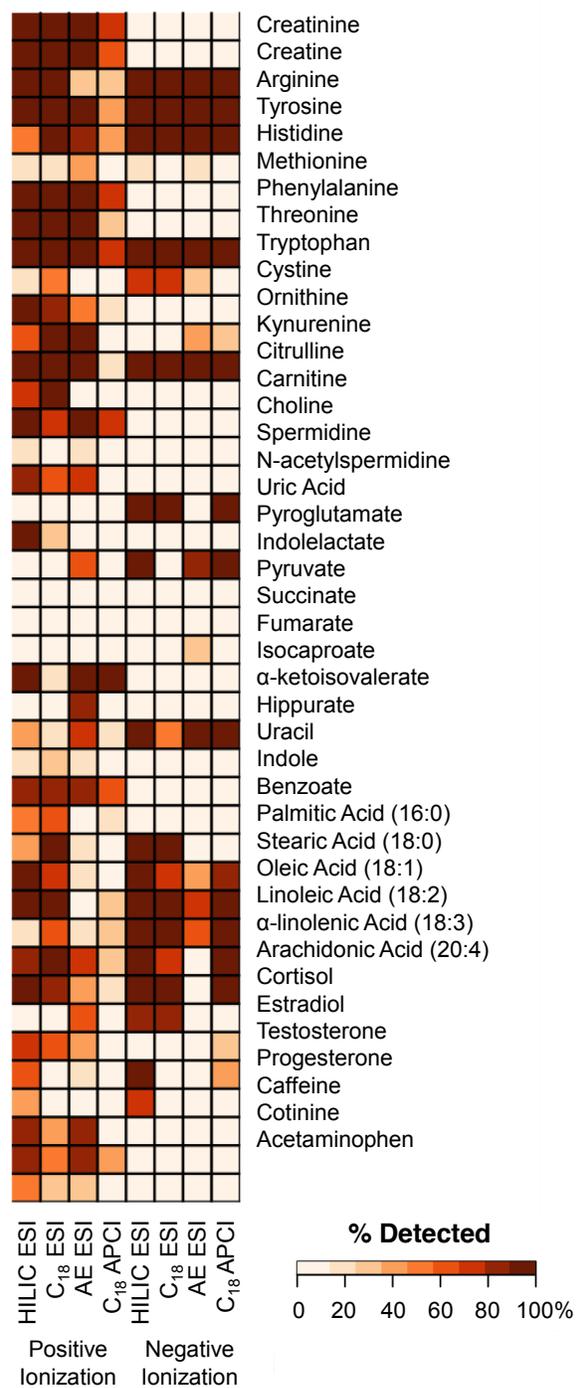


Figure 2.9 Representative chemical detection with eight HRM platforms.

2.4 Discussion

Management of risks associated with environmental exposures and their biologic effects depend upon tools to detect and predict adverse effects. Currently used clinical chemistry and toxicological analyses provide information about a limited number of markers, and there is an ongoing need to improve capabilities. HRM provides a strategy to measure thousands of known and unknown chemicals to enable routine, generalized assessment of exposures and biologic responses. These metabolic assessments can be used within an integrative framework for health evaluation in military personnel and support National Precision Medicine initiatives. In this study, we sought to identify HRM configurations that could increase ion detection and chemical database matches, and thereby increase the likelihood of detecting the maximal number of known and unknown chemicals. We evaluated different chromatographic and ionization strategies for analyzing serum samples from the DoDSR to identify optimal strategies for metabolic assessments of military personnel.

Chromatography facilitates the analysis of complex mixtures using mass spectrometry, as it separates potential interfering ions from ions of interest, improving quantification and limit of detection. Orthogonal dual chromatographic approaches prior to MS analysis, such as pairing lipophilic selectivity obtained with C₁₈ chromatography with polar selectivity obtained from either AE or HILIC analysis, have been applied in previous studies. In Soltow et al. (6), the use of dual AE and C₁₈ chromatography with positive ionization on a Thermo LTQ-FT mass spectrometer increased overall ion detection by 23-26%, yielding a total of up to 7,000 ions per sample (6). More recently, Ivanisevic et al. (26) performed a similar analysis of human plasma, comparing chemical detection with C₁₈ and HILIC using positive and negative electrospray ionization and obtained 9,709 ions with HILIC/ESI⁺ and 15,263 ions with C₁₈/ESI⁺. In negative

mode, they obtained 8,122 ions with HILIC/ESI- and 7,742 ions with C₁₈/ESI-. Our HRM platform yielded similar numbers of *m/z* features detected with HILIC /ESI+ (13,404 ions) and C₁₈/ESI+ (10,722) and with HILIC/ESI- (17,824) and C₁₈/ESI-: (16,777 ions). The Ivanisevic study also demonstrated that the combination of HILIC/ESI+ with C₁₈/ESI+ increased ions with [M+H] chemical matches by 28%, compared to using C₁₈/ESI+ alone. In negative mode, the combination of HILIC/ESI- with C₁₈/ESI- increased ions with [M-H] chemical matches by 43%, compared to using C₁₈/ESI- alone. Contrepolis et. al (49) also compared ions detected by HILIC and C₁₈ chromatography with positive and negative ESI, and found that the addition of HILIC chromatography to a C₁₈ metabolomics analysis resulted in a 68% increase in ions detected with ESI+, and a 148% increase in ESI-. In our study, we found that pairing HILIC chromatography with C₁₈ chromatography for HRM increased detected ions with chemical matches by 99% using ESI-, and a 90% increase using ESI+, compared to using C₁₈ chromatography alone. Our findings are consistent with other studies evaluating combined analytical strategies for metabolic profiling, showing that dual chromatographic approaches increase the number of ions detected with non-redundant chemical matches.

Metabolites may preferentially form positive or negative ions for detection, and may ionize more efficiently using either ESI or APCI. Because producing gas phase ions from metabolites is critical for mass spectral detection, we also examined the use of different ionization strategies (ESI vs. APCI, ESI+ vs. ESI-) for increasing potential chemical coverage for HRM. Previous studies have performed similar comparisons for metabolomics. Nordström et al. showed that pairing ESI+ with ESI- increased unique, non-redundant ion detection by 90% compared to using ESI+ alone, and that the use of APCI increases unique ion detection by 20% compared to ESI alone (30). Ivanisevic et al. showed that the combination of HILIC with ESI-

and C₁₈ with ESI⁺ provided the maximum amount of biological information relating to lipid and central carbon metabolism (26). Contrepois et al. also compared HILIC and C₁₈ chromatography with positive and negative ESI and demonstrated the use of dual chromatography with positive and negative ionization increased chemical detected (49). In our study, we show similar increases in chemical coverage when using positive and negative ionization for HRM. In contrast to our extraction with a volume of 2:1 acetonitrile to plasma and analysis with an acidic pH, Ivanisevic et al. extracted with a volume of 1:1 methanol/acetonitrile to plasma and the chromatographic gradients were buffered at basic pH; a direct comparison of extraction and solvent conditions, as well as column stability, will be needed to determine which conditions are most suitable for routine use.

The number of database matches provides an estimate of metabolic coverage. As indicated above, targeted MS/MS confirmation of database matches in our previous studies (3; 6; 48) showed that between 60% and 80% of matches are correct identifications. Other limitations to interpretation, namely that a single chemical can give rise to multiple adducts and that multiple chemicals have the same mass, were addressed in the experimental design. Specifically, we estimated the fraction of ions with putative identification by counting all ions with matches, and we eliminated redundancies due to multiple adducts of the same chemical by counting database matches in terms of the number of KEGG IDs. The limitations were also addressed by analysis of a subset of metabolites with confirmed identities (Fig 8). Thus, despite the limitations of the approach, the results provide rough estimates that about half of the KEGG database can be captured in routine analyses and that about half of the ions detected are present in the KEGG database.

Analysis of matches only using H⁺ and H⁻ forms shows that even with this simplification, an effective coverage of nearly 1000 metabolites is obtained. Confirmation of these metabolites by MS/MS for H⁺ and H⁻ forms is straightforward and, along with reference standardization (8), could provide a way to quantify up to 1000 metabolites in a routine and affordable manner. Furthermore, this approach could allow integration of data from ultra-high resolution instruments and data from Q-TOF and other instruments so that data is readily interchangeable between different laboratories.

Contrepois et. al noted that while chemicals may be detected by multiple analytical configurations, peak quality may vary based on differences in ionization or chromatography. We noted in our study that different HRM strategies resulted in differential detection of representative chemicals. Overall, our study results are in agreement with previous studies and others (22; 25; 31; 40; 50-53), as we conclude that employing dual HRM with positive and negative ionization and orthogonal chromatography increases chemical coverage of human serum metabolites and that the quality and quantification of detected chemicals may vary depending on HRM conditions. Increased chemical coverage facilitates the use of non-targeted metabolomics pathway enrichment software i.e. *mummichog*, which performs metabolic pathway enrichment from ranked spectral features (18).

Limitations of the Study

The computational metabolomics methods used in the present study provide estimates of coverage of metabolism and detection of dietary and environmental chemicals based upon accurate mass matches to chemicals in the KEGG database. A limitation is that these analyses do not provide absolute chemical identity for most of the ions detected. Confirmation of chemical identities with tandem mass spectral (MS/MS) analysis and co-elution with authentic

standards has been performed for several hundred chemicals in other studies using these methods (9-13; 15; 17; 21; 48; 54)) and provide confidence that the conclusions reached are valid, even though many individual ion matches are incorrect.

In addition to confirming chemicals with tandem mass spectrometry, future studies can explore combining NMR/GCMS technologies to increase chemical detection (such as performed by Psychogios et. al (42)), optimization of analytical gradients or the use of ion mobility MS to increase chemical coverage. Additionally, application of clustering algorithms to identify isomers and adducts arising from a single chemical are likely to decrease chemical noise. Application of complementary analytical methodologies have greatly increased the confidence of chemical matches and reduced chemical noise and artefacts from MS analysis (24; 49; 55; 56).

Conclusion

This study compared chemical coverage obtained with eight different HRM platforms. The results show that orthogonal chromatography and polarity for ionization (e.g., HILIC/ESI+ and C₁₈/ESI-) provides the best dual chromatography platform for metabolic coverage of serum metabolites and also includes an extensive number of matches to dietary and environmental chemicals. The results show that metabolites detected by multiple platforms can be used for internal validation and that little pathway information is lost by restricting analysis only to adducts formed by gain or loss of H⁺. The results support the use of DoDSR serum samples for metabolic assessments of military service personnel. Furthermore, these results show the general applicability of HRM for detection of known and unidentified metabolites. Analysis of standard chemical libraries coupled to collection of tandem MS/MS spectra

References

1. Jones DP. 2016. Sequencing the exposome: A call to action. *Toxicol Rep* 3:29-45
2. Yu T, Jones DP. 2014. Improving peak detection in high-resolution LC/MS metabolomics data using preexisting knowledge and machine learning approach. *Bioinformatics* 30:2941-8
3. Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, et al. 2013. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics* 14:15
4. Yu T, Park Y, Johnson JM, Jones DP. 2009. apLCMS--adaptive processing of high-resolution LC/MS data. *Bioinformatics* 25:1930-6
5. Park YH, Lee K, Soltow QA, Strobel FH, Brigham KL, et al. 2012. High-performance metabolic profiling of plasma from seven mammalian species for simultaneous environmental chemical surveillance and bioeffect monitoring. *Toxicology* 295:47-55
6. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. 2011. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. *Metabolomics* 9:132-43
7. Go YMW, D.; Yongliang, L.; Soltow, Q.; Tran, V.; Jones, D.P. et al. 2015. Quantitative High-Resolution Metabolomics for Exposome Research. *Toxicol Sci* Accepted September 2015
8. Go YM, Walker DI, Liang Y, Uppal K, Soltow QA, et al. Reference Standardization for Mass Spectrometry and High-resolution Metabolomics Applications to Exposome Research.
9. Roede JR, Uppal K, Park Y, Lee K, Tran V, et al. 2013. Serum metabolomics of slow vs. rapid motor progression Parkinson's disease: a pilot study. *PLoS One* 8:e77629
10. Frediani JK, Jones DP, Tukvadze N, Uppal K, Sanikidze E, et al. 2014. Plasma metabolomics in human pulmonary tuberculosis disease: a pilot study. *PLoS One* 9:e108854
11. Cribbs SK, Park Y, Guidot DM, Martin GS, Brown LA, et al. 2014. Metabolomics of bronchoalveolar lavage differentiate healthy HIV-1-infected subjects from controls. *AIDS Res Hum Retroviruses* 30:579-85
12. Osborn MP, Park Y, Parks MB, Burgess LG, Uppal K, et al. 2013. Metabolome-wide association study of neovascular age-related macular degeneration. *PLoS One* 8:e72737
13. Neujahr DC, Uppal K, Force SD, Fernandez F, Lawrence C, et al. 2014. Bile acid aspiration associated with lung chemical profile linked to other biomarkers of injury after lung transplantation. *Am J Transplant* 14:841-8
14. Obiany O, Liang Y, Burnham EL, Mehta A, Park Y, et al. 2015. Metabolic Consequences of Chronic Alcohol Abuse in Non-Smokers: A Pilot Study. *PLoS One* 10:e0129570
15. Go YM, Roede JR, Orr M, Liang Y, Jones DP. 2014. Integrated redox proteomics and metabolomics of mitochondria to identify mechanisms of Cd toxicity. *Toxicol Sci* 139:59-73
16. Uppal K, Soltow QA, Promislow DEL, Wachtman LM, Quyyumi AA, Jones DP. 2015. MetabNet: an R package for metabolic association analysis of high-resolution metabolomics data. *Frontiers in Bioengineering and Biotechnology* 3

17. Go YM, Walker DI, Soltow QA, Uppal K, Wachtman LM, et al. 2015. Metabolome-wide association study of phenylalanine in plasma of common marmosets. *Amino Acids* 47:589-601
18. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, et al. 2013. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol* 9:e1003123
19. Wishart DS, Jewison T, Guo AC, Wilson M, Wilson M, Knox C, Liu Y, et al. 2013. HMDB 3.0--The Human Metabolome Database in 2013.
20. Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, et al. 2011. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc* 6:1060-83
21. Go Y-M, Walker DI, Liang Y, Uppal K, Soltow QA, et al. 2015. Reference Standardization for Mass Spectrometry and High-Resolution Metabolomics Applications to Exposome Research. *Toxicological Sciences*
22. Cruickshank-Quinn C, Quinn KD, Powell R, Yang Y, Armstrong M, et al. 2014. Multi-step preparation technique to recover multiple metabolite compound classes for in-depth and informative metabolomic analysis. *Journal of visualized experiments : JoVE*
23. Kirkwood JS, Maier C, Stevens JF. 2013. Simultaneous, untargeted metabolic profiling of polar and nonpolar metabolites by LC-Q-TOF mass spectrometry. *Curr Protoc Toxicol* Chapter 4:Unit4 39
24. Chalcraft KR, McCarry BE. 2013. Tandem LC columns for the simultaneous retention of polar and nonpolar molecules in comprehensive metabolomics analysis. *J Sep Sci* 36:3478-85
25. Yang Y, Cruickshank C, Armstrong M, Mahaffey S, Reisdorph R, Reisdorph N. 2013. New sample preparation approach for mass spectrometry-based profiling of plasma results in improved coverage of metabolome. *J Chromatogr A* 1300:217-26
26. Ivanisevic J, Zhu ZJ, Plate L, Tautenhahn R, Chen S, et al. 2013. Toward 'omic scale metabolite profiling: a dual separation-mass spectrometry approach for coverage of lipid and central carbon metabolism. *Anal Chem* 85:6876-84
27. Patti GJ. 2011. Separation strategies for untargeted metabolomics. *J Sep Sci* 34:3460-9
28. Büscher JM, Czernik D, Ewald JC, Sauer U, Zamboni N. 2009. Cross-Platform Comparison of Methods for Quantitative Metabolomics of Primary Metabolism. *Analytical Chemistry* 81:2135-43
29. Lu W, Bennett BD, Rabinowitz JD. 2008. Analytical strategies for LC-MS-based targeted metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci* 871:236-42
30. Nordström A, Want E, Northen T, Lehtiö J, Siuzdak G. 2008. Multiple Ionization Mass Spectrometry Strategy Used To Reveal the Complexity of Metabolomics. *Analytical Chemistry* 80:421-9
31. Theodoridis G, Gika HG, Wilson ID. 2008. LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. *TrAC Trends in Analytical Chemistry* 27:251-60
32. Psychogios N, Hau DD, Peng J, Guo AC, Mandal R, et al. 2011. The Human Serum Metabolome. *PLoS ONE* 6:e16957
33. Dunn WB, Lin W, Broadhurst D, Begley P, Brown M, et al. 2015. Molecular phenotyping of a UK population: defining the human serum metabolome. *Metabolomics* 11:9-26

34. Dunn WB, Ellis DI. 2005. Metabolomics: Current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry* 24:285-94
35. Peterman SM, Duczak N, Jr., Kalgutkar AS, Lame ME, Soglia JR. 2006. Application of a linear ion trap/orbitrap mass spectrometer in metabolite characterization studies: examination of the human liver microsomal metabolism of the non-tricyclic anti-depressant nefazodone using data-dependent accurate mass measurements. *J Am Soc Mass Spectrom* 17:363-75
36. Dunn WB, Broadhurst D, Brown M, Baker PN, Redman CW, et al. 2008. Metabolic profiling of serum using Ultra Performance Liquid Chromatography and the LTQ-Orbitrap mass spectrometry system. *J Chromatogr B Analyt Technol Biomed Life Sci* 871:288-98
37. Hu Q, Noll Rj Fau - Li H, Li H Fau - Makarov A, Makarov A Fau - Hardman M, Hardman M Fau - Graham Cooks R, Graham Cooks R. The Orbitrap: a new mass spectrometer.
38. Piraud M, Vianey-Saban C, Petritis K, Elfakir C, Steghens JP, et al. 2003. ESI-MS/MS analysis of underivatized amino acids: a new tool for the diagnosis of inherited disorders of amino acid metabolism. Fragmentation study of 79 molecules of biological interest in positive and negative ionisation mode. *Rapid Commun Mass Spectrom* 17:1297-311
39. Yanes O, Tautenhahn R, Patti GJ, Siuzdak G. 2011. Expanding coverage of the metabolome for global metabolite profiling. *Anal Chem* 83:2152-61
40. Want EJ, Masson P, Michopoulos F, Wilson ID, Theodoridis G, et al. 2013. Global metabolic profiling of animal and human tissues via UPLC-MS. *Nat Protoc* 8:17-32
41. Patti GJ, Tautenhahn R, Siuzdak G. 2012. Meta-analysis of untargeted metabolomic data from multiple profiling experiments. *Nat Protoc* 7:508-16
42. Psychogios NH, D.D.; Peng, J.; Guo, A.C.; Mandal, R.; Bouatra, S.; Sinelnikov, I.; Krishnamurthy, R.; Eisner, R.; Gautam, B.; Young, N.; Xia, J.; Knox, C.; Dong, E.; Huang, P.; Hollander, Z.; Pedersen, T.L.; Smith, S.R.; Bamforth, F.; Greiner, R.; McManus, B.; Newman, J.W.; Goodfriend, T.; Wishart, D.S. 2011. The Human Serum Metabolome. *PLoS One* 6:1-23
43. Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 28:27-30
44. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42:D199-205
45. Perdue CL, Cost AAE, Rubertone MV, Lindler LE, Ludwig SL. 2015. Description and Utilization of the United States Department of Defense Serum Repository: A Review of Published Studies, 1985-2012. *PLoS ONE* 10:e0114857
46. Simón-Manso Y, Lowenthal MS, Kilpatrick LE, Sampson ML, Telu KH, et al. 2013. Metabolite Profiling of a NIST Standard Reference Material for Human Plasma (SRM 1950): GC-MS, LC-MS, NMR, and Clinical Laboratory Analyses, Libraries, and Web-Based Resources. *Analytical Chemistry* 85:11725-31
47. Hulsen T, Vlieg J, Alkema W. 2008. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9:1-6
48. Johnson JM, Yu T, Strobel FH, Jones DP. 2010. A practical approach to detect unique metabolic patterns for personalized medicine. *Analyst* 135:2864-70

49. Contrepois K, Jiang L, Snyder M. 2015. Optimized Analytical Procedures for the Untargeted Metabolomic Profiling of Human Urine and Plasma by Combining Hydrophilic Interaction (HILIC) and Reverse-Phase Liquid Chromatography (RPLC)–Mass Spectrometry. *Molecular & Cellular Proteomics : MCP* 14:1684-95
50. Want EJ, Wilson ID, Gika H, Theodoridis G, Plumb RS, et al. 2010. Global metabolic profiling procedures for urine using UPLC-MS. *Nat Protoc* 5:1005-18
51. Yuan M, Breitkopf SB, Yang X, Asara JM. 2012. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nat Protoc* 7:872-81
52. Scalbert A, Brennan L, Fiehn O, Hankemeier T, Kristal BS, et al. 2009. Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics* 5:435-58
53. Tian H, Bai J, An Z, Chen Y, Zhang R, et al. 2013. Plasma metabolome analysis by integrated ionization rapid-resolution liquid chromatography/tandem mass spectrometry. *Rapid Commun Mass Spectrom* 27:2071-80
54. Walker DI, Go Y-M, Liu K, Pennell KD, Jones DP. 2016. Chapter 7 - Population Screening for Biological and Environmental Properties of the Human Metabolic Phenotype: Implications for Personalized Medicine A2 - Holmes, Elaine. In *Metabolic Phenotyping in Personalized and Public Healthcare*, ed. JK Nicholson, AW Darzi, JC Lindon:167-211. Boston: Academic Press. Number of 167-211 pp.
55. Kuhl C, Tautenhahn R, Bottcher C, Larson TR, Neumann S. 2012. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal Chem* 84:283-9
56. O'Hagan S, Dunn Wb Fau - Knowles JD, Knowles Jd Fau - Broadhurst D, Broadhurst D Fau - Williams R, Williams R Fau - Ashworth JJ, et al. Closed-loop, multiobjective optimization of two-dimensional gas chromatography/mass spectrometry for serum metabolomics.

Chapter 3 : Reference standardization for quantification and harmonization of large-scale metabolomics

This chapter uses the methods established in chapter 2 and continues the development of a chemical detection methodology for identification and applies reference standardization for quantification of large numbers of metabolites. In the previous chapter, I show approximately half of the metabolites detectable by HRM have not yet been identified and the need to verify metabolite annotations with analysis of authentic chemical standard libraries. The research described in this chapter validates the detection of 428 metabolites and provides quantified values for 212 metabolites in human plasma. This facilitates the identification and quantification of previously unidentified metabolites and provides a foundation for the research described in chapter 4.

Reference standardization for quantification and harmonization of large-scale metabolomics

Ken H. Liu, Mary Nellis, Karan Uppal, Chunyu Ma, Ravneet Kaur, Yonggang Liang, ViLinh Tran, Douglas Walker, Dean P. Jones

Abstract

Reference standardization was developed to address quantification and harmonization challenges for high-resolution metabolomics (HRM) data collected across different studies or analytical methods. RS relies on the concurrent analysis of calibrated pooled reference samples at pre-defined intervals and enables a single-step batch-correction and quantification for high-throughput metabolomics. Here, we provide quantitative measurements for approximately 200 metabolites for each of three pooled reference materials (214 metabolites for Qstd3, 198 metabolites for NIST1950, and 205 metabolites for CHEAR) and show that application of this approach for quantification supports harmonization of metabolomics data collected from 3,677 human samples in 17 separate studies analyzed by two complementary HRM methods over a 17-month period. The results establish reference standardization as a method suitable for harmonizing large-scale metabolomics data and extending capabilities to quantify large numbers of known and unidentified metabolites detected by high-resolution mass spectrometry.

3.1 Introduction

High-resolution metabolomics (HRM), based on the collection of spectrochemical profiles obtained by liquid chromatography coupled to Fourier-transform mass spectrometry (FTMS), detects both known and unidentified metabolites in biological samples. These measurements provide extensive coverage of metabolites from endogenous metabolic pathways, diet, therapeutics, xenobiotics, and the microbiome and can be used for development of cumulative metabolomics databases suitable for personalized medicine, systems pharmacology, and exposome research.

Harmonization of metabolomics data remains a challenge because of differences in study design, pre-analytical (sample collection, analyte stability, etc.), analytical (sample preparation, analytical methodology, use of internal/external reference standards, instrument drift/batch effects, etc.), or post-analytical (data extraction, alignment, normalization, identification, quantification, statistical and functional analysis) factors. From a non-targeted metabolomics experiment, data are typically reported as mass-to-charge ratios (m/z) with retention times (RT) and associated peak intensities (which may be mathematically transformed values, or converted to other arbitrary units). The translation of spectral feature peak intensities to identified metabolites with estimated concentrations would facilitate cross-study and cross-method comparisons and development of harmonized platform-agnostic cumulative metabolomics databases. Although several strategies have been proposed for data normalization (1-5) and/or use of universally labeled isotopes to correct for matrix effects and ion suppression (6; 7), only half of spectral features detected by HRM correspond to previously characterized metabolites for which authentic standards may be available (8). The remaining half of spectral features detected

by HRM are not well-annotated or characterized; many show significant associations with disease (9; 10).

Reference standardization provides a practical and community-based solution for harmonizing data collected from non-targeted HRM studies (11; 12). In principle, this approach corrects for systematic technical errors by normalizing metabolite spectral peak intensities to metabolite concentrations relative to a concurrently analyzed calibrated reference sample. An ideal reference should exist in sufficient quantities for long-term routine use with every batch of samples analyzed and be representative of the biochemical composition of study samples. In practice, this can be achieved for individual laboratories by creation of a pooled reference that is calibrated against a widely available reference, such as National Institutes of Standards and Technology Standard Reference Material-1950 (NIST SRM1950) (13).

Because ion abundances detected by HRM are generally proportional to metabolite concentrations, these properties would enable use of the instrument response obtained for a metabolite with a stable, known concentration in the reference for estimating the concentration of the same metabolite detected in study samples. Since most metabolites in reference materials (e.g. NIST1950) are stable long-term when stored at -80°C (11), this approach enables retrospective quantification of metabolites as additional metabolites are characterized in study samples provided 1) the same metabolite is identified in a concurrently analyzed reference sample and 2) the ratio of metabolite peak intensities between a reference and a study sample is consistently measured across several studies (32). Thus, a thorough examination of reference “metabolomes” and extensive characterization and reporting of metabolite identifications and concentrations in one or more reference samples would provide a practical and scalable strategy

to estimate concentrations and harmonize data for large numbers of metabolites detectable by HRM.

To evaluate reference standardization as a strategy for harmonizing metabolite measurements across multiple studies collected on the same analytical platform in our laboratory, concentrations of approximately 200 metabolites were measured in three plasma reference materials (NIST1950, Qstd3, Chear). We tested the reproducibility of this approach by comparing inter-reference metabolite ratios over multiple analytical batches. We tested applicability of this approach to quantify metabolites in heparin plasma from EDTA plasma, and also compared calculated reference values against compiled ranges in HMDB. Finally, we tested use of these values to harmonize metabolomics data on representative metabolites collected from 3,677 human plasma samples analyzed over a 17 month period across 17 different studies on two complementary HRM methods.

3.2 Methods

Reference plasma materials:

EDTA, Heparin, and Citrate are commonly used anticoagulants for preparation of blood plasma. Two EDTA plasmas and a lithium heparin plasma prepared by NIST were used as reference materials for this study. Qstd3: Pooled EDTA plasma obtained from 50 healthy donors purchased from Equitech-Bio (SHP45) without information on drug use or fasting status.

CHEAR: Pooled EDTA plasma obtained from 100 adults (50 males and 50 females) purchased from BioreclamationIVT without information on drug use or fasting status. NIST1950 (13; 14): Pooled lithium heparin plasma obtained from 100 healthy volunteers intended for use as a healthy reference human plasma metabolome. Plasma was collected from fasted individuals with no documented drug use 72 hours prior to sample collection.

Standards and standard curve preparation:

Authentic chemical standards used for preparation of standard curves were from the commercially available Sigma Aldrich MSMLS library or individually purchased with stated purities of >95%. Mixtures of these standards were prepared row-by-row into stock solutions and used for preparation of standard curves in saline (3 concentrations and blank) and Qstd3 (6 concentrations and unspiked Qstd3). In total, 730 chemical standards were analyzed for this study. CHEAR and Qstd3 were analyzed with every batch and NIST1950 was analyzed every 4 rows of chemical standard mixture analyzed.

Sample preparation for LC-FTMS analysis:

50 μL of sample (plasma or saline) was mixed with 100 μL of acetonitrile containing a mixture of 9 stable isotope internal standards (15). Sample mixtures were incubated on ice for 30 min, and centrifuged for 10 min at 14,000 rpm at 4°C to pellet proteins. Supernatants were

transferred to autosampler vials and immediately loaded onto a chilled 4°C autosampler for analysis.

Instrumental analysis:

Five μL aliquots of sample extracts were analyzed using liquid chromatography and Fourier Transform high-resolution mass spectrometry (Dionex Ultimate 3000 UHPLC, HF Q-Exactive, Thermo Scientific). A dual pump configuration on the chromatographic system enabled parallel analyte separation and column flushing (16). Sample extracts were injected and analyzed using hydrophilic interaction liquid chromatography (HILIC) with positive electrospray ionization (ESI+) and reverse phase (C18) chromatography with negative electrospray ionization (ESI-). Analyte separation for HILIC was performed with a Waters XBridge BEH Amide XP HILIC column (2.1 mm x 50 mm, 2.6 μm particle size) and gradient elution with mobile phases A: water, B: acetonitrile, C: 2% formic acid. The initial 1.5 minute period consisted of 22.5% A, 75% B, and 2.5% C, followed by a linear increase to 75% A, 22.5% B, and 2.5% C at 4 min and a final hold of 1 min. C18 chromatography was performed on an end-capped C18 column (Higgins Targa C18 2.1 mm x 50 mm, 3 μm particle size) with mobile phases A: water, B: acetonitrile, C: 10 mM ammonium acetate. The initial 1 min period consisted of 60% A, 35% B, and 5% C followed by a linear increase to 0% A, 95% B, and 5% C at 3 min and held for the remaining 2 min. For both methods, the mobile phase flow rate was 0.35 mL/min for the first minute, and increased to 0.4 mL/min for the final 4 min. The FTMS was operated at 120k resolution and MS1 spectra were collected from 85-1,275 m/z . Tune parameters for sheath gas were 45 for ESI+ and 30 for ESI-. Auxiliary gas was set at 25 for ESI+ and 5 for ESI-. Spray voltage was set at 3.5 kV for ESI+ and -3.0 kV for ESI-. Ion dissociation spectra were collected

using parallel-reaction monitoring (PRM) mode with targeted inclusion lists for expected ions in HCD mode with normalized collision energy of 30%.

Reference standardization:

When the same metabolite is detected in a study sample (or other reference) as the calibrated reference, the ratio between the instrument response for a metabolite (MS1 peak intensity or MS2 diagnostic fragment peak intensity at a known retention time) detected in study samples (or other reference) to the instrument response associated with a known metabolite in a calibrated reference was multiplied by the metabolite concentration in the calibrated reference. Raw data were not normalized, autoscaled, or otherwise transformed prior to use of reference standardization.

Metabolite Identification in Reference Samples:

Spectral peaks associated with potentially formed adducts (M+H, M+Na, M+K, M+2Na-H, M-H₂O+H, etc. in ESI+; M-H, M+Cl, M+FA-H, M+CH₃COO, M-H₂O-H, M-2H, etc. in ESI-) were examined per metabolite using a ± 3 ppm mass window in xCalibur Qualbrowser software. When multiple adducts were detected, the most reproducible (technical replicate CV) and quantifiable (exhibiting the most predictable relationship between analyte concentration and peak intensity with an unweighted Pearson's linear model) MS1 adduct was selected for quantification. If MS1 was not sufficient to distinguish between isobaric species, a diagnostic MS2 fragment ion was used for quantification in the reference. Metabolite peak intensities were integrated in Qualbrowser. Metabolite retention times (RT), MS1 and MS/MS spectra obtained in unspiked reference samples were matched with RT and spectral information obtained from analysis of authentic chemical standards added to plasma or saline.

Metabolite Quantification:

Method of standards addition to calibrate Qstd3 metabolite concentrations: An unweighted linear regression line was plotted (x-axis = metabolite concentration, y-axis = metabolite peak intensity) and the negative of the x-intercept was the estimated concentration of a metabolite in Qstd3. When the addition of standard decreased the existing peak intensity for a metabolite or if no signal was observed in the reference, the metabolite was not quantified. Qstd3 values were used as the reference for estimation of metabolite concentrations in CHEAR and NIST1950. For 30 of the 207 metabolites, this approach was not useful because a positive x-intercept was obtained or the calculated concentrations differed more than ten-fold of expected concentrations in HMDB (17); for these cases, NIST1950 values were used to calibrate Qstd3 and CHEAR. Unless otherwise noted, no normalization or internal standard corrections were used for metabolite quantification.

Reference standardization of representative metabolites in 3,677 human plasma samples:

The data used to evaluate reference standardization were derived from 17 studies, comprising 3,677 samples analyzed over a 17 month period. These include fully de-identified human samples from a range of studies and were without demographic or health information. Thus, their comparison provides a blinded analysis in which the same instruments, methods and personnel analyzed the samples, but the sample collection and characteristics of the samples were independent of the analytical laboratory. HRM data for the 17 studies were aligned using apLCMS (18; 19). Qstd3 reference was analyzed with six technical replicates every ten samples. NIST1950 reference was analyzed at the beginning and end of every study. CHEAR reference was used in 4 studies. Reference standardization was performed batch-wise using Qstd3 values (**Appendix 1**) using a customized R-script.

3.3 Results

We validated the detection of 441 individual chemicals on one or both of the platforms. Information and examples for classification as detectable, identifiable, and quantifiable is provided in **Appendix 1 and 2**. For those detected, metabolites were identified in the reference material using accurate mass MS1 signal (± 3 ppm from theoretical mass for respective adduct form), co-elution with authentic standard within 3 s, and ion dissociation spectra (MS²/MSⁿ) matching authentic standard (**Figure 3.1A, B**). These criteria fulfill level 1 identification according to the Metabolomics Standards Initiative (20) and a level 1 confidence score according to Schymanski (21). For some metabolites, concentrations in the reference were too low to obtain useful MS/MS spectra and additional criteria were used for quantification as described below. Some metabolites with identical molecular formulas could be distinguished from one another by chromatographic separation or use of a diagnostic MS/MS fragment (e.g. valine/betaine); many were not easily distinguishable, however, and were not considered further for quantification. Coeluting isobaric metabolites that could not be reported as a single entity are reported as mixtures (e.g. leucine/isoleucine, glucose-1-phosphate/glucose-6-phosphate) or as a generic isomer which encompasses all possible isomers (hexane hexol for galactitol/mannitol/sorbitol). For some co-eluting metabolites (alanine, beta-alanine) where the major component is >85% of the total, the data are reported as the major metabolite. Not all identified metabolites in Qstd3 could be quantified. An identified metabolite was considered quantifiable if addition of authentic standard produced an increasing linear response and an extrapolatable negative X-intercept with response characteristics similar to that observed for the pure standard in saline.

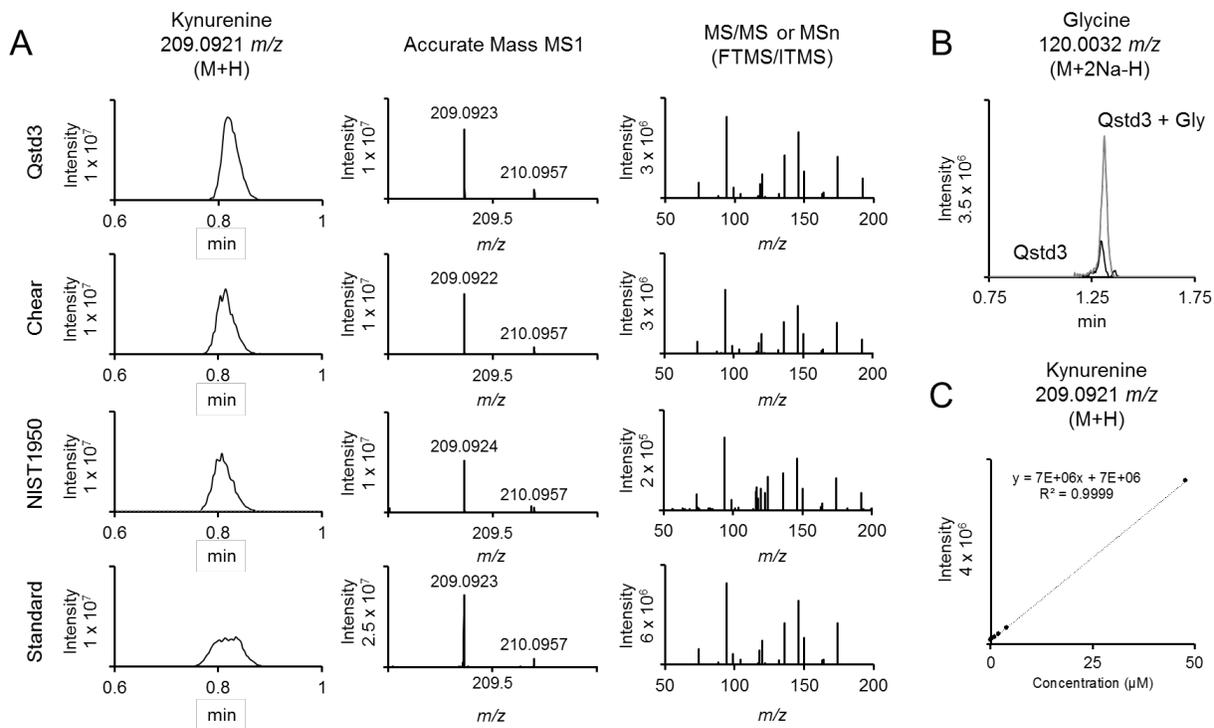


Figure 3.1 General workflow for metabolite identification and quantification. Metabolites were identified in reference samples by A) co-elution of an authentic standard and matching MS1 and MS/MS spectra or at a minimum by B) co-elution of an authentic standard with matching MS1. Metabolites were quantified in Qstd3 reference using external calibration with a method of standards addition in Qstd3 reference (C).

Adduct selection for metabolite identification

A single metabolite can generate multiple ions including adducts, isotopes, and source fragments. Thus, selection of the most reproducibly detected spectral feature is an important consideration for metabolite quantification using reference standardization. The linear response (**Figure 3.1C**) of multiple predicted adducts per standard using both HILIC/ESI+ (M+H, M+2Na-H, M+Na, M-H₂O+H, M+K, M+2H, M+3H, M-2H₂O+H, 2M+H, and 2M+ACN+H) and C18/ESI- (M-H, M+Cl, M+CH₃CO₂, M+HCO₂, M-2H, 2M-3H) were evaluated. Over 80% of detected chemical standards produced an increase in instrument response as M+H or M-H ions. In HILIC/ESI+, some organic acids formed quantitative M+2Na-H adducts and carbohydrates tended to form M+Na or M+K adducts. In C18/ESI-, carbohydrates tended to form quantitative M+Cl adducts. CoA species detected in C18/ESI- were detected as M-2H ions. These data are helpful for annotating metabolites forming non-(+H/-H) adducts. Overall, approximately one-third of metabolites in the MSMLS library were quantifiable in the reference samples (**Figure 3.2**).

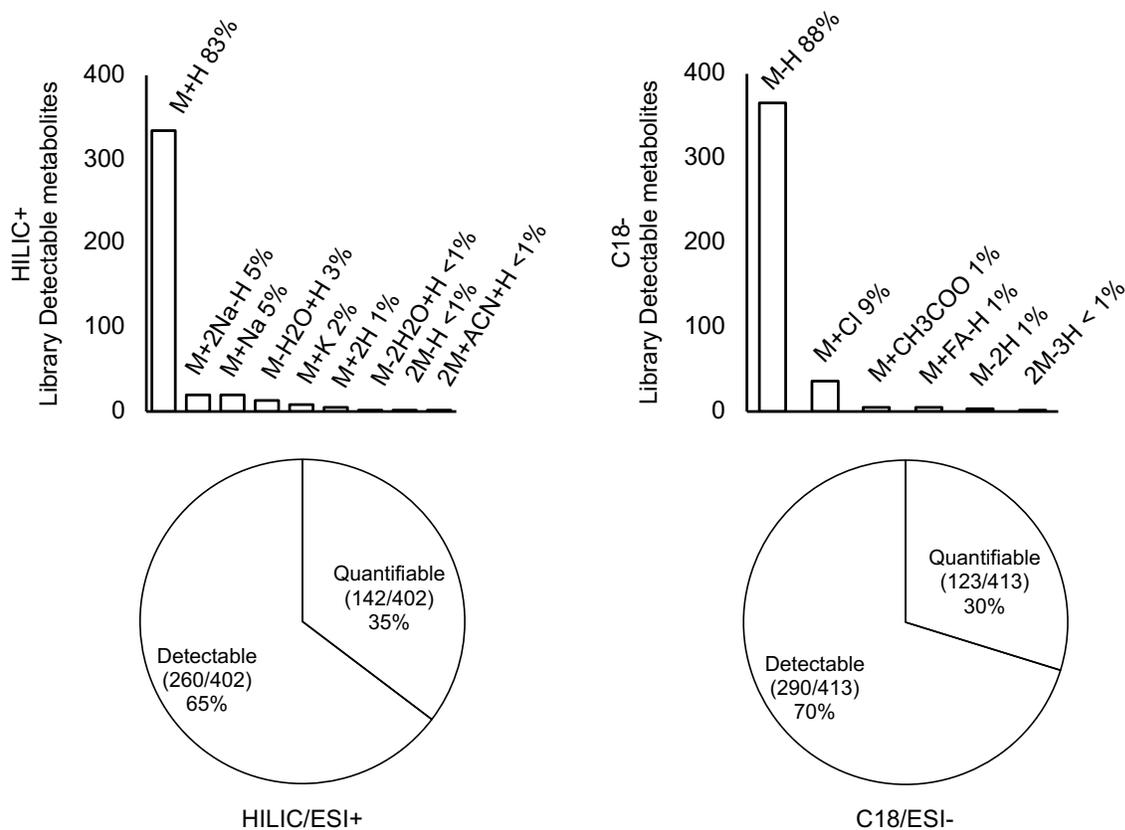


Figure 3.2 *Top*: The majority (>80%) of detectable chemical standards analyzed produced quantitative M+H or M-H ions. For metabolites ionizing as non M+H or M-H adducts as the primary form, each individual annotation should be referenced against adduct distributions obtained from chemical standard analysis to determine the probability of forming an alternative adduct. *Bottom*: Approximately one-third of detectable metabolites were quantifiable by method of standards addition, internal standardization, or external calibration to NIST1950 values.

To evaluate the reproducibility of detection of specific adducts, we calculated the relative standard deviations (%CV) of different inter-reference signal intensities of different ion forms generated by the same metabolite (**Figure 3.3**). The data show that the M+Cl adduct (130% CV NIST:Qstd3, 74% CV NIST:CHEAR, 86% CV CHEAR:Qstd3) has more variance in quantified values compared to an M-H adduct (14% NIST:Qstd3, 19% CV NIST:CHEAR 10% CV CHEAR/Qstd3) or an M+CH₃COO adduct (15% CV NIST:Qstd3, 17% CV NIST:CHEAR, 13% CV CHEAR:Qstd3). Quantified values for ion forms with the lowest standard deviations in inter-reference metabolite ratios for the assay conditions used are provided in **Appendix 1**.

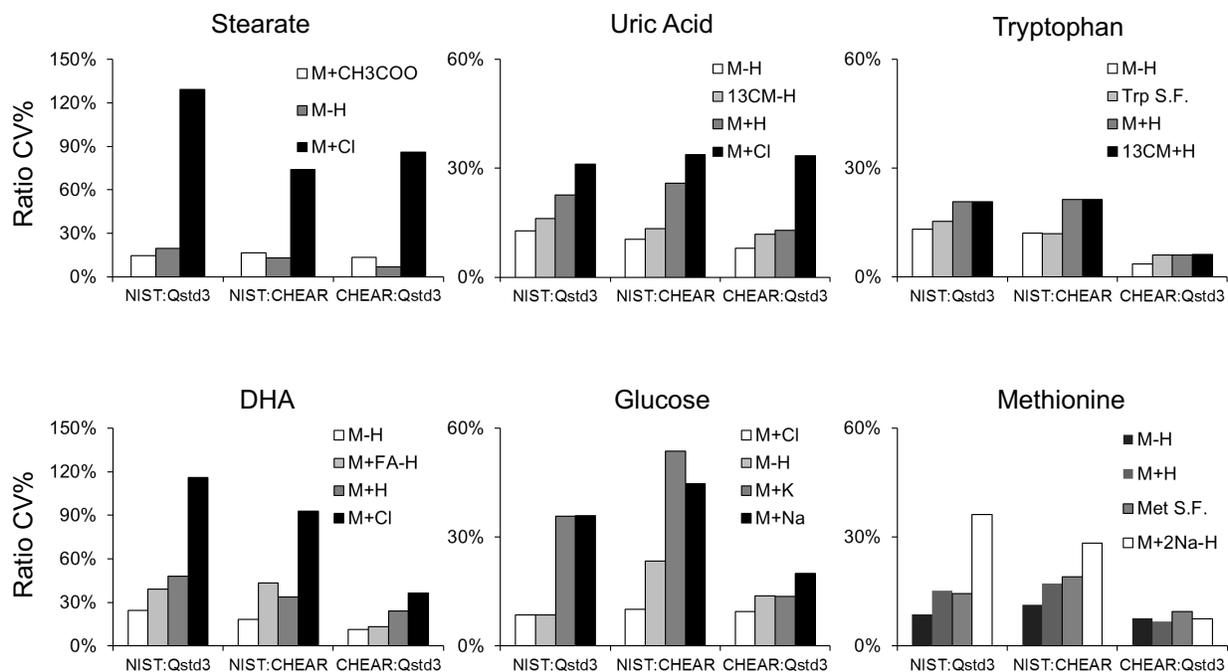


Figure 3.3 Selection of consistently detected adducts for quantification. Evaluation of inter-reference ion ratios of multiple ion forms shows selection of ion form used is important for reference standardization reproducibility. Selection of ion forms with low inter-reference ratios (<20% CV) is ideal for reference standardization. Data shown are representative of eight batches from the beginning to the end of the chemical validation period.

Inter-reference metabolite signal intensity ratios are reproducible for preferred adducts

For clinical non-targeted metabolomics assays, reportable metabolites are recommended to have QC sample relative standard deviations (RSD%) < 30% (22). An analogous metric for reference standardization is the RSD of inter-reference metabolite peak intensity ratios across several analytical batches. Previous studies showed ratios between NIST1950 to Qstd3 peak signal intensities for a panel of amino acids were stable over thirteen months of routine analysis (11). We calculated ratios of metabolite signal intensities between Qstd3:NIST1950, Qstd3:Chear, and CHEAR:NIST1950 for 8 batches spanning the 2 month period during which chemical standards were being analyzed. For nine metabolites with diverse characteristics (**Figure 3.4**), 27/27 of the inter-reference CV% ratios were less than 30%, and 26/27 were less than 20%. The NIST:CHEAR ratio CV% for carnitine was 24%, but the NIST:Qstd3 and Chear:Qstd3 ratios were less than 15%. These results support use of RSD of inter-reference metabolite peak intensity ratios across analytical batches as means to evaluate the consistency of quantitative information for reference standardization of metabolomics data.

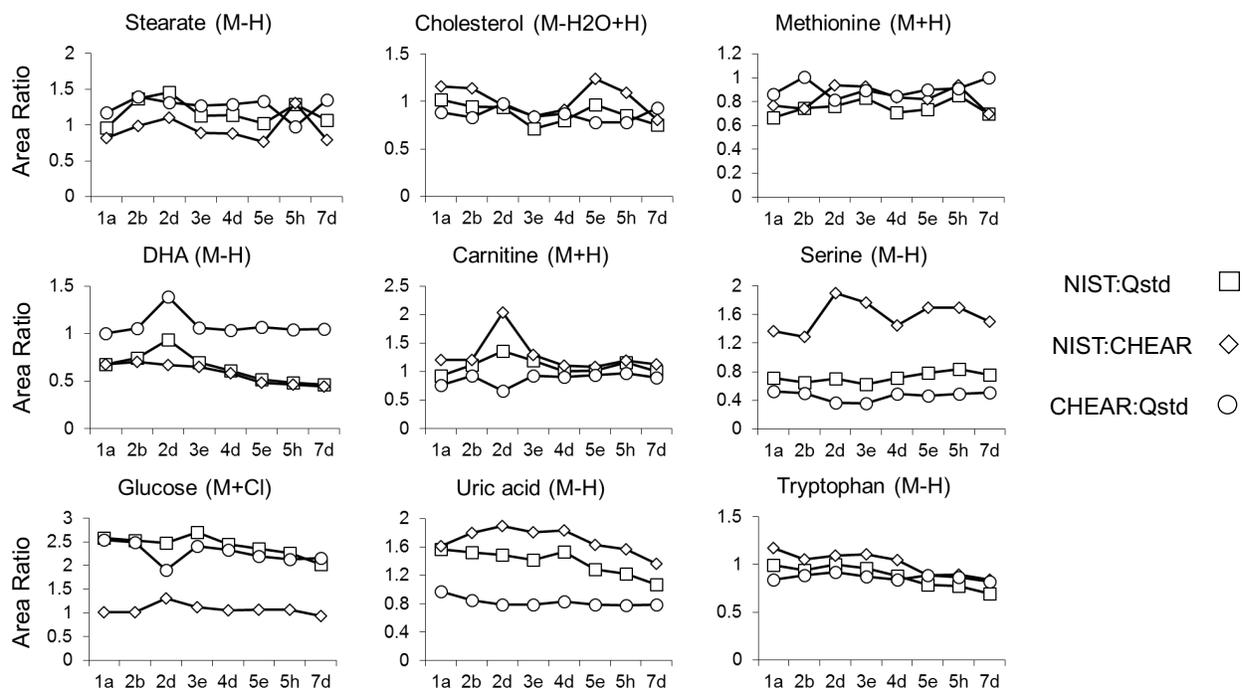


Figure 3.4 Reproducibility of inter-reference metabolite ratios for preferred adducts. Use of consistently detected adducts allows measured inter-reference metabolite signal intensity ratios to be consistent from batch to batch and allows quantification of representative metabolites from different chemical classes using reference standardization. Each data-point represents the ratio of metabolite peak intensities between two reference samples over eight batches collected over a two month time period. The average inter-sample metabolite ratio CV was 14% for the data shown here.

Estimated concentrations of metabolites in NIST1950, Chear, Qstd3

We used the most reproducibly detected adduct for each quantifiable metabolite to generate the list of detectable and identified metabolites with quantified values for NIST1950 (198), Chear (205), and Qstd3 (214) provided in **Appendix 1**. The list also includes metabolites that were detected, identified, and quantified in saline but not detected or quantified in reference samples; this information could be useful for future metabolite identification in other tissues or reference samples. Overall, the list of 441 detectable metabolites provides coverage of 264 pathways and 186 modules in KEGG. These are organized by metabolite class including amino acids and derivatives, organic acids, lipids, sterols, carnitines, vitamins and cofactors, nucleotide-related metabolites, biogenic amines, and metabolites derived from diet, xenobiotics, or other sources.

Consistency of metabolite concentrations with NIST1950 and HMDB values

We compared the estimated metabolite concentrations obtained by reference standardization for NIST1950 against previously published values and expected ranges compiled in HMDB. Previous studies showed metabolite quantification in plasma varies depending on the choice of anticoagulant (23). NIST1950 was heparinized plasma and Qstd3 and CHEAR plasma are EDTA plasma. This analysis showed that 75% of estimated amino acids concentrations were within $\pm 25\%$ of certified reference values for NIST1950 (**Figure 3.5**).

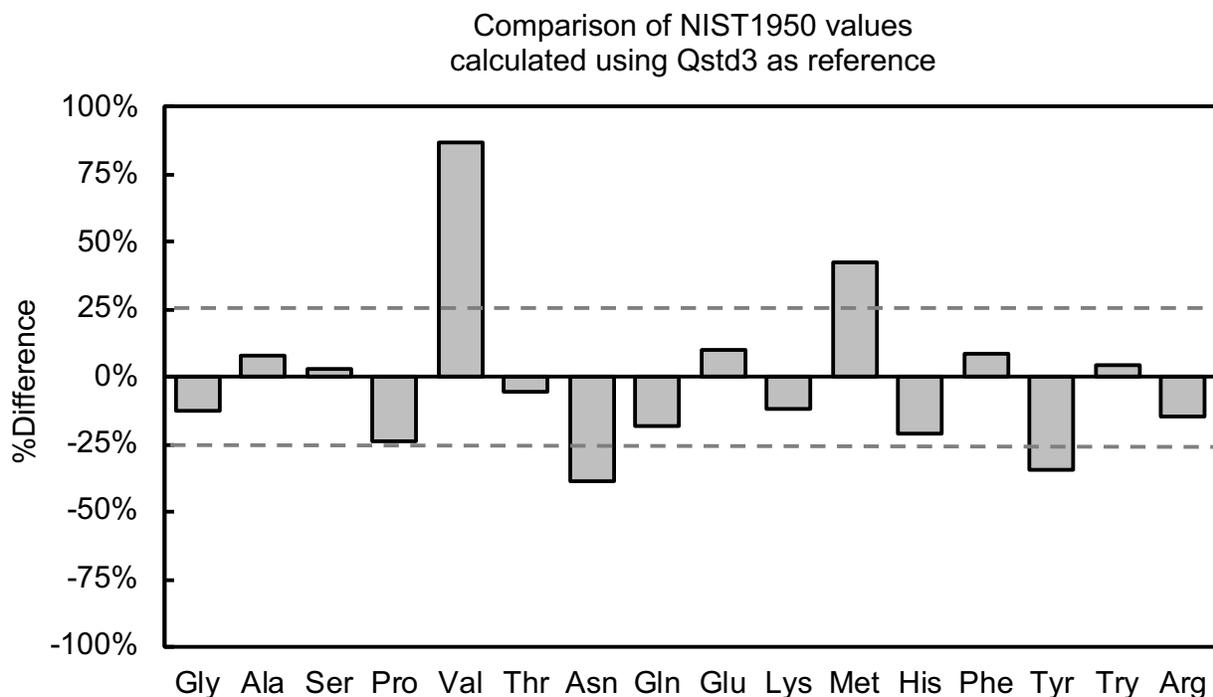


Figure 3.5 Using non-targeted HRM methods, most (75%) amino acid metabolites in heparinized plasma NIST1950 could be quantified accurately (within $\pm 25\%$ of published NIST1950 values) using EDTA plasma Qstd3 calibration and reference standardization. 15 of 16 amino acid concentrations in NIST1950 calculated from Qstd3 values were within $\pm 50\%$ of NIST1950 values. Asp and Cys not shown because NIST1950 values are not available. Leu and Iso not shown because the analytical method used did not resolve the two isomers.

Comparison of representative metabolites in Qstd3, NIST1950 and Chear HMDB (Figure 3.6) showed that estimated concentrations of metabolites in NIST1950, Qstd3, and CHEAR were within expected HMDB ranges. More than two-thirds of the 137 quantified metabolites in Qstd3 with reported blood ranges were within HMDB ranges. Thus, these estimated concentrations for NIST1950, Qstd3, and CHEAR plasma extend capabilities to test utility of reference standardization to quantify and harmonize data from different studies.

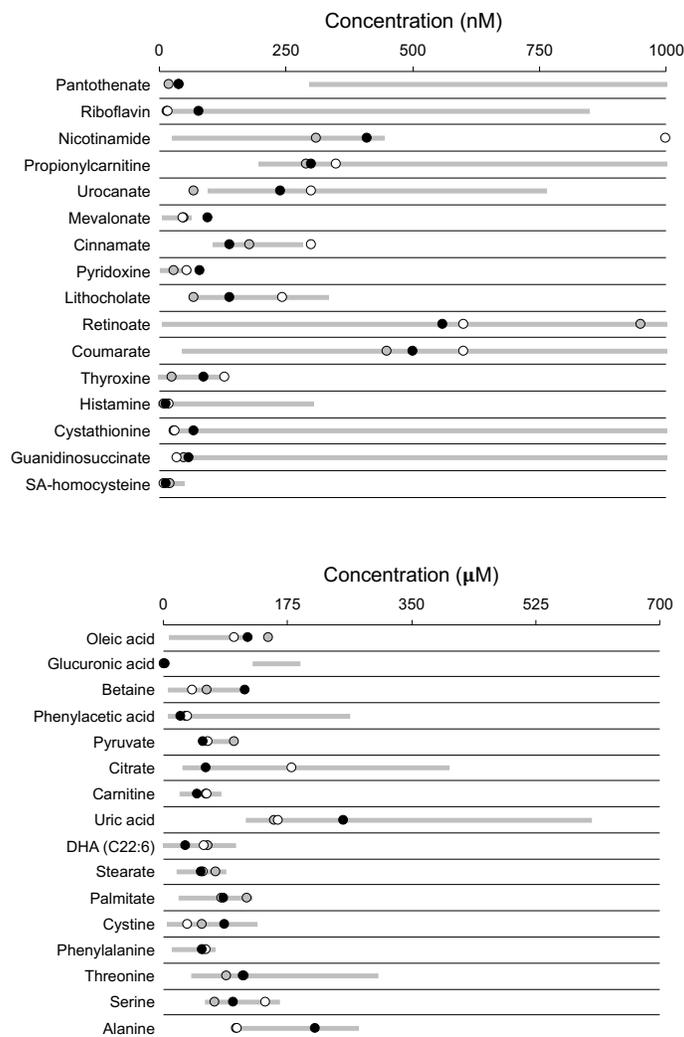


Figure 3.6 Using reference standardization with either Qstd3 or NIST1950 as the reference, 80% of representative metabolite concentrations in reference materials detected at nM and μM concentrations in NIST1950 (●), Chear (◐), or Qstd3 (○) produced values within HMDB ranges (grey rectangles). Overall, approximately two-thirds of metabolite concentrations in Qstd3 fell within previously published ranges in human plasma.

Reference Standardization for quantification and harmonization of large-scale metabolomics data

For metabolomics data to be “harmonized” or comparable between studies, technical variation due to batch effects need to be minimized and metabolite peak intensities need to be converted to units that can be reproduced across studies or analytical methods. Analytical workflows often consider these points independently. For instance, data from large-scale metabolomics studies cannot generally be combined without adjusting for batch effects (24). Several strategies have been proposed to correct for batch effects based on the use of scaling factors, quality control samples, internal standards, or use of statistical batch effect correction strategies (2; 25; 26). In this section, we tested use of reference standardization with the Qstd3 calibrated reference plasma to perform a single step correction for systematic technical errors and subsequent quantification for large-scale metabolomics. This was tested with six technical replicates of the Qstd3 reference at the beginning, middle, and end of every 20-sample batch. During the two month period analytical standards were analyzed, our data show the median drift in metabolite retention time was less than 8 seconds and mass spectral drift was less than 2 ppm using the HFQE (data not shown). These results show that accurate mass MS1 and retention times for metabolites with confirmed identity and quantification in the reference, as described above, are sufficiently stable over time for quantification of respective metabolites in study samples.

The total variance associated with metabolite measurements is the sum of biological variance and technical variance, with the biological variance represented by the median relative standard deviation (Med-RSD) of study samples and the technical variance represented by the mean relative standard deviation (RSD) of reference samples. Conversion of sample metabolite

peak intensities to concentrations minimizes technical variation (decreased Med-RSD) so that the data better represent biological variation (**Figure 3.7**). For example, comparing the batch 2 methionine median peak intensity versus methionine reference standardized concentration shows a reduction of batch-wise median variance. For about 10% of metabolites, however, use of reference standardization increased the Med-RSD by more than ten percentage points (not shown, perhaps reflecting non-linear responses, ion suppression or other factors affecting quantification by mass spectrometry). Taken together, these data show that batch-wise reference standardization can be used to correct for batch effects for most metabolites in large multi-batch studies.

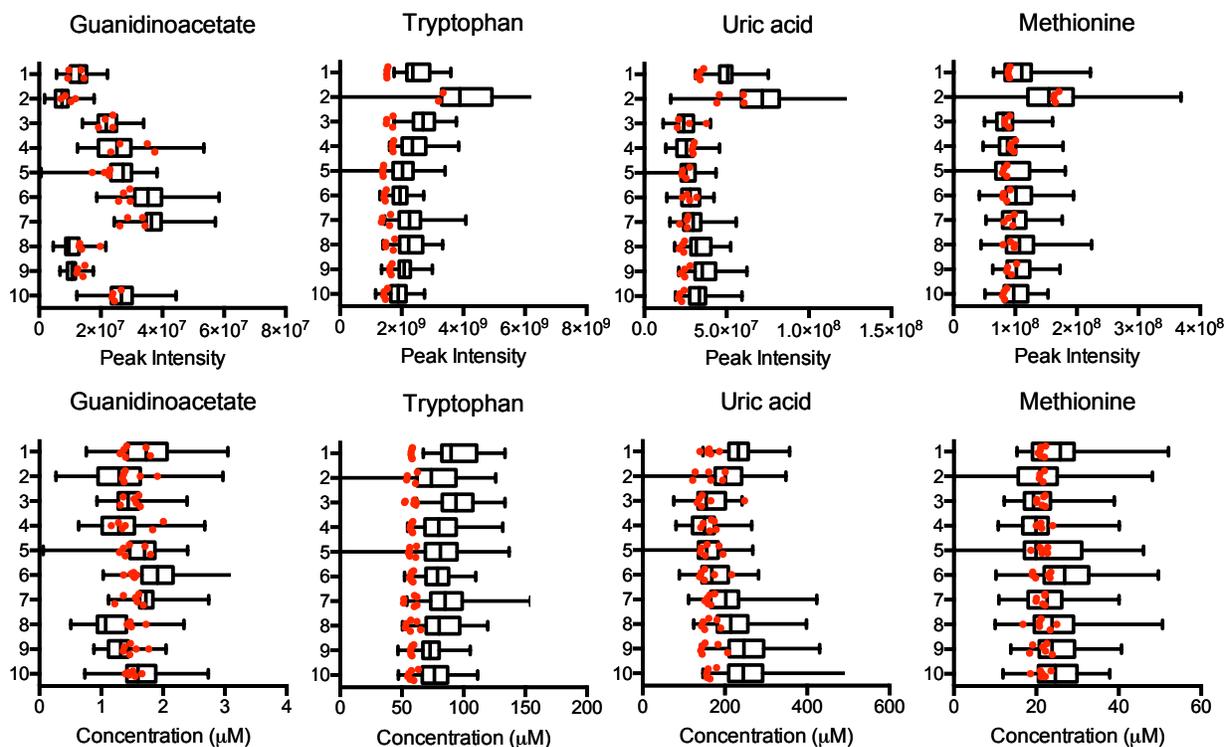


Figure 3.7 Application of reference standardization to 200 study samples analyzed across a 10 batch study over a one week period. Red dots represent Qstd3 reference samples superimposed over study sample box-and whisker plots depicting the range, quartiles, and medians of 20 study samples. Use of batch-wise reference standardization can reduce median RSD (Med-RSD) for metabolite measurements across a multi-batch study – Guanidinoacetate peak intensity Med-RSD: 50%; reference standardized Med-RSD: 18%. Tryptophan peak intensity Med-RSD: 25%; reference standardized Med-RSD: 8%. Uric acid peak intensity Med-RSD: 42%; reference standardized Med-RSD: 19%. Methionine peak intensity Med-RSD: 21%; reference standardized Med-RSD: 12%.

To test utility of reference standardization for harmonization of metabolomics data from multiple studies, we calculated metabolite concentrations for 8 metabolites with diverse characteristics and intensities in 3,677 plasma samples across 17 separate studies analyzed with the same analytical methodology over a 17 month time period. Data for each sample were quantified relative to concurrently analyzed Qstd3 samples using reference standardization (**Figure 3.8**). Results showed that the distributions were more uniform for most of the metabolites and studies but that unique characteristics were retained. For instance, uric acid and tryptophan achieved relatively normalized distributions while retinoic acid retained high skewed distributions. Unique characteristics for some metabolites and studies were evident, i.e., one study each had very distinct distributions for bilirubin and putrescine, and caffeine showed highly heterogenous distribution before and after conversion to concentration values.

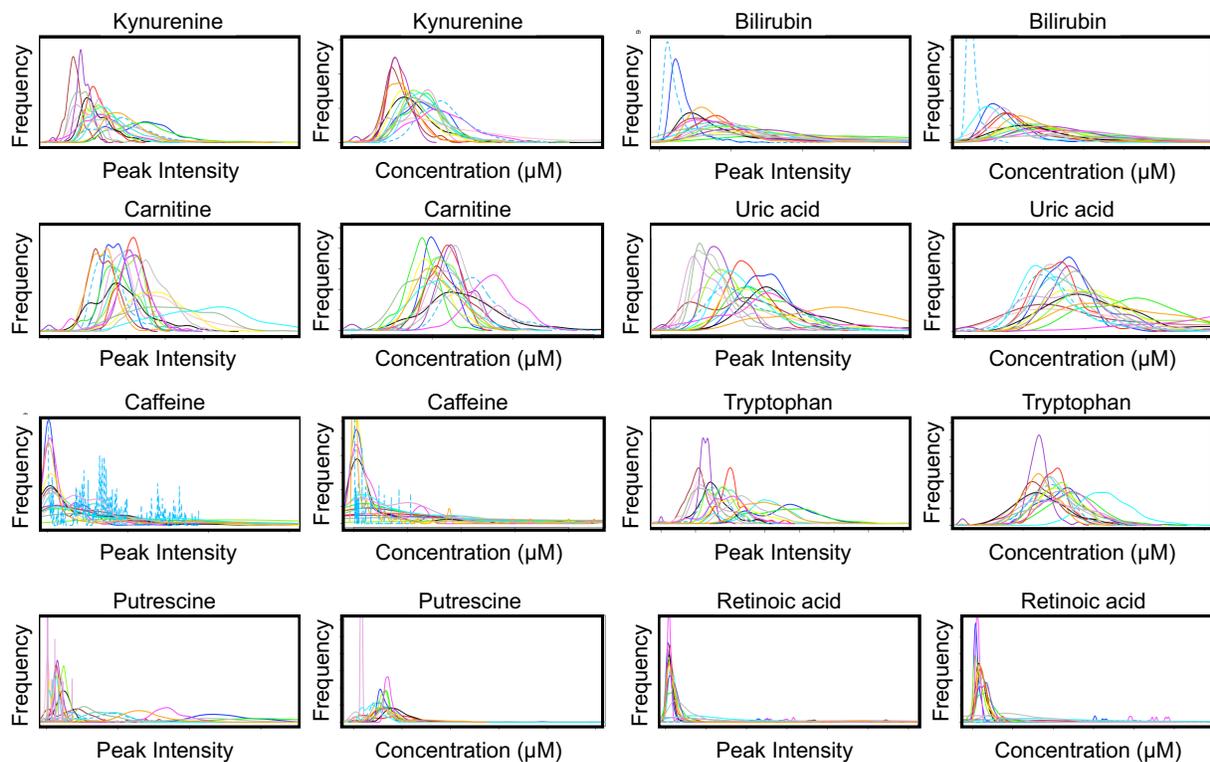


Figure 3.8 Application of reference standardization to 3,677 study samples analyzed across 17 studies over a 17 month period. Density histograms for 17 separate studies are shown for selected metabolites using raw peak intensities or reference standardized concentration decreased the variability of metabolite measurements (study median RSD) across studies.

3.4 Discussion

Many metabolites are useful health indicators and routinely quantified in targeted mass spectrometry-based assays. In contrast, non-targeted HRM methods were developed to capture as broad of a chemical space as possible and rely on computational methods to discover associations of metabolites, metabolic pathways and networks with health and disease phenotypes. In this study, we expand the list of quantified metabolites to support quantification of 207 metabolites with Qstd3, 189 metabolites with NIST1950, and 199 metabolites with CHEAR pooled reference materials using the reference standardization method. Results show that inter-sample metabolite peak intensity ratios for validated adducts are consistently measured and demonstrate the ability to provide metabolite concentrations in agreement with previously established NIST1950 values and HMDB reference ranges. The analyses further show ability to use this approach to harmonize metabolite measurements collected over a 17 month period for more than 3600 individual study samples. The results suggest a generalizable approach to use reference standardization to develop cumulative metabolomics databases suitable for personalized medicine.

Metabolite identification remains a bottleneck in non-targeted metabolomics, and community-driven efforts to describe the biochemical composition of reference materials and detectability of chemical standard library metabolites by LC-HRMS are important to facilitate progress. In this study, we analyzed over 700 analytical standards and validated the detection of close to 400 metabolites covering endogenous, microbiome, dietary, drug, and environmental-exposure metabolites on a dual liquid chromatography HILIC/ESI+ and C18/ESI- HRM platform. We delineated metabolites that are detected and quantifiable in three pooled plasma materials from those that were not confidently detected or quantifiable in the reference samples.

The increasing number of identified and quantified metabolites, with adduct forms and retention times, increases opportunity to improve interoperable metabolomics databases to support efforts for precision medicine.

A limitation of reference standardization is that metabolites can only be quantified if they are present and quantified in the reference. Some metabolites not detected may not be present in these reference materials, and others will probably require complementary methods with improved sensitivity and/or selectivity. Use of reference materials with a wide spectrum of exogenous metabolites (i.e. NIST1957/1958 SRM for organic contaminants in human serum or NIST968f, 971a, 972a, 1951c, 1955, 2378, 2973, 3949, 3950 for various endogenous lipids, nutrients, and hormones) will further enable reference standardization for metabolites that are not commonly detected in each individual. Pooled reference materials suited for specific purposes will facilitate more targeted analyses (10; 27) and enable more global assessments of the spectrum of exogenous chemicals derived from food, xenobiotics, environmental and occupational exposures, personal care products, supplements, and drugs and their biotransformation products.

NIST1950 plasma was developed as a reference material for use with metabolomics studies to facilitate identification and quantification of metabolites (13). Most of the estimated metabolite concentrations presented in this study are consistent with previously published values for NIST1950 and within expected HMDB ranges. Reference ranges for metabolite concentrations vary between laboratories depending on methodology. For example, estimated concentrations for several lipids in the NIST1950 are outside of expected HMDB ranges. Previous studies have shown differences in lipid and amino acid quantification depending on the anticoagulant used for plasma samples (23). Whether such differences in metabolite

quantification observed in this study are due to anticoagulant (NIST1950 is lithium heparin plasma, Chear and Qstd3 are EDTA plasma) or differences in sample preparation or analysis are not clear.

Reference standardization assumes a linear relationship between analyte concentration and instrument response, which was validated for most of the chemicals reported here; exceptions were those calibrated relative to previously published NIST1950 value, where single-point calibration was used. For automated quantification, as used in **Figure 3.8**, the method uses single-point calibration and performs comparably to quantification relative to internal standard, which also involves single-point calibration (32). By using a pooled reference material, the quantification is biased toward average human values and results can be expected to be less quantitatively reliable for relatively high and low values. Also, if the biochemical composition of the reference does not reflect the study samples (different intensity range observed in reference versus study samples), then the estimated concentrations for study sample metabolites will be associated with high error and other analytical (additional QC samples, kit methods, use of internal standards, improved separation) or post-analytical considerations (normalization, non-linear models, or use of other scaling factors) would be needed to address potential sources of error.

Targeted methods for clinical analysis typically require <20% CV on QC samples and accuracies within 10% of an accepted central value. A recent lipidomics harmonization effort across 31 laboratories using diverse analytical methodologies showed that less than a quarter of lipid species (339/1,527) identified by a single laboratory could be detected by more than 5 laboratories in the NIST1950 SRM (14). After filtering quantified lipids with less than 40% coefficient of dispersion (COD – a measure analogous to relative standard deviation), 259/339

(75%) of identified lipids remained. Another interlaboratory comparison of six laboratories using the Biocrates AbsoluteIDQ p180 kit for targeted metabolomics analysis shows that approximately 80% of metabolites measured had accuracies within 20% of consensus interlaboratory values (28). Similar to these values, our data show approximately 60-80% of identified metabolites could be quantified in the reference and applied for use in study samples. .

The present study shows that reference standardization can be used in an automated workflow to facilitate harmonization of metabolomics data collected across multiple studies. By correcting for batch effects on a chemical-by-chemical basis, reference standardization facilitates harmonization of large-scale metabolomics studies without the trade-off of most batch normalization methods in which improved quality of some metabolites is offset by corruption of data for other metabolites. By converting raw peak intensities to concentrations, reference standardization allows comparison of results obtained from multiple analytical platforms based upon estimates of absolute concentration. Intra- and inter-laboratory validation and proficiency testing are well developed for targeted clinical assays and will need to be developed and implemented to enable of untargeted metabolomics to medical and other uses.

In conclusion, reference standardization using calibrated reference samples analyzed at pre-defined intervals with study samples provides a practical and simple method for data normalization and estimation of metabolite concentrations in high-throughput applications. In principle, adoption of this technique could allow non-targeted metabolomics data to be comparable across studies and laboratories. The approach is scalable as additional metabolites are characterized in other pooled reference materials, thereby expanding capabilities to harmonize metabolomics for clinical research and other practical applications.

References

1. De Livera AM, Dias DA, De Souza D, Rupasinghe T, Pyke J, et al. 2012. Normalizing and integrating metabolomics data. *Anal Chem* 84:10768-76
2. De Livera AM, Sysi-Aho M, Jacob L, Gagnon-Bartsch JA, Castillo S, et al. 2015. Statistical methods for handling unwanted variation in metabolomics data. *Anal Chem* 87:3606-15
3. Boysen AK, Heal KR, Carlson LT, Ingalls AE. 2018. Best-Matched Internal Standard Normalization in Liquid Chromatography-Mass Spectrometry Metabolomics Applied to Environmental Samples. *Anal Chem* 90:1363-9
4. Kamleh MA, Ebbels Tm Fau - Spagou K, Spagou K Fau - Masson P, Masson P Fau - Want EJ, Want EJ. Optimizing the use of quality control samples for signal drift correction in large-scale urine metabolic profiling studies.
5. Li B, Tang J, Yang Q, Li S, Cui X, et al. 2017. NOREVA: normalization and evaluation of MS-based metabolomics data. *Nucleic Acids Res* 45:W162-W70
6. You L, Zhang B, Tang YJ. 2014. Application of stable isotope-assisted metabolomics for cell metabolism studies. *Metabolites* 4:142-65
7. Weindl D, Wegner A, Jager C, Hiller K. 2015. Isotopologue ratio normalization for non-targeted metabolomics. *J Chromatogr A* 1389:112-9
8. Liu KH, Walker DI, Uppal K, Tran V, Rohrbeck P, et al. 2016. High-Resolution Metabolomics Assessment of Military Personnel: Evaluating Analytical Strategies for Chemical Detection. *J Occup Environ Med* 58:S53-61
9. Uppal K, Walker DI, Liu K, Li S, Go YM, Jones DP. 2016. Computational Metabolomics: A Framework for the Million Metabolome. *Chem Res Toxicol* 29:1956-75
10. Jones DP. 2016. Sequencing the exposome: A call to action. *Toxicol Rep* 3:29-45
11. Go YM, Walker DI, Liang Y, Uppal K, Soltow QA, et al. 2015. Reference Standardization for Mass Spectrometry and High-resolution Metabolomics Applications to Exposome Research. *Toxicol Sci* 148:531-43
12. van der Greef J, Martin S Fau - Juhasz P, Juhasz P Fau - Adourian A, Adourian A Fau - Plasterer T, Plasterer T Fau - Verheij ER, et al. The art and practice of systems biology in medicine: mapping patterns of relationships.
13. Phinney KW, Ballihaut G, Bedner M, Benford BS, Camara JE, et al. 2013. Development of a Standard Reference Material for metabolomics research. *Analytical chemistry* 85:11732-8
14. Bowden JA, Heckert A, Ulmer CZ, Jones CM, Koelmel JP, et al. 2017. Harmonizing lipidomics: NIST interlaboratory comparison exercise for lipidomics using SRM 1950-Metabolites in Frozen Human Plasma. *J Lipid Res* 58:2275-88
15. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. 2013. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. *Metabolomics* 9:S132-S43
16. Soltow QA, Strobel FH, Mansfield KG, Wachtman L, Park Y, Jones DP. 2011. High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. *Metabolomics* 9:132-43
17. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, et al. HMDB 4.0: the human metabolome database for 2018.

18. Yu T, Jones DP. Improving peak detection in high-resolution LC/MS metabolomics data using preexisting knowledge and machine learning approach.
19. Yu T, Park Y, Johnson JM, Jones DP. 2009. apLCMS--adaptive processing of high-resolution LC/MS data. *Bioinformatics (Oxford, England)* 25:1930-6
20. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, et al. 2007. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 3:211-21
21. Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, et al. 2014. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol* 48:2097-8
22. Broadhurst D, Goodacre R, Reinke SN, Kuligowski J, Wilson ID, et al. 2018. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 14:72
23. Khadka M, Todor A, Maner-Smith KM, Colucci JK, Tran V, et al. 2019. The Effect of Anticoagulants, Temperature, and Time on the Human Plasma Metabolome and Lipidome from Healthy Donors as Determined by Liquid Chromatography-Mass Spectrometry. *Biomolecules* 9
24. Johnson WE, Li C, Rabinovic A. 2006. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8:118-27
25. Fernández-Albert F, Llorach R, Garcia-Aloy M, Ziyatdinov A, Andres-Lacueva C, Perera A. 2014. Intensity drift removal in LC/MS metabolomics by common variance compensation. *Bioinformatics* 30:2899-905
26. Kuligowski J, Sánchez-Illana Á, Sanjuán-Herráez D, Vento M, Quintás G. 2015. Intra-batch effect correction in liquid chromatography-mass spectrometry using quality control samples and support vector regression (QC-SVRC). *Analyst* 140:7810-7
27. Jones DP, Park Y, Fau - Ziegler TR, Ziegler TR. Nutritional metabolomics: progress in addressing complexity in diet and health.
28. Siskos AP, Jain P, Römisch-Margl W, Bennett M, Achaintre D, et al. 2017. Interlaboratory Reproducibility of a Targeted Metabolomics Platform for Analysis of Human Serum and Plasma. *Analytical Chemistry* 89:656-65

Chapter 4: Microbiome-mitochondria cross-talk: Inhibition of mitochondrial fatty acid oxidation by the microbiome-derived metabolite, delta-valerobetaine

This chapter applies the validated chemical analysis platform to identify mitochondrial metabolites produced by the microbiome and characterization of the effects of microbial metabolites on mitochondrial function and cellular energy metabolism.

Microbiome-mitochondria cross-talk: inhibition of mitochondrial fatty acid oxidation by the microbial metabolite delta-valerobetaine

Ken H. Liu, Bejan Saeedi, Catherine Cioffi, Moriah Bellissimo, Joshua A. Owens, Samuel Druzak, Michael Orr, Xin Hu, Jolyn Fernandes, Mary Catherine Camacho, Sarah Hunter-Chang, Chunyu Ma, Trevor Darby, Gregory Tharp, Thota Ganesh, Samantha Yeligar, Karan Uppal, Young-Mi Go, Jessica Alvarez, Miriam Vos, Thomas Ziegler, Rheinallt Jones, Eric Ortlund, Andrew Neish, Dean P. Jones

Abstract

The intestinal microbiome is linked to an epidemic of human metabolic disorders involving energy metabolism. Here, we identify the microbial-derived small molecule, δ -valerobetaine (VB), a structural analog of the carnitine precursor γ -butyrobetaine. VB is absent in germ-free mice and derived liver mitochondria but is present in conventionalized mice and mitochondria. When supplemented to cells *in vitro*, VB decreased mitochondrial fatty acid oxidation by lowering intracellular carnitine. *In vivo*, VB administered to fasted mice lowered carnitine, decreased fatty acid oxidation and exacerbated hepatic steatosis. In addition, circulating VB in humans was associated with the severity of hepatic steatosis, hepatic insulin resistance and central adiposity. These results show VB is a gut microbe-derived small molecule that participates in microbiome-mitochondrial cross talk to influence mitochondrial energy metabolism and can contribute to fatty liver and other metabolic disorders.

4.1 Introduction

Emerging evidence links the epidemic of human metabolic disorders to changes in the composition and activity of the intestinal microbiota (1-6). The intestinal microbiota possess remarkable metabolic functions for digestion of dietary macromolecules and synthesis of diverse metabolites that directly impact human metabolism (7; 8). For example, epidemiological and experimental evidence show that microbial products such as lipopolysaccharide (LPS) (9; 10), phenylacetic acid (11), and methylamines (12-14) contribute to the development of insulin resistance, hepatic steatosis, and cardiovascular disease. Metabolomic analyses of germ-free (GF) and colonized (C) mouse tissues reveal that approximately 10% of the circulating mammalian metabolome are metabolites of microbial origin (15); the majority of these are unidentified (16) and associated mechanisms are unknown.

Mitochondrial dysfunction is commonly associated with human metabolic disorders. An emerging paradigm is that microbiome-mitochondrial cross-talk, i.e., bi-directional communication between mitochondria and the microbiome, plays an important role in metabolic health and disease (17). Studies show microbial metabolites directly influence host mitochondrial functions and metabolism (18; 19), and alterations to mitochondrial genomes influence the composition of the microbiome (20). To investigate the effect of the microbiome on systemic mitochondrial function and to identify microbiome-derived mitochondrial metabolites, we performed integrated analysis of transcriptome and metabolome profiles of liver from germ-free (GF) and conventionalized (CV) mice. Our analysis revealed that the microbiome altered hepatic gene expression related to mitochondria and lipid homeostasis. In addition, the top discriminatory metabolite detected in liver and liver mitochondria of CV mice was δ -valerobetaine (VB). VB is absent in GF mice and structurally resembles γ -butyrobetaine,

the immediate biosynthetic precursor to carnitine, which is required for mitochondrial fatty acid oxidation.

To elucidate the role of VB on mitochondrial function and hepatic lipid metabolism, we performed dose-response experiments with VB in cultured human cell lines and mice. These experiments show VB treatment decreases cellular carnitine, inhibits mitochondrial fatty acid oxidation and increases hepatic lipid accumulation. The results establish VB as a link between the gut microbiome and mitochondrial functions related to fatty acid oxidation. Furthermore, we found that circulating VB was positively correlated with central adiposity and hepatic steatosis in humans. Taken together, our data show that VB participates in a microbiome-mitochondrial cross-talk which could be a determinant of clinical phenotypes linked to adiposity and hepatic steatosis.

4.2 Methods

Animals

Germ-free (GF) male and female Swiss-Webster mice (Taconic Biosciences) were raised in germ-free isolators and fed sterilized mouse chow (Envigo 2019S Teklad Global 19% Protein Extruded Rodent Diet) and sterilized water *ad libitum* at the Emory University Gnotobiotic Animal Research Facility. Conventionalized (CV) mice were transferred to bedding from conventional Swiss-Webster (S) mice at three weeks of age, and maintained for 3 weeks while paired GF mice were maintained in germ-free isolators. After three weeks, luminal contents from the cecum and colon, portal vein serum, and intact livers were collected for mitochondrial isolation, RNA-sequencing and metabolomics. Data for GF and conventional C57BL6J (jax.org/strain/000664) mice (Jackson laboratories, identical housing as Swiss-Webster) are provided as indicated.

RNA-Sequencing

RNA-Seq analyses were conducted at the Yerkes NHP Genomics Core on GF (n = 6) and CV liver (n = 6). RNA was collected and extracted from PAXgene tubes using on-column DNase digestion as described previously (38) and assessed for integrity and quantity using an Agilent Bioanalyzer (Agilent Technologies) and a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). Libraries were prepared using the Illumina TruSeq mRNA stranded kit. Briefly, 500–1,000 ng of globin-depleted RNA was used for library preparation. ERCC synthetic spike-in controls 1 or 2 (Ambion) were added to each total RNA sample and processed in parallel. Amplified libraries were validated using the Agilent 4200 TapeStation and quantified using a Qubit fluorometer. Libraries were normalized and pooled, followed by clustering on a HiSeq 3000/4000 flowcell using the Illumina cBot. The clustered flowcell was then sequenced on the

Illumina HiSeq 3000 system employing a single-end 101-cycle run, with multiplexing to achieve approximately 20 million reads per sample. Gene IDs were annotated for 41,128 gene IDs (ENSEMBLID) and mapped to 23,568 genes (*mus musculus*) in DAVID. Transcript abundance was estimated using htseq-count v0.6.1p1 and differential expression analyses were performed using DESeq2. A combination of bioinformatics tools was used for data analysis including the molecular signature database (MSigDB – <http://software.broadinstitute.org/gsea/msigdb/index.jsp>) and Ingenuity Pathway Analysis.

High-resolution metabolomics (HRM)

Cytosolic and mitochondrial fractions from GF/CV liver (n = 12) were prepared by differential centrifugation (21). Serum and fractionated cytosol and mitochondria were mixed with 2x volume of ice-cold acetonitrile-internal standard solution. Liver and luminal contents were weighed and mixed with 10x volume of ice-cold 20% water/80% acetonitrile-internal standard solution. Tissue samples were homogenized using a pellet pestle prior to the next step. Samples were vortexed, and placed on ice for 30 minutes prior to centrifugation at 14,000g for 10 minutes at 4°C to precipitate proteins. Supernatants were transferred to autosampler vials and stored at -70°C prior to instrumental analysis.

Untargeted high-resolution mass spectrometric profiling was performed using a Dionex Ultimate 3000 UHPLC system coupled to a Thermo Scientific Velos LTQ-Orbitrap (metabolite identification studies), Thermo Scientific High-Field Q-Exactive (HFQE) (liver, liver mitochondria, liver cytosol, portal vein) or Thermo Scientific Fusion (cecum, colon, *ex vivo* incubations, isotope tracer) (22). Briefly on the Fusion and HFQE, an LC column switching method using a reversed phase C18 column (Higgins Targa C18 2.1 mm x 50 mm, 3 µM particle size) and a HILIC column (Waters XBridge BEH Amide XP HILIC column 2.1 mm x 50 mm,

2.6 μM particle size) were used for the analytical separation of extracts prior to high-resolution mass spectrometry analysis with HILIC/ESI+ and C18/ESI- at 120k resolution while of ions between 85-1,275 m/z . A similar LC column switching method 10 minute and 20 minute reversed phase C18 and HILIC separation methods with 10 cm columns with ESI+ was used on the Velos-Orbitrap for confirmation.

Data processing, statistics, and feature selection

Spectral features were extracted and aligned using optimized parameters for apLCMS (23; 24), with downstream quality control performed by xMSanalyzer (25). Each metabolic feature was characterized by its m/z ratio, retention time, and peak intensity. Statistical analysis was performed using R-packages *limma* (26) for differential expression analysis and *diffexp* for linear regression analysis in xmsPANDA (<https://github.com/kuppal2/xmsPANDA>).

Differentially expressed metabolites with a raw p-value of less than 0.05 were used for pathway enrichment analysis in *mummichog* (27). Pathway enrichment significance in *mummichog* was based permutation testing with a $p < 0.05$. Metabolite names in *mummichog* were converted to KEGG IDs using <http://csbg.cnb.csic.es/mbrole/conversion.jsp>.

Metabolite Identification and Quantification

Detected metabolites were referenced against an in-house reference library established with authentic chemical standards and matched within 5 ppm of the confirmed mass and within 10 s of the confirmed retention time. Features of interest were visually inspected to ensure chromatographic and spectral peak quality. When no standard was available, MS isotopic ratios and MS/MS spectra were referenced against online spectral libraries or *in silico* ion dissociation spectra. Tandem MS² spectra were collected for target features on either Thermo Scientific Velos-Orbitrap, High-Field Q-Exactive, or Fusion mass spectrometers using collisional induced

dissociation (CID) or high-energy collisional induced dissociation (HCD). MS2 spectra were matched to experimental online mass spectral databases mzCloud and METLIN (28) or *in silico* predicted fragmentation patterns from metfrag (29) to identify candidate structures. Standards were purchased when commercially available or synthesized as described below.

Reference standardization (30) was used for quantification of identified metabolites. A calibrated reference plasma sample analyzed concurrently with study samples is used as a single-point calibration for serum samples. Briefly, a

Synthesis of δ -valerobetaine and valine betaine

δ -Valerobetaine (VB) was synthesized as described previously (31). Briefly, 5-bromovaleric acid (Sigma-Aldrich) and 1 molar equivalent trimethylamine (20% in EtOH) were stirred for 24h under vacuum at room temperature. The precipitate containing trimethylamine HBr salt was removed by filtration and the filtrate was evaporated under vacuum and recrystallized with cold isopropanol and acetonitrile. Purity was assessed to be greater than 95% with proton NMR and elemental analysis. MS: 160.1332 *m/z* ESI+ (MS2: 101.0597, 60.0809). ¹H NMR (**Appendix 3**) (400 MHz, D₂O) δ 3.15 (multiplet, 2H), δ 2.91 (singlet, 9H), δ 2.27 (triplet, 2H), δ 1.64 (quintuplet, 2H), δ 1.46 (quintuplet, 2H). Elemental analysis (as VB-HBr salt) was C 37%, H 8%, N 5%, Br 31%. Valine betaine was synthesized as previously described (32). L-Valine and methyl iodide were stirred for 48h in anhydrous methanol in the presence of excess potassium bicarbonate. The precipitate was dried and washed with ice-cold methanol. MS: 160.1332 *m/z* ESI+ (MS2: 60.0809).

Ex-vivo fermentation

Cecal contents from conventional or GF mice were collected and immediately placed in either De Man, Rogosa, and Sharpe media (MRS broth – Oxoid, CM0359) or degassed Tryptic

Soy Broth (Millipore 22092) in oxygen evacuated headspace vials. Samples were incubated at 37°C were collected in triplicate over a 24-h time period and prepared for metabolomics analysis as above.

Ex-vivo microbial metabolism assays

Bacteria were purchased from ATCC and/or isolated in our laboratory and were cultured in specified liquid media. Samples were collected at baseline and again at 18 hours after incubation at 37°C. *Lactococcus lactis Subsp. cremoris* (LC, ATCC 19257, ATCC 14365, ATCC 11602), *Lactobacillus rhamnosus GG* (LGG, ATCC 53103), *Lactobacillus plantarum* (LP, HA-119), *Lactobacillus paracasei* (HA274), *Lactobacillus rhamnosus* (HA-114, HA-111, R0011), were grown in MRS broth (Oxoid, CM0359). *Bacillus cereus* was grown in brain heart infusion (BHI) media. *E. coli* (K12) and *Salmonella Typhidurium* (SL1344) were grown in Luria Broth (LB).

VB dose response experiments in HepG2 cells: metabolomics analysis

HepG2 cells (ATCC) were used between passages 8-15 and grown in EMEM (ATCC) supplemented with 10% FBS and 0.5% Penicillin/Streptomycin (P/S). Cells were grown to 90% confluence in 3.5 cm cell culture dishes and treated with experimental compounds (VB, meldonium, propionylcholine, carnitine) in EMEM supplemented with 0.5% FBS and 0.5% penicillin/streptomycin. Cells were washed with ice cold Hank's Buffered Salt Solution (HBSS) and harvested by scraping on ice using 200 µL of ice-cold 20:80 water:acetonitrile containing 9 stable isotope internal standards and centrifuged at 14,000g for 10 minutes at 4°C. Supernatants were transferred to autosampler vials and maintained at 4°C until instrumental analysis.

VB dose response experiments in HepG2 cells: Stable Isotope Palmitate Tracer Assay

$^{13}\text{C}_{16}$ -palmitic acid (Sigma-Aldrich 605573) was dissolved in 150 mM NaCl after heating to 70°C and slowly mixed with prewarmed (37°C) FFA-free BSA (Sigma-Aldrich A4602) solution in 150 mM NaCl with stirring to produce a 1 mM palmitate to 0.17 mM BSA ratio. HepG2 cells were grown in 12-well cell culture plates to 80% confluence. Cell media was replaced with EMEM containing 0.5% FBS for 12 h \pm VB prior to treatment with labeled palmitic acid. Cells were washed twice with pre-warmed HBSS and media was replaced with EMEM containing 0.5% FBS, 200 μM labeled palmitate, and either vehicle, 50 μM VB, 40 μM etomoxir, or 50 μM carnitine. At each time point, cells were washed once with ice cold HBSS and harvested by scraping on ice using 200 μL of ice-cold 20:80 water:acetonitrile containing 9 stable isotope internal standards and centrifuged at 14,000g for 10 min at 4°C. Supernatants were transferred to autosampler vials and maintained at 4°C until analysis. Data were analyzed using xCalibur QuanBrowser for carnitine and labeled palmitate, palmitoyl-CoA, palmitoylcarnitine, acetyl-CoA, and acetylcarnitine.

VB dose response experiments in HepG2 cells: Mitochondrial Respiration Assays

Oxygen consumption (OCR) was measured in the human hepatoma HepG2 cell line using a Seahorse XFe96 analyzer (Agilent Technologies). For assessments of respiration linked to oxidation of glucose and glutamine, cells were cultured on 96-well cell culture microplates and treated overnight with 0, 1, 3, 10, 30, 100 μM VB in DMEM with 0.5% FBS. Cells were then washed 1x with 100 μM pH 7.4 Seahorse XF DMEM (Agilent Technologies) in 0.5% FBS and media was replaced with XF DMEM supplemented with fuel substrates and VB or vehicle. The media was supplemented with 10 mM glucose, 2 mM GlutaMAX, and 1 mM pyruvate \pm VB.

For assessments of fatty acid oxidation, cells were serum-starved overnight with or without VB. After washing, media was replaced with Krebs Heinsleit Buffer (KHB) containing 200 μ M BSA-conjugated palmitate (Agilent Technologies) in the absence of supplemental glucose, GlutaMAX, pyruvate, or carnitine. KHB was prepared from pH 7.4 sterile filtered water containing 111 mM NaCl, 4.7 mM KCl, 2 mM MgSO₄, and 1.2 mM NaH₂PO₄. VB (0, 10, 50 μ M final concentration) or etomoxir (40 μ M final concentration) were added to the assay medium prior to assessments of oxygen consumption (OCR) and extracellular acidification rate (ECAR).

To evaluate mitochondrial function, a MitoStress test kit (Agilent Technologies) was used with respiration assays. Three consecutive respiration measurements were acquired every 10 minutes for every experimental condition prior to and after injections of Oligomycin (1.5 μ M final concentration), FCCP (1 μ M final concentration), and Rotenone/Antimycin A (0.5 μ M final concentration). Basal respiration (OCR prior to injection of mitochondrial inhibitors), ATP production (difference in OCR after addition of oligomycin), maximal respiration (OCR after addition of FCCP) and spare capacity (difference in OCR between basal and maximal respiration) were calculated using the Seahorse XF Cell Mito Stress Test Report Generator (<https://www.agilent.com/en/products/cell-analysis/cell-analysis-software/data-analysis/seahorse-xf-cell-mito-stress-test-report-generators>).

Animal VB model

Animal experiments were performed under approved Emory University IACUC animal protocols. For experiments with fed mice, conventional C57BL6J mice (Jackson laboratories) were injected IP with 200 μ L saline vehicle, 200 μ L 10 mM VB (10 mg/kg), or 200 μ L 100 mM VB (100 mg/kg) once a day for one week. For experiments in fasted conventional mice, animals

were treated with the same doses for three days and then fasted for 12 hours prior to sample collection. At the end of each treatment period liver, heart, brain, cecal and colonic contents and serum were collected for mass spectrometry, histology analysis, and triglyceride quantification.

Histology (Oil Red O) analysis

Oil red O staining for neutral lipids was performed on 8-10 micron mouse liver sections prepared using a Cryostat. Images were taken on a Nikon Eclipse 50i at 20x magnification and red density was quantified in ImageJ software.

Triglyceride quantification

Triglycerides were quantified in tissues using a triglyceride assay kit (Abcam 65336). 50 mg of liver tissue was homogenized in 1 mL of assay buffer and diluted 100x prior to addition of lipase. After 20 minutes, the triglyceride probe was added. After incubation for 30 minutes, the plate was read fluorometrically (535/587 nm) and quantified using the provided standard calibration curve.

4.3 Results

The effect of the microbiome on the mouse liver transcriptome and metabolome

To identify the potential impact of VB on hepatic mitochondrial function, we identified the major liver transcriptional (**Appendix 4**) and metabolic pathways changed in sex-matched littermate CV (microbiome-replete) versus germ-free (GF) mice. Metabolomics pathway analysis showed enrichment of metabolites in fatty acid oxidation pathways ($p < 0.05$) (**Figure 4.1A**) and overrepresentation analysis (ORA) of hepatic genes induced by the microbiome revealed mitochondria (GO:CC, $p\text{-FDR} < 10^{-50}$) (**Figure 4.2A**) and lipid metabolism/transport (Reactome, $p\text{-FDR} < 10^{-50}$) (**Figure 4.2B**) as the top pathways changed by the microbiome. These included gene expression for peroxisomal and mitochondrial fatty acid metabolism (Fasn, Elovl1, Hadh, Acad, Ech, Eci, Decr, Cpt1, Acaa1, Acsl, Gpat3, Agpat2/3, Dgat2, Acox1/2, Ehhadh, Acaa2), fatty acid uptake and transport (Cd36, Fatp5, Fabp), and lipoprotein export (Apob100, Apoe, Apoc2/3), all of which were increased in CV animals. Gene expression related to fatty acid biosynthesis increased in CV liver and associated with increases in short and long-chain acyl-CoA esters in the liver (**Figure 4.1B**). Upstream analysis of transcriptional regulators responsible for these changes in Ingenuity Pathway Analysis (IPA) revealed PPAR-alpha as the top transcriptional regulator targeted by the microbiota (**Figure 4.2C**). Taken together, these data highlight the widespread changes to hepatic mitochondrial energy metabolism and lipid homeostasis associated with an intact microbiome.

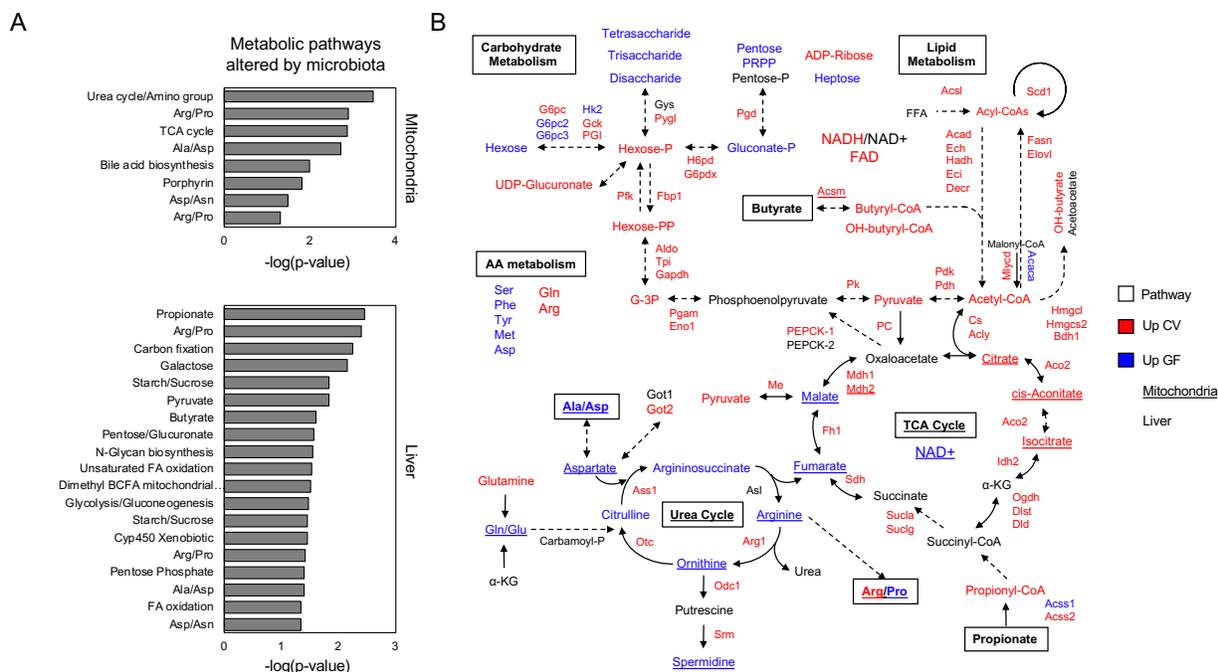


Figure 4.1 The microbiome elicits widespread changes to liver metabolism. A) Pathway enrichment analysis ($p < 0.05$, *mummichog*) of differentially expressed metabolites (raw $p < 0.05$) between germ-free and conventionalized mice show the microbiome alters mitochondrial metabolic pathways (TCA cycle), and liver metabolic pathways involved in lipid and energy metabolism (mitochondrial fatty acid oxidation). B) Integration of metabolome and transcriptome profiles shows the microbiome is associated with changes to lipid and other energy producing metabolic pathways. Red colors indicate transcripts and metabolites that were increased in CV animals. Blue colors indicate transcripts and metabolites that were increased in GF animals.

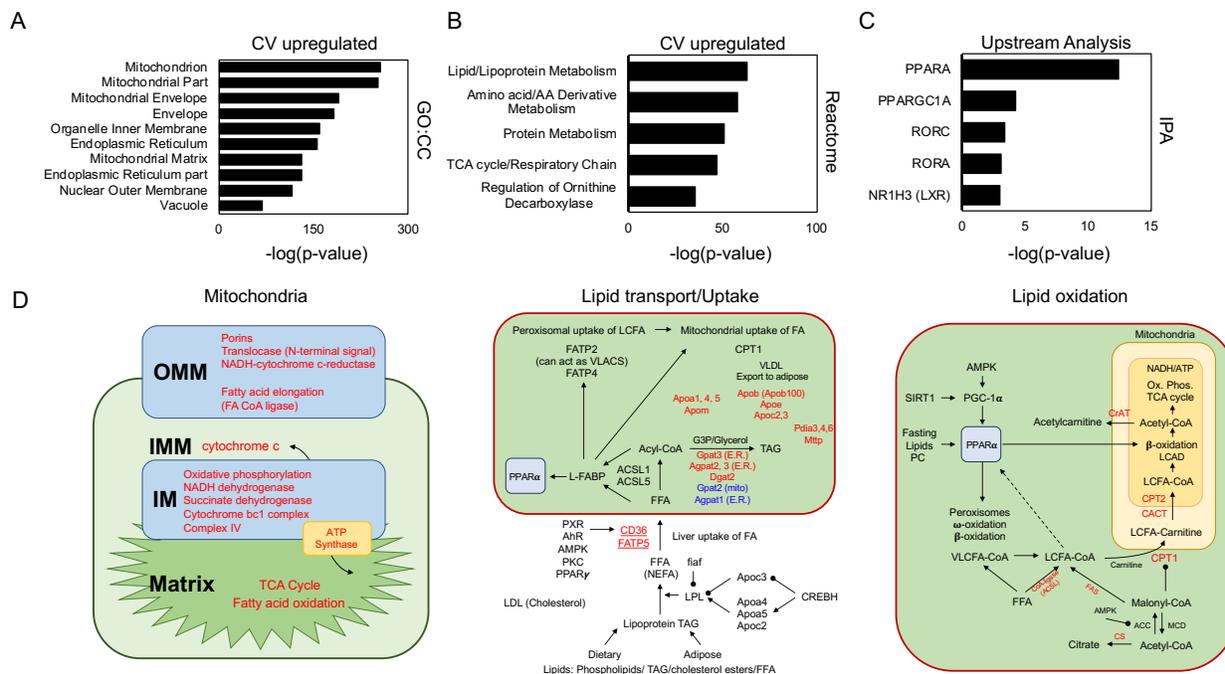


Figure 4.2 Transcriptome analysis shows the microbiome impacts gene expression related to mitochondrial functions and lipid homeostasis. A) Over-representation analysis using the Molecular Signatures Database (MSigDB – Broad Institute) shows mitochondria and related cellular compartments are the top GO:CC pathways impacted by the microbiome in the liver (p-FDR < 1e-150). B) Over-representation analysis using the Reactome database shows lipid and lipoprotein metabolism is the top Reactome pathway impacted by the microbiome in the liver (p-FDR < 1e-50). C) Upstream analysis of differentially expressed hepatic mRNA using Ingenuity Pathway Analysis predicts PPAR-alpha (p < 1e-10) as the top transcriptional regulator impacted by the microbiome. D). Illustration of increased gene expression in CV mouse liver in mitochondria (left), lipid transport/uptake (middle), and lipid oxidation (right). Red colors indicate transcripts and pathways that were increased in CV animals. Blue colors indicate transcripts that were increased in GF animals. The microbiome augments gene expression related to mitochondrial function and lipid homeostasis (uptake, export, oxidation, and lipogenesis).

Identification of the microbiome-derived mitochondrial metabolite δ -valerobetaine (VB)

To identify discrete microbial metabolites capable of altering hepatic mitochondrial function and lipid metabolism, we looked for specific metabolites which were present only in conventionalized liver and liver mitochondria. We observed that the top-ranked differentially expressed feature in both CV liver (p-FDR = 2.3e-6) and liver mitochondria (p < 0.05) was a metabolite with a m/z of 160.1332 and elemental composition $C_8H_{18}NO_2$ (**Figure 4.3, Figure 4.4A**). 160.1332 m/z was detected in portal circulation (**Figure 4.4B**), cecal (**Figure 4.4C**) and colonic contents (not shown) obtained from CV mice but not GF. Authentic standard co-elution and MS/MS experiments show the identity of 160.1332 m/z detected in the luminal contents, portal circulation, liver and liver mitochondria to be δ -valerobetaine (VB), and not propionylcholine or valine betaine (**Figure 4.5**). VB was not present in the mouse diet and not detectable in samples collected from GF mice (**Figure 4.4D**). Estimated concentrations of VB in conventional mice ranged from 51-90 ng/mg in the cecum (expressed as ng/mg tissue), 19-70 ng/mg in the colon, 9-26 μ M in the portal vein of CV animals, 12-31 ng/mg in the liver, and 2-10 μ M in peripheral serum.

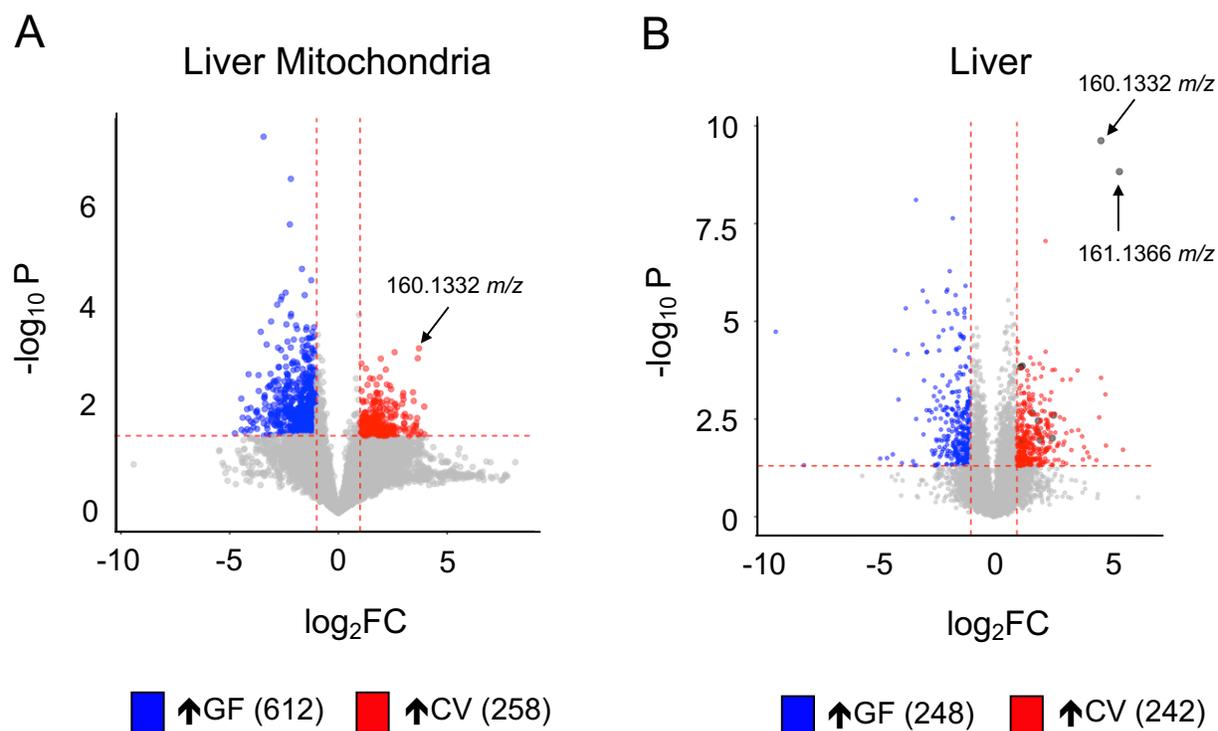


Figure 4.3 Untargeted high-resolution metabolomics analysis comparing liver and liver mitochondria of 6 germ-free (GF) and 6 conventionalized (CV) mice reveals 160.1332 m/z as the top metabolite detected in liver and liver mitochondria of conventionalized mice. A) Volcano-plot (depicting \log_2 fold-change on the x-axis and $-\log_{10}$ (FDR-p-value) on the y-axis) of germ-free (GF, blue) and conventionalized (CV, red) mouse metabolomics reveal top ranked differentially expressed feature (limma p-FDR < 0.05, \log_2 Fold-change (FC) > 2) in both A) liver mitochondria and B) liver was 160.1332 m/z with the elemental composition $C_8H_{18}NO_2$. 161.1366 m/z corresponds to a ^{13}C isotope of 160.1332 m/z . The data shown are a comparison of six GF and six conventionalized Swiss-Webster mice.

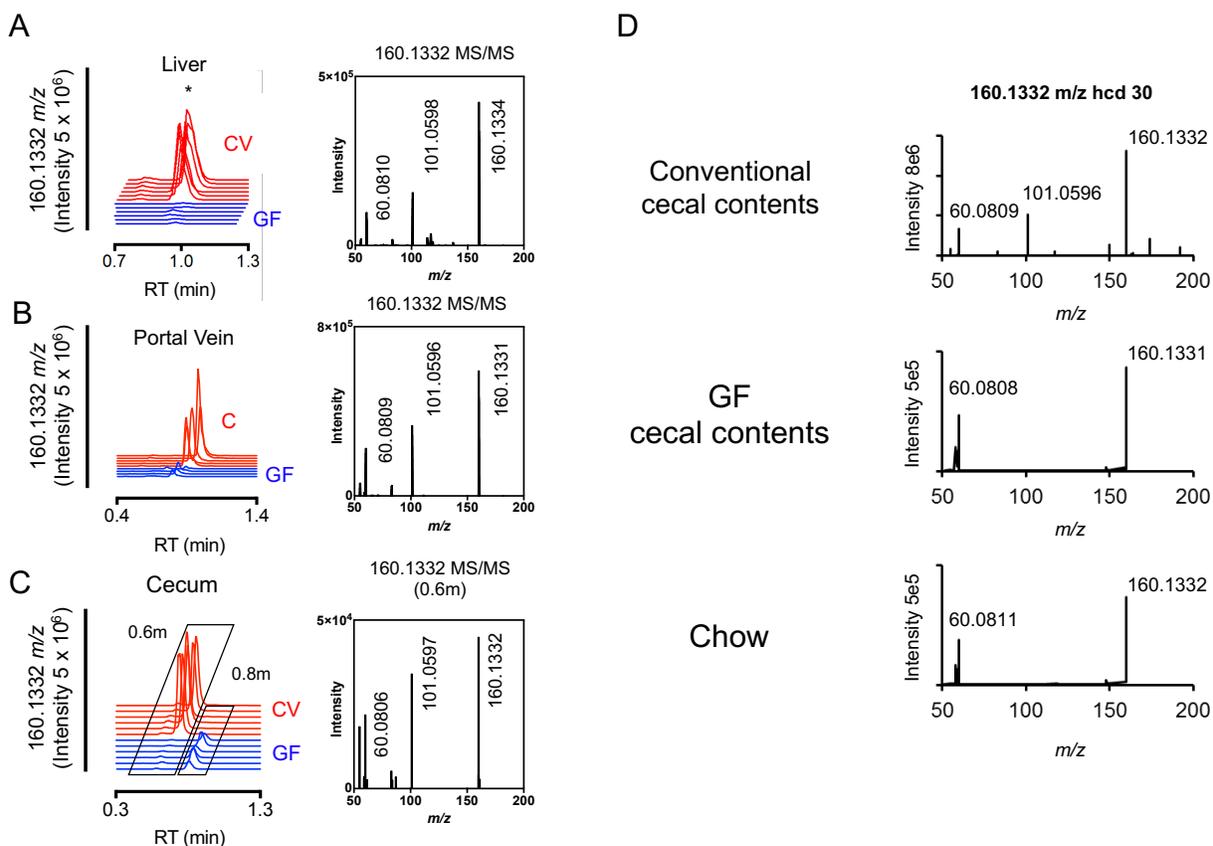


Figure 4.4 Ion dissociation spectra (MS/MS) analysis of 160.1332 m/z . Extracted ion chromatographs of 160.1332 m/z in A) liver, B) portal vein, and C) cecal contents shows the absence of 160.1332 m/z in GF animals. High-resolution FT-MS/MS analysis shows 160.1332 m/z produces characteristic 101.0600 m/z and 60.0809 m/z fragment ions. D) MS/MS spectra of 160.1332 from mouse chow, GF and CV cecal contents. GF cecal contents and mouse chow did not produce a characteristic 101.0600 m/z fragment ion (0.8 m peak in Figure 4.4C) and this signal was ten-times lower in abundance than the observed signal in conventional and CV cecal contents.

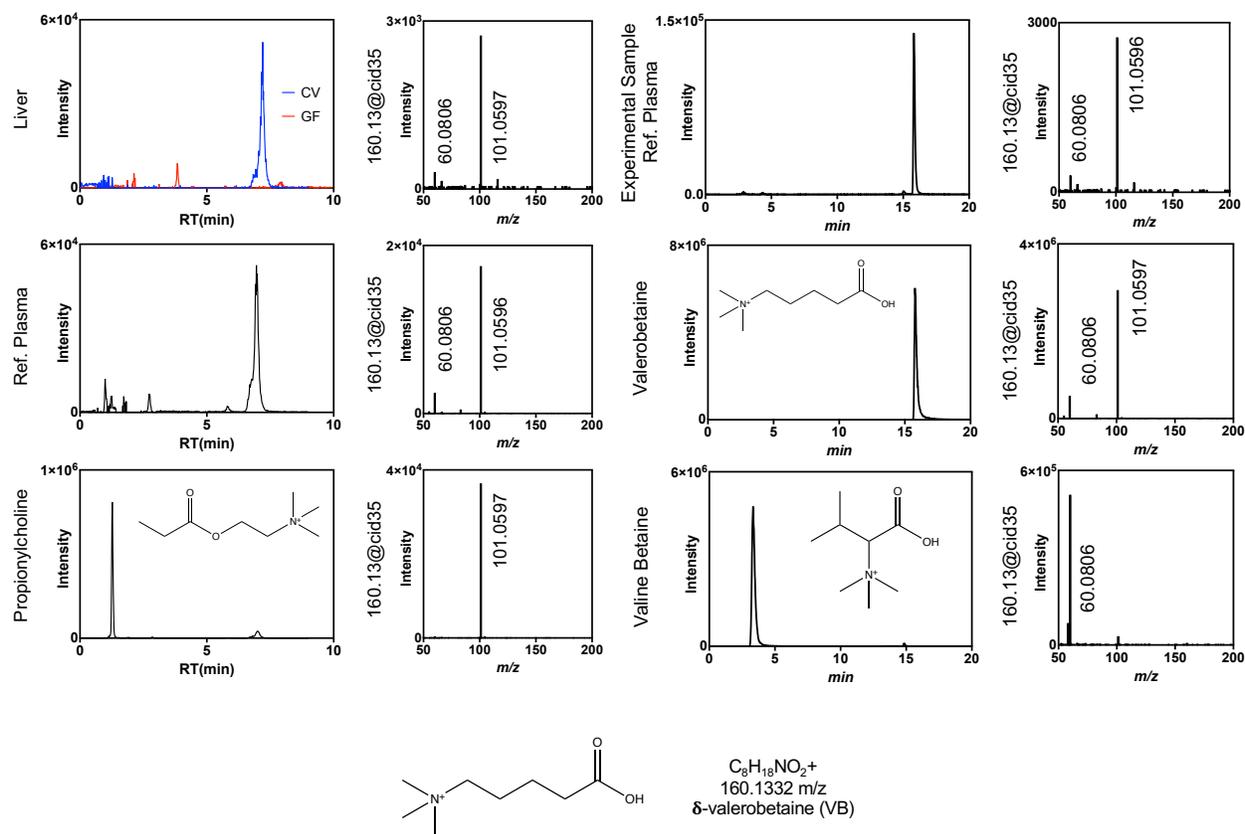


Figure 4.5 Confirmation of metabolite identity for 160.1332. Co-elution of experimental samples with authentic standard shows that 160.1332 m/z observed in conventional mouse tissues and human plasma is δ -valerobetaine (VB). Propionylcholine (retention time 2 minutes) and valine betaine (retention time 4 minutes) standards did not co-elute with the observed peak that is present only in CV animals (retention time 7-8 minutes in left panels and 16 minutes in right panels). Additionally, MS/MS analysis can differentiate propionylcholine (no 60.0809 m/z fragment ion) and valine betaine (no 101.0600 m/z fragment ion) from VB. Valerobetaine standard (middle right) produced a peak that co-eluted with the experimental sample (top right) and produced the identical MS/MS spectra as the experimental sample.

To determine whether microbial activity was required for intestinal production of VB in mice, we performed *ex vivo* incubations of cecal contents from conventional and GF mice (**Figure 4.6A**). LC-MS/MS analysis of these samples show only conventional cecal contents produced VB whereas GF cecal contents did not. A metabolic screen of several candidate commensal and pathogenic microbes shows several bacteria (*LGG*, *E. Coli*, and *Salmonella Typhimurium* (SL1344)) are able to produce VB *ex vivo* (**Figure 4.6B**). These represent diverse taxa of gram positive and gram negative pathogenic and commensal bacteria. Taken together, these data show intestinal production of VB requires microbial activity and therefore, the presence of VB in conventional animals reflects microbial production, followed by absorption and distribution into mouse tissues.

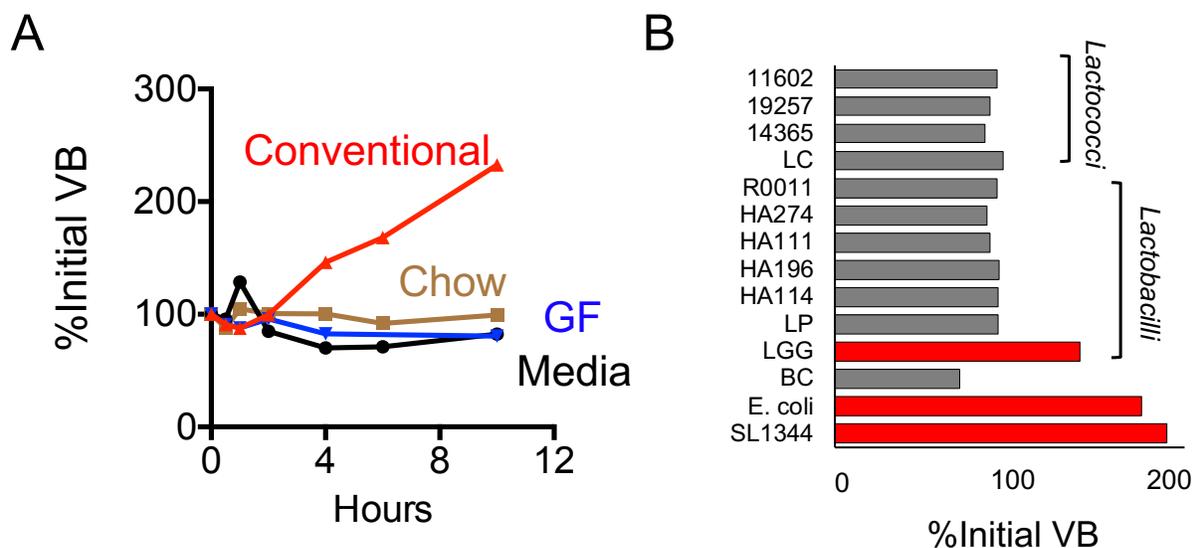


Figure 4.6 Microbial activity is required to produce VB. A) *Ex vivo* incubation of cecal contents from GF and conventional mice shows that the microbiome is required for intestinal production of 160.1332 *m/z*. VB peak area only increased over time in cecal contents from conventional mice. VB was not present in GF mouse cecal contents and was not produced over time. B) Analysis of bacterial cultures at baseline and after 18 hours of growth show several candidate bacteria (*LGG*, *E. Coli*, *SL1344* (*Salmonella Typhimurium*)) are able to produce VB *ex vivo*.

δ-valerobetaine (VB) decreases cellular carnitine in cultured HepG2 cells

To identify the impact of VB on mitochondrial functions and fatty acid metabolism, we performed a dose-response study in human HepG2 cells. Cells were treated for 12 hours with VB, and changes to the intracellular metabolome caused by VB-treatment were assessed with metabolomics analysis. The top metabolic pathway changed by VB-treatment was carnitine-shuttle metabolism (**Figure 4.7A**). More specifically, VB-treatment caused a dose-dependent decrease in cellular carnitine, short-chain and long-chain acylcarnitines (**Figure 4.7B**). At 10 μM , the lowest concentration of VB observed in the conventional mouse portal circulation, the observed carnitine peak intensity was half of the observed carnitine peak intensity in control cells. The effect was similar to effect observed with meldonium, a competitive inhibitor of membrane carnitine-reuptake and inhibitor of BBOX-mediated carnitine biosynthesis (**Figure 4.7C**). Propionylcholine, an ester with the identical elemental composition as VB, did not elicit a carnitine-lowering effect in HepG2 cells (**Figure 4.7C**). Cellular carnitine was restored by addition of carnitine to the cell media (**Figure 4.7D**). VB treatment also decreased cellular carnitine in CaCo2 cells (**Figure 4.7D**).

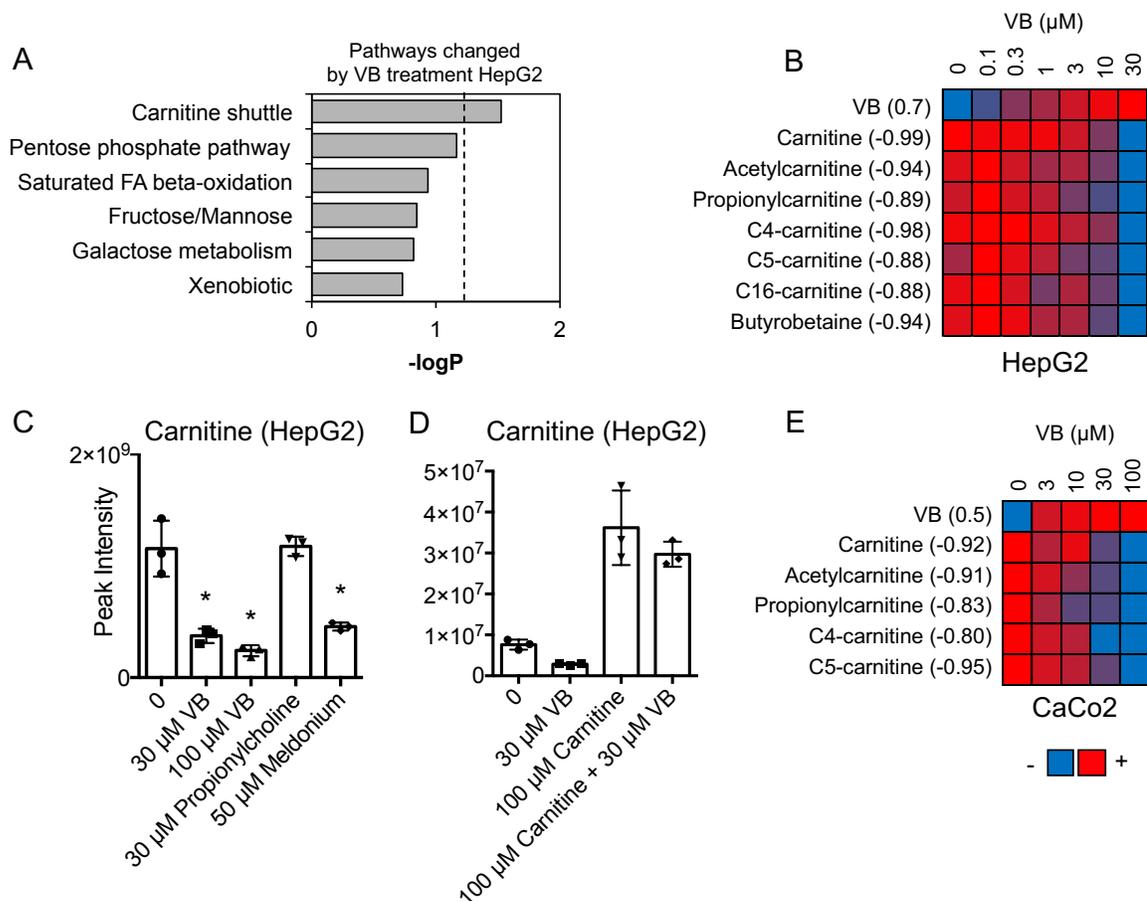


Figure 4.7 12 hour dose response study of VB in HepG2 cells. A) Untargeted metabolomic profiling of HepG2 cells treated with physiologically relevant concentrations of VB (0-30 μM , $n = 4$ biological replicates) shows VB treatment alters cellular carnitine shuttle metabolism (pathway level significance $p > 0.05$). B) VB treatment decreases cellular carnitine and acylcarnitines and the carnitine precursor butyrobetaine (lmreg $p > 0.05$). Statistical analyses were performed using the lmreg function in xMSpanda. Pearson regression coefficients are provided next to metabolite names (testing for association between treatment and metabolite abundance). C) VB decreases cellular carnitine ($n = 3$ biological replicates) in a similar fashion as the inhibitor of carnitine biosynthesis (BBOX) and uptake (OCTN2), meldonium. Propionylcholine, an ester with the same molecular weight as VB, did not elicit a carnitine lowering effect. Statistical analyses were performed using one-way ANOVA with Tukey's multiple comparison's test ($F = 38.81$, $p < 0.0001$). *Indicates different from control. Data are displayed as mean \pm standard deviation. D) VB-mediated loss of carnitine ($n = 3$ biological replicates) is reversible with added carnitine. Data are displayed as mean \pm standard deviation. E). VB treatment decreases cellular carnitine and acylcarnitines in CaCo2 cells ($n = 3$ biological replicates).

δ -valerobetaine (VB) decreases palmitate-dependent mitochondrial respiration in HepG2 cells

Carnitine is required for transport of long-chain fatty acyl chains into mitochondria for fatty acid oxidation. To test the effects of VB on mitochondrial function, we assessed the effects of VB treatment on basal respiratory rate, ATP production, spare capacity of HepG2 cells. Under culture conditions where cells were provided palmitate as a fuel source after overnight glucose, glutamine, and pyruvate deprivation, VB elicited a dose-dependent decrease in palmitate-dependent respiration following injection of the uncoupling agent FCCP (**Figure 4.8A**). The observed decrease in spare capacity (**Figure 4.8B**) for fatty acid oxidation in HepG2 cells was similar to that observed with etomoxir, an irreversible inhibitor of CPT1. Under culture conditions where glucose, glutamine, and pyruvate were present in the culture media, VB did not elicit major changes to the basal respiratory rate, ATP-production, spare capacity, or non-mitochondrial respiration (**Figure 4.8C, D**). This shows that the effect of VB on fatty acid oxidation in HepG2 cells is dependent on the availability of other fuel substrates.

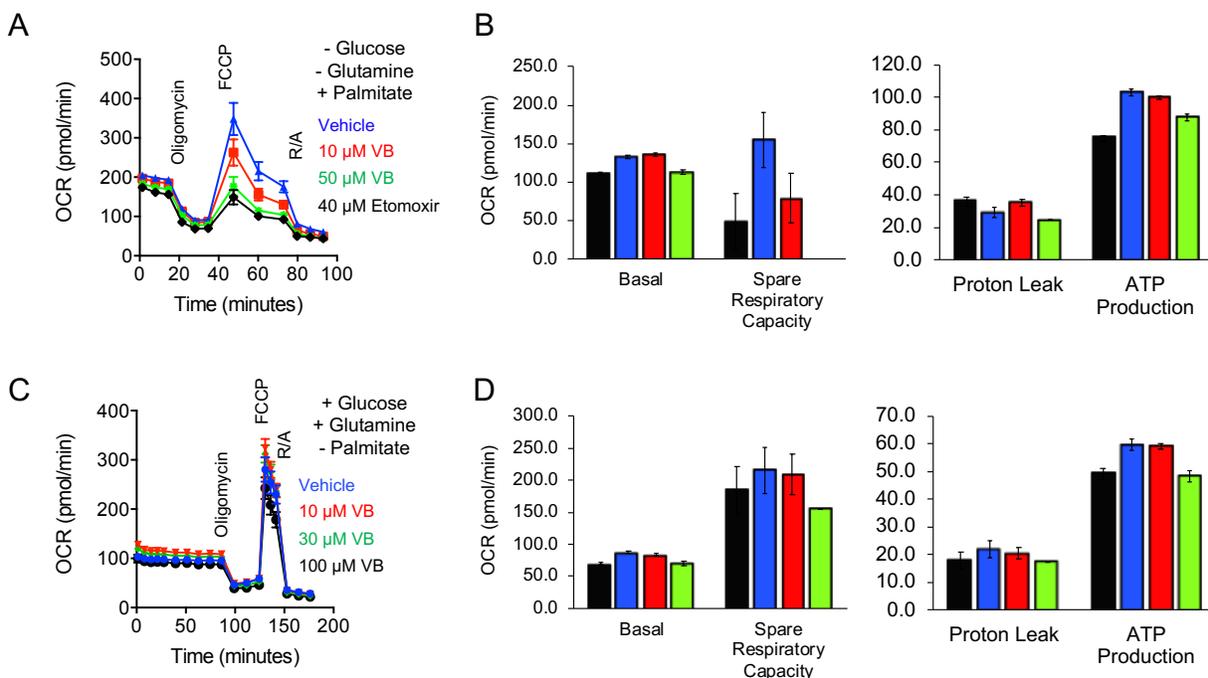


Figure 4.8 Mitochondrial respiration after 12 hour treatment with VB in HepG2 cells. A) VB elicits a concentration-dependent decrease in palmitate dependent mitochondrial respiration after FCCP treatment. B) Measures of basal respiration (oxygen consumption rate (OCR) prior to injection of test substrates), spare respiratory capacity (maximal OCR following FCCP injection – basal respiration), ATP production (basal respiration OCR – OCR post-oligomycin treatment), and proton leak (minimum OCR after oligomycin – non-mitochondrial OCR). VB treatment elicits a dose-dependent decrease in spare capacity in HepG2 cells. C) The oxygen consumption rate in serum starved HepG2 cells supplemented with glucose, glutamine and pyruvate is not altered by VB treatment. D) VB treatment does not elicit concentration-dependent effects on mitochondrial respiratory parameters in cells supplemented with glucose and glutamine. Each data point represents the 6-8 technical replicates \pm standard error of the mean. These data are representative of two independent experiments.

VB decreases formation of labeled acetyl-CoA production from labeled palmitic acid

Palmitate oxidation by mitochondria requires carnitine for uptake as palmitoylcarnitine, and the end product of mitochondrial beta-oxidation is acetyl-CoA. We performed a stable isotope tracer study to test whether added VB decreases the amount of $^{13}\text{C}_2$ -acetyl-CoA formed from $^{13}\text{C}_{16}$ -palmitate oxidation in HepG2 cells. 50 μM VB-pretreatment prior to the introduction of labeled palmitate decreased labeled $^{13}\text{C}_2$ -acetyl-CoA relative to vehicle (1h: 26%, 2h: 20%, 4h: 25%, 8h: 24% of vehicle). Carnitine supplementation following VB pretreatment restored cellular carnitine and restored the formation of labeled acetyl-CoA relative to VB pretreatment alone. VB added concurrently with labeled palmitate led to a moderate decrease in formation of labeled acetyl-CoA compared to vehicle treated cells (1h: 70%, 2h: 70%, 4h: 65%, 8h: 82% of vehicle). Etomoxir blocked formation of labeled acetyl-CoA (1h: 3%, 2h: 2%, 4h: 4%, 8h: 4% of vehicle). Thus, these results VB alone is sufficient to decrease carnitine-mediated mitochondrial fatty acid oxidation compared to cells without added VB. VB did not alter the uptake of cellular $^{13}\text{C}_{16}$ -palmitate or the formation of $^{13}\text{C}_{16}$ -palmitoyl-CoA (**Figure 4.9**).

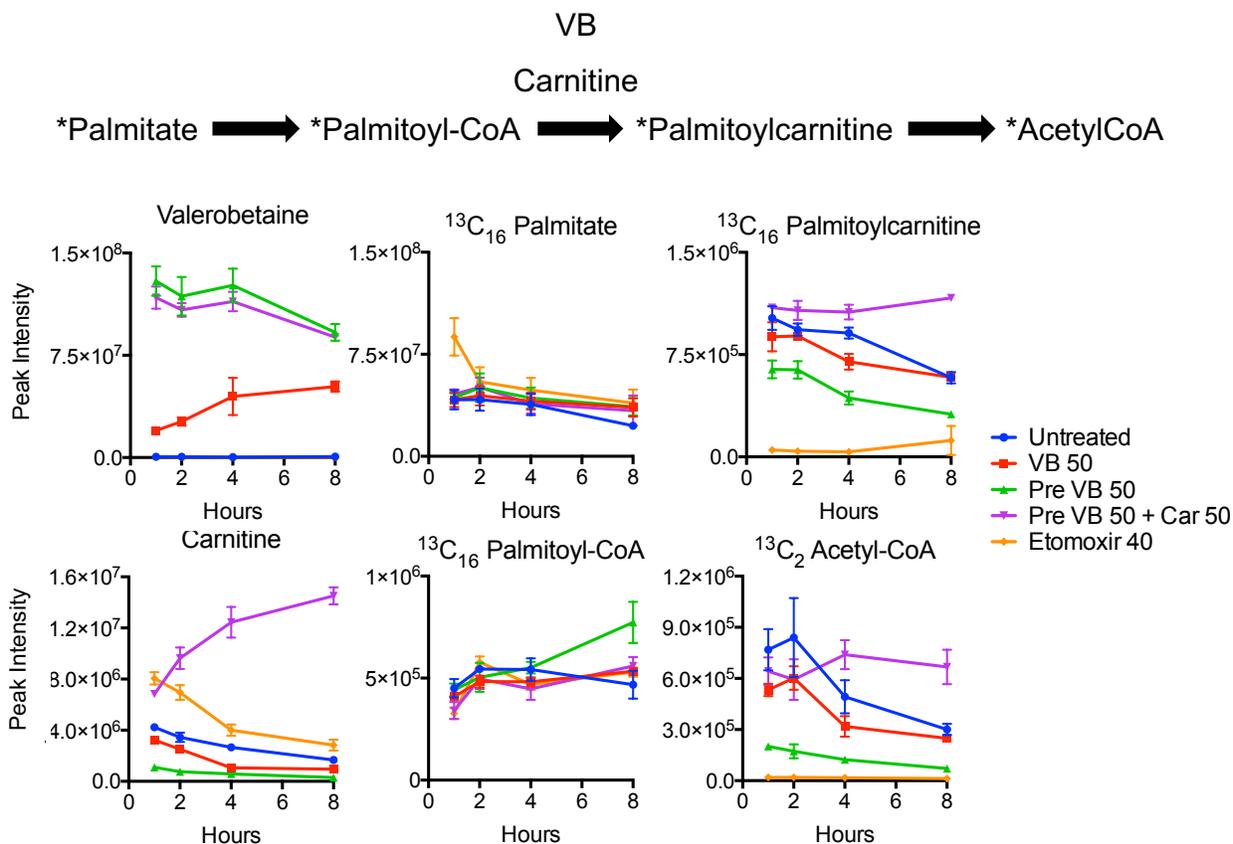


Figure 4.9 Stable isotope tracer study to examine the immediate effect of VB treatment on cellular mitochondrial fatty acid oxidation. 12 hour pretreatment with VB (green) decreased the formation of labeled acetyl-CoA (bottom right) by approximately 75% compared to vehicle (blue). Addition of carnitine back to cells pretreated with VB for 12 hours (purple) restored the carnitine-dependent formation of mitochondrial acetyl-CoA. Co-treatment of VB with the addition of isotope-labeled palmitate (red) decreased the formation of labeled acetyl-CoA by approximately 25% compared to vehicle. Data for VB, carnitine, and other metabolites are shown to illustrate VB treatment does not affect uptake of labeled $^{13}\text{C}_{16}$ palmitate (top middle), or the conjugation of labeled palmitate to CoA (bottom middle). VB treatment decreased carnitine approximately 20% after one hour (bottom left) and these changes drive decreased formation of labeled palmitoylcarnitine (top right) and labeled acetyl-CoA (bottom right). Each data point represents the average of 3 biological replicates \pm standard deviation.

VB decreases carnitine and decreases mitochondrial fatty acid oxidation in conventional mice

We used conventional mice to test whether supplemental VB inhibited mitochondrial fatty acid oxidation *in vivo*. First, we tested whether VB could deplete circulating and cellular carnitine in fed mice. After one week dosing (saline, 10 mg/kg VB (200 μ L of 10 mM), 100 mg/kg VB (200 μ L of 100 mM)), serum, liver, and urine were collected for mass spectrometry analysis. Results show that one-week at 10 mg/kg VB decreased serum carnitine in male mice from $71 \pm 13 \mu\text{M}$ (mean \pm standard deviation) to $52 \pm 6 \mu\text{M}$ and 100 mg/kg VB further decreased circulating carnitine to $17 \pm 4 \mu\text{M}$. In female mice, 10 mg/kg decreased serum carnitine from $56 \pm 11 \mu\text{M}$ (in vehicle treated mice) to $39 \pm 14 \mu\text{M}$ and 100 mg/kg VB further decreased serum carnitine to $17 \pm 4 \mu\text{M}$ (**Figure 4.10A**). In fasted mice, similar decreases in circulating and hepatic carnitine were observed (**Figure 4.11A**).

In fed mice, 10 mg/kg VB increased hepatic VB in male (2.5 fold-change) and female (4.7 fold-change) mice and 100 mg/kg VB increased VB in male (6.6 fold-change) and female (7.6 fold-change) mice compared to vehicle treatment. In fed mice, 10 mg/kg VB decreased hepatic carnitine in male (0.7 fold-change [not statistically significant]) and female (0.7 fold-change) mice and 100 mg/kg VB decreased hepatic carnitine in male (0.65 fold-change [not statistically significant]) and female (0.27 fold-change) mice compared to vehicle treatment (**Figure 4.10A**). In fasted mice, 10 mg/kg VB increased hepatic VB in male (2.5 fold-change) and female (4.7 fold-change) mice and 100 mg/kg VB increased VB in male (6.6 fold-change) and female (7.6 fold-change) mice compared to vehicle treatment. In fasted mice, 10 mg/kg VB decreased hepatic carnitine in male (0.5 fold-change) and female (0.4 fold-change) mice and 100 mg/kg VB decreased hepatic carnitine in male (0.41 fold-change) and female (0.32 fold-change) mice compared to vehicle treatment (**Figure 4.11A**). Oil Red O analysis showed a trending

difference in lipid accumulation with VB treatment under fed conditions (**Figure 4.10B, C**) but VB increased neutral lipid accumulation in both male and female mice after fasting (**Figure 4.11B, C**). Liver triglycerides were increased (60% in male mice and 100% in female mice) in fasted mice after VB treatment (**Figure 4.12A**). Circulating and hepatic beta-hydroxybutyrate were decreased in mice treated with VB after fasting (**Figure 4.12B**). Taken together, these data show that VB decreases circulating and hepatic carnitine, decreases hepatic fatty acid oxidation, causing a significant increase in hepatic triglycerides, and decrease in the formation of ketone bodies during fasting.

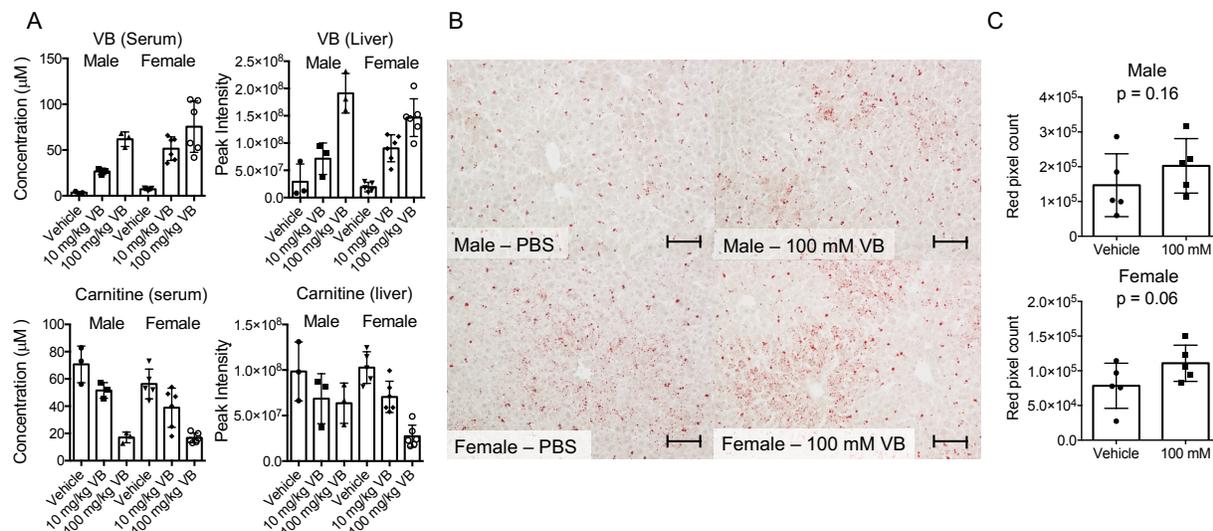


Figure 4.10 VB dose-response study in fed mice. A) VB treatment increases circulating and hepatic VB [Kruskal-Wallis analysis to determine group-wise differences with Dunn's multiple comparisons test (serum VB males – KW statistic 7.2, $p = 0.0036$; serum VB females – KW statistic 10.84, $p = 0.0008$; liver VB males – KW statistic 6.489, $p = 0.0107$; liver VB females – KW statistic 12.94, $p < 0.0001$)] and decreases circulating and hepatic carnitine in male ($n = 3$ per treatment) and female mice ($n = 5$ vehicle, 3 10 mg/kg, 6 100 mg/kg) after 1 week intraperitoneal administration (serum carnitine males – KW statistic 6.489, $p = 0.01$; serum carnitine females – KW statistic 11.42, $p = 0.0002$; liver carnitine males – KW statistic 1.689, $p = 0.51$; liver carnitine females – KW statistic 10.98, $p = 0.0004$). Control vs. 100 mg/kg was different for all comparisons, except for male liver carnitine (which was not significantly different). B) Representative Oil Red O stains of fed mouse liver with vehicle or VB treatment. Scale bar represents 100 μm under 20x magnification. C) Quantification of red pixel count in Oil Red O staining in fed mice ($n = 5$ per treatment). Data were analyzed using a one-tailed t-test ($p = 0.16$ for males, $p = 0.06$ for females) to test whether VB treatment increased neutral lipid accumulation in the liver under fed conditions. VB treatment does not worsen or improve lipid accumulation in the liver under fed conditions.

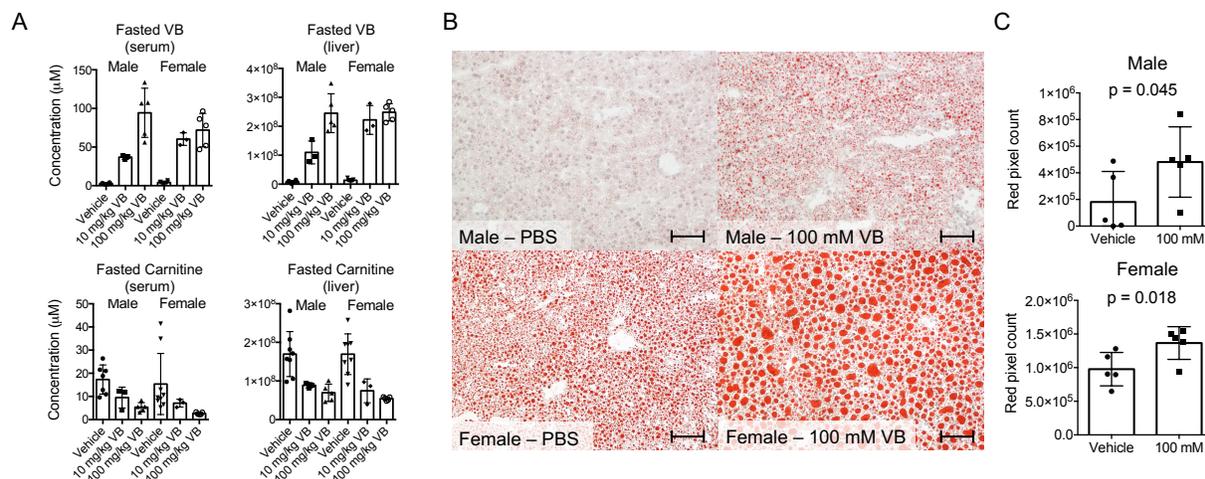


Figure 4.11 VB dose-response study in fasted mice. A) VB treatment increases circulating and hepatic VB and decreases circulating and hepatic carnitine in male ($n = 8$ vehicle, $n = 3$ 10 mg/kg, $n = 8$ 100 mg/kg) and female mice ($n = 8$ vehicle, $n = 3$ 10 mg/kg, $n = 8$ 100 mg/kg) after 3 days intraperitoneal administration. *Kruskal-Wallis analysis was used to determine group-wise differences and Dunn's multiple comparisons test was used to identify pairwise differences – male serum carnitine (KW statistic 9.021, $p = 0.0029$); female serum carnitine (KW statistic 10.98, $p = 0.0004$); male liver carnitine (KW statistic 11, $p = 0.0003$); female liver carnitine (KW statistic 10.83, $p = 0.0004$). Control vs 100 mg/kg was significantly different for all comparisons.* B) Representative Oil Red O stains of fasted mouse liver with vehicle or VB treatment. Scale bar represents 100 μm under 20x magnification. C) Quantification of red pixel count in Oil Red O staining in fed mice ($n = 5$ per treatment). Data were analyzed using a one-tailed t-test ($p = 0.045$ for males, $p = 0.018$ for females) to test whether VB treatment increased neutral lipid accumulation in the liver under fed conditions. VB treatment increases neutral lipid accumulation in the fasted state.

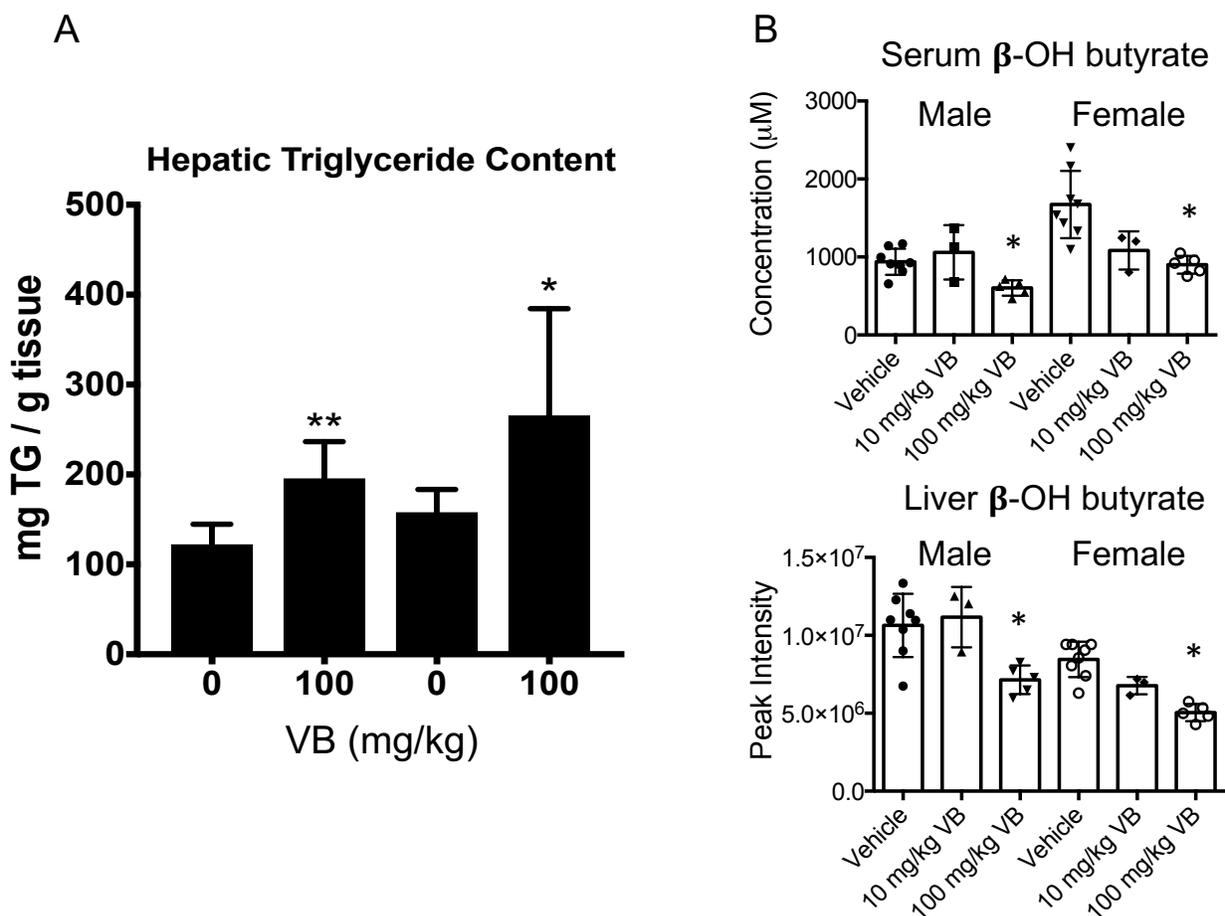


Figure 4.12 VB increases hepatic triglycerides after fasting and decreases ketone body formation during fasting in male and female mice. A) VB increases liver triglycerides after fasting ($n = 5$, one-tailed t-test $p < 0.05$ for males and females). B) Circulating and hepatic beta-hydroxybutyrate are decreased during fasting after VB treatment in male and female mice. Kruskal-Wallis with Dunn's multiple comparisons test was used for serum and liver analyses ($n = 8$ vehicle, $n = 3$ 10 mg/kg, $n = 5$ 100 mg/kg for male and female). Male serum KW statistic 7.864, $p < 0.0085$; female serum KW statistic 10.46, $p < 0.0006$. Male liver KW statistic 7.864, $p < 0.0085$; female liver KW statistic 11.73, $p < 0.0001$. Control vs. 100 mg/kg was significantly different for all comparisons.

VB is associated with the severity of hepatic steatosis in adolescents and with central adiposity in adults

We tested whether circulating VB is positively correlated with hepatic steatosis and central adiposity (visceral adipose tissue (VAT) weight) in humans. Data from two separate studies was used for this analysis – one focused on identification of plasma metabolites associated with the severity of hepatic steatosis (mild (<33% steatosis), moderate (33-66% steatosis), to severe (>66% steatosis)) in adolescents (*manuscript in submission*) with NAFLD and the other re-examined a previously published dataset (33) focused on identification of plasma metabolites associated with body composition parameters and hepatic insulin resistance (HOMA-IR) in fasted, healthy adults. The former study was from a cohort of 74 individuals (samples collected from 2007-2015) from the Emory University Pediatric Liver Biopsy Biorepository with clinically diagnosed NAFLD via liver biopsy. The latter study was comprised of individuals enrolled in the Emory-Georgia Tech Center for Health Discovery and Well-Being (CHDWB) cohort (samples collected from 2007-2013) for whom baseline plasma HRM, insulin resistance and HOMA by fasting blood glucose and insulin were measured by HOMA-IR (34) and body composition by dual energy X-ray absorptiometry (DEXA). In adolescents with clinically diagnosed NAFLD (n = 74, mean age 14, Males = 54, Females = 20), VB was positively correlated with severe steatosis ($\beta = 0.345$, $p < 0.02$). In fasted adults (n = 179, mean age 50 years, Males = 63, Females = 116), plasma VB was positively correlated with VAT ($\beta = 3.7E+04 \pm 1.1E+04$, $p = 0.0006$), independent of age, race, and sex. This relationship held when additionally controlling for total body fat ($\beta = 1.6E+04 \pm 7.2E+03$, $p = 0.03$). Plasma VB levels were significantly, positively related to HOMA-IR ($\beta = 2.4E4 \pm 8.6E3$, $p = 0.007$),

independent of age, race, and sex (**Figure 4.13**). Furthermore, plasma VB was positively correlated with age (data not shown).

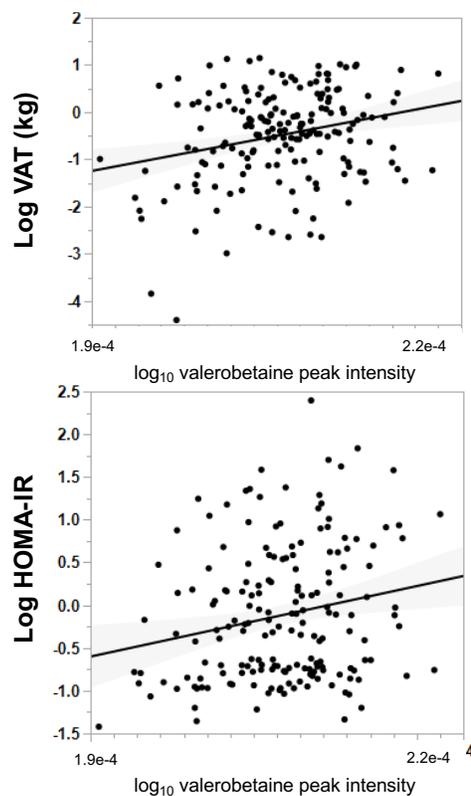
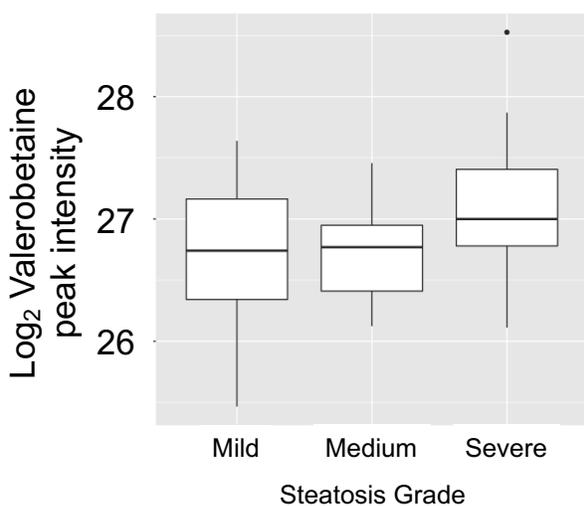


Figure 4.13 Clinical phenotypes associated with circulating VB. *Left*) Plasma VB is correlated with the severity of hepatic steatosis ($n = 74$, $\beta = 0.345$, $p < 0.02$) in adolescents. *Right*) Plasma VB is correlated with increased central adiposity ($n = 179$, $p < 0.0006$), and hepatic insulin resistance (HOMA-IR) ($p < 0.0001$).

4.4 Discussion

Phylogenetic and evolutionary evidence indicate that eukaryotic mitochondria are descended from an ancient bacterial endosymbiont (35), and there is increasing recognition of the holobiont – the eukaryotic host and its associated microbiome – function as the biological unit subject to evolutionary pressures. The present results suggest that VB is a microbiome-derived metabolite through which intestinal microbes communicate with the host mitochondria to influence collective bioenergetics. VB regulates fatty acid oxidation through modulation of systemic and hepatic carnitine abundance, which decreases long-chain fatty acid transport and metabolism by mitochondria. This may allow enhanced utilization of short-chain fatty acids by the eukaryotic host since short-chain fatty acids do not require the carnitine shuttle to be oxidized by host mitochondria (36). This, in turn, could benefit members of the intestinal microbiota since fermentation byproducts of the microbiome need to be removed to maintain microbial energetics, and accumulation of which has been shown to be toxic to microbes (37). Additionally, evolutionarily conserved benefits of microbiome-mitochondria cross-talk could involve energy conservation. In calorie-restricted states, microbial production of VB could promote survival of the holobiont/symbiont by functioning as a brake on host fatty acid oxidation, preserving collective nutritional resources. In support of this, germ-free mice succumb more quickly following prolonged fasting periods compared to conventional mice (38). Thus, these results establish a novel molecular communication between the microbiome and mitochondria, which could underlie aspects of microbiome-host symbiosis.

A decrease in fatty acid oxidation and resultant lipid accumulation mediated by VB could also drive an adaptive response termed mitohormesis (39), a process by which a minor impairment increases gene expression of mitochondrial systems which improve metabolic function of mitochondria. The comparative analysis of GF/CV mouse liver transcriptomes shows the microbiome, possibly due to the activity of VB, increases the expression of many genes for mitochondrial bioenergetic and lipid homeostasis (**Figure 4.2**). In support of this, lipid accumulation drives activity of SIRT1, PGC-1alpha and PPAR-alpha pathways (40) which control carnitine homeostasis (41-43), mitochondrial bioenergetics, and lipid metabolism (44). Because PPAR-alpha expression declines with age (45), this adaptive response could be attenuated with aging and underlie the observed associations of VB with central adiposity in adults.

Recent studies also show VB can be obtained from dietary sources (46-48), and recently published reports describe potential beneficial effects of VB on human health. VB is produced by ruminant microbes in water buffalo and may contribute to increased acylcarnitines and nutritional value of water buffalo milk (48; 49). Whole grain diets, which are associated with decreased risk of cardiometabolic disorders, type 2 diabetes, and weight gain (50), increased circulating VB in humans (46). Whole grains do not contain VB but have pre-biotic effects on gut microbial composition (50) (e.g. increasing *Bacteroidetes/Firmicutes* ratio) and increases in *Lactobacillus* and *Bifidobacterium*. These changes are negatively correlated with markers of obesity and dyslipidaemia, and suggest that microbial products (such as VB) produced by beneficial diets and probiotic-associated bacteria could elicit beneficial effects on human metabolic health.

VB regulation of carnitine pools could also impact health because alterations to carnitine homeostasis are common to metabolic disorders of humans. Carnitine is a central regulator of glucose and lipid metabolism (51; 52), and previous studies show chemicals that decrease cellular carnitine cause mitochondrial dysfunction, decreases in hepatic fatty acid oxidation, and hepatic toxicities including steatosis (53-57). Previous studies show a decline in free carnitine were accompanied by an accumulation of medium and long-chain acylcarnitines in muscle and associated with obesity in aging (58). Furthermore, carnitine supplementation has been shown to be beneficial for obesity and fatty liver and increase glucose utilization in humans and mice (59-63). However, excessive fatty acid oxidation and the accumulation of incompletely oxidized acylcarnitines are also associated with obesity and insulin resistance (64). For example, excessive fatty acid oxidation (and decreased glucose utilization) is observed in insulin-resistant heart tissues (65). Limiting carnitine-dependent fatty acid oxidation under these conditions can improve metabolic flexibility by driving an increase in glucose utilization (the Randle cycle). While VB was observed in brains, lungs, and hearts of conventional mice, the contribution of the microbial metabolite VB to extrahepatic metabolism is not currently known.

In conclusion, the present results show that δ -valerobetaine, a structural analogue of the carnitine precursor γ -butyrobetaine, is a gut microbe-derived small molecule that orchestrates microbiome-mitochondrial cross talk to influence mitochondrial energy metabolism. A key focus on prebiotic and probiotic management of VB production is needed to guide management of the critical microbiome-host symbiotic relationship that is needed to conquer the current epidemic of human metabolic disorders.

References

1. Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. Dysbiosis of the gut microbiota in disease.
2. Holmes E, Li JV, Marchesi JR, Nicholson JK. 2012. Gut microbiota composition and activity in relation to host metabolic phenotype and disease risk. *Cell Metab* 16:559-64
3. Dumas MB, R. H.; Toye, A.; Cloarec, O.; Blancher, C.; Rothwell, A.; Fearnside, J.; Tatoud, R.; Blanc, V.; Lindon, J. C.; Mitchell, S. C.; Holmes, E.; McCarthy, M. I.; Scott, J.; Gauguier, D., Nicholson, J. K. 2006. Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *PNAS* 103:12511-6
4. Rosenbaum M, Knight R, Leibel RL. 2015. The gut microbiota in human energy homeostasis and obesity. *Trends Endocrinol Metab* 26:493-501
5. Ignacio A, Fernandes MR, Rodrigues VA, Groppo FC, Cardoso AL, et al. 2016. Correlation between body mass index and faecal microbiota from children. *Clin Microbiol Infect* 22:258 e1-8
6. Gonzalez FJ, Jiang C, Patterson AD. 2016. An Intestinal Microbiota-Farnesoid X Receptor Axis Modulates Metabolic Disease. *Gastroenterology* 151:845-59
7. Holmes E, Li JV, Athanasiou T, Ashrafian H, Nicholson JK. 2011. Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. *Trends Microbiol* 19:349-59
8. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. 2016. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14:273-87
9. Cani PD, Amar J, Iglesias MA, Poggi M, Knauf C, et al. 2007. Metabolic Endotoxemia Initiates Obesity and Insulin Resistance. *Diabetes* 56:1761-72
10. Cani PD, Bibiloni R, Knauf C, Waget A, Neyrinck AM, et al. 2008. Changes in Gut Microbiota Control Metabolic Endotoxemia-Induced Inflammation in High-Fat Diet-Induced Obesity and Diabetes in Mice. *Diabetes* 57:1470-81
11. Hoyles L, Fernandez-Real JM, Federici M, Serino M, Abbott J, et al. 2018. Molecular phenomics and metagenomics of hepatic steatosis in non-diabetic obese women. *Nat Med* 24:1070-80
12. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, et al. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472:57-63
13. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, et al. 2013. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 19:576-85

14. Koeth RA, Levison BS, Culley MK, Buffa JA, Wang Z, et al. 2014. gamma-Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab* 20:799-812
15. Wikoff WRA, A. T.; Liu, J.; Schultz, P. G.; Lesley, S. A.; Peters, E. C.; Siuzdak, G. 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *PNAS* 106:3698-703
16. Peisl BYL, Schymanski EL, Wilmes P. 2018. Dark matter in host-microbiome metabolomics: Tackling the unknowns-A review. *Anal Chim Acta* 1037:13-27
17. Bajpai P, Darra A, Agrawal A. 2018. Microbe-mitochondrion crosstalk and health: An emerging paradigm. *Mitochondrion* 39:20-5
18. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, et al. 2011. The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* 13:517-26
19. Mollica MP, Mattace Raso G, Cavaliere G, Trinchese G, De Filippo C, et al. 2017. Butyrate Regulates Liver Mitochondrial Function, Efficiency, and Dynamics in Insulin-Resistant Obese Mice. *Diabetes* 66:1405-18
20. Yardeni T, Tanes CE, Bittinger K, Mattei LM, Schaefer PM, et al. 2019. Host mitochondria influence gut microbiome diversity: A role for ROS. *Science Signaling* 12:eaaw3159
21. Go YM, Uppal K, Walker DI, Tran V, Dury L, et al. 2014. Mitochondrial metabolomics using high-resolution Fourier-transform mass spectrometry. *Methods in molecular biology* 1198:43-73
22. Liu KH, Walker DI, Uppal K, Tran V, Rohrbeck P, et al. 2016. High-Resolution Metabolomics Assessment of Military Personnel: Evaluating Analytical Strategies for Chemical Detection. *J Occup Environ Med* 58:S53-61
23. Yu T, Jones DP. 2014. Improving peak detection in high-resolution LC/MS metabolomics data using preexisting knowledge and machine learning approach. *Bioinformatics* 30:2941-8
24. Yu T, Park Y, Johnson JM, Jones DP. 2009. apLCMS--adaptive processing of high-resolution LC/MS data. *Bioinformatics* 25:1930-6
25. Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, et al. 2013. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC bioinformatics* 14:15
26. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, et al. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 43:e47-e
27. Li S, Park Y, Duraisingham S, Strobel FH, Khan N, et al. 2013. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol* 9:e1003123
28. Smith CA, O'Maille G, Want EJ, Want EJ, Qin C, Qin C, Trauger SA, Trauger SA, Brandon TR, et al. METLIN: a metabolite mass spectral database.
29. Ruttkies C, Schymanski EL, Wolf S, Hollender J, Neumann S. 2016. MetFrag relaunched: incorporating strategies beyond in silico fragmentation. *J Cheminform* 8:3-
30. Go Y-M, Walker DI, Liang Y, Uppal K, Soltow QA, et al. 2015. Reference Standardization for Mass Spectrometry and High-Resolution Metabolomics Applications to Exposome Research. *Toxicological Sciences*

31. Tars K, Leitans J, Kazaks A, Zelencova D, Liepinsh E, et al. 2014. Targeting Carnitine Biosynthesis: Discovery of New Inhibitors against γ -Butyrobetaine Hydroxylase. *Journal of Medicinal Chemistry* 57:2213-36
32. C. M. Chen F, Leo Benoiton N. 2011. *A new method of quaternizing amines and its use in amino acid and peptide chemistry*. 3310-1 pp.
33. Bellissimo MP, Cai Q, Ziegler TR, Liu KH, Tran PH, et al. 2019. Plasma High-Resolution Metabolomics Differentiates Adults with Normal Weight Obesity from Lean Individuals. *Obesity* 27:1729-37
34. Wallace TM, Levy, J., Matthews, D.R. 2004. Use and Abuse of HOMA Modeling. *Diabetes Care* 27:1487-95
35. Sagan L. 1967. On the Origin of Mitosing Cells. *J. Theoret. Biol.* 14:225-74
36. Schönfeld PW, L. 2016. Short- and medium-chain fatty acids in the energy metabolism - the cellular perspective. *J Lipid Res* 57:943-54
37. Royce LA, Liu P, Stebbins MJ, Hanson BC, Jarboe LR. 2013. The damaging effects of short chain fatty acids on Escherichia coli membranes. *Appl Microbiol Biotechnol* 97:8317-27
38. Tennant B Fau - Malm OJ, Malm Oj Fau - Horowitz RE, Horowitz Re Fau - Levenson SM, Levenson SM. Response of germfree, conventional, conventionalized and E. coli monocontaminated mice to starvation.
39. Barcena C, Mayoral P, Quiros PM. Mitohormesis, an Antiaging Paradigm.
40. Khan SA, Sathyanarayan A, Mashek MT, Ong KT, Wollaston-Hayden EE, Mashek DG. 2015. ATGL-Catalyzed Lipolysis Regulates SIRT1 to Control PGC-1 α /PPAR- α Signaling. *Diabetes* 64:418
41. Rupp H, Rupp TP, Maisch B. 2005. Fatty acid oxidation inhibition with PPARalpha activation (FOXIB/PPARalpha) for normalizing gene expression in heart failure? *Cardiovasc Res* 66:423-6
42. van Vlies N, Ferdinandusse S, Turkenburg M, Wanders RJA, Vaz FM. 2007. PPAR α -activation results in enhanced carnitine biosynthesis and OCTN2-mediated hepatic carnitine accumulation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1767:1134-42
43. Ringseis R, Wen G, Eder K. 2012. Regulation of Genes Involved in Carnitine Homeostasis by PPARalpha across Different Species (Rat, Mouse, Pig, Cattle, Chicken, and Human). *PPAR Res* 2012:868317
44. Montagner A, Polizzi A, Fouche E, Ducheix S, Lippi Y, et al. Liver PPARalpha is crucial for whole-body fatty acid homeostasis and is protective against NAFLD.
45. Iemitsu M, Miyauchi T Fau - Maeda S, Maeda S Fau - Tanabe T, Tanabe T Fau - Takanashi M, Takanashi M Fau - Irukayama-Tomobe Y, et al. Aging-induced decrease in the PPAR-alpha level in hearts is improved by exercise training.
46. Karkkainen O, Lankinen MA, Vitale M, Jokkala J, Leppanen J, et al. Diets rich in whole grains increase betainized compounds associated with glucose metabolism.
47. Yasumoto T, Shimuzu N. 1977. Identification of δ -Valerobetaine and Other Betaines in the Ovary of a Shellfish *Callista brevisiphonata*. *NIPPON SUISAN GAKKAISHI* 43:201-6

48. Servillo L, D'Onofrio N, Giovane A, Casale R, Cautela D, et al. 2018. Ruminant meat and milk contain delta-valerobetaine, another precursor of trimethylamine N-oxide (TMAO) like gamma-butyrobetaine. *Food Chem* 260:193-9
49. Servillo L, D'Onofrio N, Neglia G, Casale R, Cautela D, et al. 2018. Carnitine Precursors and Short-Chain Acylcarnitines in Water Buffalo Milk. *J Agric Food Chem* 66:8142-9
50. Ye EQ, Chacko Sa Fau - Chou EL, Chou El Fau - Kugizaki M, Kugizaki M Fau - Liu S, Liu S. Greater whole-grain intake is associated with lower risk of type 2 diabetes, cardiovascular disease, and weight gain.
51. Ringseis R, Keller J, Eder K. 2012. Role of carnitine in the regulation of glucose homeostasis and insulin sensitivity: evidence from in vivo and in vitro studies with carnitine supplementation and carnitine deficiency. *Eur J Nutr* 51:1-18
52. Fritz IB. 1959. Action of carnitine on long chain fatty acid oxidation by liver. *Am. J. Physiol.* 197:297-304
53. Spaniol M, Kaufmann P, Beier K, Wüthrich J, Török M, et al. 2003. Mechanisms of liver steatosis in rats with systemic carnitine deficiency due to treatment with trimethylhydraziniumpropionate. *Journal of Lipid Research* 44:144-53
54. Spaniol MB, H.; Auer, L.; Zimmermann, A.; Solioz, M.; Stieger, B.; Krähenbühl, S. 2001. Development and characterization of an animal model of carnitine deficiency. *Eur. J. Biochem.* 268:1876-87
55. Krähenbühl SM, G.; Kupferschmidt, H.; Meier, P. J.; Krause, M. 1995. Plasma and hepatic carnitine and coenzyme A pools in a patient with fatal, valproate induced hepatotoxicity. *Gut* 37:140-3
56. Luef GJ, Waldmann M Fau - Sturm W, Sturm W Fau - Naser A, Naser A Fau - Trinkä E, Trinkä E Fau - Unterberger I, et al. Valproate therapy and nonalcoholic fatty liver disease.
57. Lheureux PE, Hantson P. Carnitine in the treatment of valproic acid-induced toxicity.
58. Noland RC, Koves Tr Fau - Seiler SE, Seiler Se Fau - Lum H, Lum H Fau - Lust RM, Lust Rm Fau - Ilkayeva O, et al. Carnitine insufficiency caused by aging and overnutrition compromises mitochondrial performance and metabolic control.
59. Bianchi PBL, D. C.; Davis, A. T. 1996. Carnitine Supplementation Ameliorates the Steatosis and Ketosis Induced by Pivalate in Rats. *J Nutr* 126:2873-9
60. Asai T, Okumura K, Takahashi R, Matsui H, Numaguchi Y, et al. 2006. Combined therapy with PPARalpha agonist and L-carnitine rescues lipotoxic cardiomyopathy due to systemic carnitine deficiency. *Cardiovasc Res* 70:566-77
61. Ishikawa H, Takaki A, Tsuzaki R, Yasunaka T, Koike K, et al. 2014. L-carnitine prevents progression of non-alcoholic steatohepatitis in a mouse model with upregulation of mitochondrial pathway. *PLoS One* 9:e100627
62. Jiang F, Zhang Z, Zhang YI, Wu J, Yu LI, Liu SU. 2016. L-carnitine ameliorates the liver inflammatory response by regulating carnitine palmitoyltransferase I-dependent PPARγ signaling. *Molecular Medicine Reports* 13:1320-8
63. Libert DM, Nowacki AS, Natowicz MR. 2018. Metabolomic analysis of obesity, metabolic syndrome, and type 2 diabetes: amino acid and acylcarnitine levels change along a spectrum of metabolic wellness. *PeerJ* 6
64. Koves TR, Ussher JR, Noland RC, Slentz D, Mosedale M, et al. 2008. Mitochondrial overload and incomplete fatty acid oxidation contribute to skeletal muscle insulin resistance. *Cell Metab* 7:45-56

65. Lopaschuk GD. 2016. Fatty Acid Oxidation and Its Relation with Insulin Resistance and Associated Disorders. *Ann Nutr Metab* 68 Suppl 3:15-20

Chapter 5: Conclusions

This dissertation focused on the identification of chemical signals produced by the gut microbiome that interact with targets in the liver to regulate systemic energy balance. In this dissertation, I provided an introduction to the health issues of microbiome-host symbiosis. I optimized (chapter 2) and validated (chapter 3) high-throughput chemical analyses for identifying metabolites from the microbiome with potential to impact host energy metabolism. In chapter 4, I identified, synthesized, and characterized a previously uncharacterized mitochondrial metabolite which is produced by the microbiome. Here, in the concluding chapter, I summarize my findings and provide perspective for future work in this area.

5.1 - Summary

This research improves understanding of how the microbiome communicates with host metabolism in balancing fatty acid and energy metabolism. I developed an experimental pipeline (**Figure 5.1**) to identify microbiome-derived metabolites and characterize their impact on host mitochondrial function and energy metabolism. Development of optimized metabolite detection methods and extension of quantification methodology with a library of identified metabolites provided a foundation for the identification and quantification of unidentified metabolites. Application of validated methods and the ability to use mechanistic information derived from other systems biology layers allows hypothesis-driven experiments to provide experimental validation for interaction networks that are mapped out between the microbiome and host. Finally, use of these interaction networks aids interpretation of observed associations between specific microbial metabolites and clinical phenotypes.

In **Figure 5.2**, I summarize the results from chapter 4 to provide a VB interaction network with liver mitochondria and hepatic energy metabolism. Integrated analysis of liver transcriptomics with metabolomics with experimental results from chapter 4 shows the top transcriptional and metabolic pathways changing in response to the microbiome could be changing as a result of microbial production of VB. As discussed in Chapter 4, symbiosis of microbes and host involves mutually beneficial outcomes; in the context of energy metabolism, VB appears to provide an important message to mitochondria to manage use of long-chain and short-chain fatty acids for the benefit of the symbiotic microbiome-host unit.

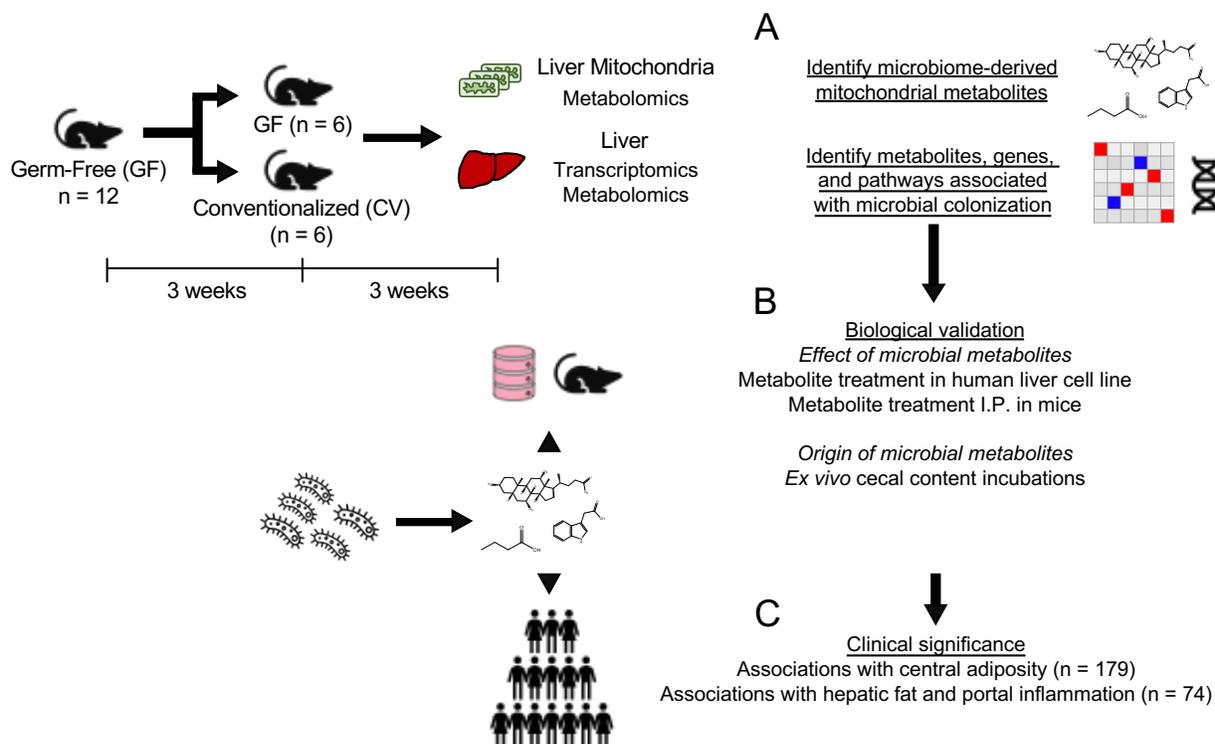


Figure 5.1 A simplified experimental pipeline for understanding interaction networks between the microbiome and host. A) Identification of microbial metabolites and associated biological responses B) Identify the specific contribution of individual microbial metabolites to the global systems level response observed in host tissues C) Test for associations of microbiome-derived metabolites with clinical phenotypes.

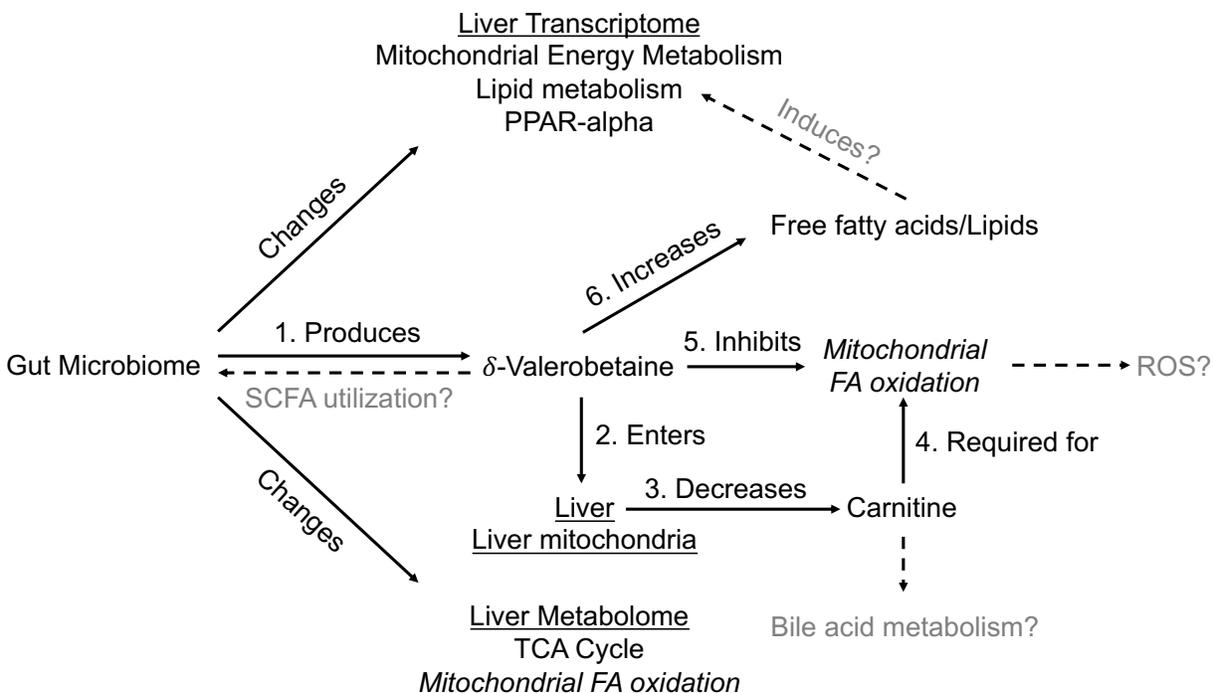


Figure 5.2 Working model of the VB-mitochondrial interaction network. The microbiome elicits major changes to the hepatic metabolome and transcriptome relating to mitochondrial energy metabolism and lipid metabolism. The partial inhibition of FA oxidation elicited by the microbiome-derived mitochondrial metabolite VB increases hepatic lipids after fasting and could increase the expression of genes related to mitochondrial energy metabolism and lipid metabolism indirectly through activation of the transcription factor PPAR-alpha. Preliminary data show that VB-treatment led to a 2-fold increase in PPRE-linked luciferase signal. Other proposed mechanisms of bi-directional crosstalk with the microbiome include promoting the utilization of short-chain fatty acids over long-chain fatty acids in the intestinal lumen. Previous studies show carnitine decreases the bile acid pool which can impact microbiome-host cross talk mediated by FXR in the intestinal lumen.

5.2 – Preliminary data and future directions

Metabolic fate of VB

There are no data describing the metabolic fate of VB in the literature. Because VB is a 5-carbon analogue of the carnitine precursor butyrobetaine, we hypothesized that VB could be a precursor for a 5-carbon analogue of carnitine, homocarnitine. A 5-carbon analogue of carnitine could serve as an additional carrier of acyl groups in and out of mitochondria. To characterize the metabolic fate of VB, I analyzed extracts with LC-HRMS/MS from liver microsomal fractions (S9) incubated with VB, as well as drosophila and mice treated with VB (**figure 5.3**). My data show that 176.1281 m/z , the theoretical mass for homocarnitine, increases in a dose-dependent manner in flies and mice and is produced by liver S9 fractions in the presence of alpha-ketoglutarate, iron, and ascorbate. Ion dissociation spectra of 176.1281 m/z shows a mass shift of +14.0152 m/z from the 103.0395 m/z observed for carnitine to 117.0547 m/z for homocarnitine corresponding to the additional CH₂ group. The MS/MS spectra is characteristic of a 5-carbon analogue of carnitine. However, MS/MS analysis was not sufficient to determine the position of the hydroxyl group either at the 3 or 4 position. Homocarnitine is only detected in conventional mouse liver and is absent in GF mice. Future work will need to be done to fully characterize the structure of this metabolite (e.g. use of a bioreactor to scale up S9 reactions with VB as a precursor), determine the enzymes responsible for production, whether it is detectable in humans, and to identify its biological significance.

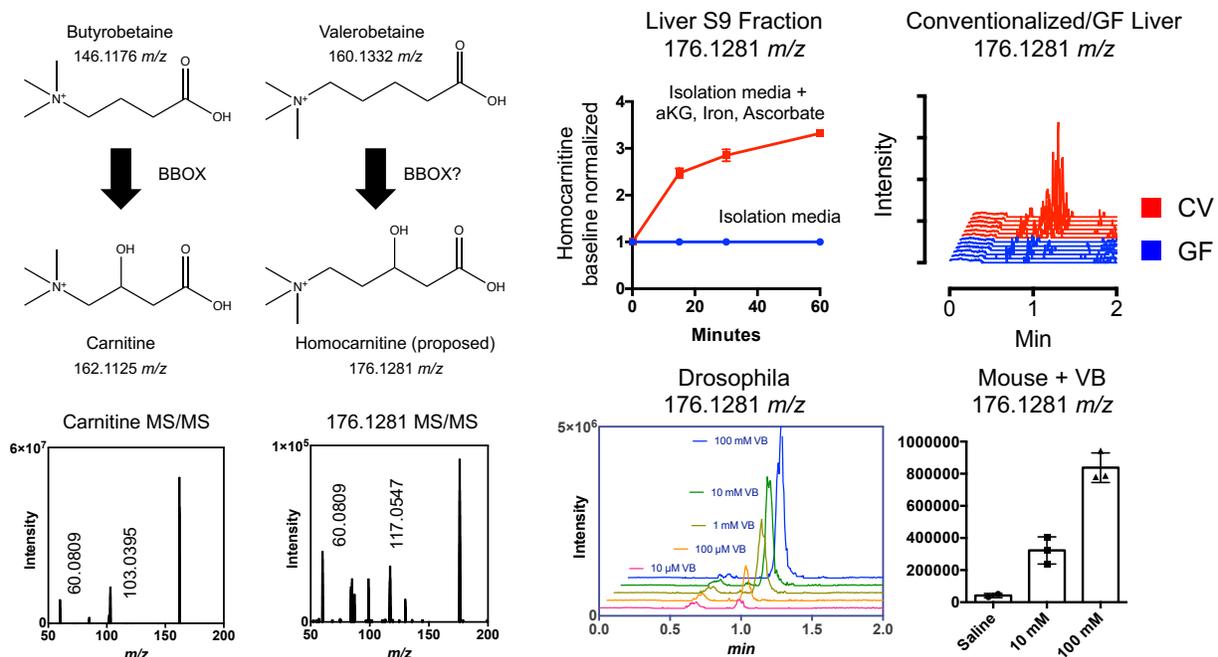


Figure 5.3 Formation of a 5-carbon analogue of carnitine, homocarnitine from VB. (Left) Incubation of liver and kidney S9 fractions with butyrobetaine produces carnitine (162.1125 m/z), a reaction mediated by gamma-butyrobetaine dioxygenase (BBOX1). Incubation of liver and kidney S9 fractions with valerobetaine produces 176.1281 m/z . The accurate mass and MS/MS spectra indicate that VB is a precursor for a 5-carbon analogue of carnitine, homocarnitine. This reaction requires alpha-ketoglutarate, iron and ascorbate and does not proceed in the absence of these co-factors. In drosophila and mice that were administered VB, we observed a dose-dependent increase in homocarnitine. Furthermore, homocarnitine is present in conventionalized mice but absent in GF mice. Further work is needed to elucidate the function of this 5-carbon analogue of carnitine that has not been previously reported in the literature.

Presence and activity of VB in other tissues

While VB was absent in GF mice, it was detected in lung, brain, and heart (data not shown) samples collected from conventional mice (**figure 5.4**). Because each organ preferentially utilizes different fuel substrates, the impact of VB in cellular metabolism in these tissues represent an additional opportunities to improve our understanding of host-microbiome interactions. Our data shows that VB decreases carnitine in brain and heart of mice. While carnitine deficiencies are also linked to hypoglycemia, myopathy, and cardiomyopathy, the consequences of decreased systemic carnitine and fatty acid oxidation elicited by VB are not currently known. In addition to these questions, future studies should examine whether VB treatment increases the size of adipose depots or increases weight gain in mice.

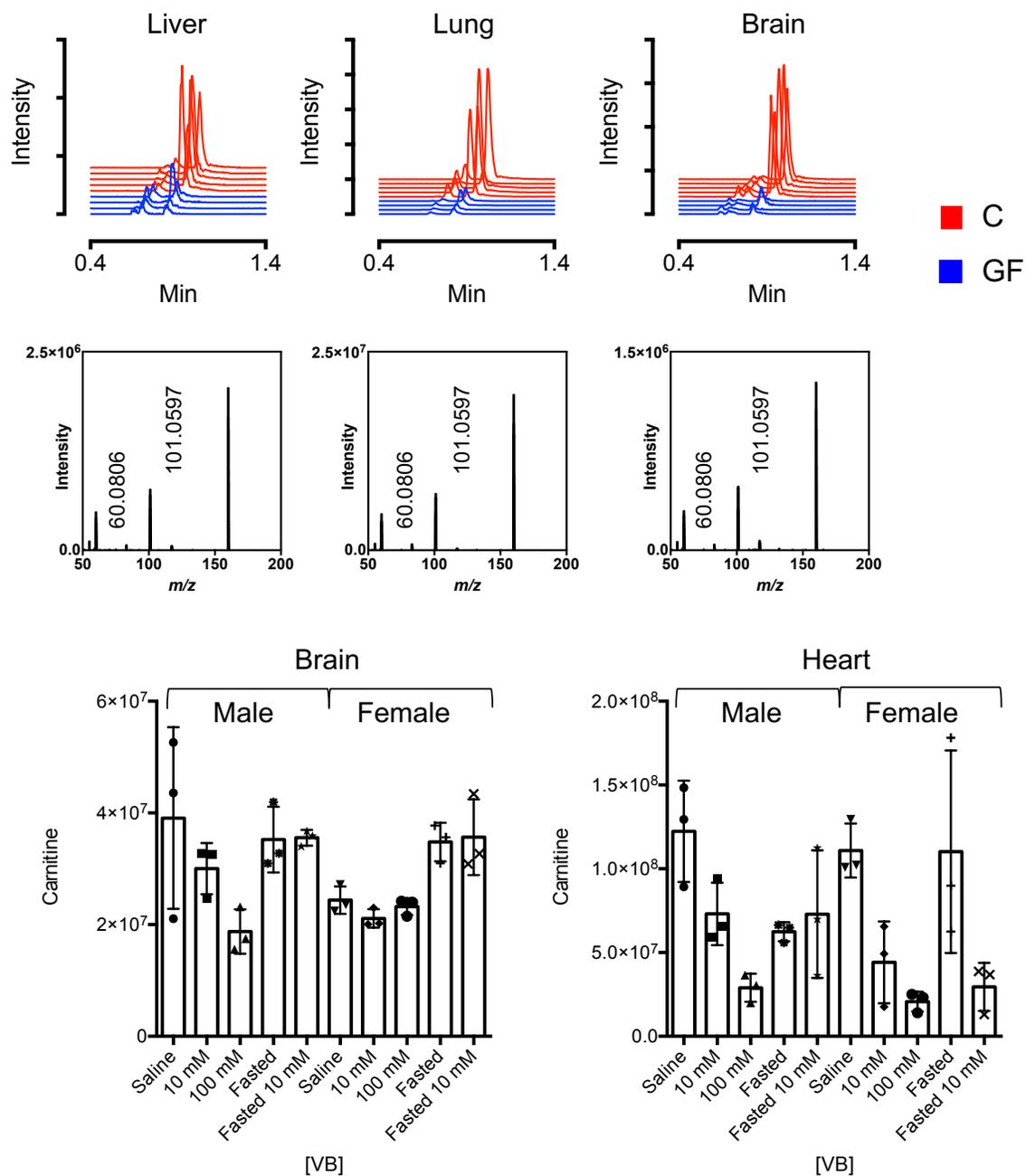


Figure 5.4 Top) VB is present in liver, lung and brain of conventional (C) mice but absent in GF mouse tissues. Bottom) VB treatment decreases brain and heart carnitine, but the function of VB in tissues other than the liver are not unknown.

5.3 – Conclusion

Overall, the work presented in this dissertation establishes a chemical analysis platform suitable for detection of known and unidentified metabolites and provides evidence for a contribution of the microbiome-derived metabolite VB to mitochondrial function in the liver and systemic energy metabolism. Experimental data shows VB decreases the rate of mitochondrial fatty acid oxidation in the liver and can increase the accumulation of hepatic fat. Epidemiologic surveys show that plasma VB is moderately associated with the severity of hepatic steatosis in adolescents and central adiposity in adults. Taken together, this work suggests that microbial production of VB may be a contributor to hepatic steatosis and the eventual development of fatty liver disease. The results of this dissertation are strongly suggestive of a causative link between VB and the characteristic accumulation of fat in host tissues associated with hepatic steatosis and other obesity-related disorders. Because of the importance of identifying novel therapeutic strategies to combat the epidemic of obesity-related metabolic disorders, the insights gained from this dissertation and future studies on this topic could drive the development of targeted therapies directed towards microbiome-mitochondrial signaling to improve metabolic health in humans.

Appendix 1 (from Chapter 3). Validated metabolites (441) and estimated concentrations of metabolites in Qstd3 (214), NIST1950 (198), and Chear (205) pooled plasma reference materials. Concentrations were estimated by either (1) Reference standardization (RS) to calibrated (by method of standards addition) Qstd3 accurate mass MS1 (2) RS to NIST1950 published value (3) Method of standards addition to Qstd3 using diagnostic fragment ion (4) RS using most abundant expected metabolite in plasma (5) Relative peak intensities.

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
1	Glycine	120.0032 (M+2Na-H)	75	215 ± 7*	103 ± 23*	105 ± 13*	1	HMDB0000123	C00037
2	Alanine	90.0550 (M+H)	75	258 ± 17	330 ± 9	439 ± 8	1	HMDB0000161	C00041
3	Beta-alanine*	90.0550 (M+H)	75	-	-	-	-	HMDB0000056	C00099
4	Sarcosine*	90.0550 (M+H)	75	-	-	-	-	HMDB0000271	C00213
5	Serine	106.0499 (M+H)	87	99 ± 5	73 ± 6	144 ± 8	1	HMDB0000187	C00065
6	Proline	116.0706 (M+H)	65	128 ± 2	90 ± 2.4	90 ± 1.1	2	HMDB0000162	C00148
7	Valine	118.0863 (M+H)	52	159.7 ± 2	100 ± 2	124 ± 1.2	2	HMDB0000883	C00183
8	Threonine	120.0655 (M+H)	77	113 ± 4	90 ± 1	114 ± 5	1	HMDB0000167	C00188
9	Leucine	132.1018 (M+H)	48	100	-	-	-	HMDB0000687	C00123
10	Isoleucine	132.1018 (M+H)	48	56	-	-	-	HMDB0000172	C00407
[9,10]	Leu/Iso	132.1018 (M+H)	48	156 ± 4	144 ± 3	202 ± 1	2	-	-
11	Asparagine	133.0608 (M+H)	91	33.4 ± 1	9 ± 0.1	12 ± 0.3	2	HMDB0000168	C00152
12	Aspartate	134.0448 (M+H)	85	1.2 ± 0.1	1.9 ± 0.1	4.6 ± 0.02	1	HMDB0000191	C00049
13	Glutamine	147.0764 (M+H)	87	301 ± 14	147 ± 6	140 ± 3	1	HMDB0000641	C00064
14	Glutamate	146.0458 (M-H)	20	76 ± 3	89 ± 4	148 ± 5	1	HMDB0000148	C00025
15	Lysine	145.0983 (M-H)	69	123 ± 9	51 ± 2	113 ± 10	1	HMDB0000182	C00047
16	Methionine	150.0583 (M+H)	53	31.7 ± 1	23 ± 0.3	21.3 ± 1	1	HMDB0000696	C00073
17	Histidine	154.0622 (M-H)	23	60 ± 1	61 ± 1	74 ± 7	2	HMDB0000177	C00135
18	Phenylalanine	166.0863 (M+H)	48	54.8 ± 1.4	55.7 ± 1.2	60.7 ± 1.7	1	HMDB0000159	C00079
19	Tyrosine	182.0812 (M+H)	49	58.9 ± 0.9	54.3 ± 1	68.3 ± 0.54	2	HMDB0000158	C00082
20	Tryptophan	203.0826 (M-H)	26	48 ± 0.8	50.3 ± 2.4	58 ± 2.3	1	HMDB00030396	C00078
21	Arginine	173.1044 (M-H)	23	69 ± 4	78 ± 4.5	49 ± 2	1	HMDB0000517	C00062
22	Cysteine*	-	-	-	-	-	-	HMDB0000574	C00097
23	Cystine	241.0311 (M+H)	211	87 ± 13	55 ± 4	34.2 ± 3.5	1	HMDB0000192	C00491
24	Creatine	132.0768 (M+H)	55	32 ± 2	43 ± 1	61 ± 7	2	HMDB0000064	C00300
25	Creatinine	114.0662 (M+H)	40	60 ± 2	57 ± 1	44 ± 0.3	2	HMDB0000562	C00791
26	Creatine Phosphate	212.0431 (M+H)	111	-	0.87 ± 0.15	0.12 ± 0.07	1	HMDB0001511	C02305
27	Guanidinoacetate	118.0611 (M+H)	70	1.7 ± 0.06	2.2 ± 0.08	1.5 ± 0.1	1	HMDB0000128	C00581
28	Guanidinosuccinate	176.0666 (M+H)	74	0.06 ± 0.003	0.05 ± 0.005	0.035 ± 0.003	1	HMDB0003157	C03139
29	Guanidinobutanote	146.0924 (M+H)	38	0.48 ± 0.02*	0.3 ± 0.03*	0.139 ± 0.03	1	HMDB0003464	C01035
30	Taurine	126.0219 (M+H)	63	32.4 ± 0.6	10 ± 0.3	39 ± 3	2	HMDB0000251	C00245
31	Hypotaurine	110.0270 (M+H)	77	0.9 ± 0.04*	0.4 ± 0.01	3.1 ± 0.08	1	HMDB0000965	C00519
32	Pyroglutamate (Oxoproline)	130.0499 (M+H)	40	16 ± 2	11 ± 0.4	37 ± 1	1	HMDB0000267	C01879
33	Methionine Sulfoxide	166.0532 (M+H)	88	1 ± 0.04	1 ± 0.04	1.4 ± 0.1	1	HMDB0002005	C02989
34	S-adenosylmethionine*	399.1445 (M+H)	133	-	-	-	-	HMDB0001185	C00019
35	S-adenosylhomocysteine	385.1294 (M+H)	80	0.015 ± 0.001	0.023 ± 0.005	0.01 ± 0.002	1	HMDB0000939	C00021
36	Cysteamine*	153.0515 (2M-H)	76	-	-	-	-	HMDB0002991	C01678
37	Cystathionine	223.0752 (M+H)	197	0.069 ± 0.01	0.029 ± 0.002	0.032 ± 0.003	1	HMDB0000099	C02291
38	Homocysteine*	136.0432 (M+H)	138	-	-	-	-	HMDB0000742	C00155
39	Homocystine*	269.0625 (M+H)	146	-	-	-	-	HMDB0000676	C01817
40	N-acetylcysteine*	164.0376 (M+H)	43	-	-	-	-	HMDB0001890	C06809
41	Nitrotyrosine	227.0663 (M+H)	53	-	-	-	-	HMDB0001904	-
42	Methionine sulfoximine	181.0642 (M+H)	91	-	-	-	-	HMDB0029430	C03510
43	Formyl methionyl peptide	178.0532 (M+H)	34	-	-	-	-	-	-
44	Ophthalmic acid	290.1347 (M+H)	80	0.005	0.004	0.004	1	HMDB0005765	-
45	Glutathione*	308.0911 (M+H)	80	-	-	-	-	HMDB0000125	C00051

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
46	Glutathione disulfide*	613.1593 (M+H)	270	-	-	-	-	HMDB0003337	C00127
47	Asymmetric dimethylarginine (ADMA)*	203.1503 (M+H)	79	-	-	-	-	HMDB0001539	C03626
48	Symmetric dimethylarginine (SDMA)*	203.1503 (M+H)	79	-	-	-	-	HMDB0003334	-
[47,48]	Dimethylarginine*	203.1503 (M+H)	79	86096714	47976032	60016920	5	-	-
49	Agmatine*	131.1296 (M+H)	74	-	-	-	-	HMDB0001432	C00179
50	Carnosine	227.1139 (M+H)	107	-	0.52 ± 0.04	0.34 ± 0.05	1	HMDB0000033	C00386
51	Anserine*	241.1295 (M+H)	102	-	0.014 ± 0.002	0.087 ± 0.01	1	HMDB0000194	C01262
52	1-Methylhistidine*	170.0924 (M+H)	92	-	-	-	-	HMDB0000001	C01152
53	3-Methylhistidine*	170.0924 (M+H)	102	-	-	-	-	HMDB0000479	C01152
[52,53]	Methylhistidine	170.0924 (M+H)	97	9 ± 0.8	4.6 ± 0.5	3.7 ± 0.1	4(52)	-	C01152
54	Histidinol	142.0975 (M+H)	74	-	-	-	-	HMDB0003431	C00860
55	Histamine	112.0869 (M+H)	74	0.015 ± 0.0007	0.01 ± 0.001*	0.02 ± 0.002	1	HMDB0000870	C00388
56	1-Methylhistamine	126.1026 (M+H)	73	-	-	-	-	HMDB0000898	C05127
57	3-Methylhistamine	126.1026 (M+H)	73	-	-	-	-	HMDB0001861	-
[56,57]	Methylhistamine*	126.1026 (M+H)	73	-	-	-	-	-	-
58	Urocanate	139.0502 (M+H)	38	0.24 ± 0.02	0.07 ± 0.003	0.3 ± 0.03	1	HMDB0000301	C00785
59	Thyrotropin Releasing Hormone*	361.1630 (M-H)	24	-	-	-	-	HMDB0060080	-
60	Diiodothyronine*	523.8861 (M-H)	31	-	-	-	-	HMDB0000582	-
61	(3,5)-Diiodotyrosine*	433.8744 (M+H)	47	-	-	-	-	HMDB0003474	C01060
62	Triiodothyronine*	651.7973 (M+H)	38	-	-	-	-	HMDB0000265	C02465
63	Thyroxine	775.6794 (M-H)	72	0.089 ± 0.003	0.026 ± 0.005	0.13 ± 0.015	1	HMDB0000248	C01829
[62,63]	3-Methoxytyrosine	212.0917 (M+H)	55	0.0024 ± 0.0011	0.001 ± 0.0001	0.0004 ± 0.0001	1	HMDB0001434	-
64	Tyramine	138.0913 (M+H)	31	-	-	-	-	HMDB0000306	C00483
65	Phenylethanolamine	138.0913 (M+H)	32	-	-	-	-	HMDB0001065	C02735
[64,65]	Tyramine/Phenylethanolamine	138.0913 (M+H)	32	0.045 ± 0.0014*	0.032 ± 0.003*	0.036 ± 0.005*	1	-	-
66	3-Methoxytyramine*	168.1018 (M+H)	30	-	-	-	-	HMDB0000022	C05587
67	4-Methoxytyramine*	168.1018 (M+H)	30	-	-	-	-	HMDB0012162	-
[66,67]	Methoxytyramine (3 or 4)	168.1018 (M+H)	30	0.13 ± 0.005	0.11 ± 0.02	0.084 ± 0.006	1	-	-
68	Normetanephrine*	184.0968 (M+H)	32	-	-	-	-	HMDB0000819	C05589
69	Epinephrine*	184.0968 (M+H)	32	-	-	-	-	HMDB0000068	C00788
[68,69]	Epinephrine/Normetanephrine	184.0968 (M+H)	32	0.09 ± 0.02*	0.03 ± 0.008*	0.06 ± 0.01*	1	-	-
70	Octopamine	154.0863 (M+H)	33	0.06 ± 0.006*	0.07 ± 0.007*	0.05 ± 0.006*	1	HMDB0004825	C04227
71	Dopamine*	154.0863 (M+H)	33	-	-	-	-	HMDB0000073	C03758
[70,71]	Octopamine/Dopamine	154.0863 (M+H)	33	0.06 ± 0.006*	0.07 ± 0.007*	0.05 ± 0.006*	1	-	-
72	L-dopa*	198.0761 (M+H)	62	-	-	-	-	HMDB0000181	C00355
73	Acetylphenylalanine	208.0968 (M+H)	30	0.6 ± 0.024	0.34 ± 0.035	0.19 ± 0.1	1	HMDB0000512	C03519
74	Kynurenine	209.0921 (M+H)	50	1.07 ± 0.03	1.01 ± 0.04	1.29 ± 0.03	1	HMDB0000684	C00328
75	Hydroxykynurenine*	225.0870 (M+H)	55	-	-	-	-	HMDB0000732	C02794
76	Kynurenic acid	190.0499 (M+H)	48	0.025 ± 0.006*	0.014 ± 0.002*	0.01 ± 0.008*	1	HMDB0000715	C01717
77	Xanthurenic acid*	206.0448 (M+H)	50	-	-	-	-	HMDB0000881	C02470
78	5-Hydroxytryptophan	221.0921 (M+H)	59	-	-	-	-	HMDB0000472	C01017
79	Serotonin	177.1022 (M+H)	31	-	-	-	-	HMDB0000259	C00780
80	N-Acetylserotonin	219.1128 (M+H)	30	0.86 ± 0.08	1.02 ± 0.07	0.4 ± 0.14	1	HMDB0001238	C00978
81	5-Hydroxyindoleacetic acid	192.0657 (M+H)	31	0.017 ± 0.005*	0.067 ± 0.003	0.057 ± 0.009	1	HMDB0000763	C05635
82	Indole-3-acetaldehyde	158.0611 (M-H)	60	1 ± 0.02	0.2 ± 0.008	0.21 ± 0.01	1	HMDB0001190	C00637
83	Indoleacetate	174.0561 (M-H)	27	1.2 ± 0.12	1.3 ± 0.09	0.97 ± 0.03	1	HMDB0000197	C00954

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
84	Indole 3-acetamide	175.0866 (M+H)	31	-	-	-	-	HMDB0029739	C02693
85	Methyl indoleacetate	190.0863 (M+H)	29	-	-	-	-	HMDB0029738	-
86	Indole propionate	190.0863 (M+H)	-	-	-	-	-	HMDB0002302	-
[85,86]	MIA/IPA	190.0863 (M+H)	29	4.9 ± 0.4	2.8 ± 0.25	3.2 ± 0.2	4(84)	-	-
87	Acetyltryptophan	247.108 (M+H)	30	0.377 ± 0.016	0.333 ± 0.08	0.289 ± 0.03	1	HMDB0013713	-
88	Melatonin	233.1285 (M+H)	29	0.0002 ± 6e-5	0.00006 ± 4e-6	0.00004 ± 6e-6	1	HMDB0001389	C01598
89	N ₆ ,N ₆ ,N ₆ -Trimethyllysine	189.1598 (M+H)	107	0.36 ± 0.01	0.35 ± 0.02	0.39 ± 0.02	1	HMDB0001325	C03793
90	Hydroxylysine	163.1077 (M+H)	145	0.32 ± 0.03	0.17 ± 0.003	0.19 ± 0.014	1	HMDB0000450	C16741
91	Acetyllysine	189.1234 (M+H)	76	0.06 ± 0.0005	0.16 ± 0.006	0.15 ± 0.01	1	HMDB0000446	C12989
92	Pipecolate	130.0863 (M+H)	61	13 ± 0.2	11 ± 0.1	8.3 ± 0.2	1	HMDB0000070	C00408
93	α-Aminoadipate	162.0761 (M+H)	74	0.05 ± 0.006*	2.1 ± 0.06*	1.6 ± 0.5	1	HMDB0000510	C00956
94	Ornithine	133.0972 (M+H)	110	98 ± 5	50 ± 2	86 ± 2	1	HMDB0000214	C00077
95	Citrulline	176.103 (M+H)	91	18 ± 1.5	10.7 ± 1.8	16.4 ± 2.9	1	HMDB0000904	C00327
96	5-Aminovaleric acid	118.0863 (M+H)	39	-	-	-	-	HMDB0003355	C00431
97	Diaminopimelate	189.0881 (M-H)	19	-	2.1 ± 0.06*	1.6 ± 0.5	1	HMDB0001370	C00666
98	Sulfinoalanine	154.0169 (M+H)	120	-	-	-	-	HMDB0000996	C00606
99	Hydroxyproline	132.0655 (M+H)	77	7 ± 0.2	8 ± 0.7	7 ± 0.1	1	HMDB0000725	C01157
100	5-Aminolevulinat	132.0655 (M+H)	45	-	-	-	-	HMDB0001149	C00430
101	N-Acetyllalanine	132.0655 (M+H)	36	-	-	-	-	HMDB0000766	-
102	N-Acetylglucine	118.0499 (M+H)	37	3 ± 0.14*	2.3 ± 0.5*	1.6 ± 0.8*	1	HMDB0000532	-
103	N-Acetylleucine	174.1125 (M+H)	31	0.2 ± 0.015*	0.1 ± 0.02*	0.067 ± 0.02*	1	HMDB0011756	C02710
104	N-Acetylaspartate	174.0408 (M-H)	21	0.07 ± 0.002	0.14 ± 0.008*	0.2 ± 0.027*	1	HMDB0000812	C01042
105	N-Methyl-D-aspartate	148.0604 (M+H)	77	-	-	-	-	HMDB0002393	C12269
106	O-Acetylserine	148.0604 (M+H)	84	-	-	-	-	HMDB0003011	C00979
107	N-Acetylserine	148.0604 (M+H)	42	-	-	-	-	HMDB0002931	-
108	N-Methyl-glutamate	162.0761 (M+H)	74	-	-	-	-	HMDB0062660	C01046
109	Acetylglutamate	188.0564 (M-H)	21	0.01 ± 0.001	0.001 ± 0.0003*	0.003 ± 0.0004*	1	HMDB0001138	C00624
110	2-Aminobutyrate	104.0706 (M+H)	70	1.2 ± 0.01	1 ± 0.01	1.2 ± 0.05	4	HMDB0000452	C02356
111	GABA (4-Aminobutyrate)	104.0706 (M+H)	45	0	0	0.21	3	HMDB0000112	C00334
112	3-Aminoisobutanoate	104.0706 (M+H)	45	-	-	-	-	HMDB0003911	C05145
113	Dimethylglycine	104.0706 (M+H)	-	-	-	-	-	HMDB0000092	C01026
114	Acetamidobutanoate	146.0812 (M+H)	34	0.014 ± 0.001*	0.002 ± 0.001*	0.006 ± 0.003*	1	HMDB0003681	C02712
115	Acetylmethionine	190.0543 (M-H)	22	0.002 ± 0.0008*	0.006 ± 0.003*	0.022 ± 0.006*	1	HMDB0011745	C05335
116	Selenomethionine	198.0028 (M+H)	53	-	-	-	-	HMDB0003966	C00263
117	Selenocystamine	248.9409 (M+H)	74	-	-	-	-	-	-
118	Homoserine	120.0655 (M+H)	77	-	-	-	-	HMDB0000719	C00022
119	1-Aminocyclopropane-1-carboxylate	102.0550 (M+H)	60	2 ± 0.2*	0.2 ± 0.04*	0.3 ± 0.1*	1	HMDB0036458	C01234
120	Norvaline	118.0863 (M+H)	52	-	-	-	-	HMDB0013716	C01826
121	Pyruvate	87.0087 (M-H)	20	57 ± 3	101 ± 1	64 ± 3	1	HMDB0000243	C00186
122	Lactate	89.0244 (M-H)	22	2404 ± 20	2633 ± 125	7506 ± 161	2	HMDB0001311	C00164
123	Acetoacetate	101.0244 (M-H)	22	281 ± 6*	324 ± 1*	239 ± 15*	1	HMDB0000060	C00164
124	2-Ketobutyric acid	101.0244 (M-H)	22	-	-	-	-	HMDB0000005	C00109
125	Citric acid	191.0197 (M-H)	19	191/268	-	182	1	HMDB0000094	C00158
126	Isocitric acid	191.0197 (M-H)	19	-	-	-	-	HMDB0000193	C00311
127	(Cis/Trans)-Aconitate	173.0091 (M-H)	19	7.5 ± 0.5*	124 ± 1.1	19 ± 0.24	1	HMDB0000958	C02341
128	α-Ketoglutarate (Oxoglutarate)	145.0142 (M-H)	21	19 ± 0.7	26 ± 0.5	43 ± 2	1	HMDB0000208	C00026

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
129	Succinic acid	117.0193 (M-H)	21	3 ± 0.2	4.4 ± 0.2	19 ± 2	1	HMDB0000254	C00042
130	Methylmalonic acid	117.0193 (M-H)	21	-	-	-	-	HMDB0000202	C02170
131	Fumaric acid	115.0037 (M-H)	21	11 ± 0.11	5 ± 0.3	27 ± 3	1	HMDB0000134	C00122
132	Malic acid	133.0143 (M-H)	21	0.9 ± 0.1	1.8 ± 0.1	8.4 ± 0.13	1	HMDB0000156	C00149
133	α-OH-butyrate	103.0401 (M-H)	23	-	-	-	-	HMDB0000008	C05984
134	β-OH butyrate	103.0401 (M-H)	23	260 ± 16	176 ± 2	138 ± 5	2	HMDB0000357	C01089
135	Glyceric acid	105.0193 (M-H)	20	2.2 ± 0.05	2.1 ± 0.1	2.7 ± 0.2	1	HMDB0000139	C00258
136	Oxovaleric acid	115.0401 (M-H)	20	-	-	-	-	HMDB0001865	C06255
137	α-Ketoisovalerate (Ketovaline)	115.0401 (M-H)	20	-	-	-	-	HMDB0000019	C00141
[136,137]	Oxovalerate/Ketoisovalerate	115.0401 (M-H)	20	5028167	4160375	2108007	5	-	-
138	α-Ketocaproate (Ketoleucine)	129.0557 (M-H)	24	-	-	-	-	HMDB0000695	C00233
139	α-Keto-β-methyl-valerate (Ketoisoleucine)	129.0557 (M-H)	24	-	-	-	-	HMDB0000491	C03465
[139,140]	Ketoleucine/Ketoisoleucine	129.0557 (M-H)	24	29279018	32075354	25674455	5	-	-
141	Hydroxymethylglutarate	161.0456 (M-H)	23	2.3 ± 0.03	3.8 ± 0.1	1.5 ± 0.1	1	HMDB0000355	C03761
142	Mevalonic acid	147.0663 (M-H)	22	0.096 ± 0.008*	0.049 ± 0.007*	0.048 ± 0.02*	1	HMDB0000227	C00418
143	Sorbate	113.0597 (M+H)	30	2.6 ± 0.2*	1.3 ± 0.08	1.9 ± 0.3	1	HMDB0029581	-
144	2-Methylmaleate	129.0193 (M-H)	22	4.3 ± 0.02	28 ± 0.5	4 ± 0.04	1	HMDB0000634	C02226
145	Glutaric acid	131.0350 (M-H)	25	6 ± 0.1	6 ± 0.2*	2 ± 0.3*	1	HMDB0000661	C00489
146	Ethylmalonic acid	131.0350 (M-H)	21	-	-	-	-	HMDB0000622	-
147	Adipate	145.0506 (M-H)	25	-	-	-	-	HMDB0000448	C06104
148	Oxoadipic acid	159.0299 (M-H)	21	-	-	-	-	HMDB0000225	C00322
149	Hydroxyadipic acid	161.0455 (M-H)	-	-	-	-	-	HMDB0000321	C02360
150	Dehydroshikimate	171.0299 (M-H)	20	-	-	-	-	-	C02652
151	Shikimic acid	173.0456 (M-H)	18	-	-	-	-	HMDB0003070	C00493
152	Suberate	173.0819 (M-H)	25	-	-	-	-	HMDB0000893	C08278
153	Azelaic acid	187.0976 (M-H)	25	0.012 ± 0.003*	0.04 ± 0.007*	0.01 ± 0.005*	1	HMDB0000784	C08261
154	Benzoate	121.0295 (M-H)	23	2.6 ± 0.3	44 ± 4*	12 ± 2*	1	HMDB0001870	C00180
155	2-Hydroxybenzoate (Salicylic acid)	137.0244 (M-H)	24	-	-	-	-	HMDB0001895	C00805
156	3-Hydroxybenzoate	137.0244 (M-H)	24	-	-	-	-	HMDB0002466	C00587
157	4-Hydroxybenzoate	137.0244 (M-H)	24	-	-	-	-	HMDB0000500	C00156
[155-157]	Hydroxybenzoate	137.0244 (M-H)	24	77 ± 1	17 ± 0.2	0.3 ± 0.1	4(152)	-	-
158	2,4-Dihydroxybenzoate	153.0193 (M-H)	24	-	-	-	-	HMDB0029666	C00230
159	3,5-Dihydroxybenzoate	153.0193 (M-H)	24	-	-	-	-	HMDB0013677	-
160	2,6-Dihydroxybenzoate	153.0193 (M-H)	24	-	-	-	-	HMDB0013676	-
161	2,3-Dihydroxybenzoate (2-pyrocatechuate)	153.0193 (M-H)	24	-	-	-	-	HMDB0000397	C00196
162	3,4-Dihydroxybenzoate (Protocatechuic acid)	153.0193 (M-H)	24	-	-	-	-	HMDB0001856	C00230
163	2,5-Dihydroxybenzoate (Gentisic acid)	153.0193 (M-H)	24	-	-	-	-	HMDB0000152	C00628
[158-163]	Dihydroxybenzoate	153.0193 (M-H)	24	1.3 ± 0.06	0.9 ± 0.05	1.1 ± 0.007	4(158)	-	-
164	2-Phenylacetate	135.0452 (M-H)	26	25 ± 1.4*	32 ± 2.6*	35 ± 2.3	1	HMDB0000209	C07086
165	Mandelic acid	151.0401 (M-H)	24	-	-	-	-	HMDB0000703	C01984
166	2-Methoxybenzoic acid	151.0401 (M-H)	-	-	-	-	-	HMDB0032604	-
167	3-Methoxybenzoic acid	151.0401 (M-H)	-	-	-	-	-	HMDB0032606	-
168	p-Hydroxyphenylacetic acid	151.0401 (M-H)	24	-	-	-	-	HMDB0000020	C00642
169	o-Hydroxyphenylacetic acid	151.0401 (M-H)	24	-	-	-	-	HMDB0000669	C05852
170	m-Hydroxyphenylacetic acid	151.0401 (M-H)	25	-	-	-	4(165)	HMDB0000440	C05593
171	Vanillin	151.0401 (M-H)	-	-	-	-	-	HMDB0012308	C00755

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
[166-171]	Hydroxyphenylacetate/Methoxybenzoate	151.0401 (M-H)	25	1.5 ± 0.03*	2.3 ± 0.2*	2.2 ± 0.4*	1	-	-
172	Resorscinol monoacetate	153.0546 (M+H)	30	-	-	-	-	-	C12064
173	Methylparaben	153.0546 (M+H)	31	-	-	-	-	HMDB0032572	-
174	Dihydroxyacetophenone	153.0546 (M+H)	32	-	-	-	-	-	-
175	Vanillic acid	167.0350 (M-H)	-	-	-	-	-	HMDB0000484	C06672
176	Methylvanillate	183.0652 (M+H)	30	-	-	-	-	HMDB0240266	-
177	m-Hydroxymandelic acid	167.0350 (M-H)	-	-	-	-	-	HMDB0000750	-
178	p-Hydroxymandelic acid	167.0350 (M-H)	-	-	-	-	-	HMDB0000822	C11527
179	Homogentisate	167.0350 (M-H)	25	-	-	-	-	HMDB0000130	C00544
180	Cinnamic acid	147.0448 (M-H)	30	0.14 ± 0.02*	0.18 ± 0.018*	0.3 ± 0.035*	1	HMDB0000567	C10438
181	Coumarate	163.0401 (M-H)	25	0.5 ± 0.06*	0.45 ± 0.03*	0.6 ± 0.19*	1	HMDB0002035	C00811
182	Caffeic acid	181.0495 (M-H)	31	-	-	-	-	HMDB0001964	C01481
183	4-Ethylbenzoic acid	149.0608 (M-H)	-	-	-	-	-	HMDB0002097	C00156
182	Hydrocinnamic acid	149.0608 (M-H)	36	486337	298139	399234	5	HMDB0000764	C05629
183	2-Phenylpropionate	149.0608 (M-H)	-	-	-	-	-	HMDB0011743	-
184	Hydroxyphenyllactate	181.0506 (M-H)	23	-	-	-	-	HMDB0000755	C03672
185	Homovanillate	181.0506 (M-H)	23	-	-	-	-	HMDB0000118	C05582
186	3,4-Dihydroxyphenylpropionic acid (DHPP)	181.0506 (M-H)	23	-	-	-	-	HMDB0000423	C10447
[184-186]	Hydroxyphenyllactate/Homovanillate/DHPP	181.0506 (M-H)	23	2.2 ± 0.03	1.4 ± 0.04	1 ± 0.2	4(177)	-	-
187	Hydroxyphenylpropionate	165.0558 (M-H)	25	0.11 ± 0.01	0.08 ± 0.01	0.18 ± 0.08	1	-	-
188	Vanillylmandelate	197.0456 (M-H)	24	-	-	-	-	HMDB0000291	C05584
189	Quinic acid	191.0561 (M-H)	22	1.1	0.5	1.4	1	HMDB0003072	C06746
190	Ferulic acid	195.0652 (M-H)	30	0.15	0.006	0.002	1	HMDB0000954	C01494
191	Tartaric acid	151.0237 (M+2Na-H)	73	-	-	-	-	HMDB0000956	C00898
192	Picolinate (2-Pyridinecarboxylate)	124.0393 (M+H)	41	-	-	-	-	HMDB0002243	C10164
193	Orotate	200.9883 (M+2Na-H)	60	-	-	-	-	HMDB0000226	C00295
194	Dihydroorotate	159.0406 (M+H)	144	-	-	-	-	HMDB0000528	C00337
195	Quinolate	166.0146 (M-H)	22	-	-	-	-	HMDB0000232	C03722
196	Salicylamide	138.0550 (M+H)	30	-	-	-	-	HMDB0015687	D01811
197	4-Aminobenzoate	138.0550 (M+H)	30	-	-	-	-	HMDB0001392	C00568
198	Anthranilate	138.0550 (M+H)	30	-	-	-	-	HMDB0001123	C00108
199	Hippurate	180.0655 (M+H)	31	3.6 ± 0.13	3.9 ± 0.13	4.9 ± 0.2	1	HMDB0000714	C01586
200	2-Methylhippurate	194.0812 (M+H)	30	-	-	-	-	HMDB0011723	-
201	3-Methylhippurate	194.0812 (M+H)	30	-	-	-	-	HMDB0013245	-
202	4-Methylhippurate	194.0812 (M+H)	30	-	-	-	-	HMDB0013292	-
[200-202]	O-Methylhippurate	194.0812 (M+H)	30	-	-	-	-	HMDB0000859	-
203	Methylhippurate	194.0812 (M+H)	30	-	-	-	-	-	-
204	Indoxyl sulfate	214.0169 (M+H)	42	-	-	-	-	HMDB0000682	-
205	Ethanolamine Phosphate	142.0264 (M+H)	125	6.75 ± 0.47*	0.564 ± 0.02*	4.45 ± 0.44*	1	HMDB0000224	C00346
206	Phosphoserine	186.0162 (M+H)	280	-	-	-	-	HMDB0000272	C01005
207	Phosphoenolpyruvate	166.9751 (M-H)	18	-	-	-	-	HMDB0000263	C00074
208	2-Phosphoglycerate	184.9856 (M-H)	24	-	-	-	-	HMDB0000362	-
209	3-Phosphoglycerate	184.9856 (M-H)	24	-	-	-	-	HMDB0000807	C00597
[208,209]	Phosphoglycerate	184.9856 (M-H)	24	-	-	3.6 ± 0.3*	4(200)	-	-
210	Glucuronic acid	239.0138 (M+2Na-H)	75	2.6 ± 0.2*	0.93 ± 0.2*	0.8 ± 0.006*	1	HMDB0000127	C00191
211	Gluconic acid	195.051 (M-H)	20	2.5 ± 0.24	1.83 ± 0.16	3.9 ± 0.33	1	HMDB0000625	C00257

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
212	Saccharic acid	209.0303 (M-H)	20	-	-	-	1	HMDB0000663	C00818
213	Galactarate	209.0303 (M-H)	19	-	-	-	1	HMDB0000639	C00879
[212,213]	Tetrahydroxyhexanedioic acid (Saccharate/Galactarate)	209.0303 (M-H)	20	1.6	0.4	0.35	4	-	-
214	Mercaptopyruvate	118.9808 (M-H)	20	-	-	-	-	HMDB0001368	C00957
215	Erythritol	157.0273 (M+Cl)	22	-	-	-	-	HMDB0002994	C00503
216	Threitol	157.0273 (M+Cl)	22	-	-	-	-	HMDB00004136	C16884
[215,216]	Erythritol/Threitol	157.0273 (M+Cl)	22	0.34 ± 0.23*	0.8 ± 0.2*	9 ± 3*	4(206)	-	-
217	2-Deoxyglucose	199.0379 (M+Cl)	22	4.6 ± 0.18*	3.8 ± 0.17*	3.6 ± 0.1*	1	HMDB0062477	C00586
218	Glucosamine	180.0867 (M+H)	89	0.73 ± 0.025*	0.18 ± 0.02*	0.18 ± 0.016*	1	HMDB0001514	C00329
219	Fructosamine	-	-	-	-	-	-	HMDB0002030	-
220	Glucosamine 6-sulfate	260.0435 (M+H)	110	-	-	0.049 ± 0.01	1	HMDB0000592	C04132
221	Glucosamine 6-phosphate	260.0530 (M+H)	275	-	-	-	-	HMDB0001254	C00352
222	Arabinose	149.0455 (M-H)	22	-	-	-	-	HMDB0000646	C00259
223	Xylose	149.0455 (M-H)	22	4515736	4493538	3590622	5	HMDB0000098	C00181
224	Ribose	149.0455 (M-H)	22	-	-	-	-	HMDB0000283	C00121
225	Ribose 5-phosphate	231.0264 (M+H)	150	-	-	-	-	HMDB0001548	C00117
226	Fructose	219.0265 (M+K)	60	-	-	-	-	HMDB0000660	C02336
227	Mannose	203.0526 (M+Na)	60	-	-	-	-	HMDB0000169	C00159
228	Galactose	203.0526 (M+Na)	60	-	-	-	-	HMDB0000143	C00984
229	Glucose	215.0328 (M+Cl)	22	4560 ± 93	1156 ± 56	2603 ± 141	2	HMDB0000122	C00031
230	Hexane hexol (Mannitol/Sorbitol)	205.0683 (M+Na)	60	14 ± 0.5*	3.4 ± 0.2*	2.4 ± 0.24*	1	-	-
231	Galactose 1-phosphate	261.0370 (M+H)	264	-	-	-	-	HMDB0000645	C00446
232	Glucose 6-phosphate	261.0370 (M+H)	278	-	-	-	-	HMDB0001401	C00668
233	Fructose 6-phosphate	261.0370 (M+H)	280	-	-	-	-	HMDB0000124	C00085
[231-233]	Hexose (Glucose, others) phosphate (1 or 6)	261.0370 (M+H)	269	-	-	0.49 ± 0.05	1	-	-
234	N-Acetylglucosamine	258.0385 (M+Cl)	23	0.3 ± 0.04	0.17 ± 0.03	0.77 ± 0.07	1	HMDB0000215	C01074
235	N-Acetylneuramine	308.0987 (M-H)	20	1.9 ± 0.08	0.88 ± 0.07	1.13 ± 0.12	1	HMDB0000230	C00140
236	Disaccharide (Lactose)	365.1054 (M+Na)	78	13.6 ± 0.018	14.3 ± 0.08	11.9 ± 0.3	1	HMDB0000186	C00243
237	Trisaccharide (Raffinose)	527.1583 (M+Na)	138	-	-	-	-	HMDB0003213	C00492
238	Tetrasaccharide (Glycogen, Stachyose)	689.2111 (M+Na)	250	0.073 ± 0.02	-	0.18 ± 0.05	1	HMDB0000757	C00182
239	Putrescine	89.1073 (M+H)	73	0.16 ± 0.04*	0.2 ± 0.02	0.7 ± 0.1	1	HMDB0001414	C00134
240	N-Acetylputrescine	131.1179 (M+H)	43	1.7 ± 0.2	0.9 ± 0.04	0.66 ± 0.06*	1	HMDB0002064	C02714
241	Spermidine	146.1652 (M+H)	121	0.63 ± 0.24*	9.6 ± 3.2*	1.6 ± 2.9*	1	HMDB0001257	C00315
242	N1-acetylspermidine	245.2336 (M+H)	42	-	-	-	-	-	-
243	Benzylamine	108.0808 (M+H)	31	1.4 ± 0.1*	1.4 ± 0.4*	1.03 ± 0.09*	1	HMDB0033871	C15562
244	Aminophenol (2, 3, or 4)	110.0600 (M+H)	31	0.4 ± 0.02*	0.25 ± 0.01*	0.1 ± 0.01*	1	HMDB0001169	C02372
245	Phenethylamine	122.0964 (M+H)	31	8 ± 0.4*	14 ± 1*	11 ± 1*	1	HMDB0012275	C05332
246	Cadaverine	103.123 (M+H)	76	-	-	-	-	HMDB0002322	C01672
247	Quinoline	130.0651 (M+H)	31	0.9 ± 0.09	4.3 ± 0.1	3.3 ± 0.2	1	HMDB0033731	C06413
248	2-Quinolinecarboxylic acid	174.055 (M+H)	31	0.1 ± 0.005*	0.1 ± 0.009*	0.13 ± 0.025*	1	HMDB0000842	C06325
249	Salsolinol	180.1019 (M+H)	32	0.02 ± 0.001*	0.011 ± 0.001*	0.0078 ± 0.0009	1	HMDB0005199	C09642
250	Choline	104.1070 (M+H)	40	19 ± 1	15 ± 0.2	26 ± 1	1	HMDB0000097	C00114
251	Acetylcholine	146.1176 (M+H)	32	2.1 ± 0.12*	0.73 ± 0.04	0.89 ± 0.21	1	HMDB0000895	C01996
252	Phosphocholine	184.0733 (M+H)	160	0.03 ± 0.0008	1.1 ± 0.03	0.91 ± 0.11	1	HMDB0001565	C00588
253	Glycerophosphocholine	258.1101 (M+H)	104	2.4 ± 0.2	39 ± 0.8	46 ± 1.6	1	HMDB0000086	C00670
254	PC(10:0/10:0)	566.3816 (M+H)	38	-	-	-	-	-	CA1375

#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
255	Betaine	118.0863 (M+H)	64	116 ± 6	62 ± 0.5	42 ± 1	1	HMDB0000043	C00719
256	Butyrobetaine	146.1176 (M+H)	66	2.52 ± 0.005	2.18 ± 0.07	2.17 ± 0.008	1	HMDB0001161	C01181
257	Valerobetaine	160.1332 (M+H)	42	1.5 ± 0.01	0.9 ± 0.02	1 ± 0.03	1	-	-
258	Valine Betaine	160.1332 (M+H)	-	-	-	-	-	-	-
259	Carnitine	162.1125 (M+H)	73	48 ± 1	61 ± 1	62 ± 1	1	HMDB0000062	C00318
260	Acetylcarnitine	204.123 (M+H)	55	16.5 ± 0.48	7.8 ± 0.2	6.9 ± 0.2	1	HMDB0000201	C02571
261	Propionylcarnitine	218.1387 (M+H)	48	0.3 ± 0.002	0.29 ± 0.007	0.35 ± 0.008	2	HMDB0000824	C03017
262	Butyrylcarnitine	232.1543 (M+H)	45	0.2 ± 0.005	0.17 ± 0.002	0.14 ± 0.003	2	HMDB0002013	C02862
263	Hexanoylcarnitine	260.1857 (M+H)	42	0.1 ± 0.002	0.02 ± 0.001	0.04 ± 0.003	2	HMDB0000705	-
264	Octanoylcarnitine	288.2170 (M+H)	39	0.1 ± 0.001	0.08 ± 0.001	0.05 ± 0.001	2	HMDB0000791	C02838
265	Decanoylcarnitine	316.2483 (M+H)	39	0.1 ± 0.006	0.04 ± 0.001	0.06 ± 0.001	2	HMDB0000651	-
266	Lauroylcarnitine	344.2795 (M+H)	27	0.04 ± 0.0004	0.03 ± 0.001	0.03 ± 0.002	1	HMDB0002250	-
267	Palmitoylcarnitine	400.3421 (M+H)	36	0.1 ± 0.01	0.2 ± 0.01	0.13 ± 0.001	2	HMDB0000222	C02990
268	CoA	382.5507 (M-2H)	30	-	-	-	-	HMDB0001423	C00010
269	Acetyl-CoA	403.5556 (M-2H)	32	-	-	-	-	HMDB0001206	C00024
270	Propionyl-CoA	410.5634 (M-2H)	32	-	-	-	-	HMDB0001275	C00100
271	C4-CoA	417.5713 (M-2H)	32	-	-	-	-	HMDB0001088	C00136
272	C5:1-CoA	423.5713 (M-2H)	32	-	-	-	-	HMDB0001493	C03069
273	C5-CoA	424.5791 (M-2H)	35	-	-	-	-	HMDB0001113	C02939
274	Hexanoyl-CoA	431.5869 (M-2H)	35	-	-	-	-	HMDB0002845	C05270
275	Octanoyl-CoA	445.0260 (M-2H)	35	-	-	-	-	HMDB0001070	C01944
276	Decanoyl-CoA	459.6182 (M-2H)	35	-	-	-	-	HMDB0006404	C05274
277	Myristoyl-CoA	489.6641 (M-2H)	36	-	-	-	-	HMDB0001521	C02593
278	Palmitoyl-CoA	501.6652 (M-2H)	45	-	-	-	-	HMDB0001338	C00154
279	Vitamin A (Retinol)	269.2269 (M-H2O+H)	26	6.7 ± 0.4	8.8 ± 0.9	5.4 ± 0.25	1	HMDB0000305	C00473
280	Retinoic acid	299.2017 (M-H)	226	0.56 ± 0.025	0.95 ± 0.05	0.6 ± 0.16	1	HMDB0001852	C00777
281	Thiamine (B1)	265.1118 (M+H)	60	0.01 ± 0.002	-	0.008 ± 0.005	1	HMDB0000235	C00378
282	Thiamine Pyrophosphate	425.0444 (M+H)	294	-	-	-	-	HMDB0001372	-
283	Riboflavin (B2)	377.1456 (M+H)	45	0.078 ± 0.005	0.0171 ± 0.0022	0.018 ± 0.007	1	HMDB0000244	C00255
284	Nicotinate (B3)	124.0393 (M+H)	35	-	-	0.05 ± 0.008	1	HMDB0001488	C00253
285	Nicotinamide (B3)	123.0553 (M+H)	37	0.41 ± 0.03	0.31 ± 0.01	1.81 ± 0.04	1	HMDB0001406	C00153
286	Pantothenic acid (B5)	220.1179 (M+H)	36	0.04 ± 0.006	0.02 ± 0.0007	0.04 ± 0.004	1	HMDB0000210	C00864
287	Pyridoxine (B6)	170.0812 (M+H)	52	0.08 ± 0.01	0.03 ± 0.002	0.055 ± 0.02	1	HMDB0000239	C00314
288	Pyridoxal (B6)	168.0655 (M+H)	39	0.11 ± 0.014	0.019 ± 0.006	0.02 ± 0.009	1	HMDB0001545	C00250
289	Biotin (B7)	245.0954 (M+H)	35.4	-	-	-	-	HMDB0000030	C00120
290	Folate (B9)	442.1470 (M+H)	48	-	-	-	-	HMDB0000121	C00504
291	Dihydrofolate	444.1631 (M+H)	62	-	-	-	-	HMDB0001056	-
292	Tetrahydrofolate	446.1783 (M+H)	-	-	-	-	-	HMDB0001846	-
293	5-Methyltetrahydrofolate	460.1939 (M+H)	-	-	-	-	-	HMDB0001396	-
294	Vitamin B12	678.2895 (M+2H)	120	-	-	-	-	HMDB0000607	C05776
295	Methylcobalamin	673.3021 (M+H)	96	-	-	-	-	HMDB0002274	-
296	Adenosylcobalamin	527.2267 (M+3H)	156	-	-	-	-	-	-
297	Vitamin C (Ascorbate)	175.024814 (M-H)	20	-	-	-	-	HMDB0000044	C00072
298	Vitamin D2	397.3465 (M+H)	27	-	-	-	-	HMDB0000900	C05441
299	Alpha-Tocopherol	431.3883 (M+H)	27	-	-	-	-	-	-
300	Dihydrofolate	444.1631 (M+H)	52	-	-	-	-	HMDB0001056	C00415

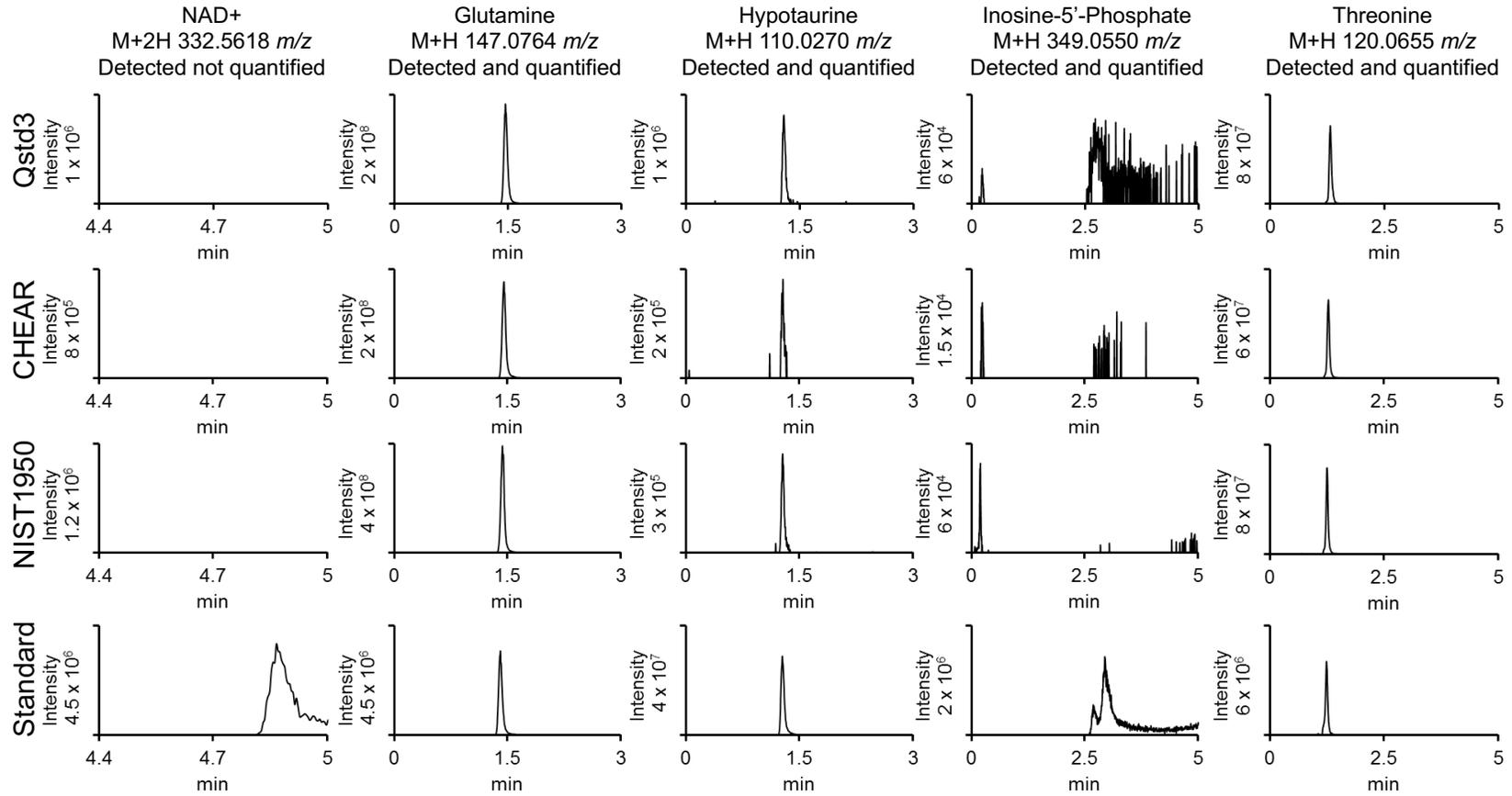
#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
301	Flavin mononucleotide	457.1119 (M+H)	110	-	-	-	-	HMDB0001520	
302	1-Methylnicotinamide	137.0709 (M+H)	50	0.2 ± 0.005	0.2 ± 0.004	0.2 ± 0.007	1	HMDB0000699	C02918
303	Trigonelline	138.055 (M+H)	68	5.4 ± 0.07	1.3 ± 0.05	1.3 ± 0.03	1	HMDB0000875	C01004
304	Flavin adenine dinucleotide (FAD)	786.1644 (M+H)	15	-	-	-	-	HMDB0001248	C00016
305	Nicotinamide mononucleotide	335.0639 (M+H)	133	-	-	-	-	HMDB0000229	C00455
306	NAD+	332.5618 (M+2H)	293	-	-	-	-	HMDB0000902	C00003
307	NADPH	372.5450 (M+2H)	294	-	-	-	-	HMDB0000221	C00005
308	Pyridoxate	184.0609 (M+H)	40	0.17	0.03	0.06	1	HMDB0000017	C00847
309	Pyridoxamine	169.0972 (M+H)	75	-	-	-	-	HMDB0001431	C00534
310	Menatretrenone - Vitamin MK4	445.3101 (M+H)	27	-	-	-	-	HMDB0030017	-
311	FA 4:0 (Butyrate/Isobutyrate)	87.0451 (M-H)	20	-	-	-	-	-	-
312	FA 5:0 (Valerate, Isovalerate, others)	101.0608 (M-H)	20	-	-	-	-	-	-
313	FA 6:0 (Caproate, Isocaproate, others)	115.0765 (M-H)	22	-	-	-	-	-	-
314	FA 7:0	129.0921 (M-H)	44	4.4 ± 0.4	5.1 ± 0.9	5.2 ± 0.5	1	HMDB0000666	-
315	FA 8:0 (Octanoate)	143.1078 (M-H)	77	110 ± 2.3	13.5 ± 0.3*	80 ± 2*	1	HMDB0000482	-
316	FA 9:0	157.1234 (M-H)	100	2 ± 0.08*	2.1 ± 0.08*	1.9 ± 0.2*	1	HMDB0000847	-
317	OH-FA10:0	187.134 (M-H)	38	0.19 ± 0.014	0.27 ± 0.004	0.33 ± 0.005	1	-	-
318	FA 12:0 (Laurate)	199.1704 (M-H)	180	2.6 ± 0.09	3.3 ± 0.05	2 ± 0.2	1	HMDB0000638	-
319	OH-FA12:0	215.1653 (M-H)	59-87	0.73 ± 0.002	0.98 ± 0.03	1.1 ± 0.03	1	HMDB0002059	-
320	FA 14:0 (Myristate)	227.2017 (M-H)	220	7.6 ± 0.07	11 ± 0.13	5.4 ± 0.07	1	HMDB0000806	-
321	FA 16:1 (Palmitoleic)	253.2173 (M-H)	228	18.7 ± 0.2	21.6 ± 0.2	11 ± 0.26	1	HMDB0003229	-
322	FA 16:0 (Palmitate)	255.233 (M-H)	252	85 ± 0.5	118 ± 2.3	82 ± 0.9	1	HMDB0000220	-
323	FA 17:0	269.2486 (M-H)	266	1.98 ± 0.006	2.5 ± 0.02	2 ± 0.02	1	HMDB0002259	-
324	FA 18:3 n-3 or n-6 (Linolenic acid)	277.2173 (M-H)	223	39.9 ± 0.3	69 ± 2	47 ± 1.1	1	-	-
325	FA 18:2 (Linoleic acid)	279.233 (M-H)	238	2858 ± 78	4805 ± 95	3678 ± 119	2	HMDB0000673	C01595
326	FA 18:1 (Oleic acid)	281.2486 (M-H)	258	120 ± 0.6	148 ± 1.3	101 ± 1.7	1	HMDB0000207	C00712
327	FA 18:0 (Stearic acid)	283.2643 (M-H)	282	54 ± 0.8	75 ± 1	56 ± 0.6	1	HMDB0000827	C01530
328	FA 20:0 (Arachidic acid)	311.2955 (M-H)	298	18 ± 0.7	17 ± 0.8	11 ± 0.2	2	HMDB0002212	C06425
329	FA 20:1 (Gondoic acid)	309.2799 (M-H)	290	12 ± 0.4	12 ± 0.1	10 ± 0.3	2	HMDB0002231	C16526
330	FA 20:2 (Eicosadienoic acid)	307.2642 (M-H)	272	19 ± 0.5	22 ± 0.3	17 ± 0.04	2	HMDB0005060	C16525
331	FA 20:3 (Homolinoleic acid)	305.2486 (M-H)	252	139 ± 3	265 ± 1	267 ± 3	2	HMDB0060039	-
332	FA 20:4 (Arachidonic acid)	303.2329 (M-H)	239	984 ± 17	2657 ± 22	2991 ± 61	2	HMDB0001043	C00219
333	FA 20:5 (Eicosapentaenoic acid)	301.2173 (M-H)	224	39 ± 1	108 ± 3	102 ± 2	2	HMDB0001999	C06428
334	FA 22:4 (Docosatetraenoic acid)	331.2642 (M-H)	258	25.5 ± 1	31 ± 0.2	31 ± 0.4	2	HMDB0002226	C16527
335	FA 22:5 n-3 or n-5	329.2486 (M-H)	240	58 ± 1	88 ± 2	79 ± 1	2	HMDB0039133	-
336	Docosahexaenoic acid (FA 22:6)	327.233 (M-H)	227	32 ± 0.4	63.7 ± 0.2	58 ± 1.4	1	HMDB0002183	C06429
337	Sphinganine	302.3054 (M+H)	26	0.008 ± 0.0006	0.024 ± 0.002	0.066 ± 0.0006	1	HMDB0000269	C00836
338	Sphingosine	300.2897 (M+H)	28	0.002 ± 0.0001	0.006 ± 0.0002	0.015 ± 0.0007	1	HMDB0000252	C00319
339	MG(14:0/0:0/0:0)	361.2596 (M+CH3COO)	226	0.48 ± 0.02	0.4 ± 0.01	0.33 ± 0.003	1	HMDB0011561	-
340	MG(18:1/0:0/0:0)	391.2621 (M+CH3COO)	251	4.7 ± 0.002	3.1 ± 0.01	1.4 ± 0.01	1	HMDB0011567	-
341	TG(16:0/16:0/16:0)	-	-	-	-	-	-	HMDB0005356	-
342	Methyl jasmonate	225.1485 (M+H)	28	1.6 ± 0.17	0.37 ± 0.07	1.06 ± 0.3	1	HMDB0036583	-
343	Cholesterol	369.352 (M-H2O+H)	27	1182 ± 141	1352 ± 29	1456 ± 80	1	HMDB0000067	C00187
344	Lanosterol	408.3834 (M-H2O+H)	30	-	-	-	-	HMDB0001251	C01724
345	25-hydroxycholesterol	385.3470 (M-H2O+H)	27	-	-	-	-	HMDB0006247	C15519
346	17/21-hydroxyprogesterone	331.2268 (M+H)	30	-	-	-	-	-	-

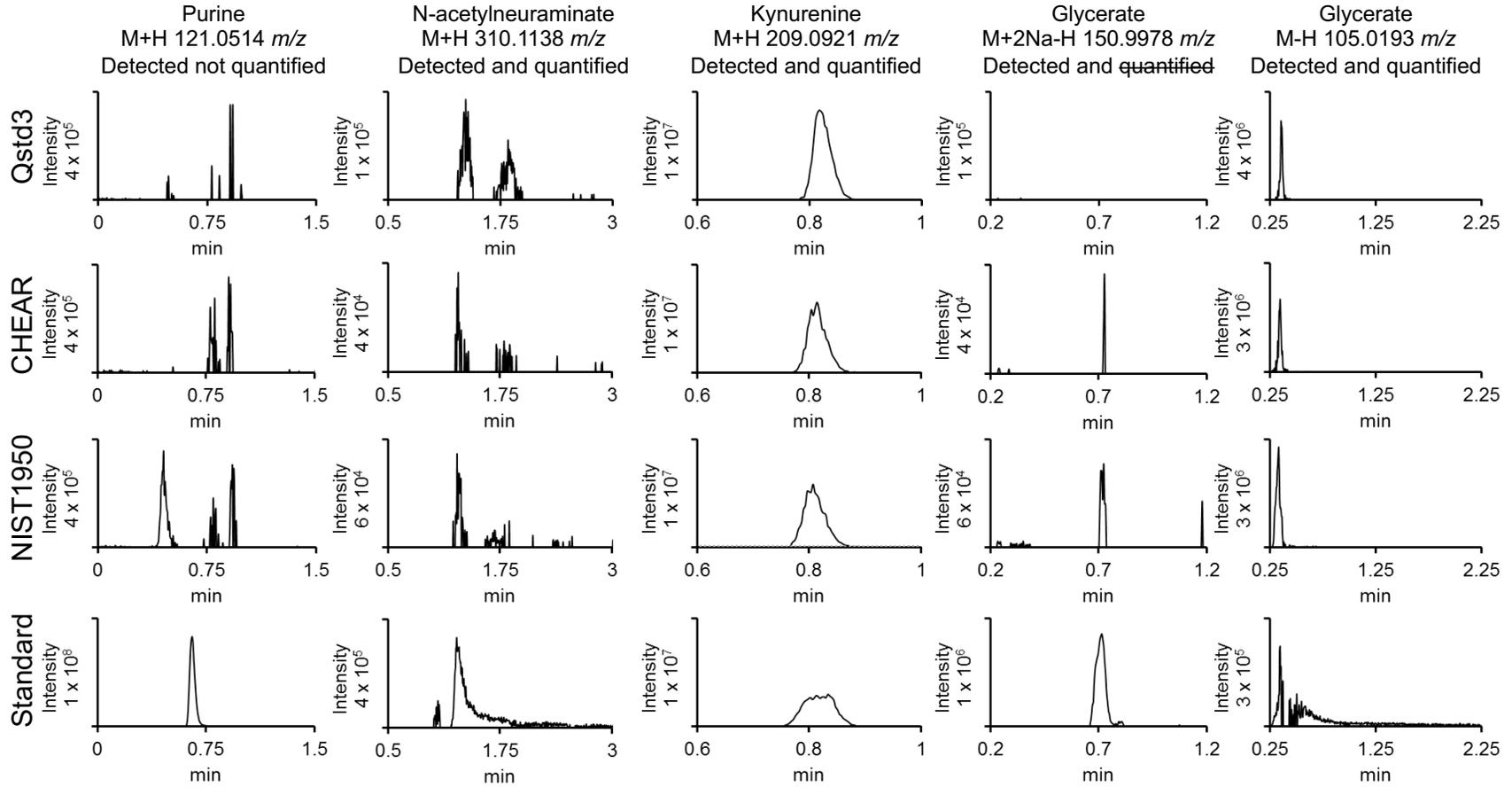
#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
347	Progesterone	315.2319 (M+H)	27	-	-	-	-		
348	Cortisol-21-acetate	405.2272 (M+H)	28	-	-	-	-		C02821
349	Cortexolone	347.2217 (M+H)	29.1	-	-	-	-		
350	Cortisol	363.2166 (M+H)	30	0.45 ± 0.04	0.26 ± 0.003	0.23 ± 0.023	1	HMDB0000063	C00735
351	Tetrahydrocortisol	349.2379 (M-H ₂ O+H)	30	-	-	-	-	HMDB0000949	C05472
352	Estriol	358.0368 (M+H)	27	-	-	-	-	HMDB0000153	
353	Aldosterone	361.2010 (M+H)	27	0	222106	132992	-		
354	Adrenosterone	271.1696 (M+H)	27	-	-	-	-	HMDB0006772	
355	Lithocholic acid	375.2905 (M-H)	196	0.14 ± 0.02	0.07 ± 0.02	0.244 ± 0.05	1	HMDB0000761	C03990
356	Chenodeoxycholate (CDCA)	391.2853 (M-H)	138	-	-	-	1	HMDB0000518	C02528
357	Deoxycholate (CDA)	391.2853 (M-H)	138	-	-	-	1	HMDB0000626	C04483
[356,357]	CDCA/CDA	391.2853 (M-H)	138	0.27 ± 0.009	0.31 ± 0.004	0.33 ± 0.016	4	-	-
358	Cholic acid	407.2803 (M-H)	90	0.12 ± 0.001	0.16 ± 0.004	0.04 ± 0.004	1	HMDB0000619	C00695
359	Glycocholate	466.3163 (M+H)	38	0.63 ± 0.01	0.71 ± 0.05	0.95 ± 0.16	1	HMDB0000138	C01921
360	Taurolithocholate	484.3090 (M+H)	42	-	-	-	-	HMDB0000722	C02592
361	Estradiol 17-alpha	271.1704 (M-H)	152	-	-	-	-	HMDB0000429	C02537
362	Urate	167.0211 (M-H)	21	254 ± 12	157 ± 10	162 ± 2	2	HMDB0000289	C00366
363	Xanthine	153.0407 (M+H)	45	0.5 ± 0.03*	0.55 ± 0.01	2.4 ± 0.1	1	HMDB0000292	C00385
364	Hypoxanthine	137.0463 (M+H)	44	4.6 ± 0.14	3.3 ± 0.12	73 ± 1.4	1	HMDB0000157	C00262
365	Purine	121.0514 (M+H)	39	-	-	-	-	HMDB0001366	C15587
366	Adenine	136.0618 (M+H)	43	0.24 ± 0.06*	0.3 ± 0.03*	0.23 ± 0.04*	1	HMDB0000034	C00147
367	Adenosine	268.1041 (M+H)	44	-	-	0.059 ± 0.017	1	HMDB0000050	C00212
368	Methyladenosine	282.1197 (M+H)	40	-	-	-	-	HMDB0003331	C02494
369	Adenosine monophosphate	348.0704 (M+H)	148	-	0.41 ± 0.13	5.4 ± 3.6	1	HMDB0000045	C00020
370	Deoxyadenosine monophosphate	332.0760 (M+H)	122	-	-	-	-	HMDB0000905	C00360
371	Adenosine diphosphate	426.0221 (M-H)	36	-	2.4 ± 0.3	1.5 ± 0.74	1	HMDB0001341	C00008
372	Inosine	267.0735 (M-H)	24	0.8 ± 0.04	0.27 ± 0.09	5.3 ± 0.5	1	HMDB0000195	C00294
373	Inosine monophosphate	349.0549 (M+H)	186	-	0.033 ± 0.001	1.5 ± 0.36	1	HMDB0000175	C00130
374	Inosine diphosphate	427.0062 (M-H)	20	-	3.7 ± 0.38	5 ± 3.7	1	HMDB0003335	C00104
375	Guanine	152.0567 (M+H)	50	-	-	-	-	HMDB0000132	C00242
376	Guanosine	284.0989 (M+H)	55	-	-	-	-	HMDB0000133	C00387
377	Deoxyguanosine	268.1041 (M+H)	50	-	-	-	-	HMDB0000085	C00330
378	8-Hydroxydeoxyguanosine	284.0989 (M+H)	-	-	-	-	-	HMDB0003333	-
379	Guanosine monophosphate	364.0653 (M+H)	256	-	0.003 ± 0.0003	0.33 ± 0.04	1	HMDB0001397	C00144
380	Thymidine	243.0975 (M+H)	37	-	-	-	-	HMDB0000273	C00214
381	Cytidine	244.0928 (M+H)	58	-	-	-	-	HMDB0000089	C00475
382	Cytosine	112.0505 (M+H)	50	0.21 ± 0.02*	0.1 ± 0.01	0.05 ± 0.03*	1	HMDB0000630	C00380
383	Deoxycytidine	228.0979 (M+H)	50	-	-	-	-		
384	5-Methylcytosine	126.0667 (M+H)	46	0.007 ± 0.001*	0.005 ± 0.001*	0.002 ± 0.001*	1	HMDB0002894	C02376
385	Cytidine monophosphate	324.0591 (M+H)	168	-	-	1.1 ± 0.27	1	HMDB0000095	C00055
386	Cytidine diphosphocholine	245.0609 (M+2H)	297	-	-	0.314 ± 0.02	1	HMDB0001413	C00307
387	2'-Deoxycytidine 5'-monophosphate	308.0642 (M+H)	147	-	-	-	-	HMDB0001202	C00239
388	Uracil	113.0346 (M+H)	38	0.02 ± 0.01	0.07 ± 0.02	0.8 ± 0.1	1	HMDB0000300	C00106
389	Dihydrouracil (5, 6)	115.0502 (M+H)	38	0.4 ± 0.03*	0.4 ± 0.02*	2 ± 1*	1	HMDB0000076	C00429
390	Uridine	245.0768 (M+H)	43	5.3 ± 0.14	3.28 ± 0.11	2.96 ± 0.14	1	HMDB0000296	C00299
391	Deoxyuridine	229.0824 (M+H)	39	-	-	-	-		

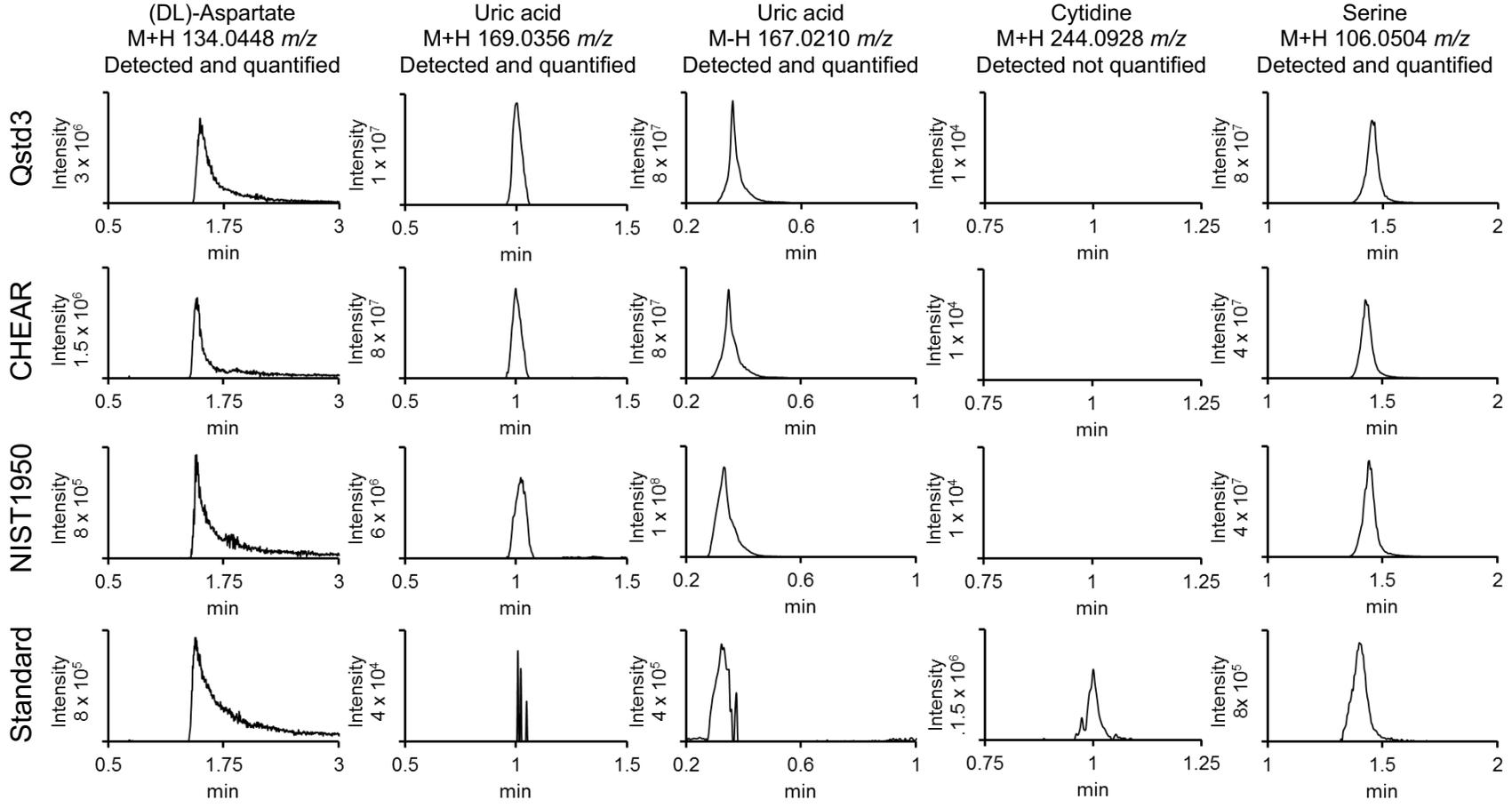
#	Metabolite	m/z (adduct)	rt (sec)	[NIST] μM ± S.E.M	[Chear] μM ± S.E.M	[Qstd] μM ± S.E.M	Quantified by	HMDB ID	KEGG ID
392	Uridine-5-Monophosphate	325.0431 (M+H)	149	-	-	-	-	HMDB0000288	C00105
393	Hydroxymethyluracil	143.0451 (M+H)	43	-	-	-	-		
394	Biliverdin	583.2551 (M+H)	28	6.2 ± 0.5	13.3 ± 0.32	7.7 ± 0.2	1	HMDB0001008	C00500
395	Bilirubin	585.2708 (M+H)	26	5.8 ± 0.4	1.6 ± 0.1	1.9 ± 0.02	2	HMDB0000054	C00486
396	Protoporphyrin	563.2653 (M+H)	28	-	-	-	-	HMDB0000241	C02191
397	Caffeine	195.0877 (M+H)	32	5 ± 0.2	3 ± 0.06	4.3 ± 0.3	1	HMDB0001847	C07481
398	Theophylline	181.0720 (M+H)	35	-	-	-	-	HMDB0001889	C07130
399	Paraxanthine	181.0720 (M+H)	36	-	-	-	-	HMDB0001860	C13747
400	Theobromine	181.0720 (M+H)	36	-	-	-	-	HMDB0002825	C07480
401	Dimethylxanthine	181.0720 (M+H)	36	-	-	-	-		
402	Nicotine	163.1230 (M+H)	37	0.18 ± 0.007	0.096 ± 0.012	0.13 ± 0.019	1	HMDB0001934	C00745
403	Cotinine	177.1022 (M+H)	31	-	-	-	-	HMDB0001046	
404	L-Homocysteine thiolactone	118.0321 (M+H)	37	-	-	-	-	HMDB0002287	-
405	Allantoin	159.0513 (M+H)	75	-	-	-	-	HMDB0000462	C01551
406	Daidzein	255.0655 (M+H)	27	-	-	-	-		
407	Metribuzin	215.0966 (M+H)	30	-	-	-	-		
408	Malathion	331.0439 (M+H)	30	-	-	-	-		
409	Benserazide	258.1084 (M+H)	66	-	-	-	-		
410	Pirimicarb	239.1493 (M+H)	24	-	-	-	-		
411	PFOS	498.9302 (M-H)	120	-	-	-	-		
412	Chlorobenzoate	154.9905 (M-H)	20	-	-	-	-		
413	Diisopropylphthalate	251.1278 (M+H)	27	-	-	-	-		
414	Aminoethylphosphonate	126.0513 (M+H)	150	-	-	-	-	HMDB0011747	C03557
415	Aminoethyldihydrogenphosphate	142.0264 (M+H)	120	-	-	-	-		
416	Methyl beta galactoside	217.0683 (M+Na)	46	-	-	-	-	HMDB0029965	C03619
417	Glucosaminat	196.0812 (M+H)	126	-	-	-	-		C03752
418	Hydroxynicotinate	140.0342 (M+H)	40	-	-	-	-	HMDB0002658	C01020
419	Anilinesulfonate	174.0219 (M+H)	45	-	-	-	-		C00292
420	Carboxymethylcysteine	180.0325 (M+H)	81	-	-	-	-	HMDB0029415	C03727
421	Ureidopropionate	133.0608 (M+H)	40	-	-	-	-	HMDB0000026	C02642
422	Formylglycine	104.0342 (M+H)	40	-	-	-	-		
423	2-Acetamido-2-Deoxy-Beta-D-Glucosylamine	221.1132 (M+H)	81	-	-	-	-		
424	Imidazoleacetate	127.0502 (M+H)	71	-	-	-	-		
425	Diaminopropionate	105.0659 (M+H)	110	-	-	-	-		
426	Tryptamine	161.1074 (M+H)	28	0.001 ± 0.0005	0.001 ± 0.0002	0.001 ± 0.0001	1	-	C00398
427	Lumichrome	243.0877 (M+H)	29	-	-	-	-		C01727
428	Tryptophanamide	204.1131 (M+H)	38	-	-	-	-	HMDB0013318	C00977
429	N-methyltryptamine	175.1230 (M+H)	29	-	-	-	-	HMDB0004370	C06213
430	Indole ethanol (Tryptophanol)	162.0914 (M+H)	30	0.030 ± 0.003	0.008 ± 0.001	0.009 ± 0.001	1	HMDB0003447	C00955
431	Trehalose	365.1065 (M+Na)	100	-	-	-	-	HMDB0000975	C01083
432	Thiopurine S-methyl ether	167.0386 (M+H)	33	-	-	-	-	HMDB00060412	C03542
433	Aminohydroxybenzoate/Hydroxyanthranilate	154.0504 (M+H)	32	-	-	-	-		
434	Hydroxypyridine	96.0444 (M+H)	43	0.27 ± 0.01	0.21 ± 0.003	0.2 ± 0.01	1	HMDB0013751	C02502
435	Isopentenyladenine	204.1244 (M+H)	31	-	-	-	-	-	C04083
436	Dimethylphenylenediamine	137.1073 (M+H)	31	0.36 ± 0.06	0.44 ± 0.04	0.47 ± 0.15	1	-	C04203
437	Methylxindole/Indolecarbinol	148.0757 (M+H)	29	0.06 ± 0.01	0.07 ± 0.001	0.02 ± 0.007	1	HMDB0004186	CA1325

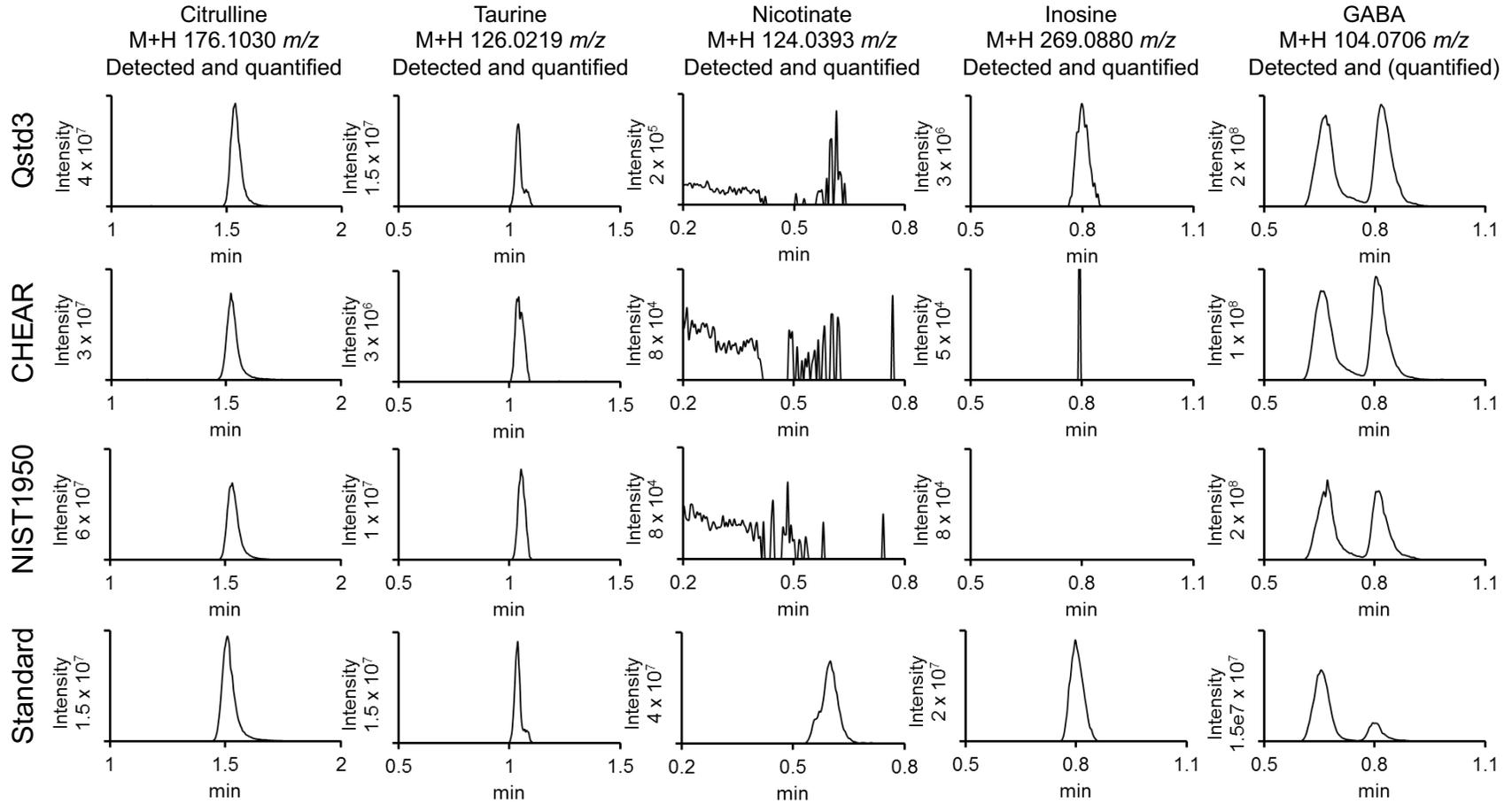
#	Metabolite	<i>m/z</i> (adduct)	<i>rt</i> (sec)	[NIST] $\mu\text{M} \pm \text{S.E.M}$	[Chear] $\mu\text{M} \pm \text{S.E.M}$	[Qstd] $\mu\text{M} \pm \text{S.E.M}$	Quantified by	HMDB ID	KEGG ID
438	Pyrrrole carboxylate	110.0247 (M-H)	20	-	-	-	-	HMDB0004230	C05942
439	Ethyl indoleacetate	204.1019 (M+H)	27	-	-	-	-	-	CA1213
440	Rosamarinic acid	359.0772 (M-H)	22	-	-	-	-	HMDB0003572	C10489
441	Reichstein's substance S	345.2071 (M-H)	98	-	-	-	-	HMDB0000015	C05488

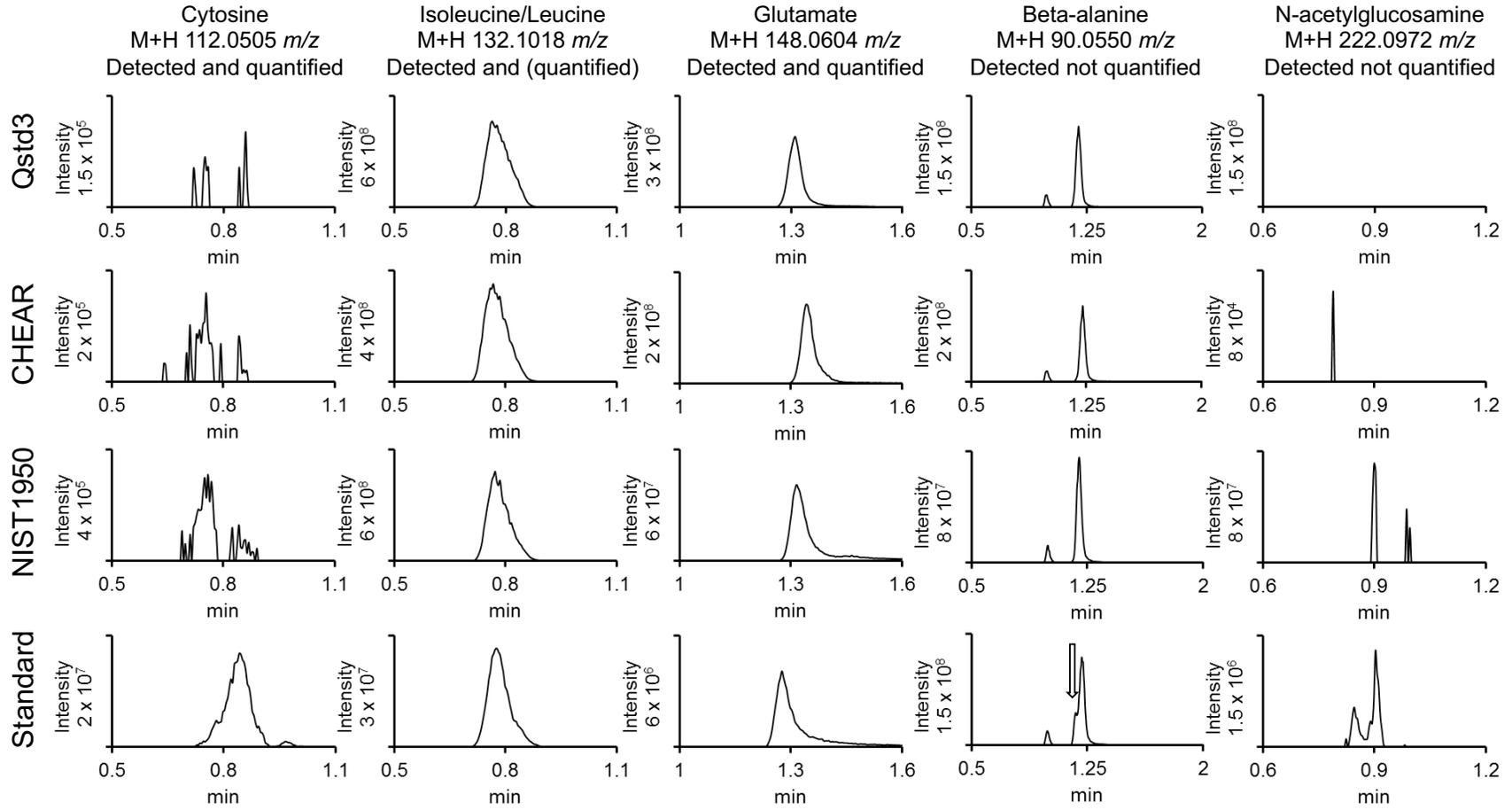
Appendix 2 (Chapter 3). Reference for how metabolites were classified as detected, identified, or quantifiable. Extracted ion chromatograms (± 3 ppm of expected mass) of detected metabolites in Qstd3, CHEAR, and NIST1950 along with authentic standard for reference. Metabolites were considered detected if addition of standard increased peak area. Metabolites were identified in the reference if addition of standard increased an existing peak area.

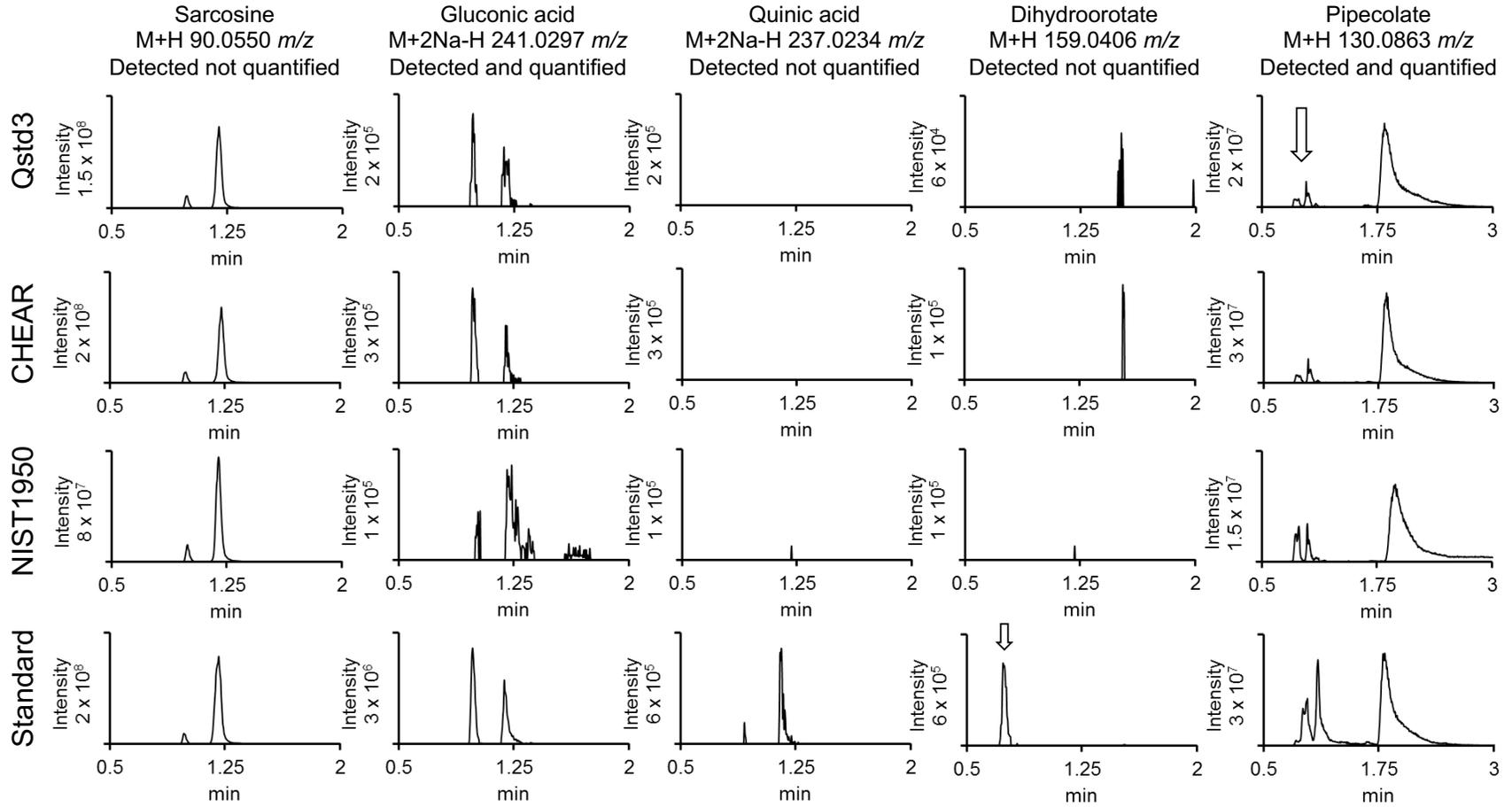


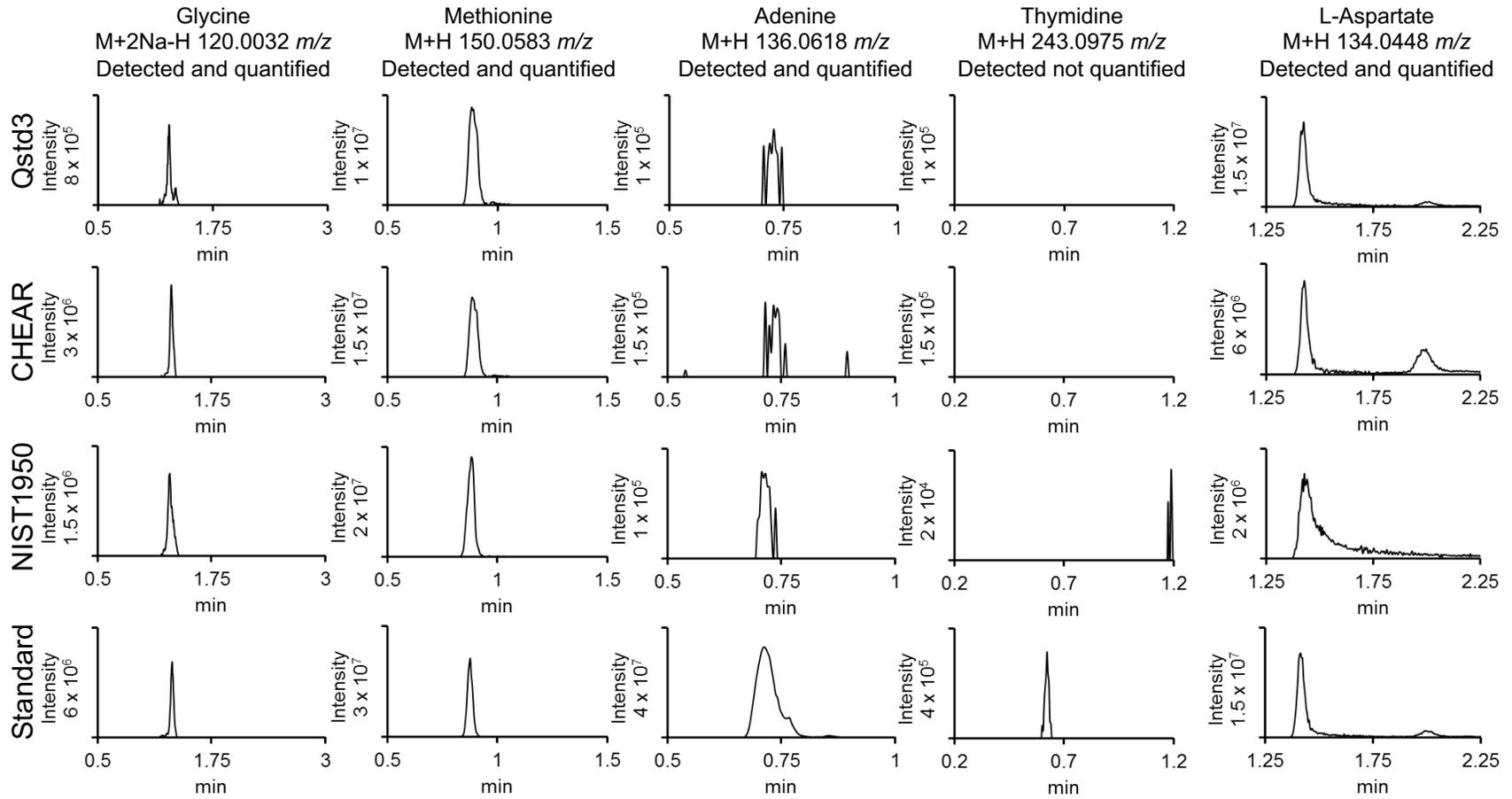


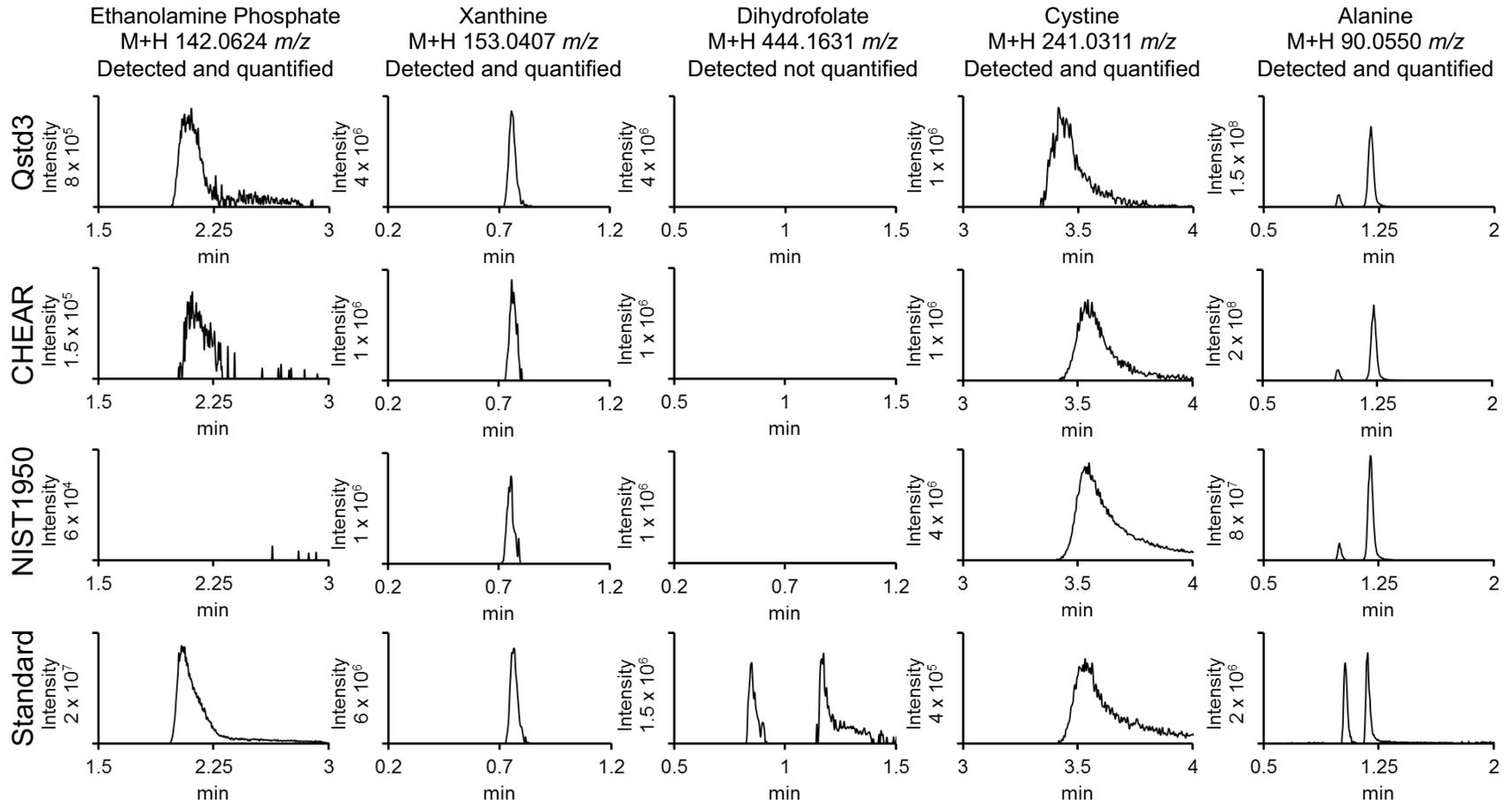


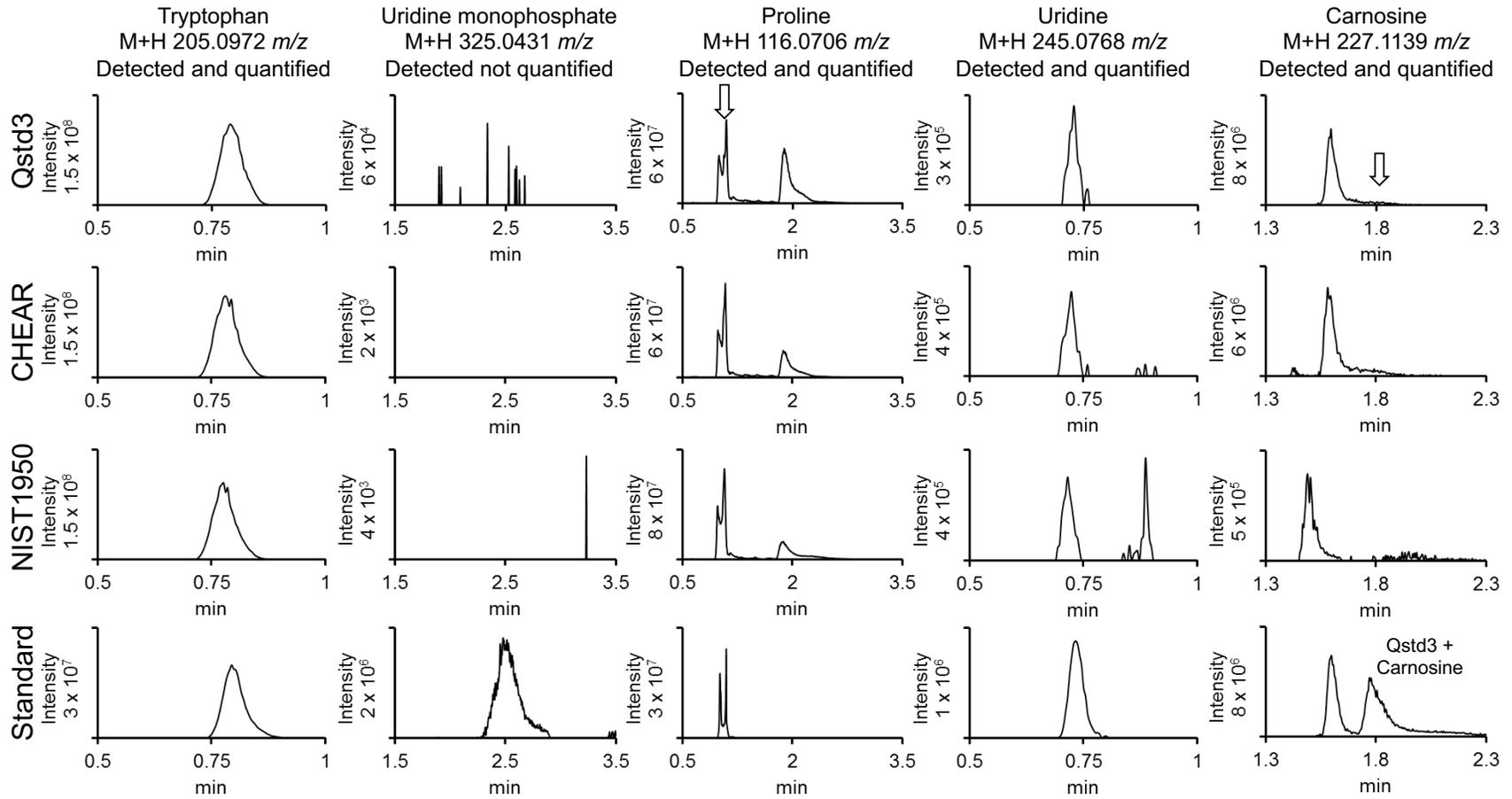


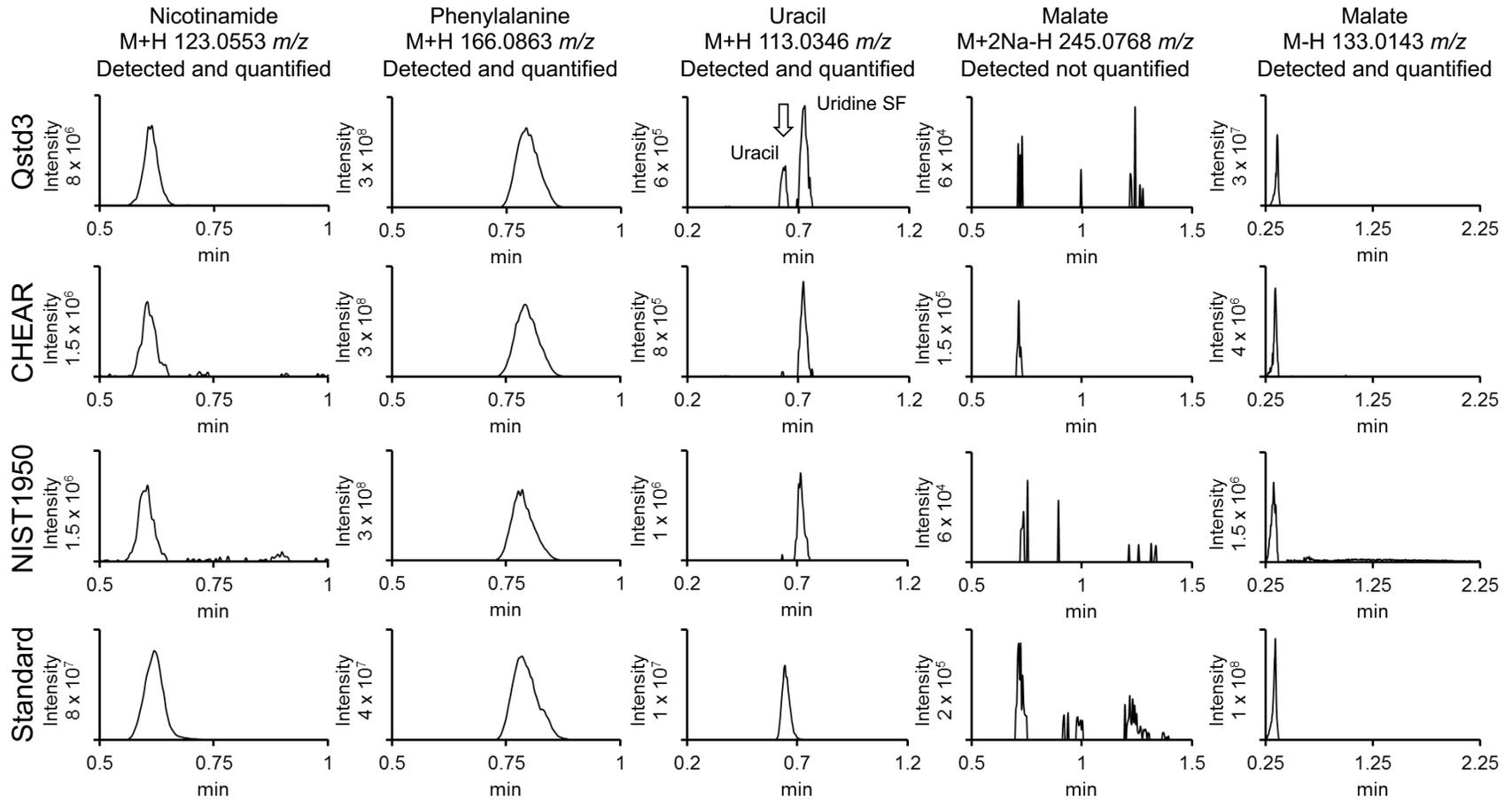


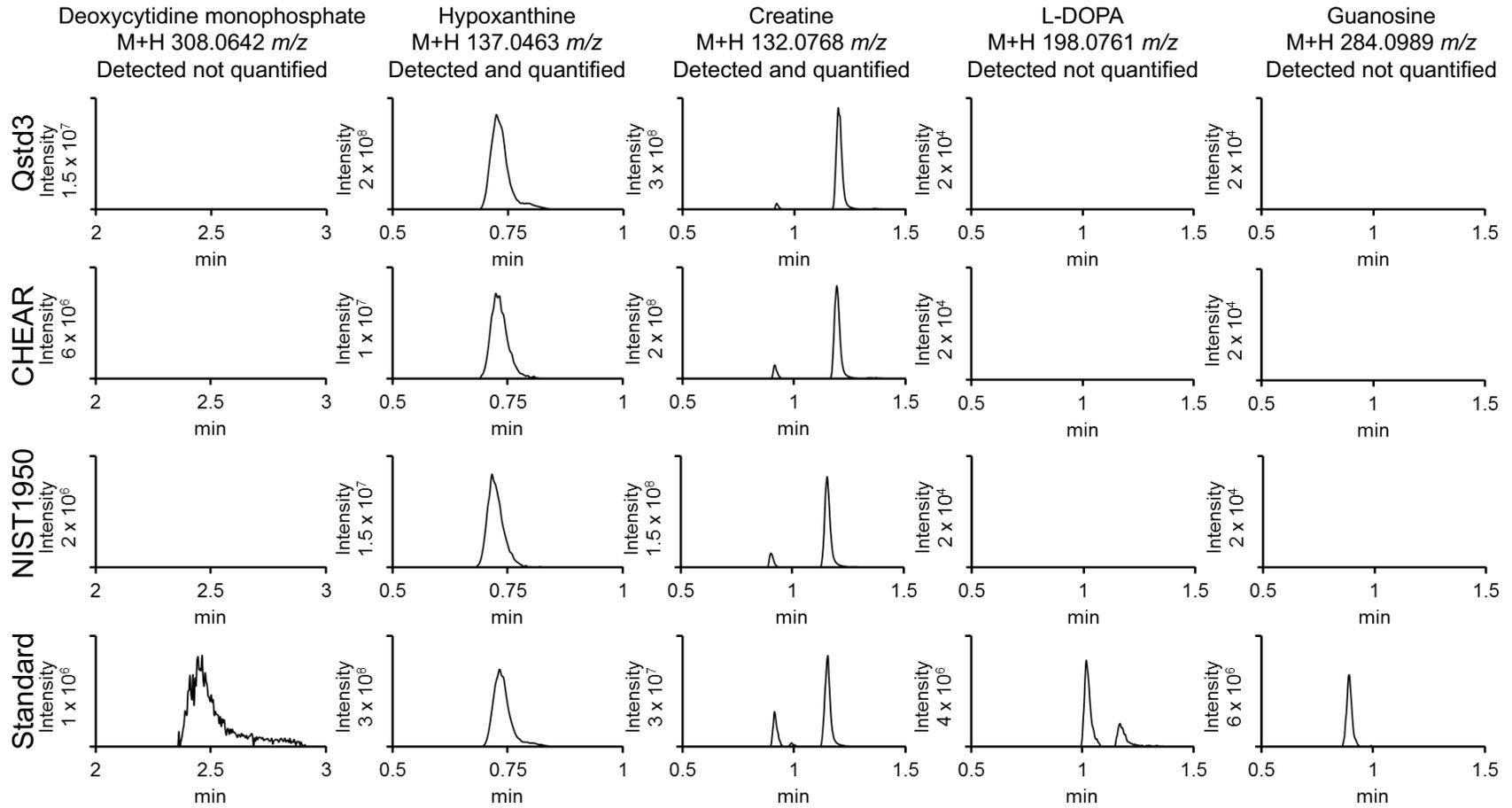


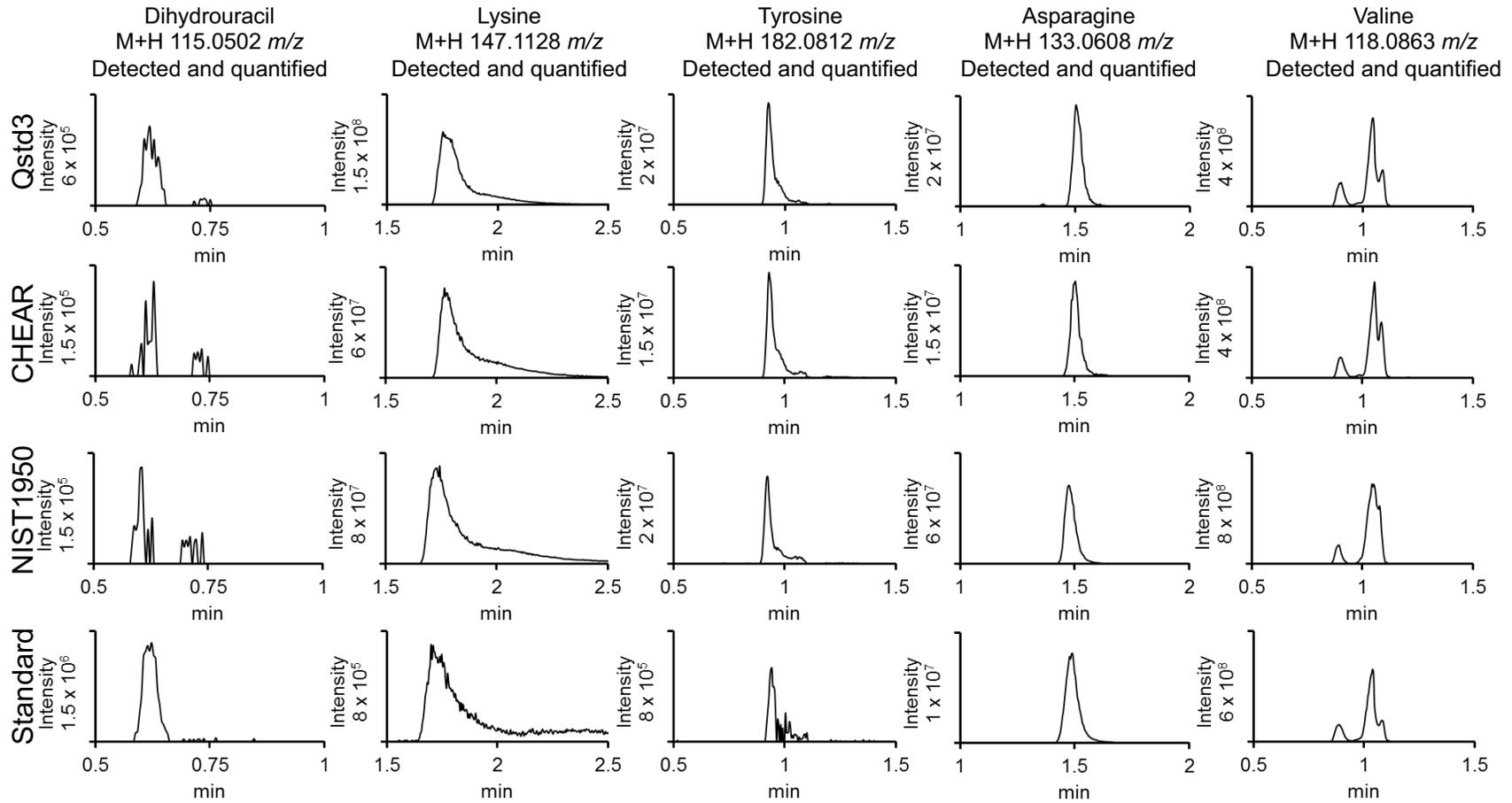


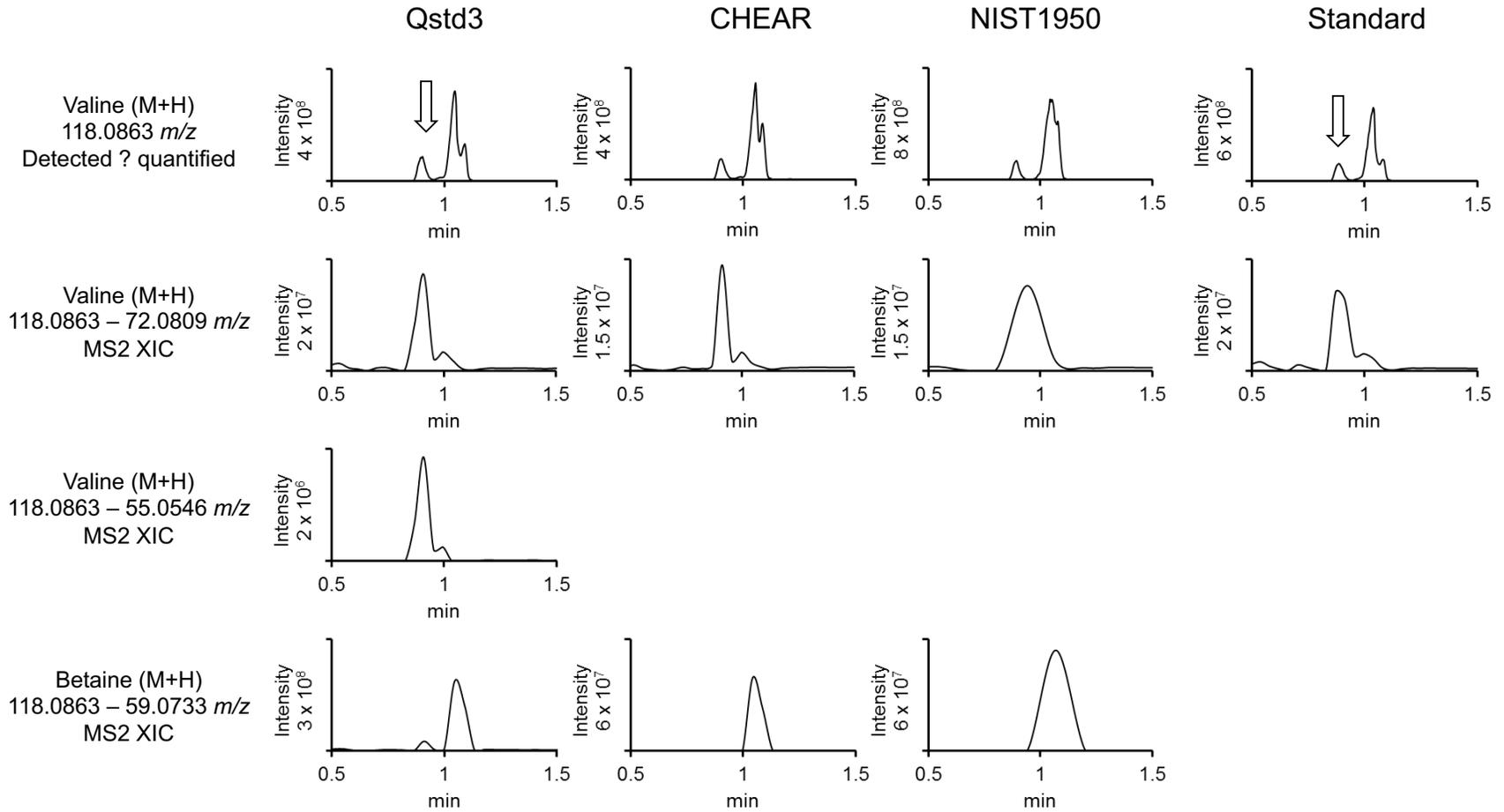


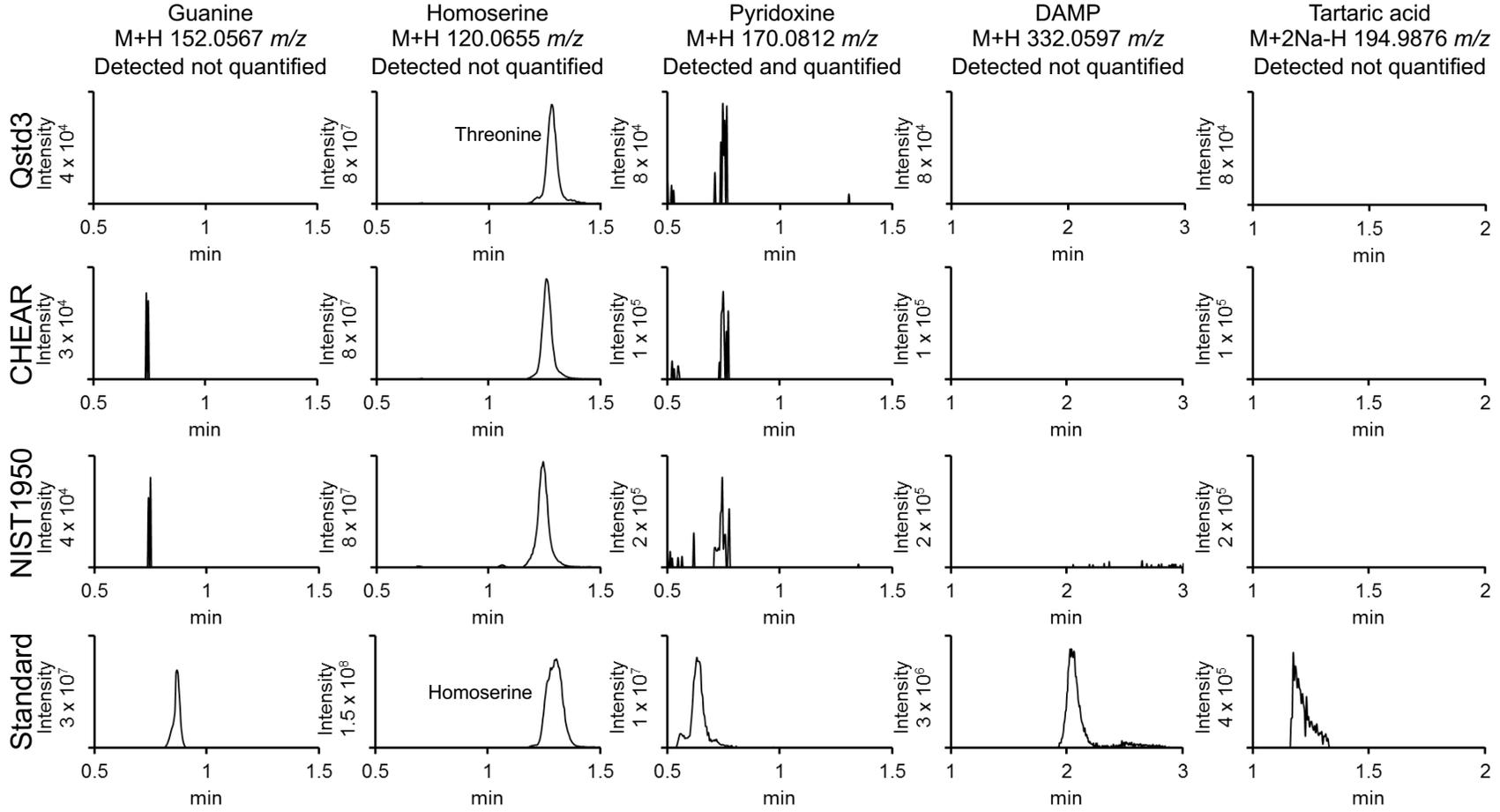


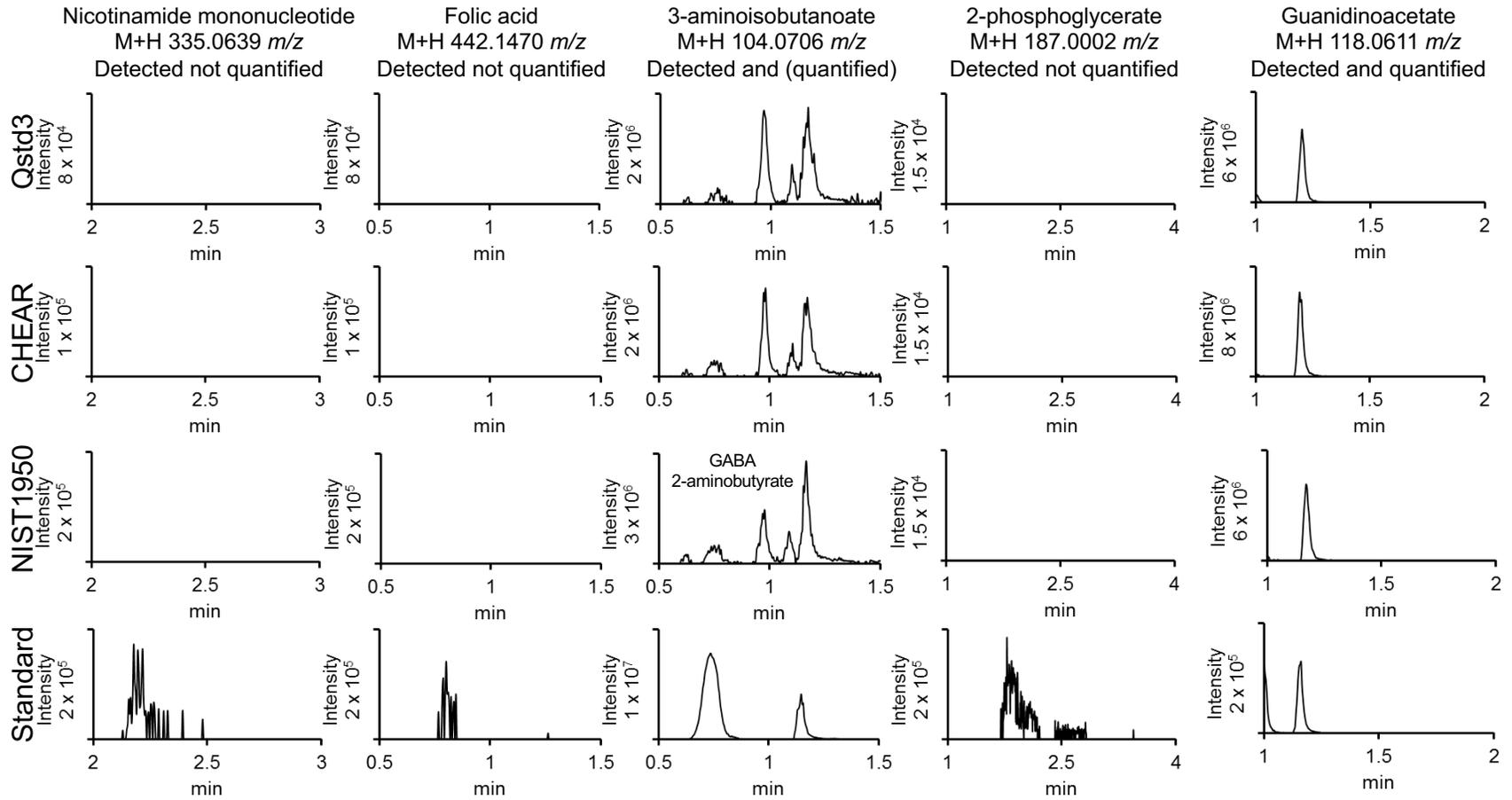


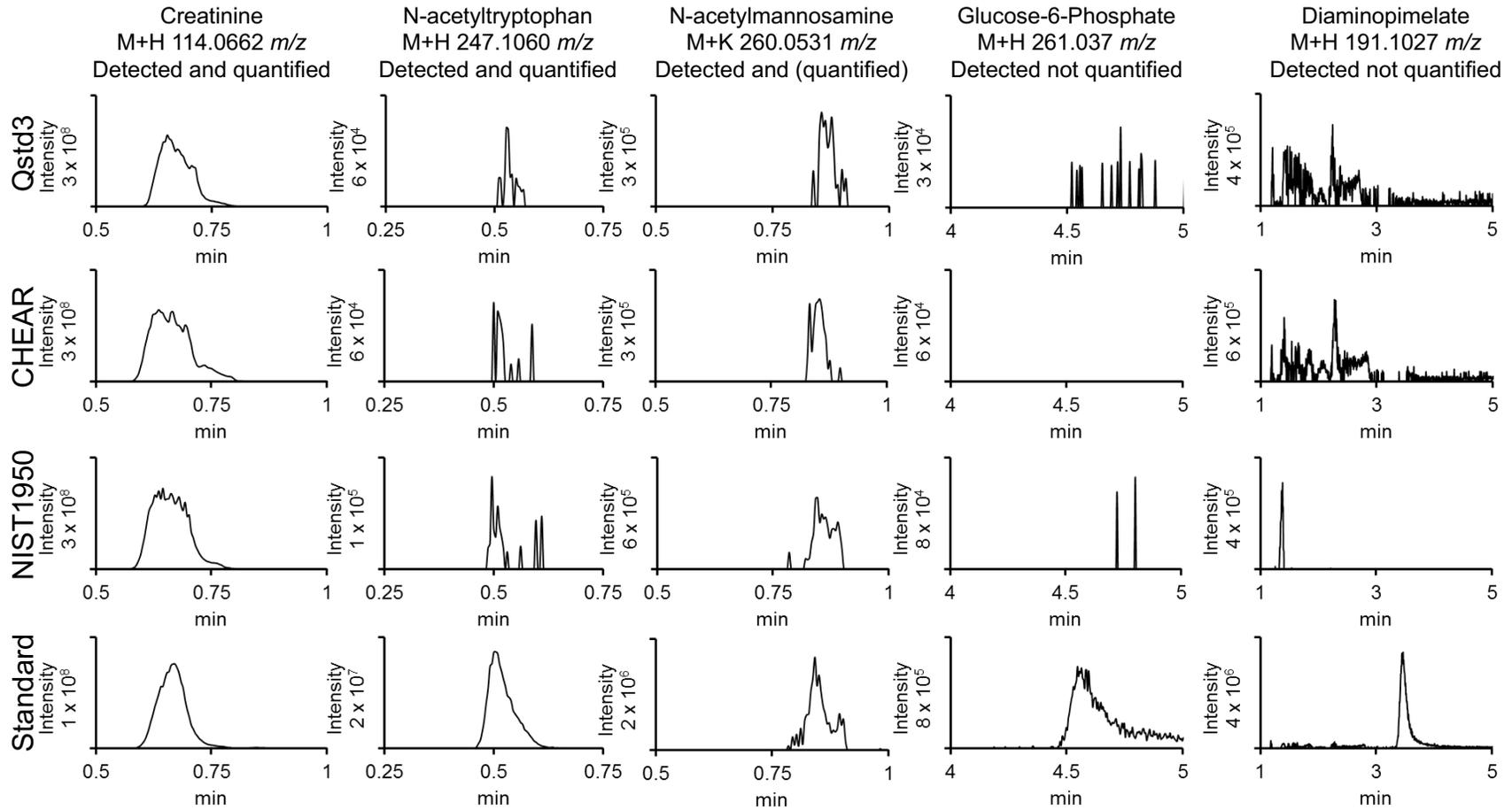


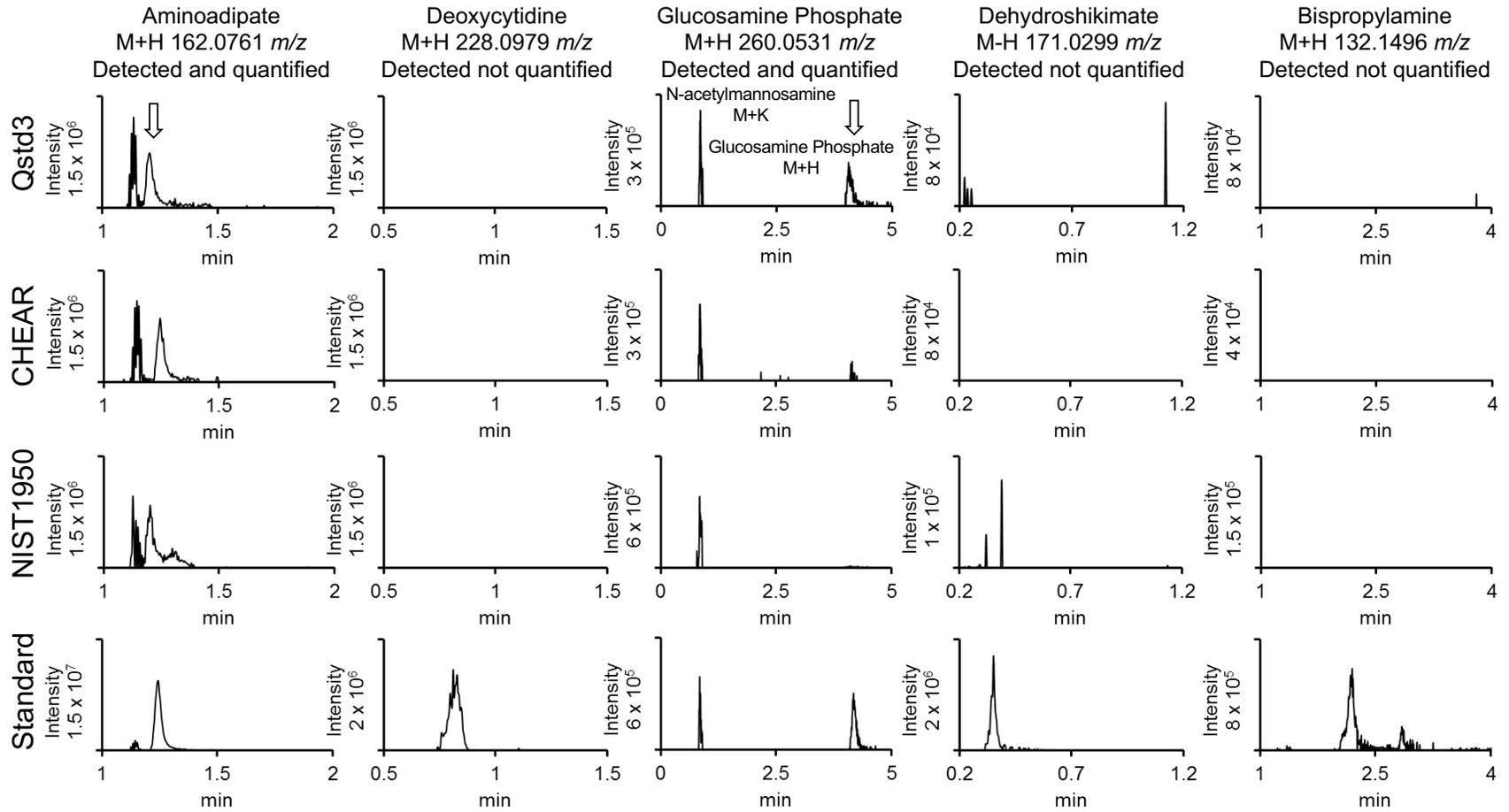


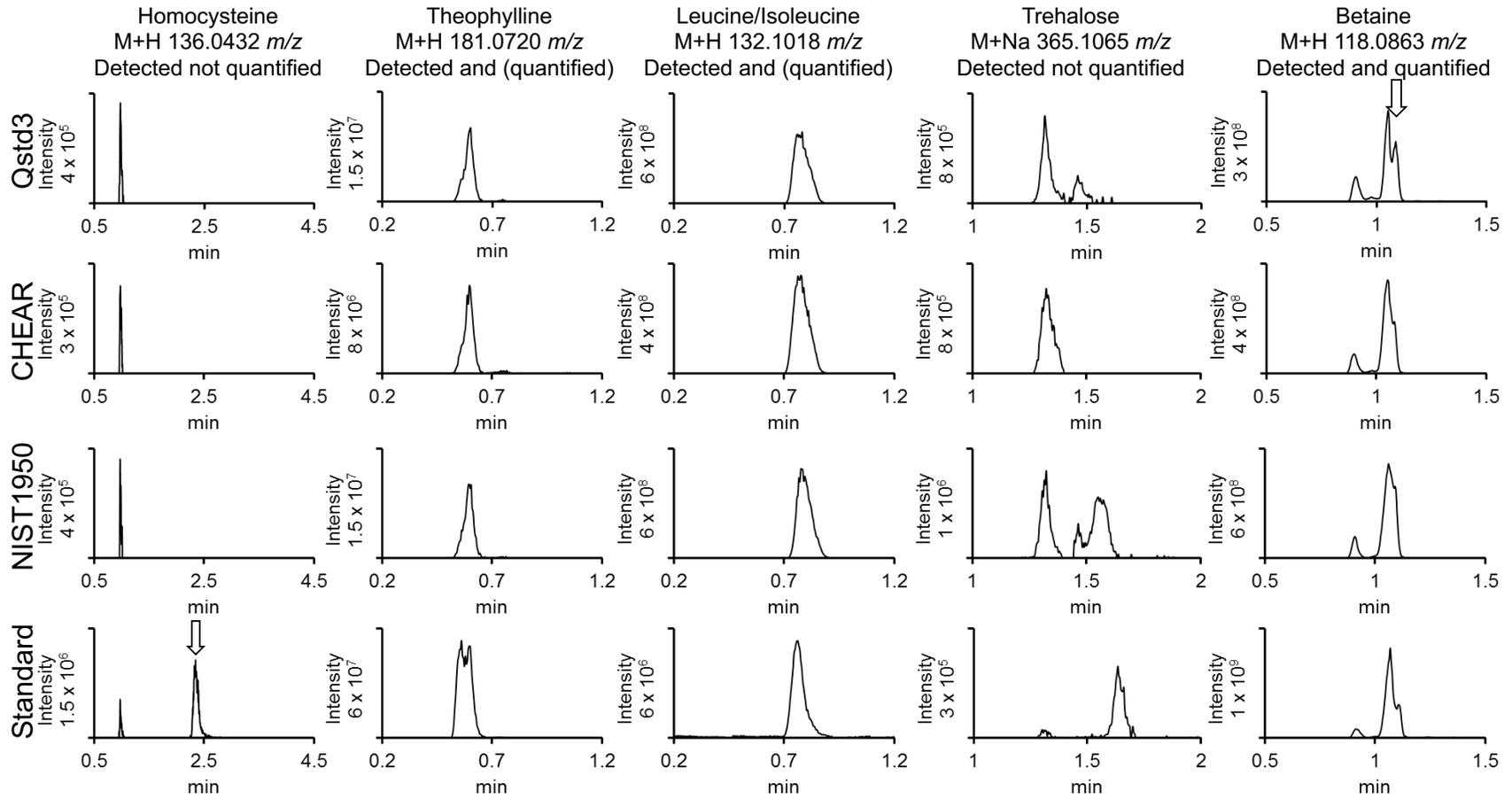


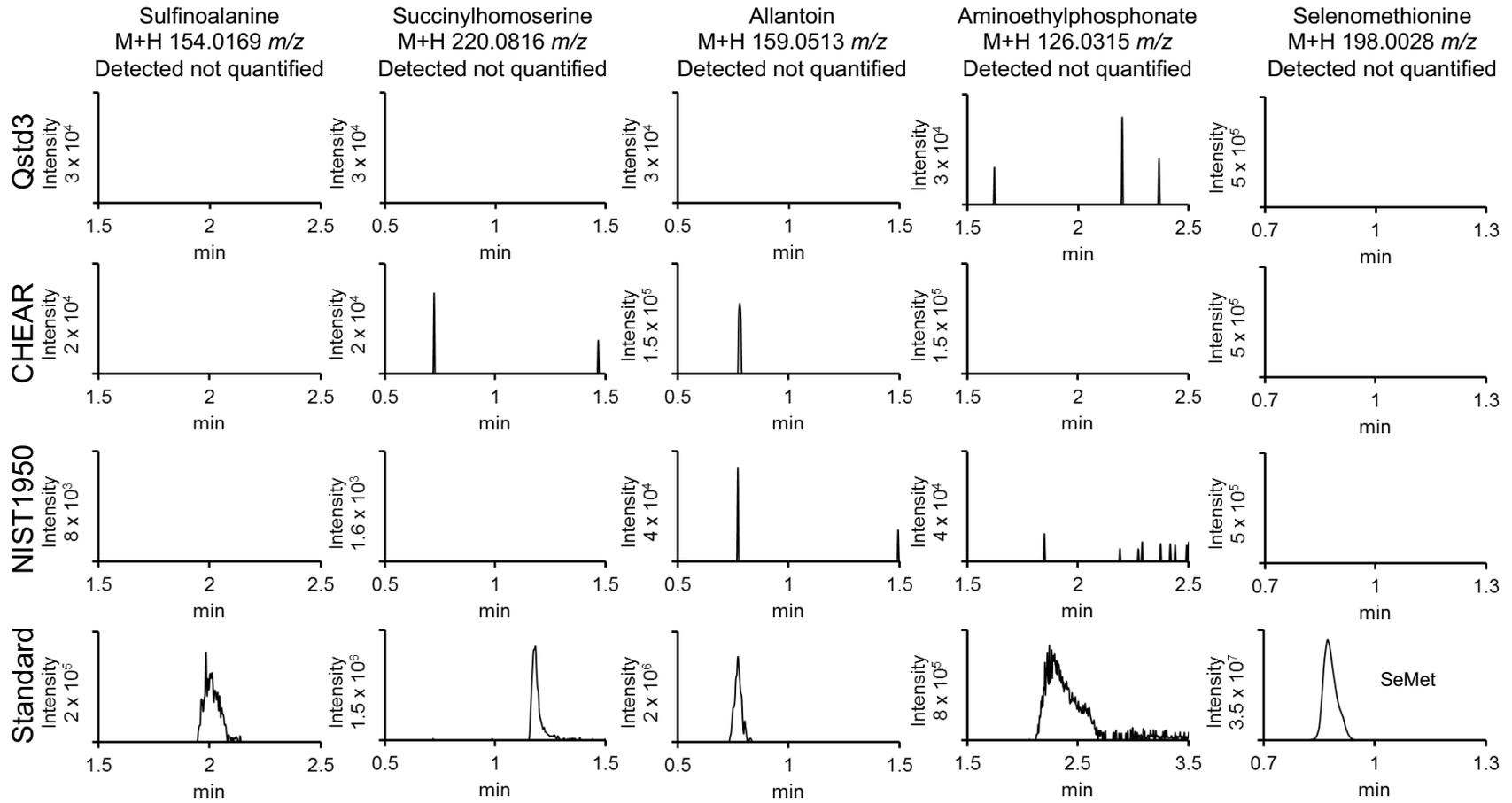


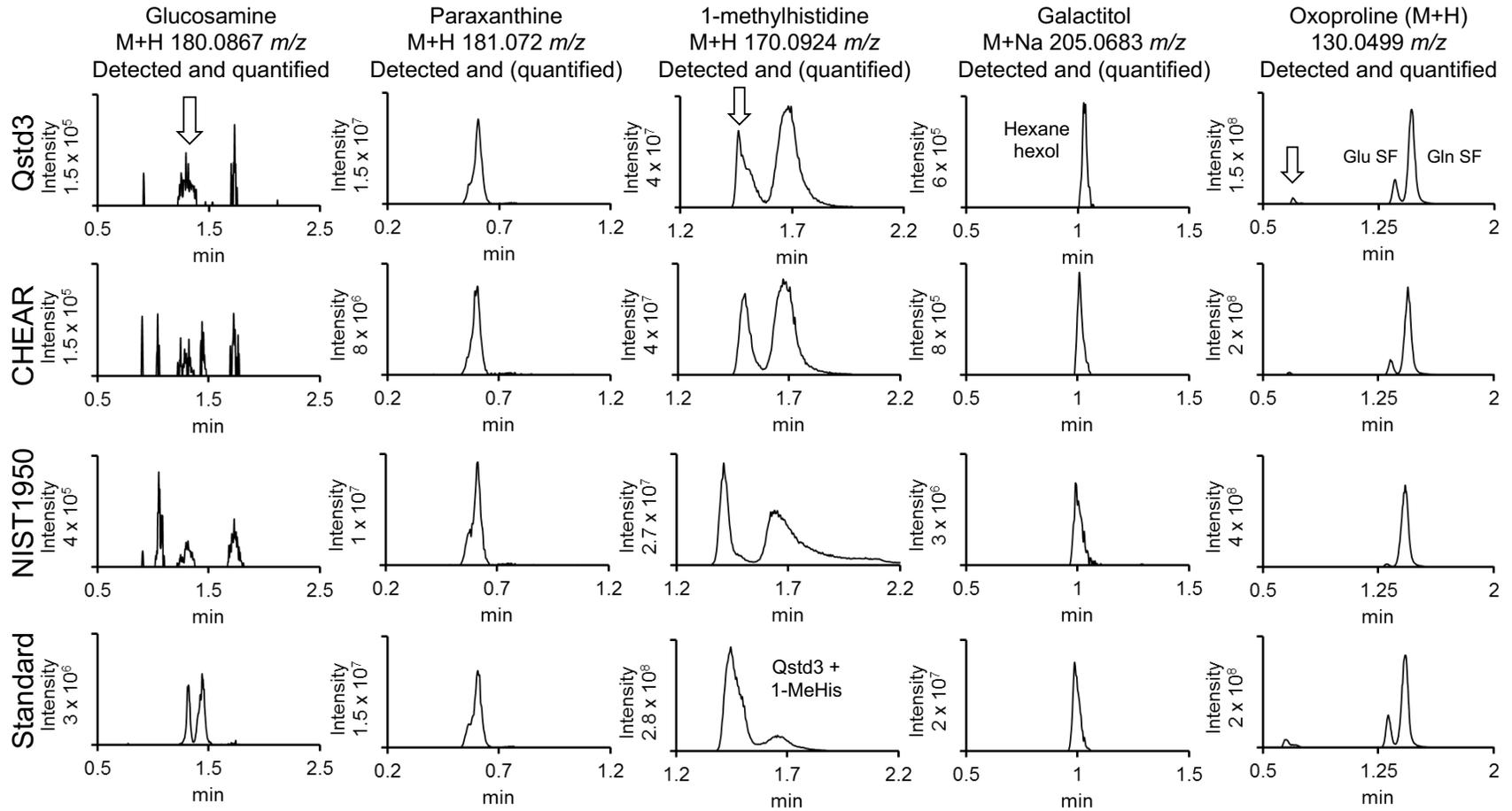


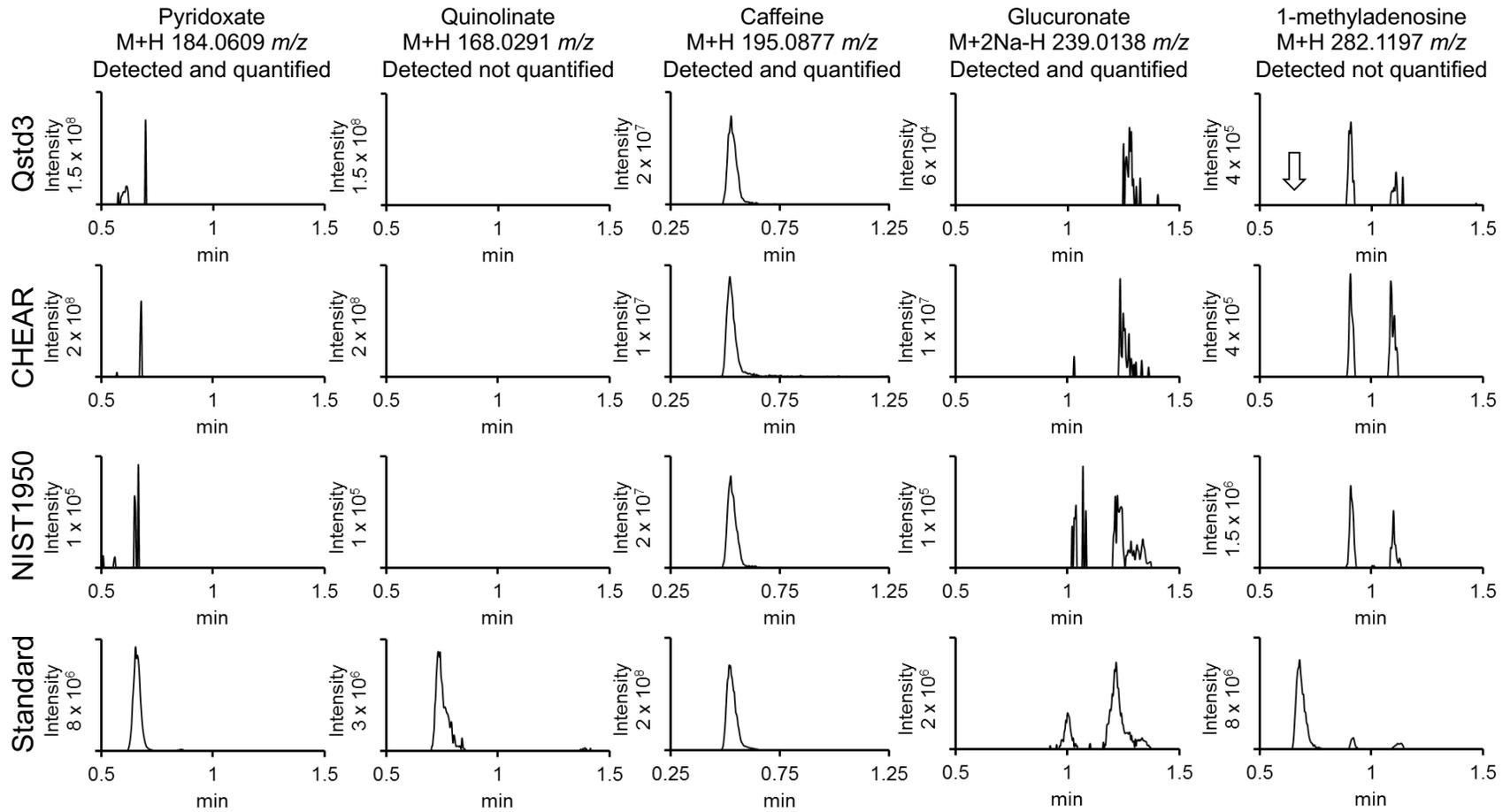


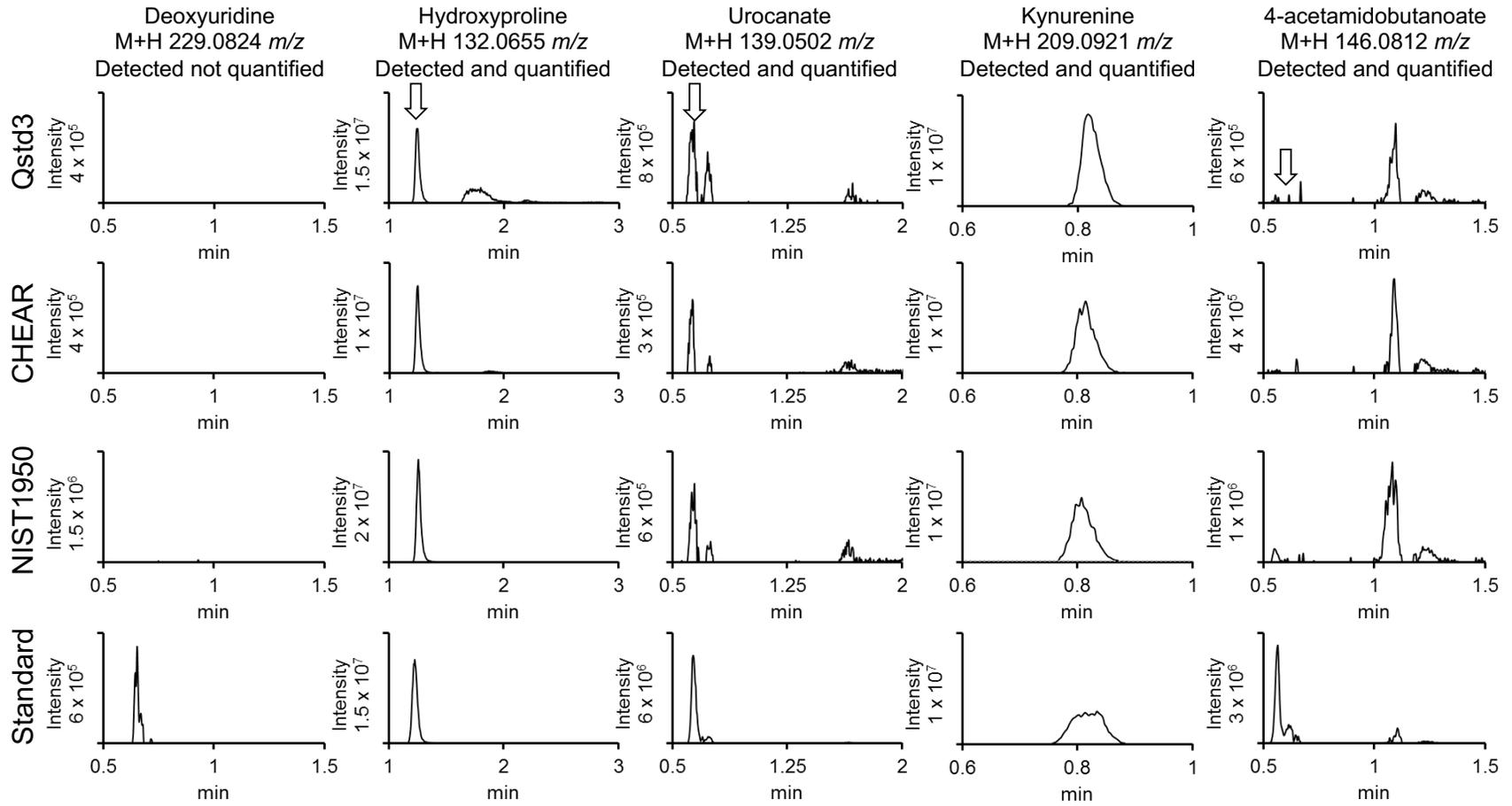


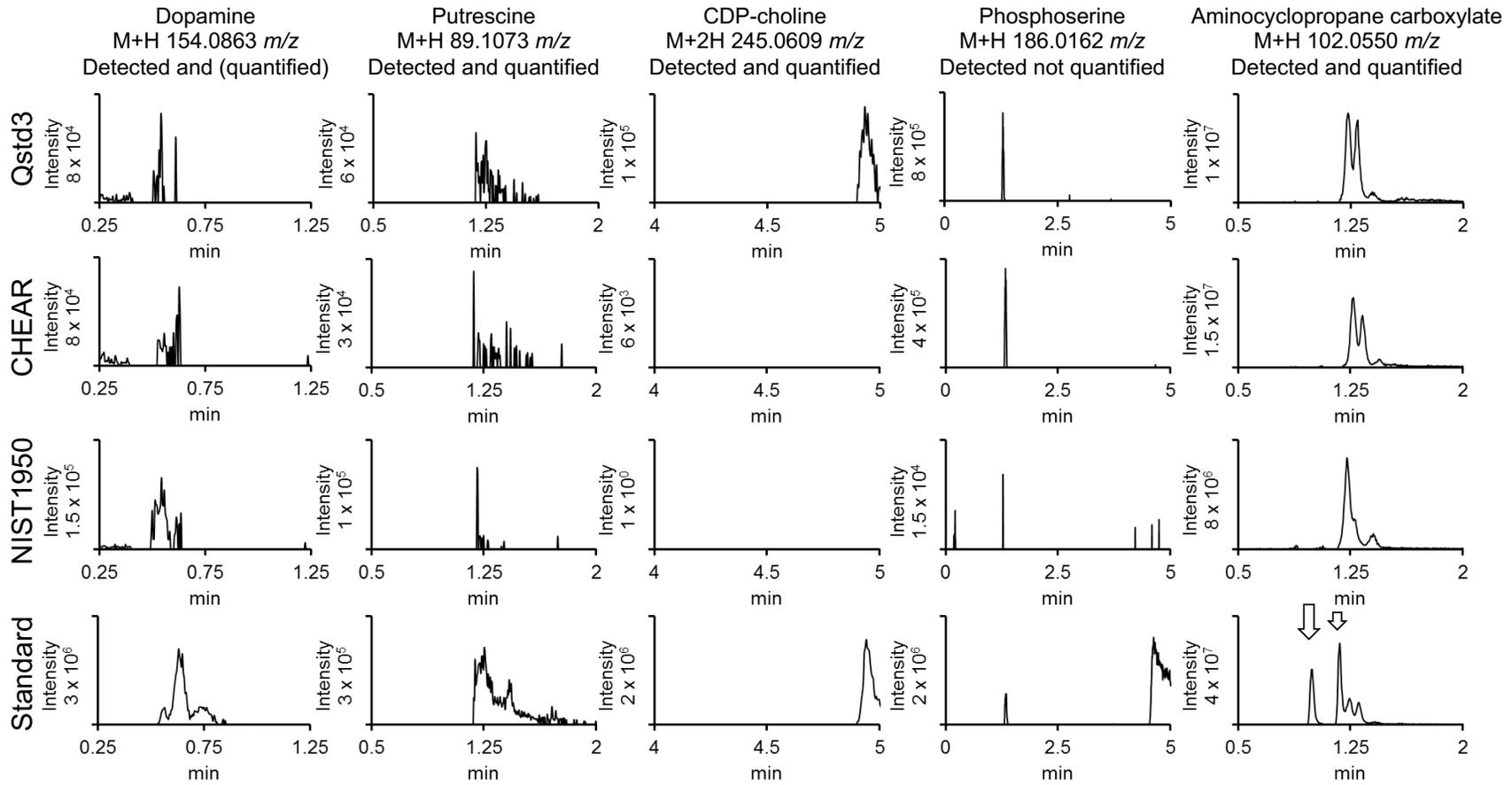


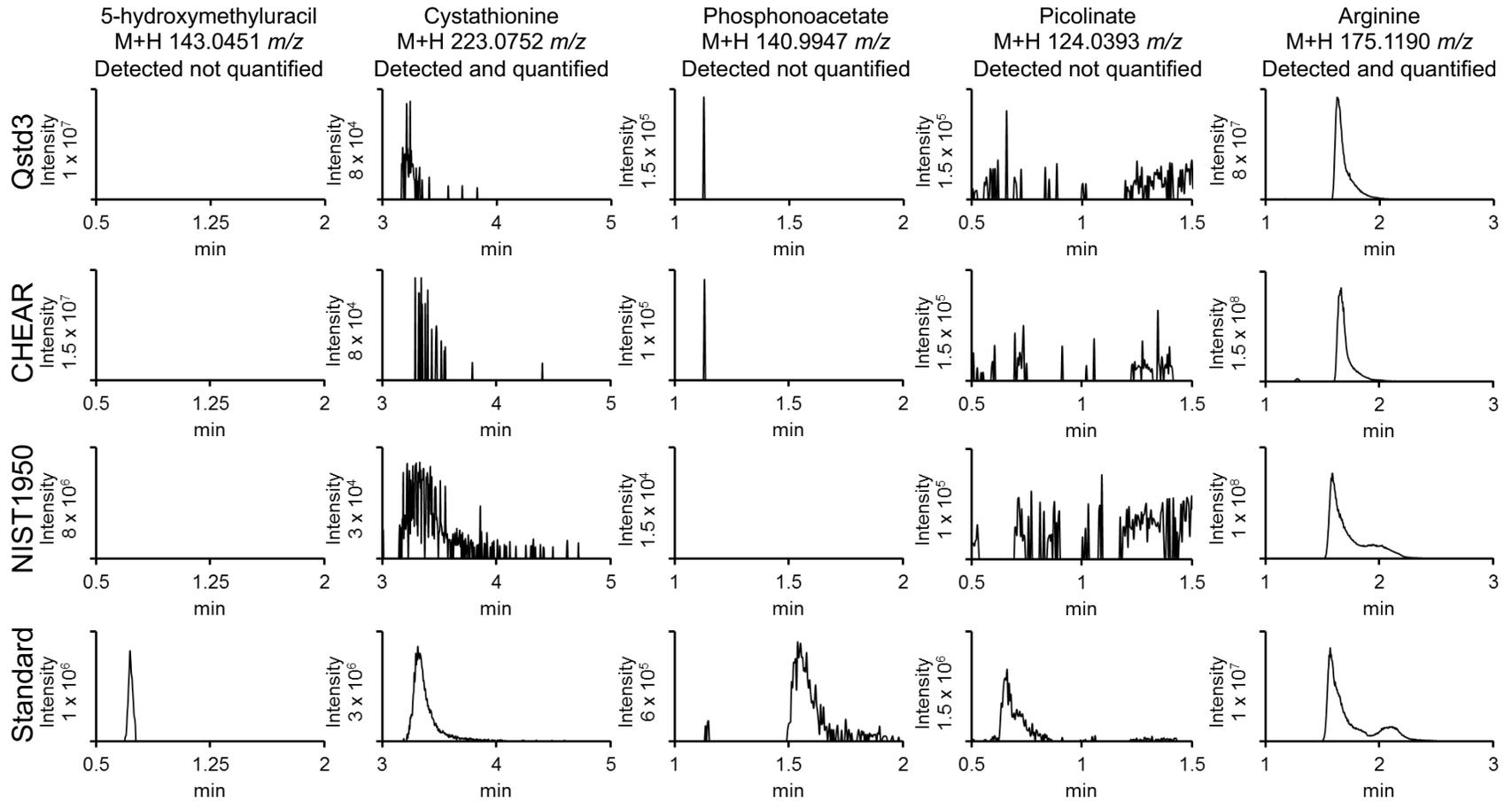


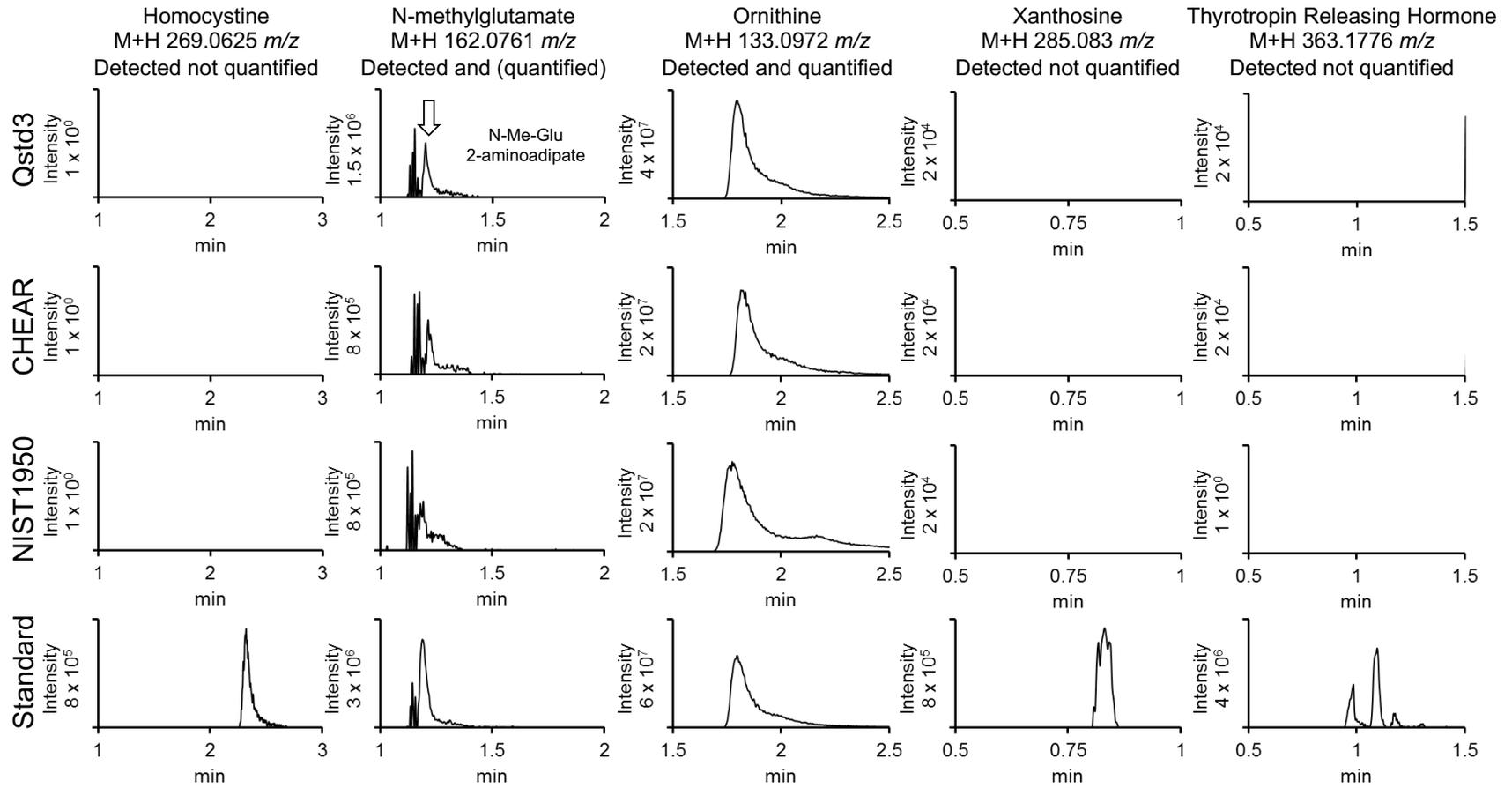


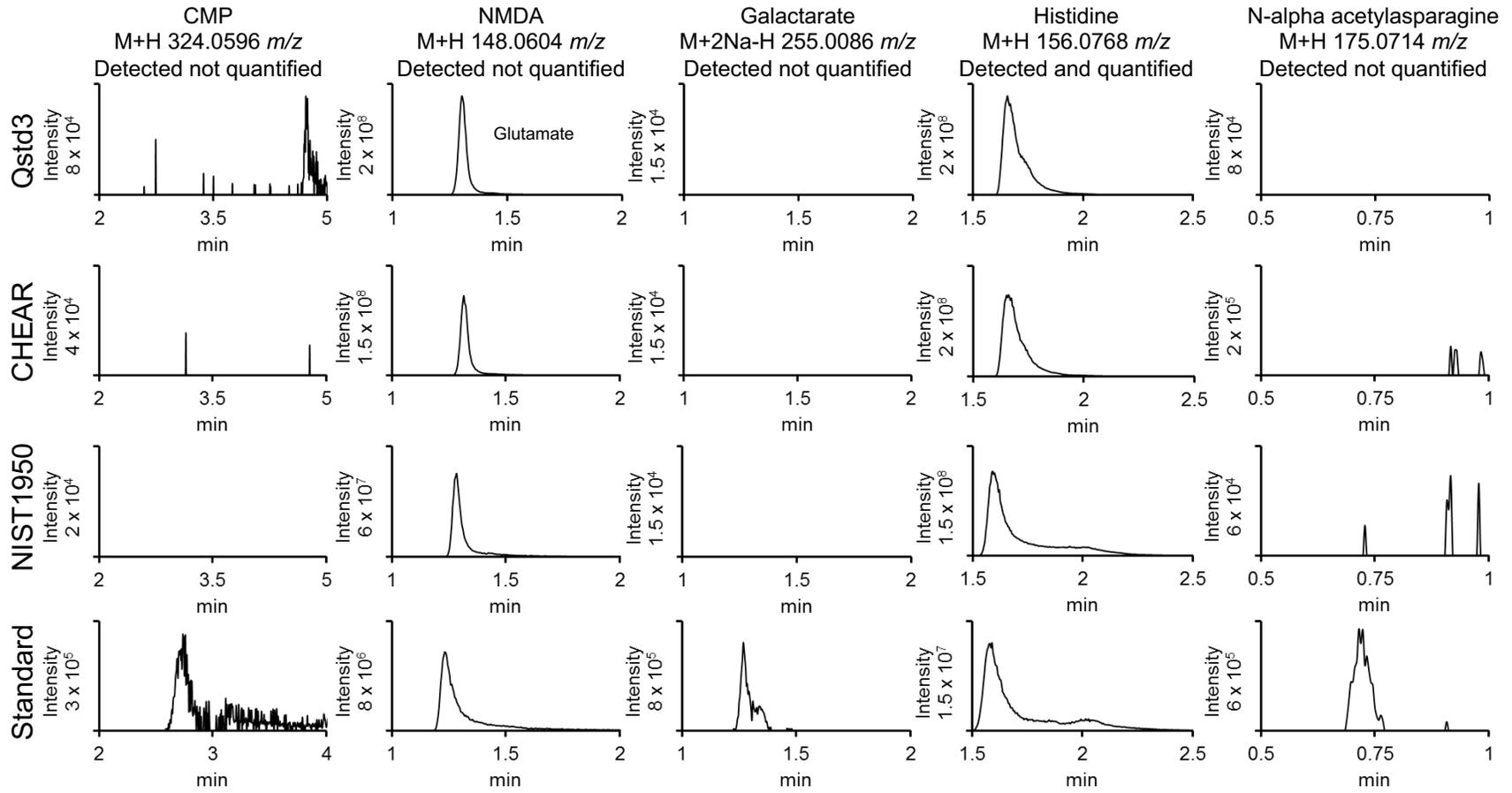


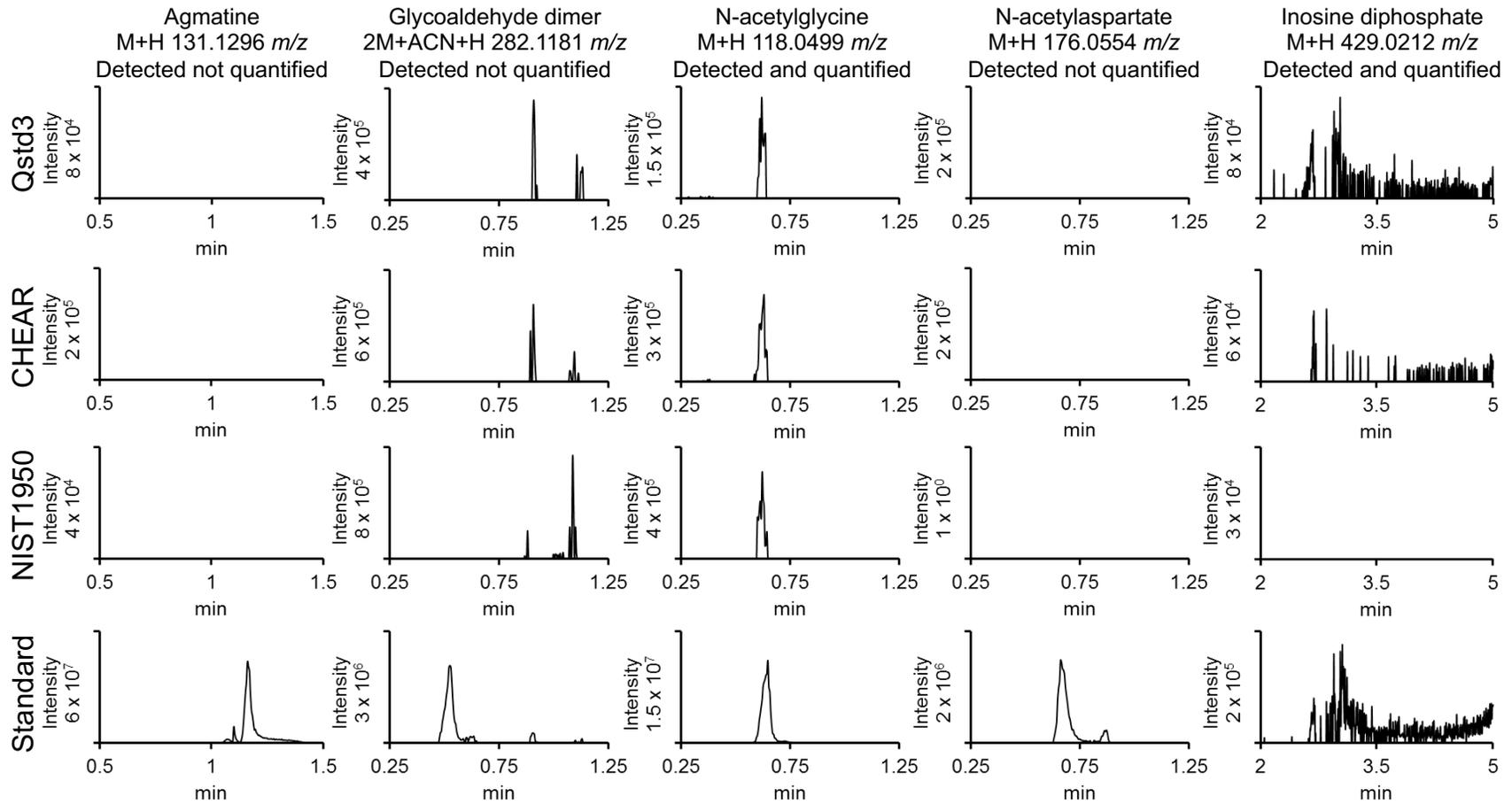


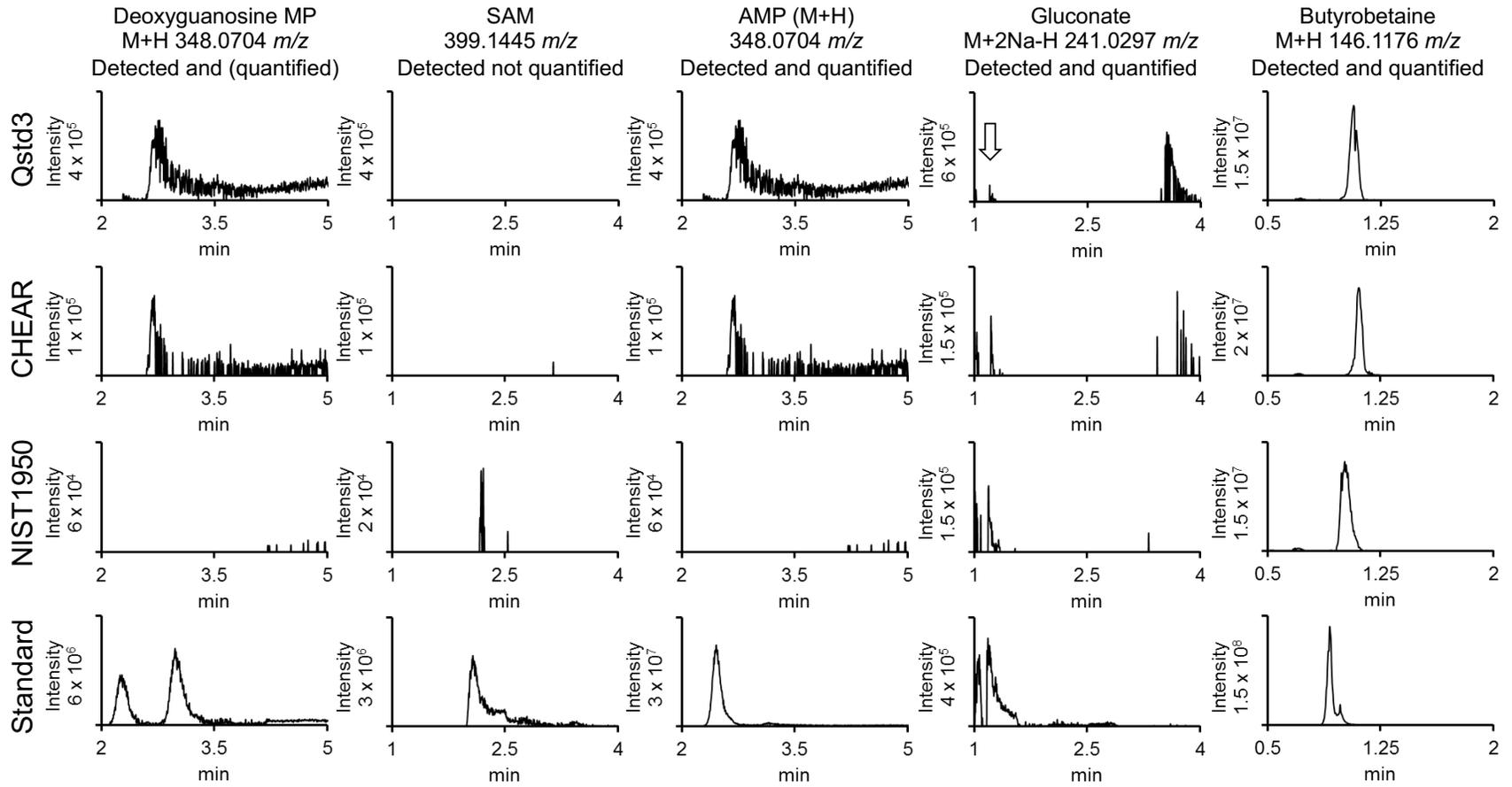


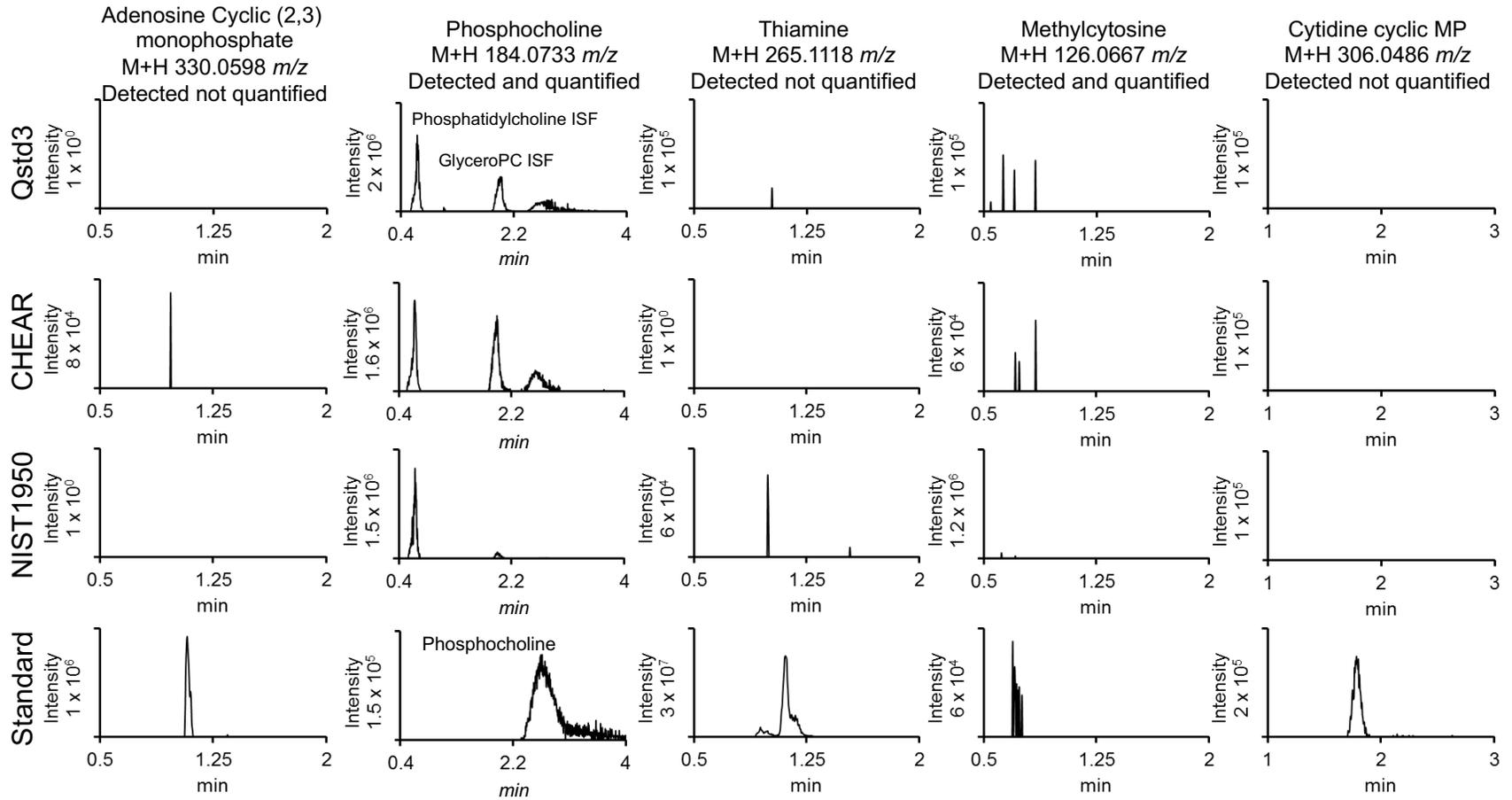


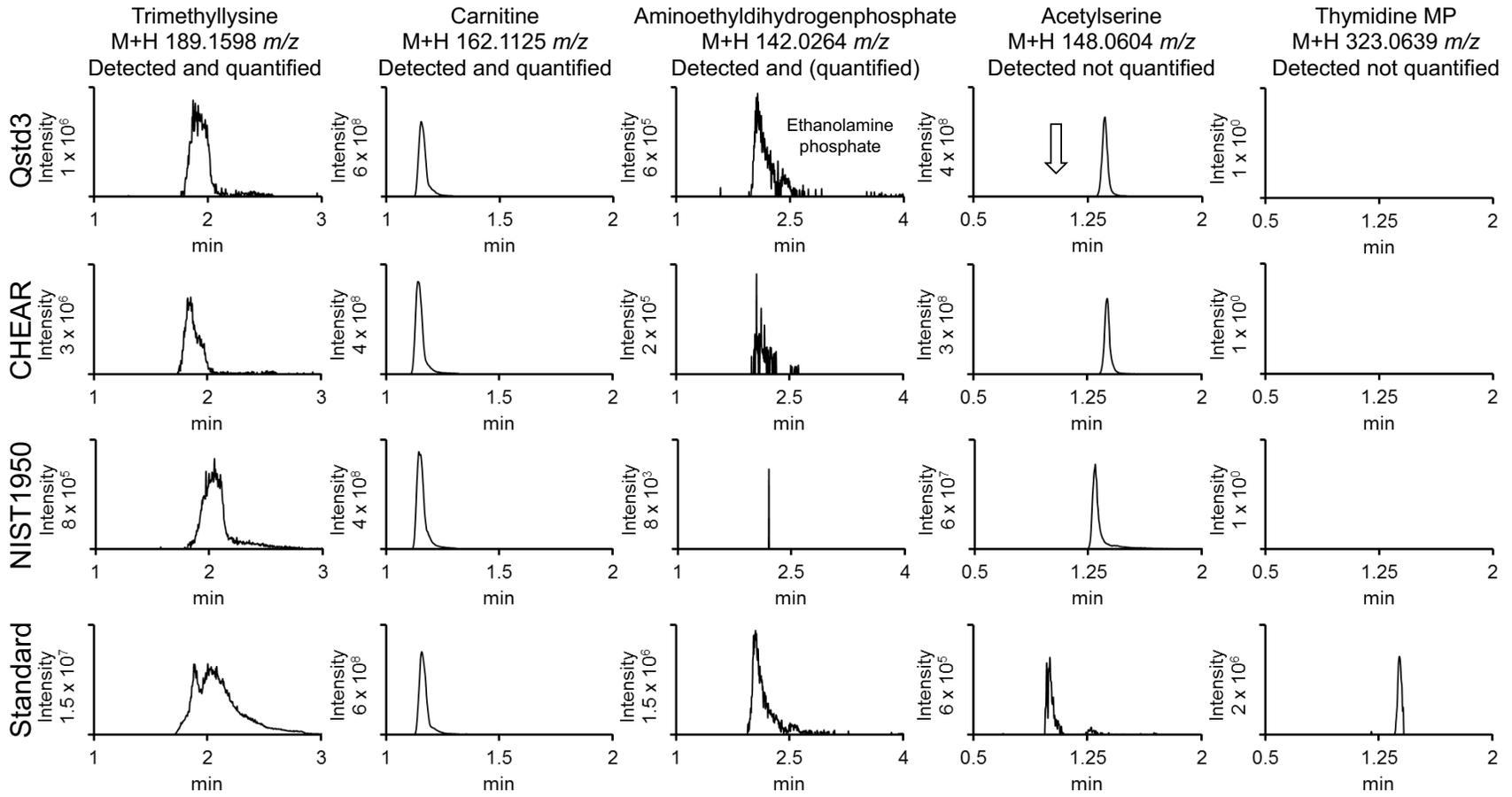


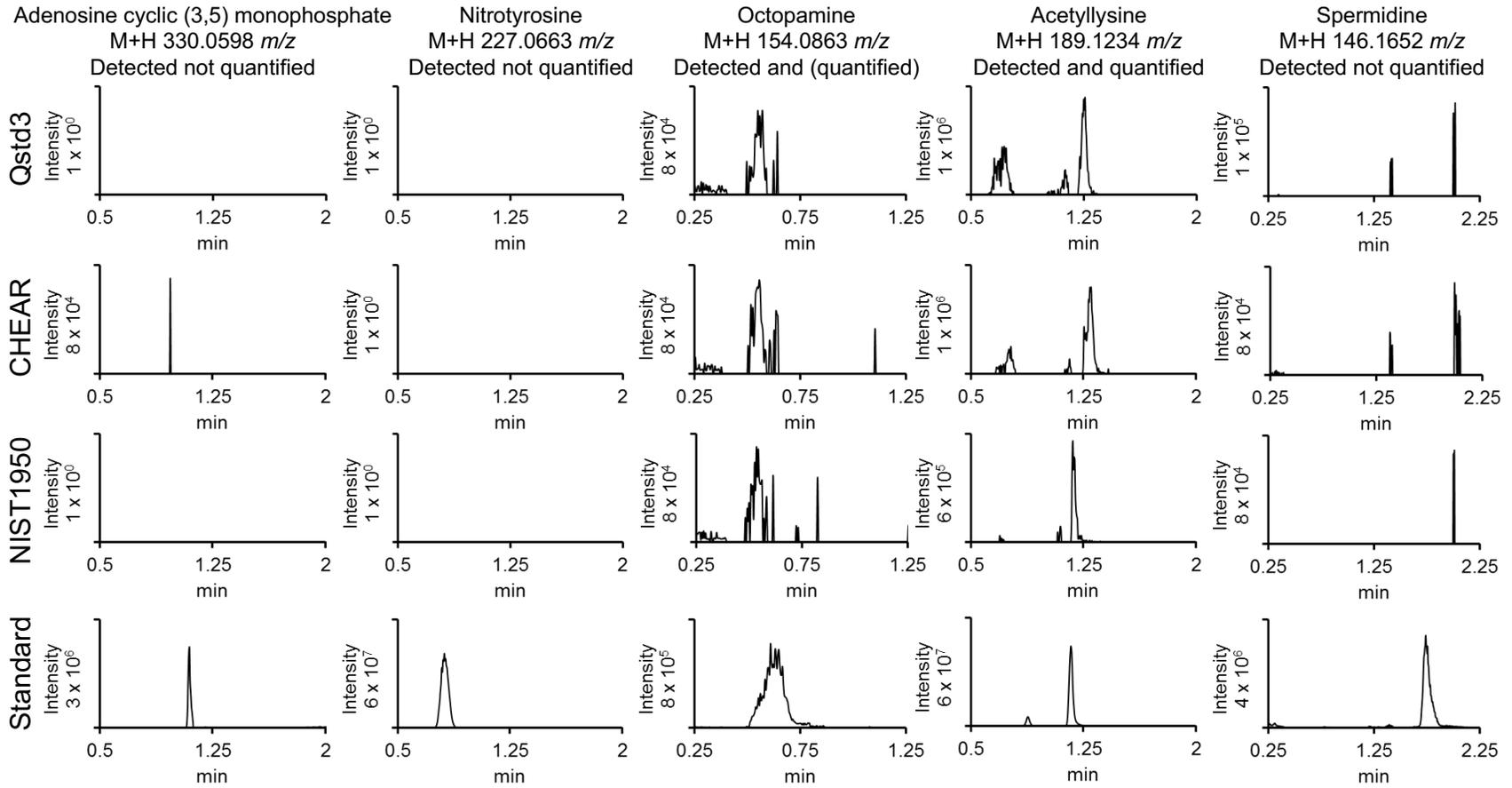


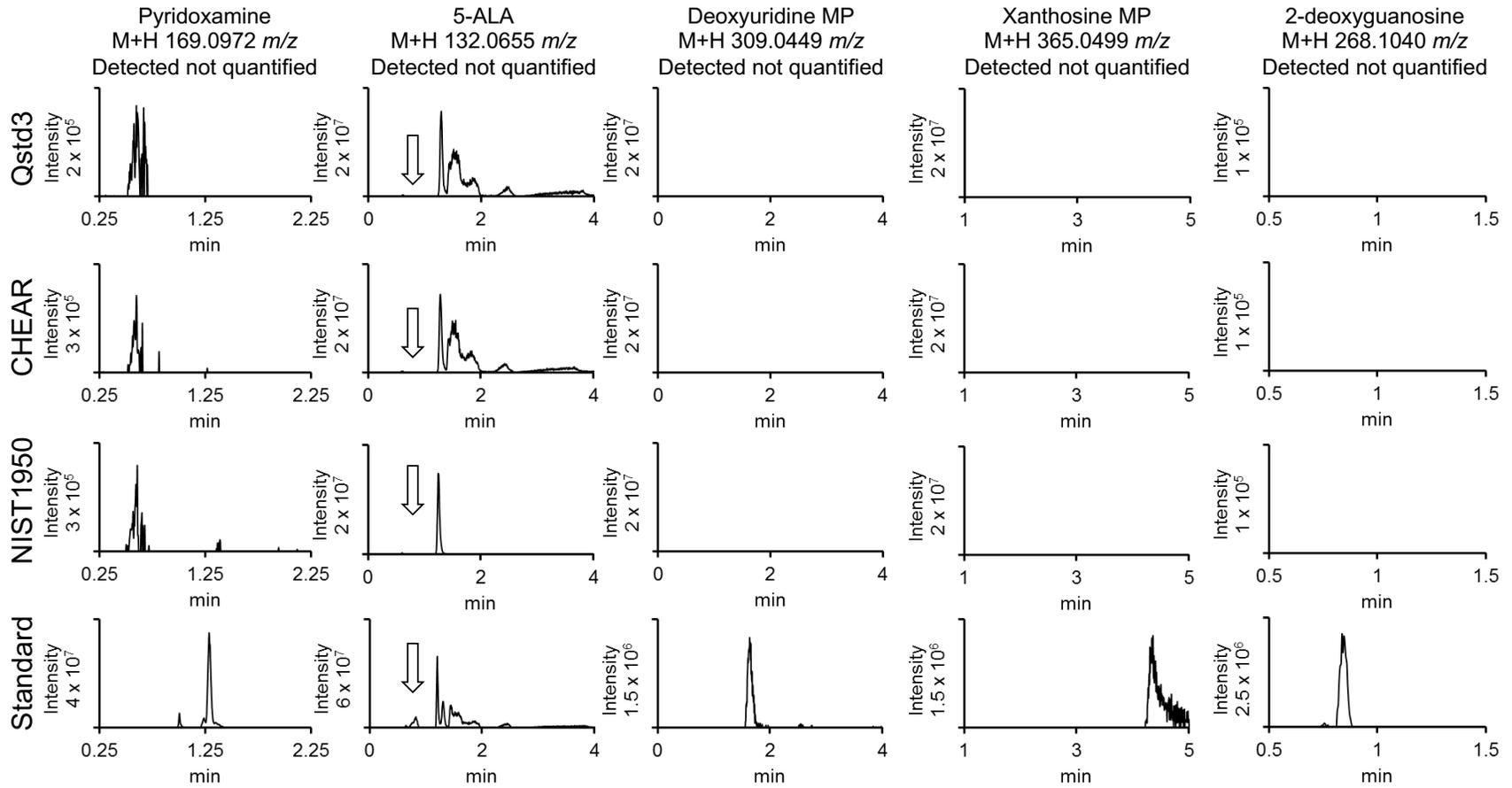


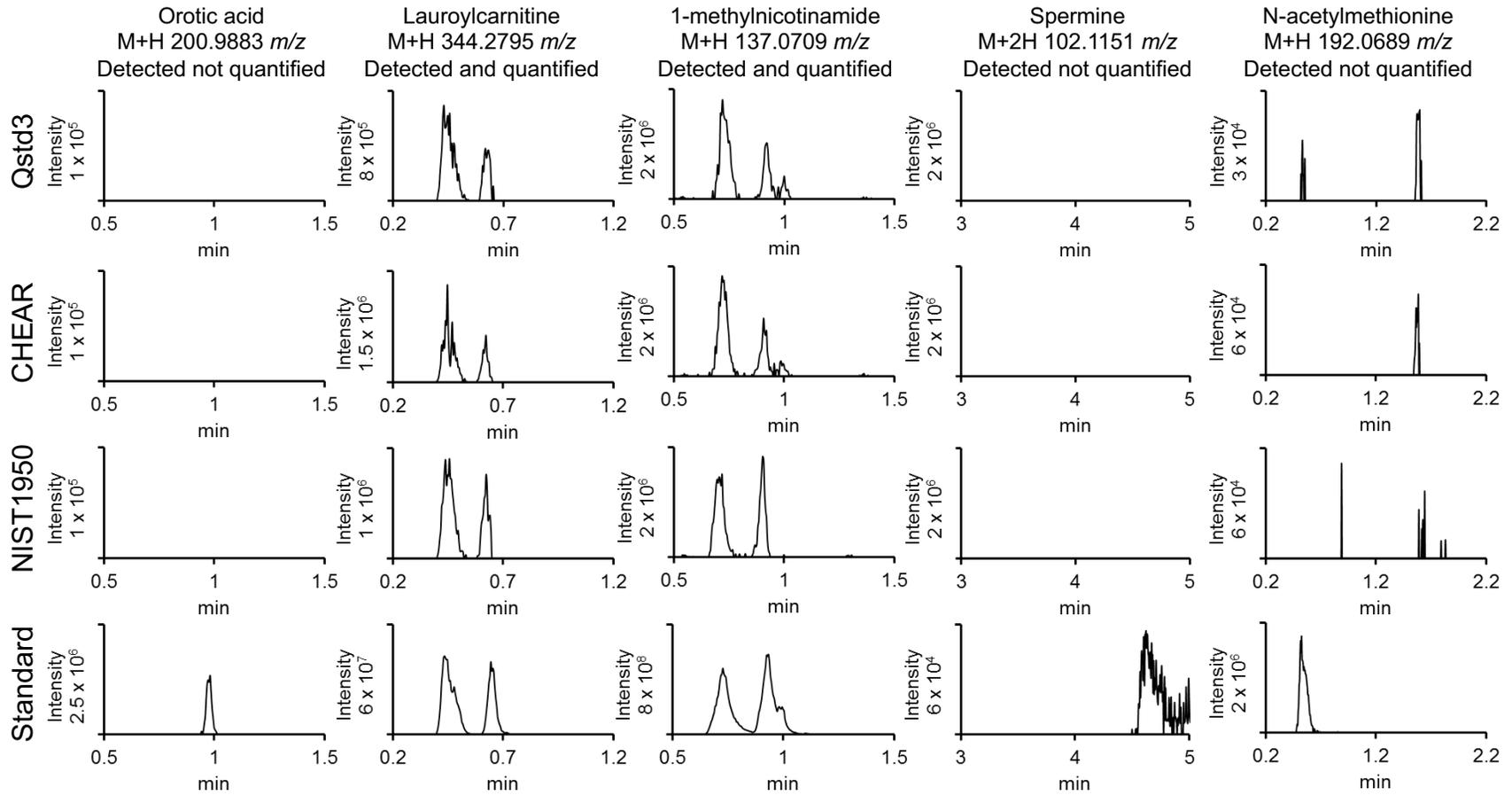


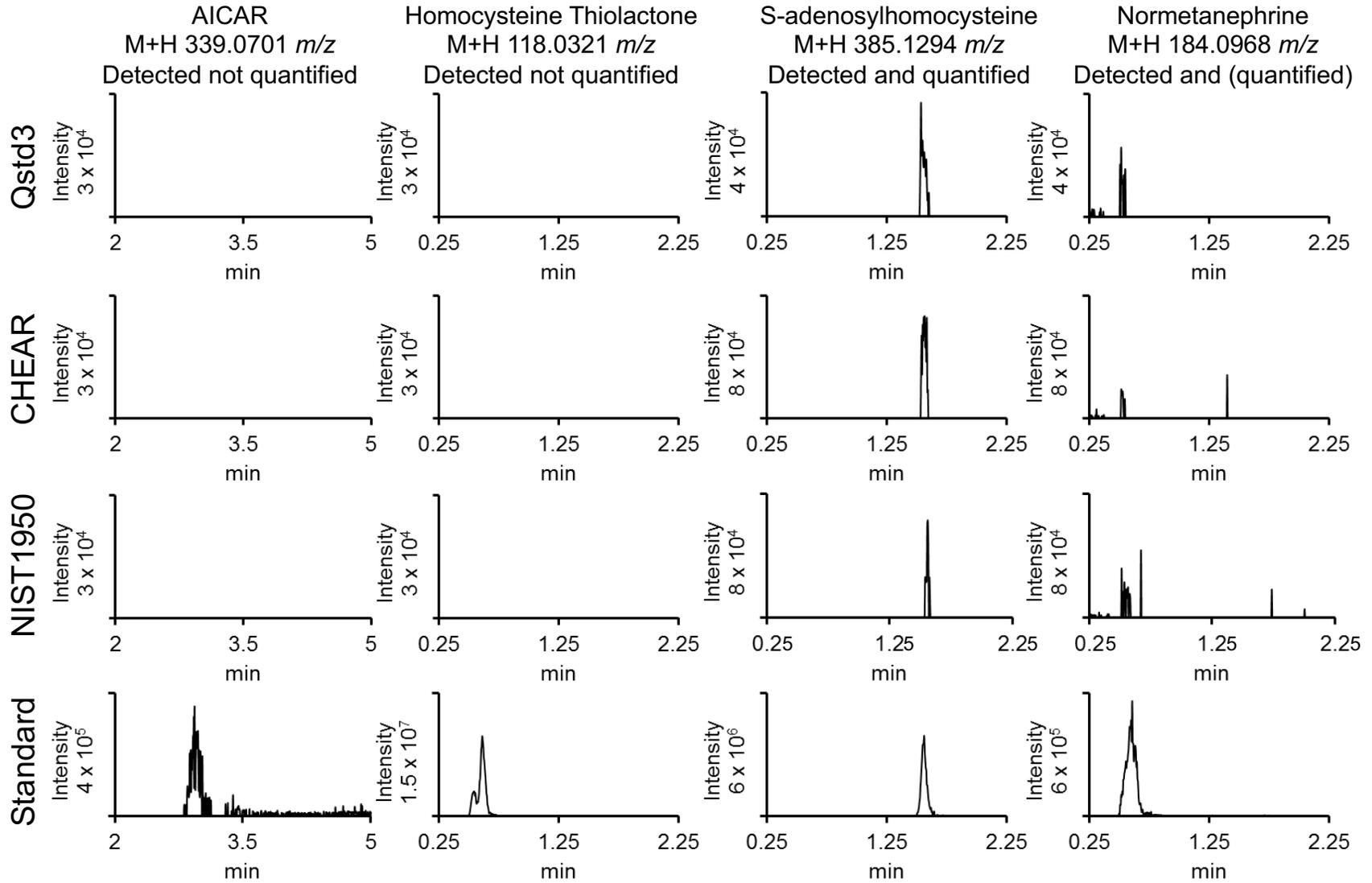


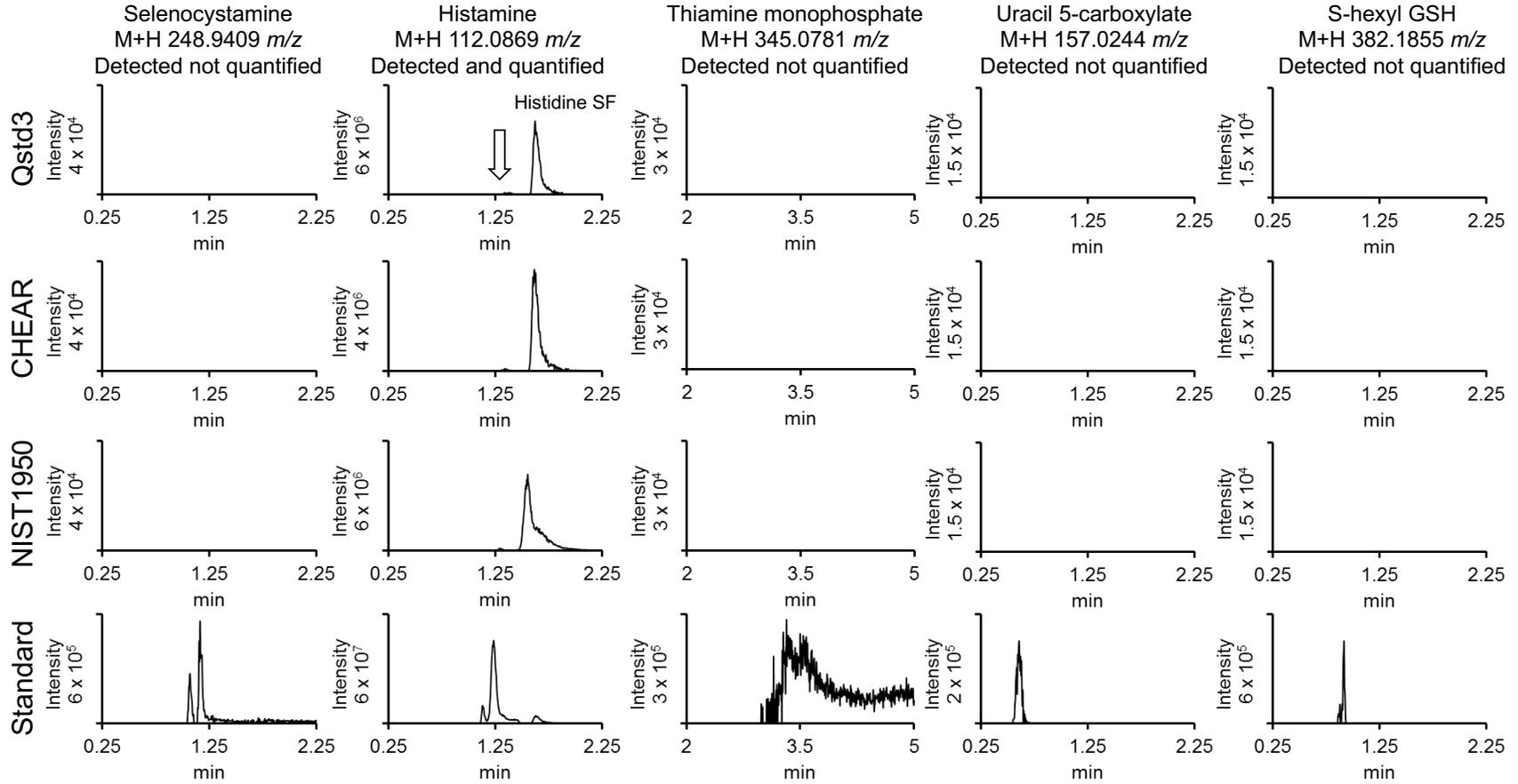


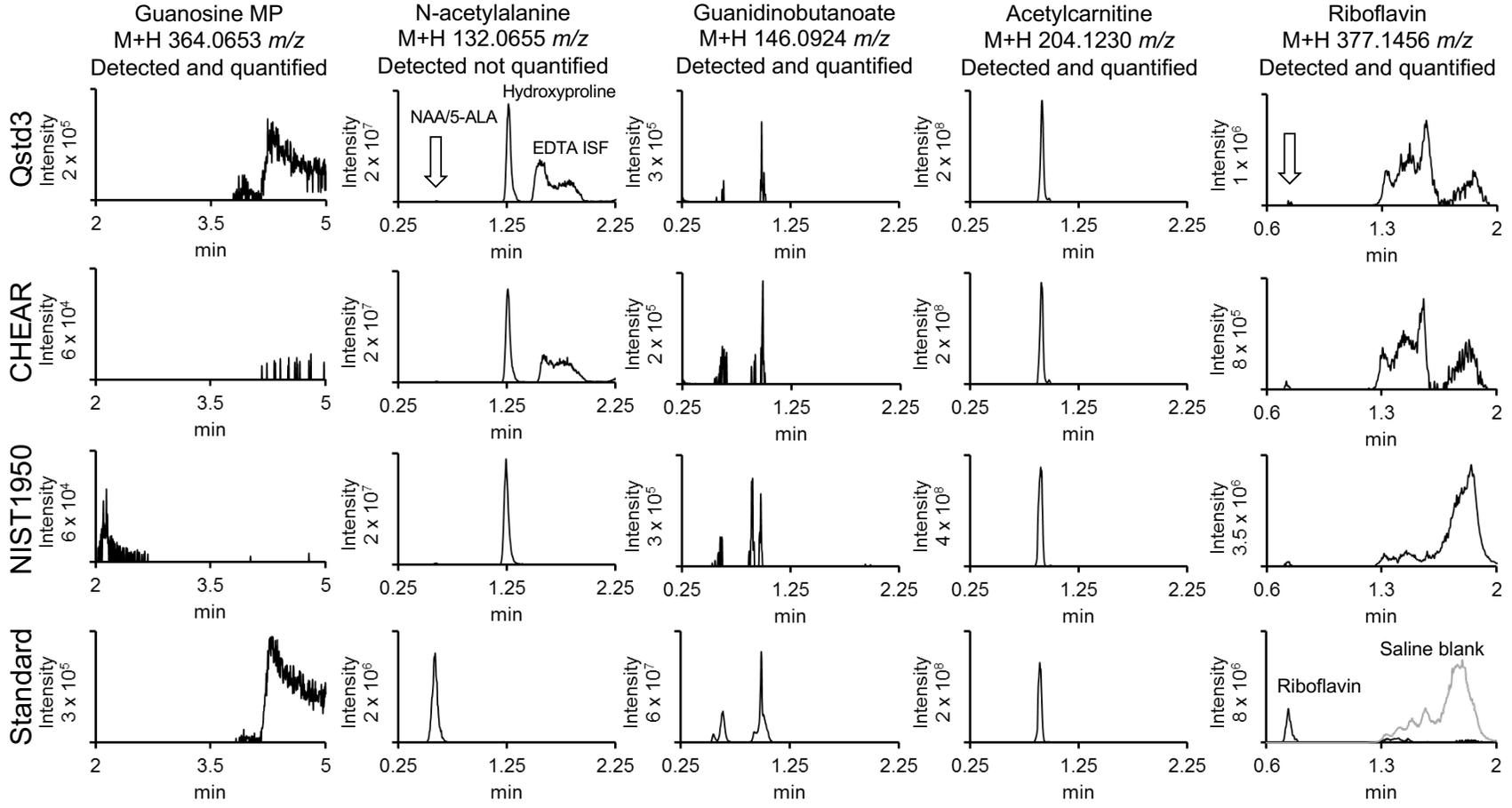


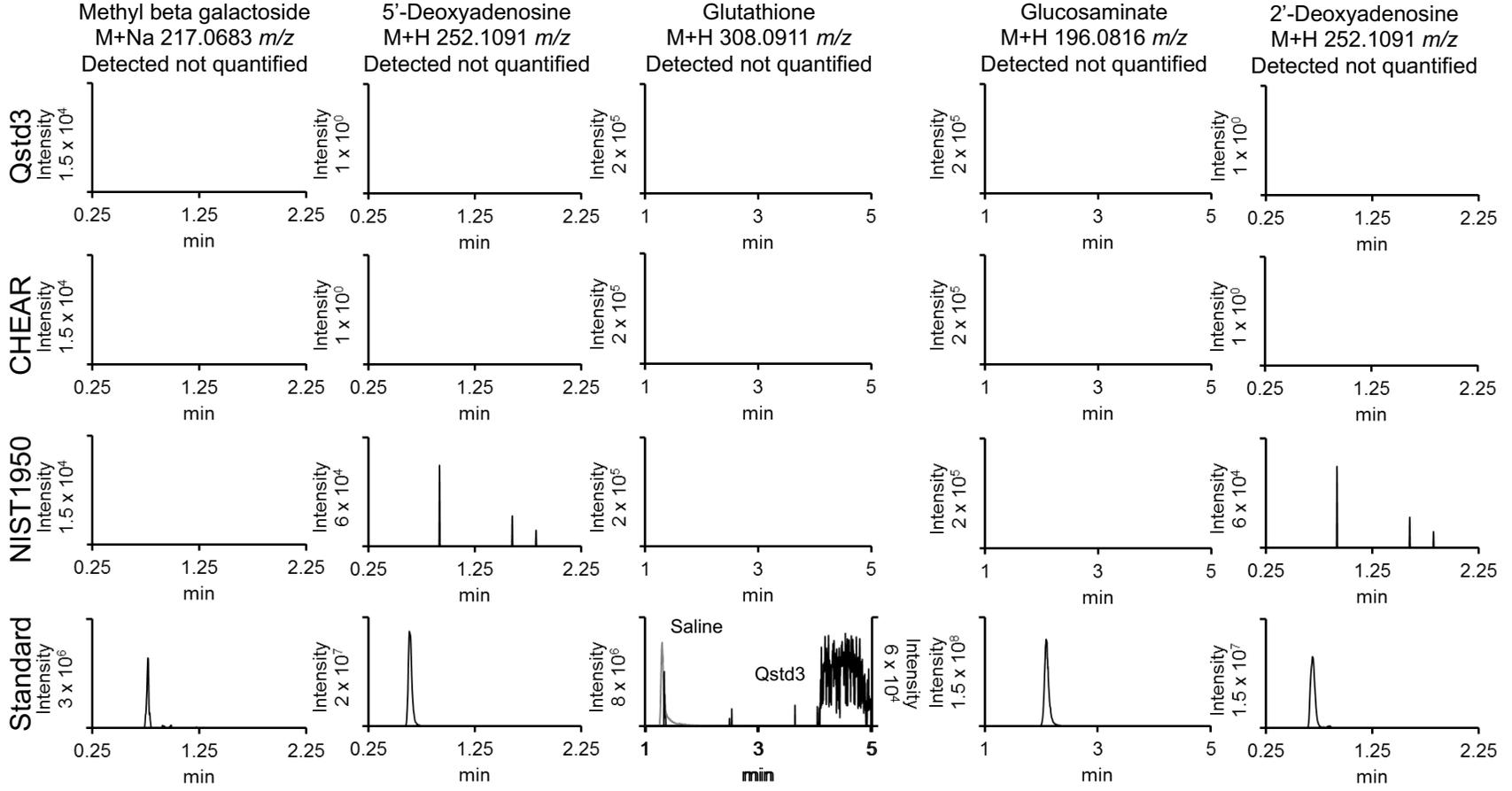


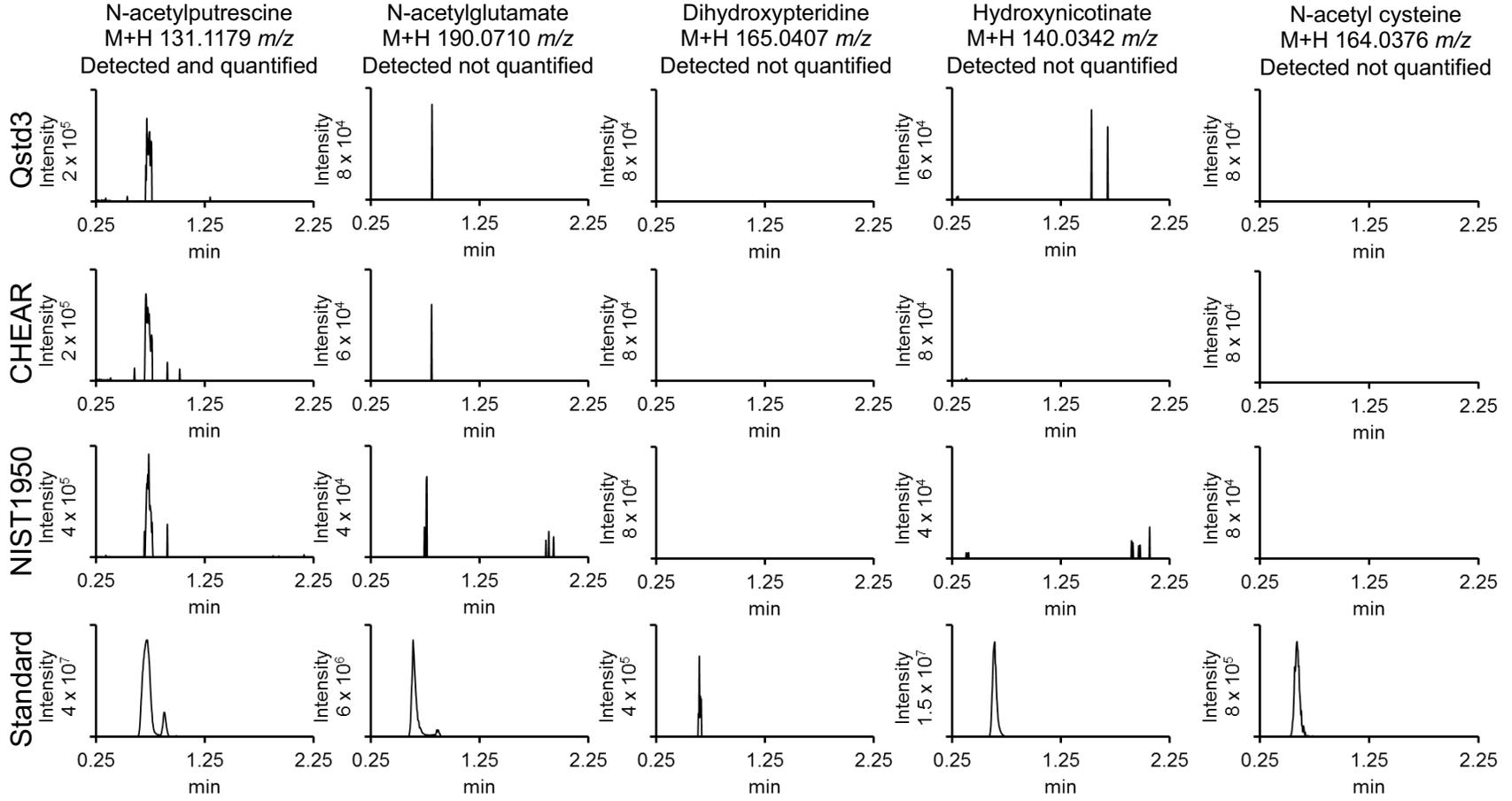


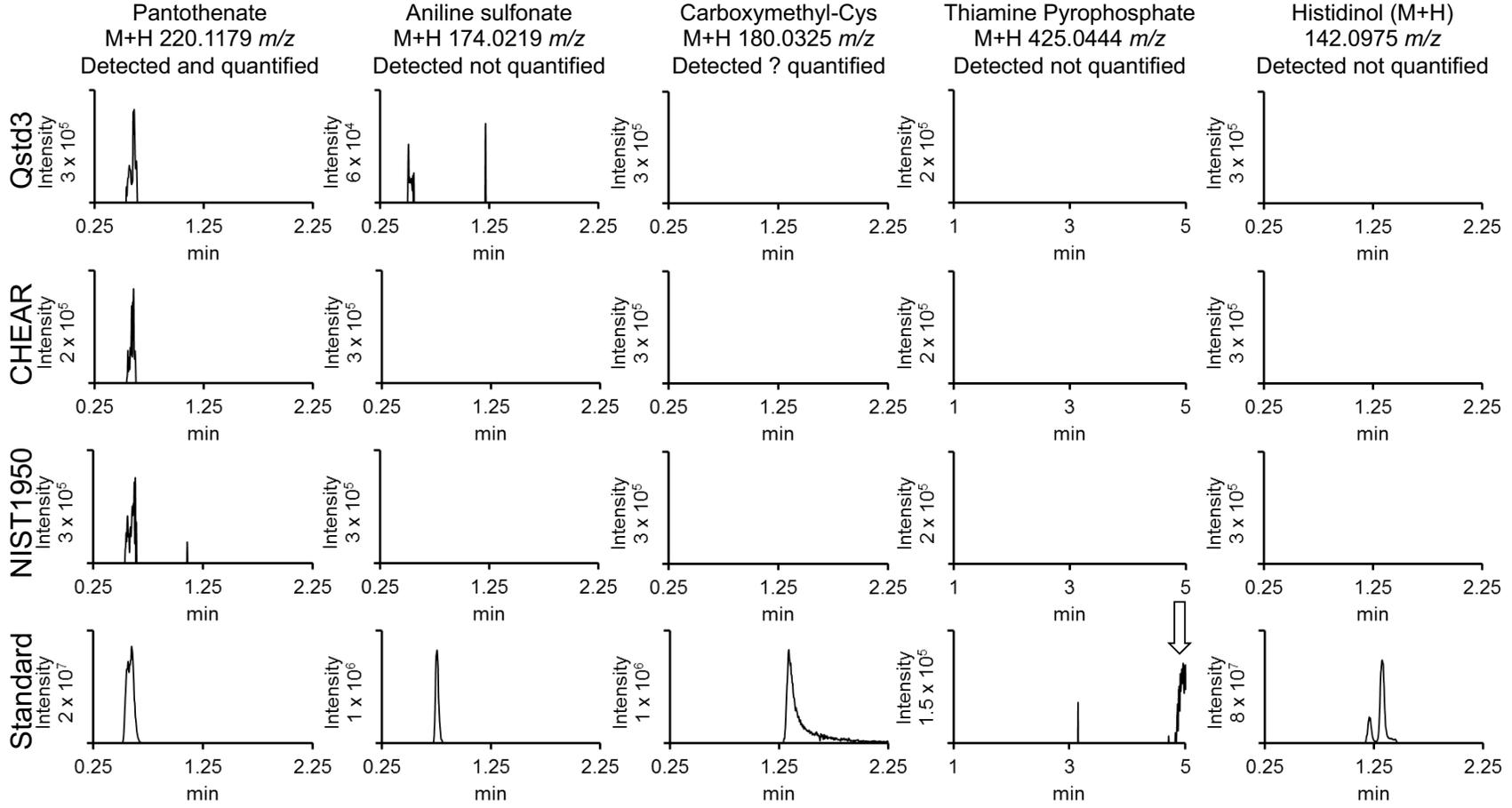


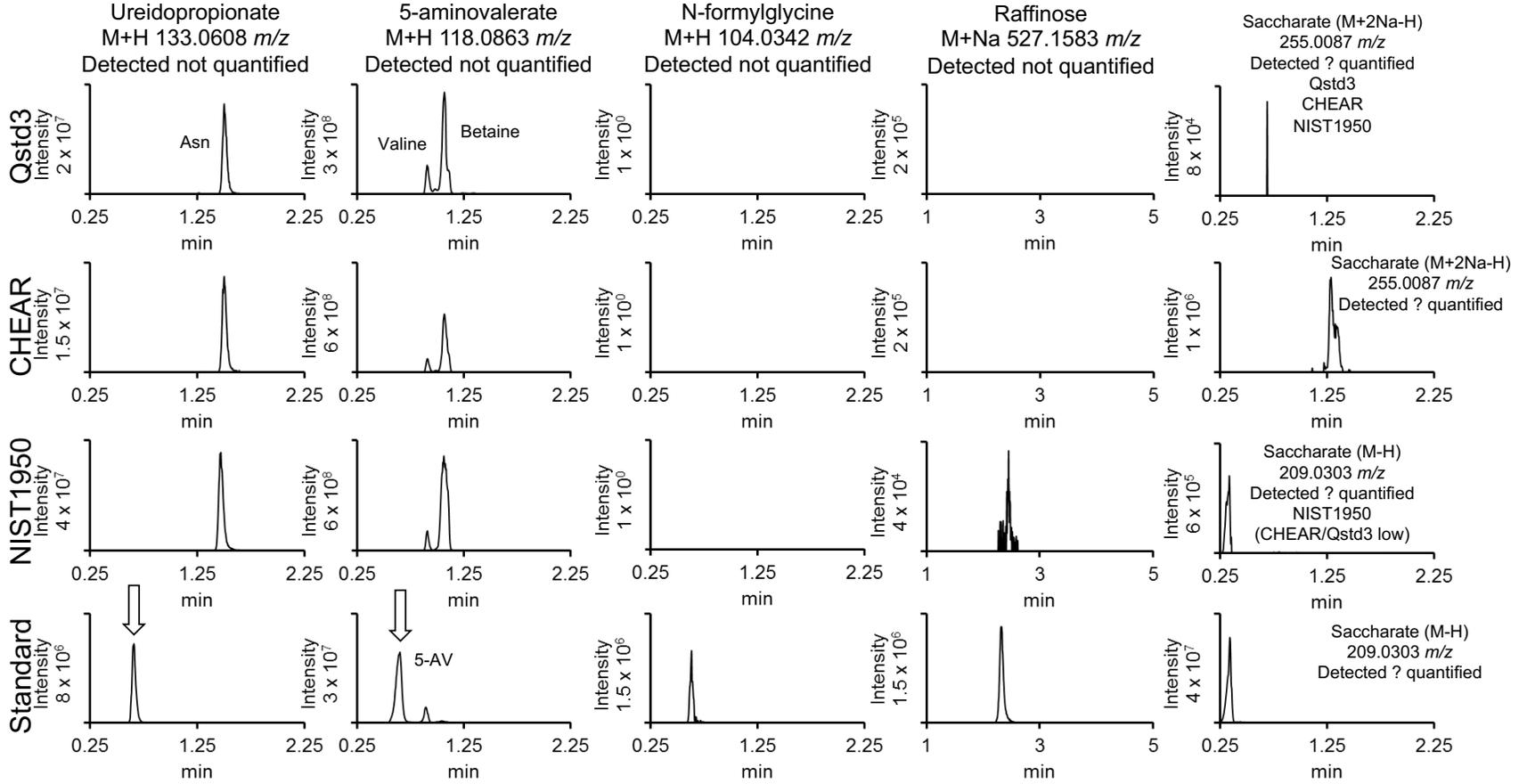


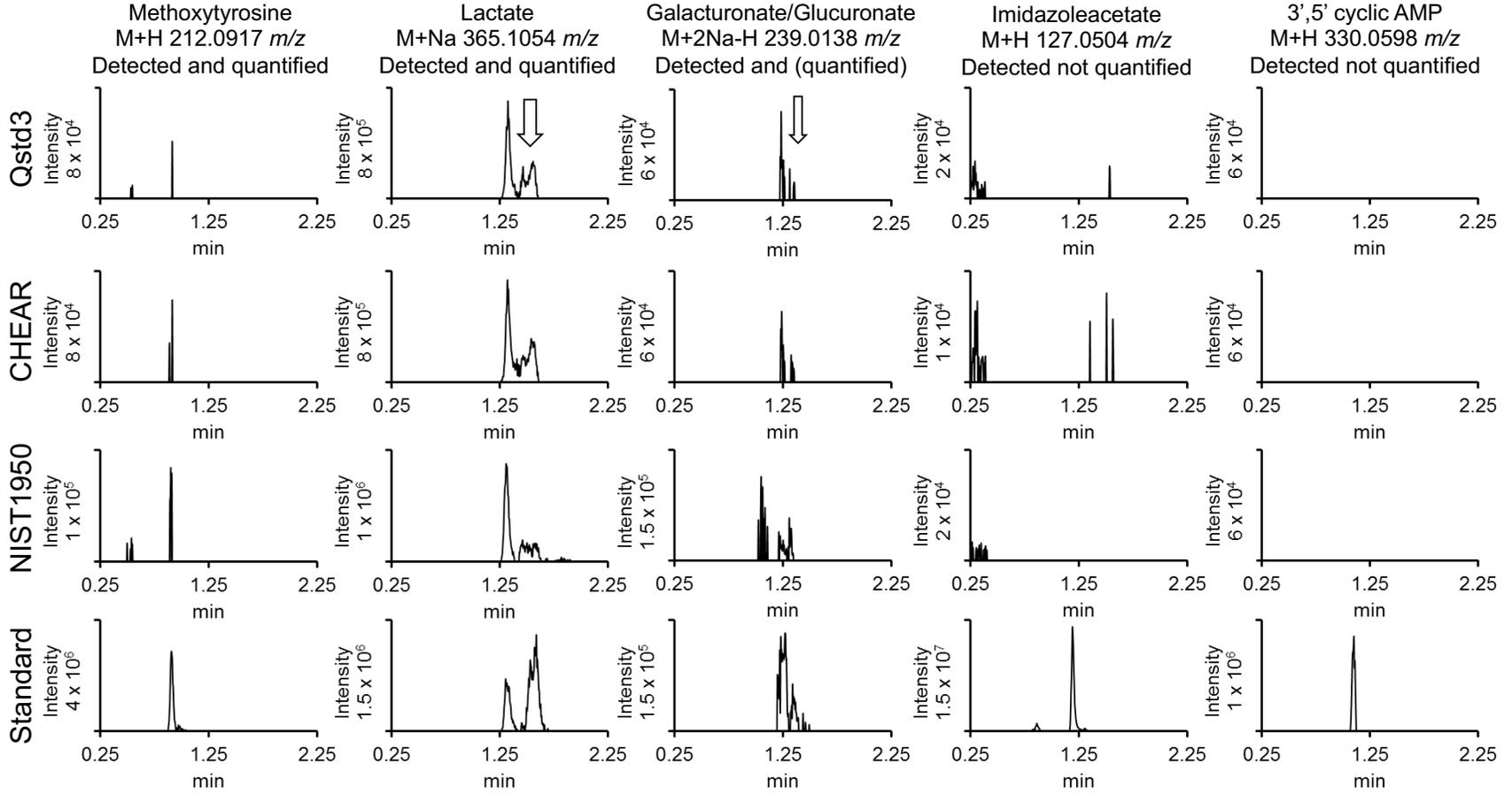


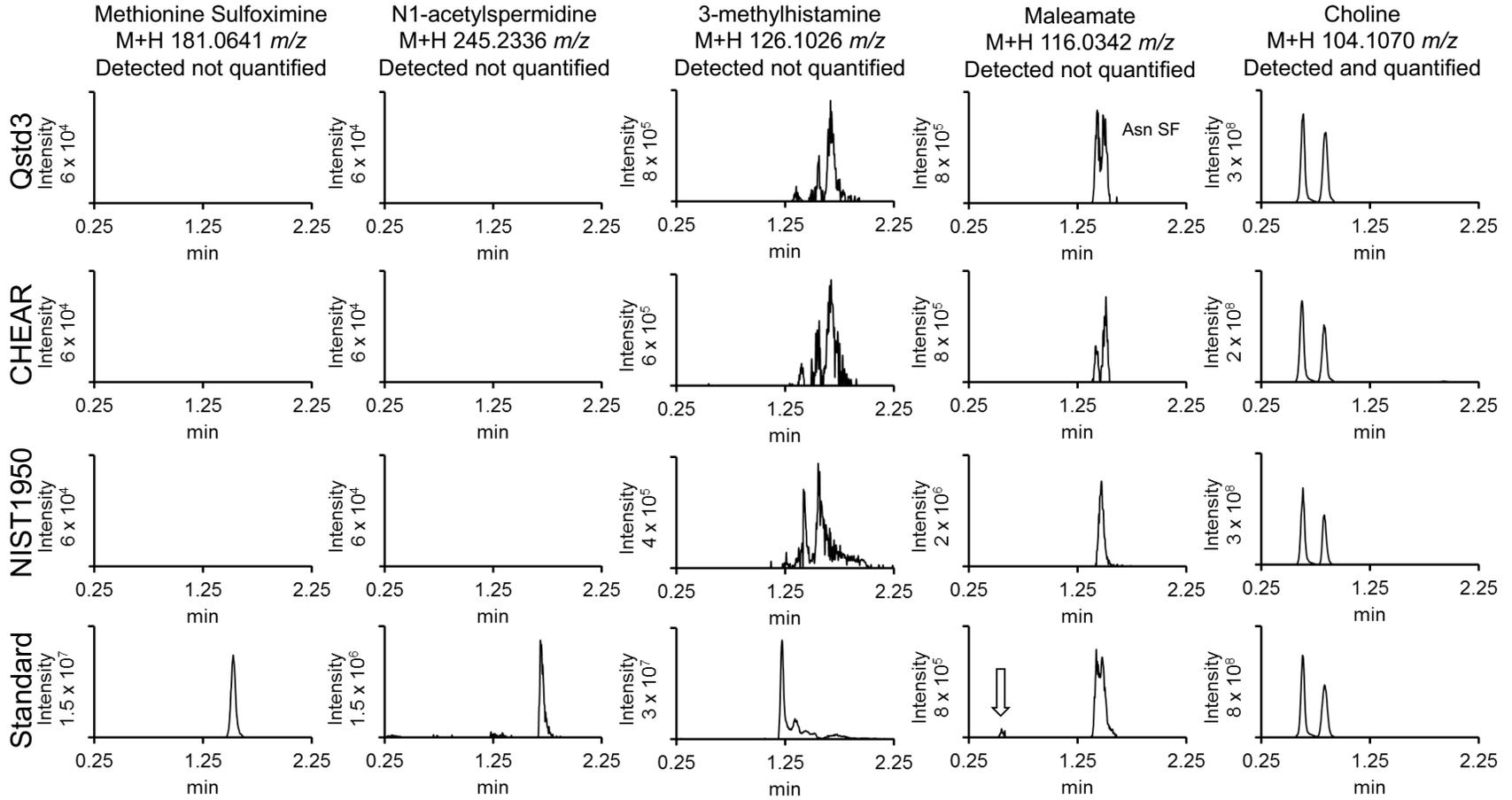


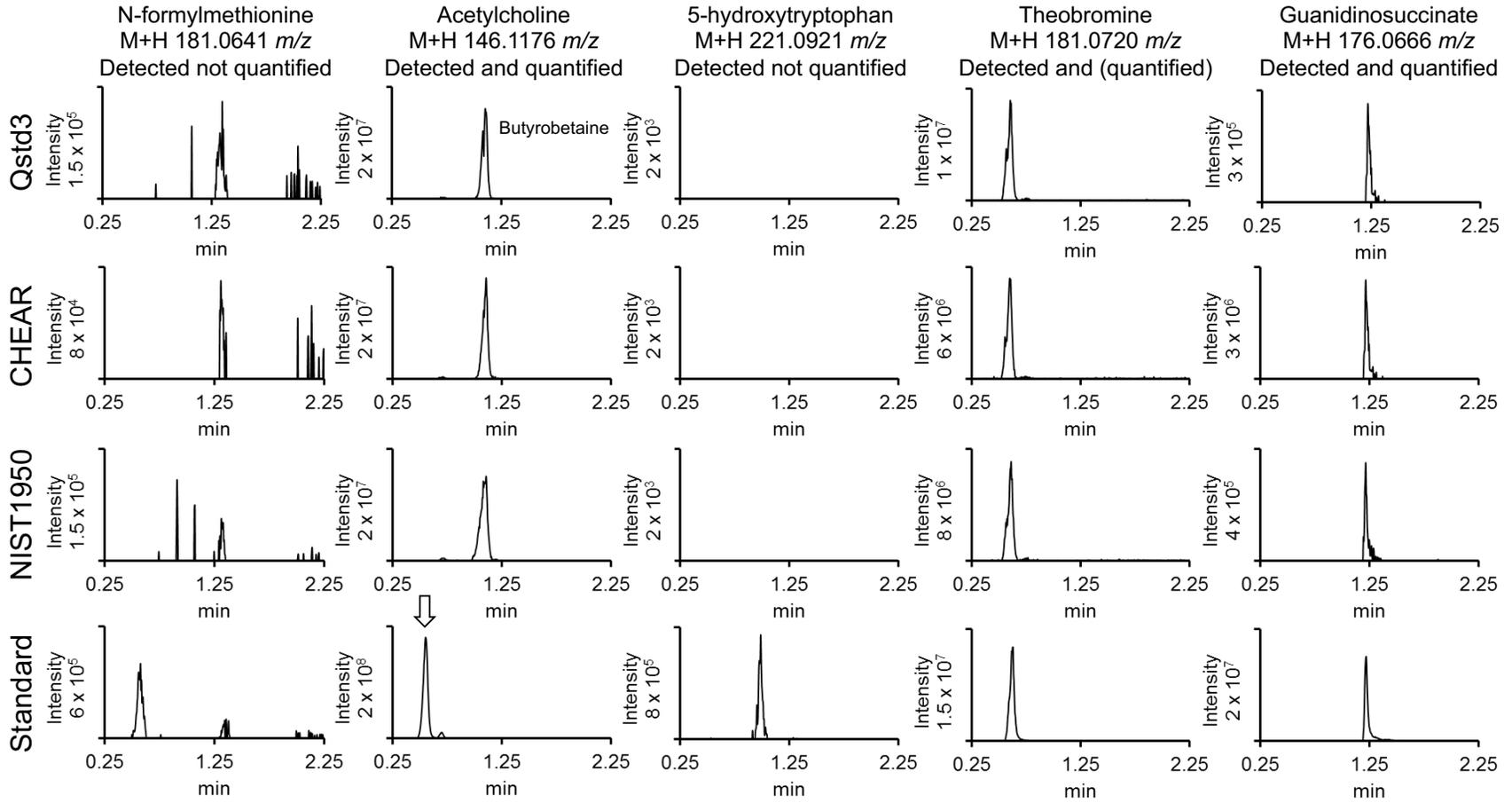


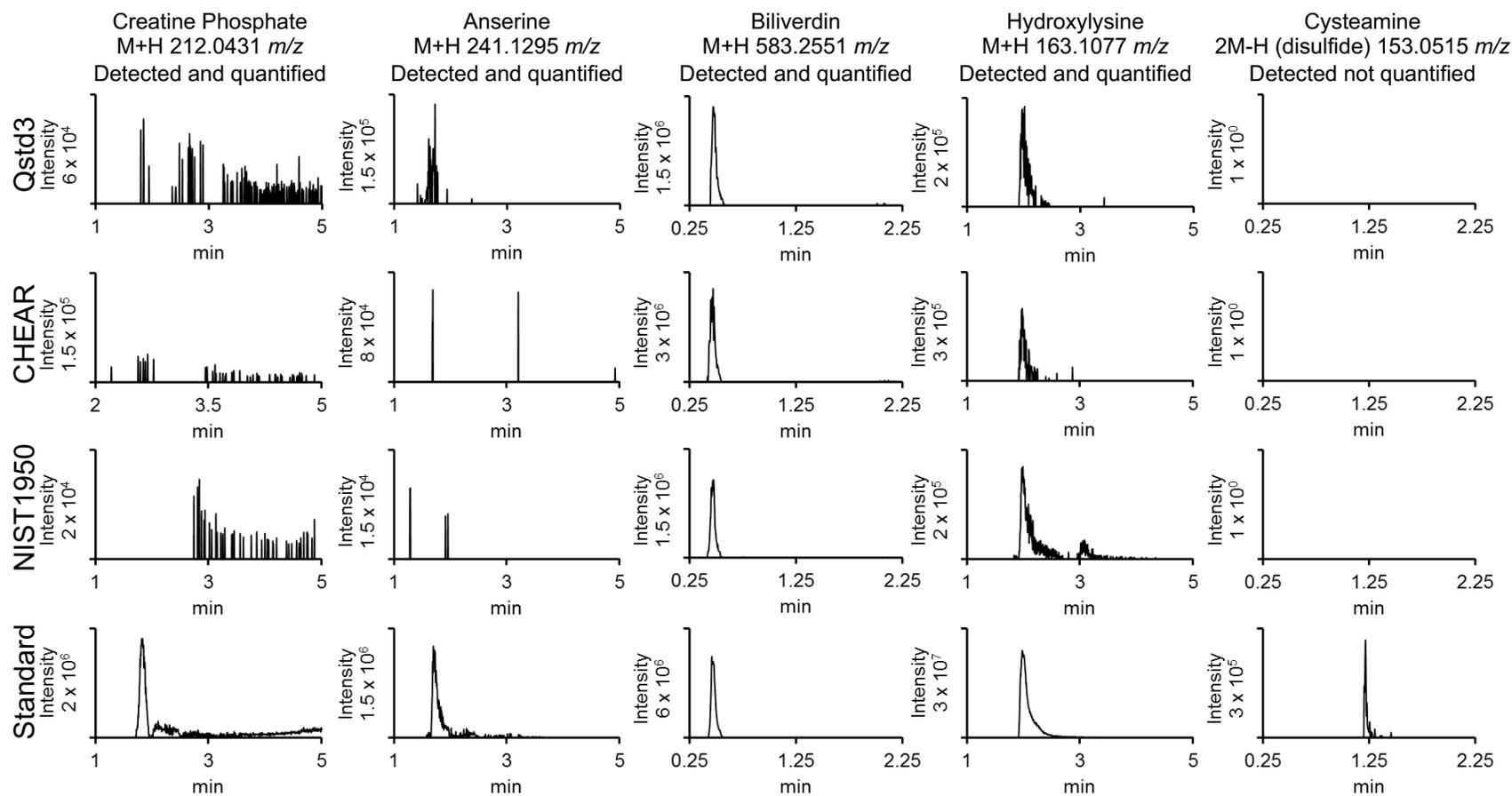


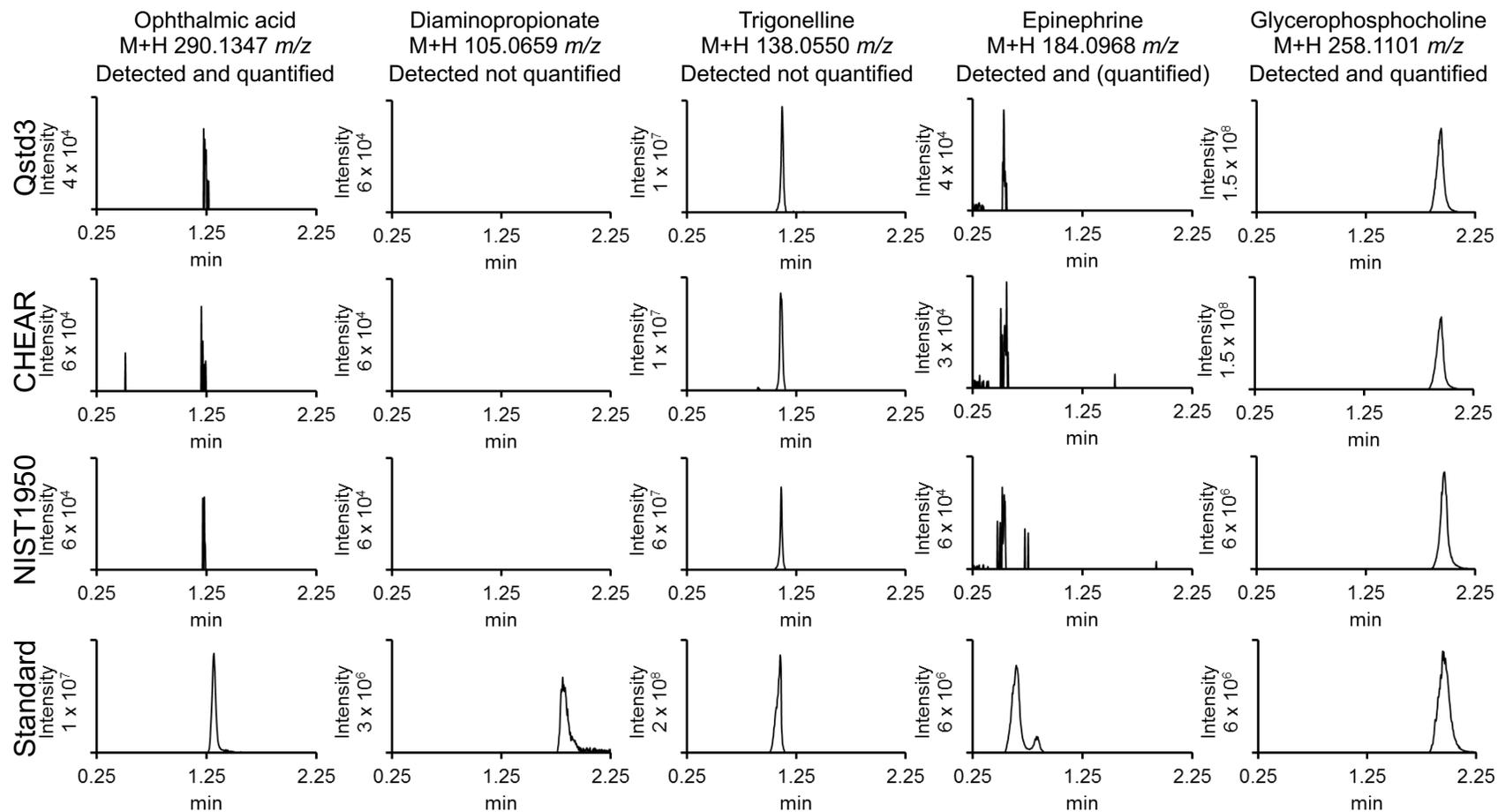


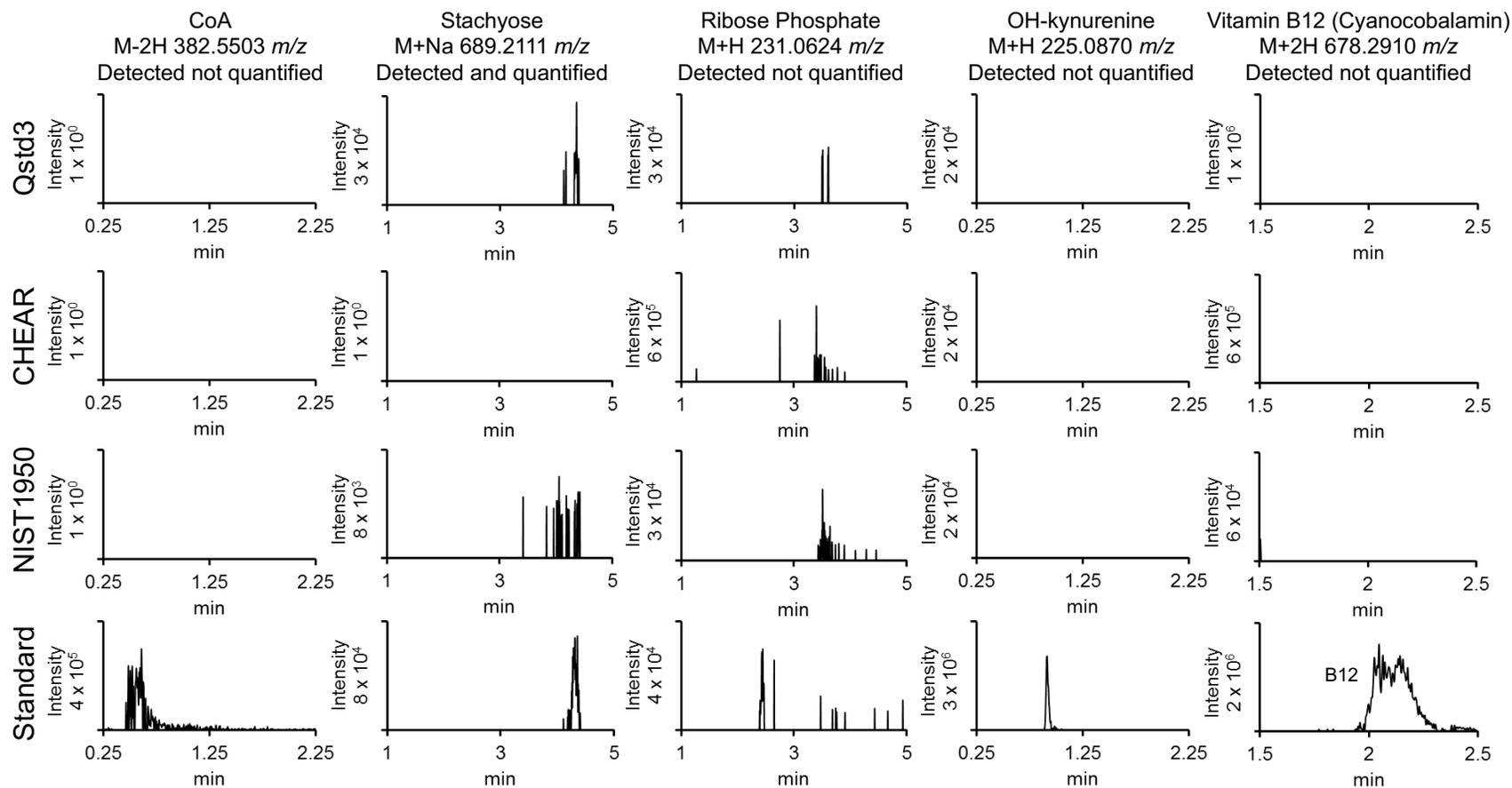


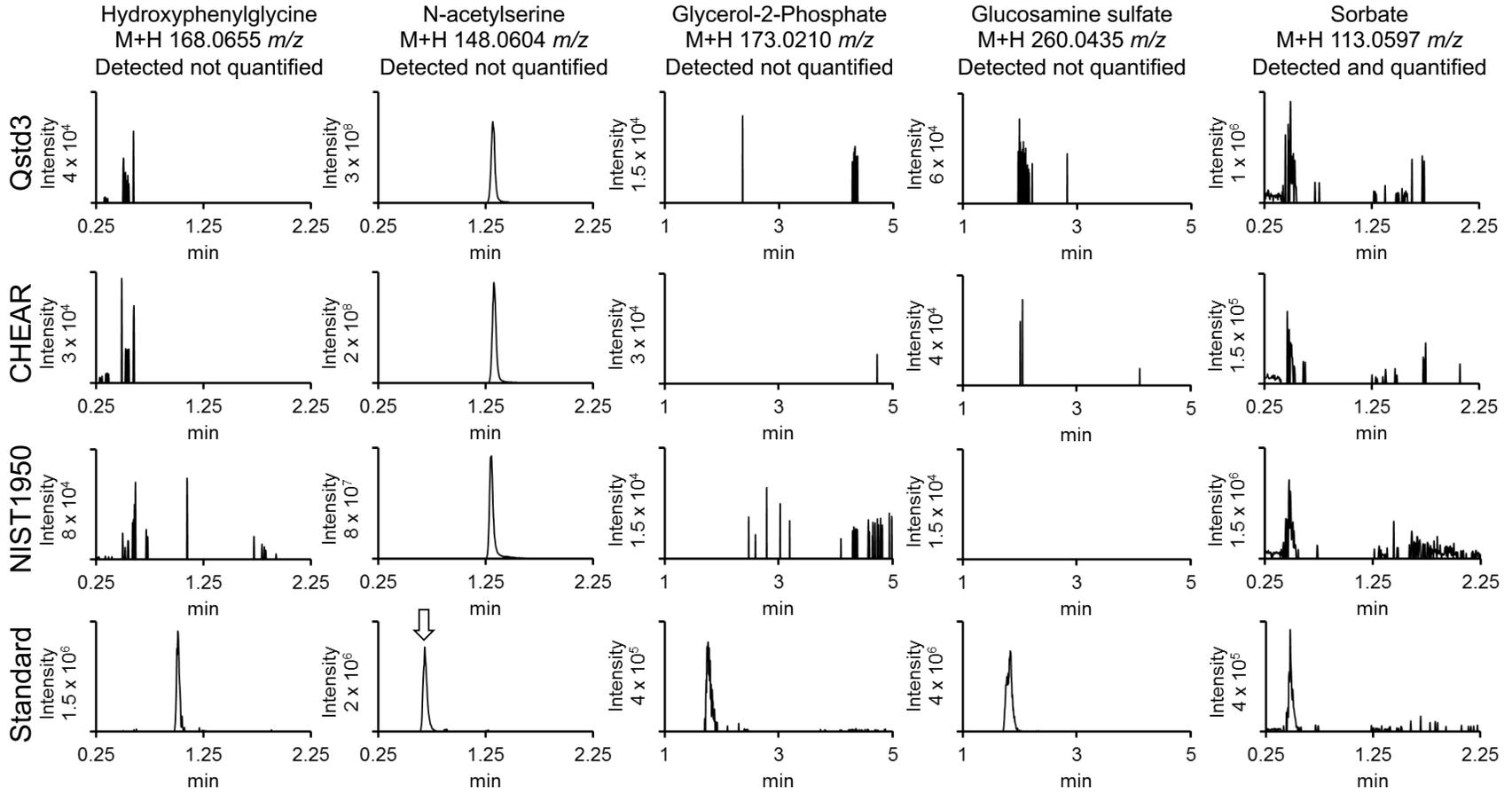


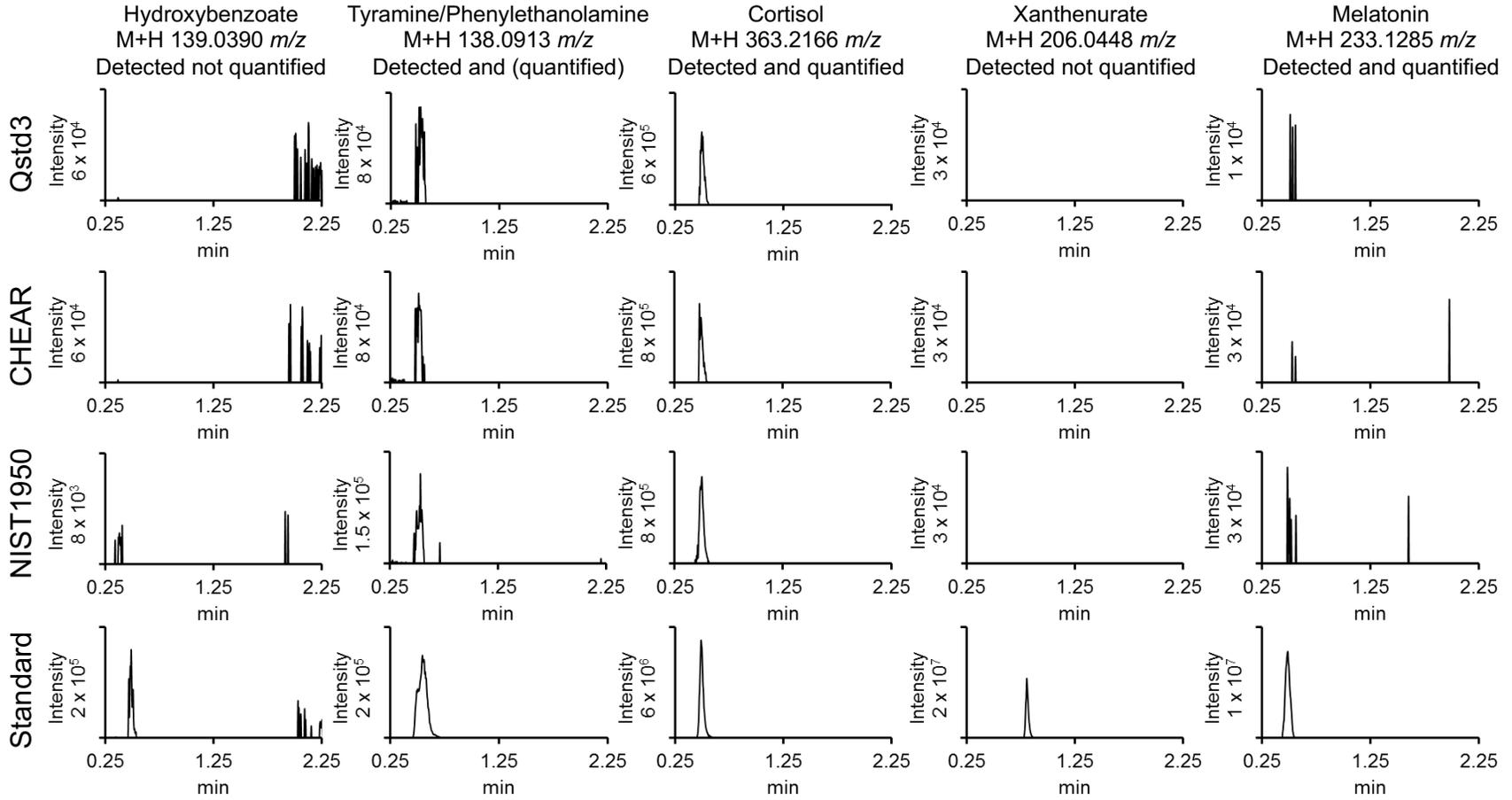


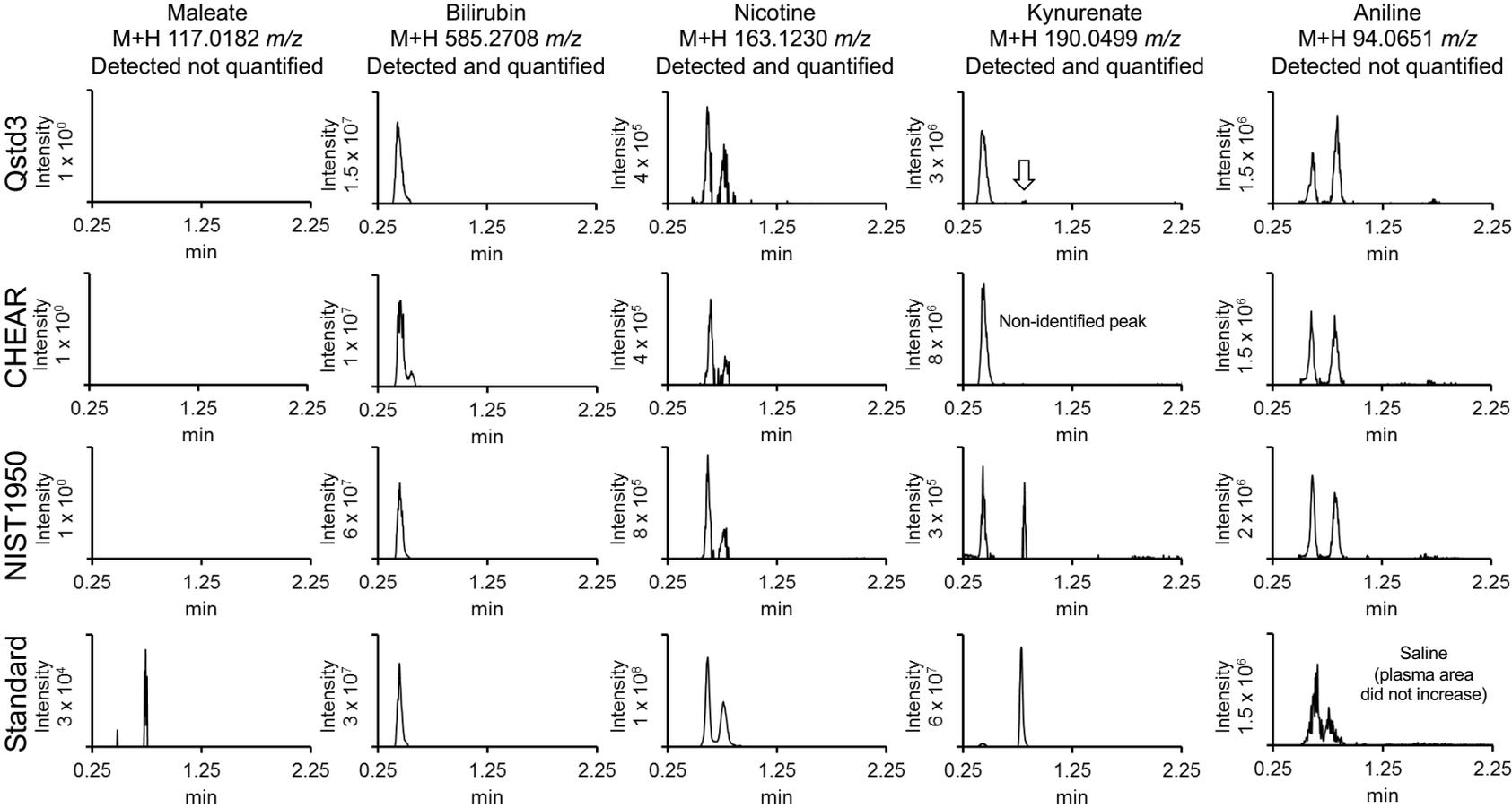


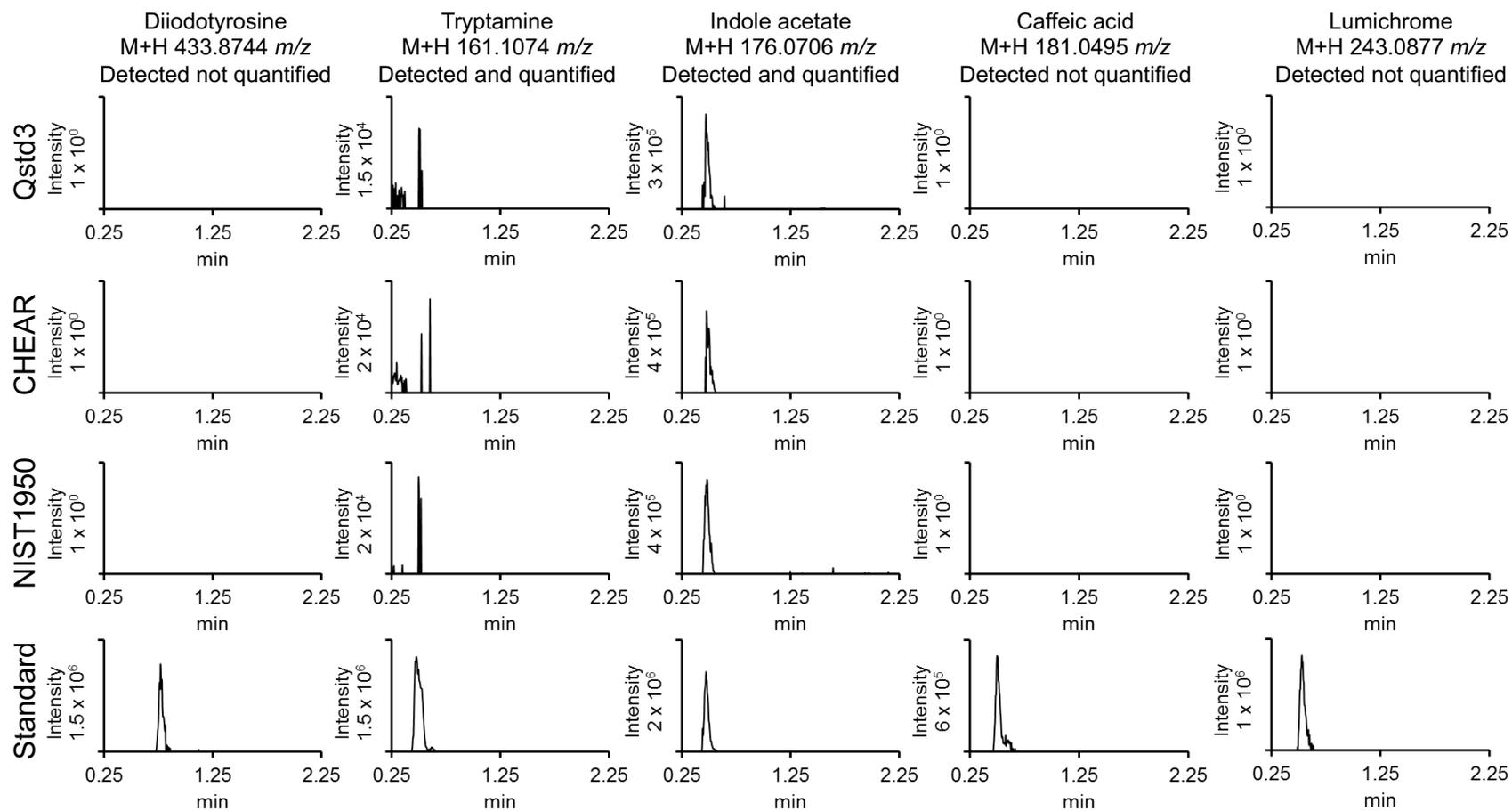


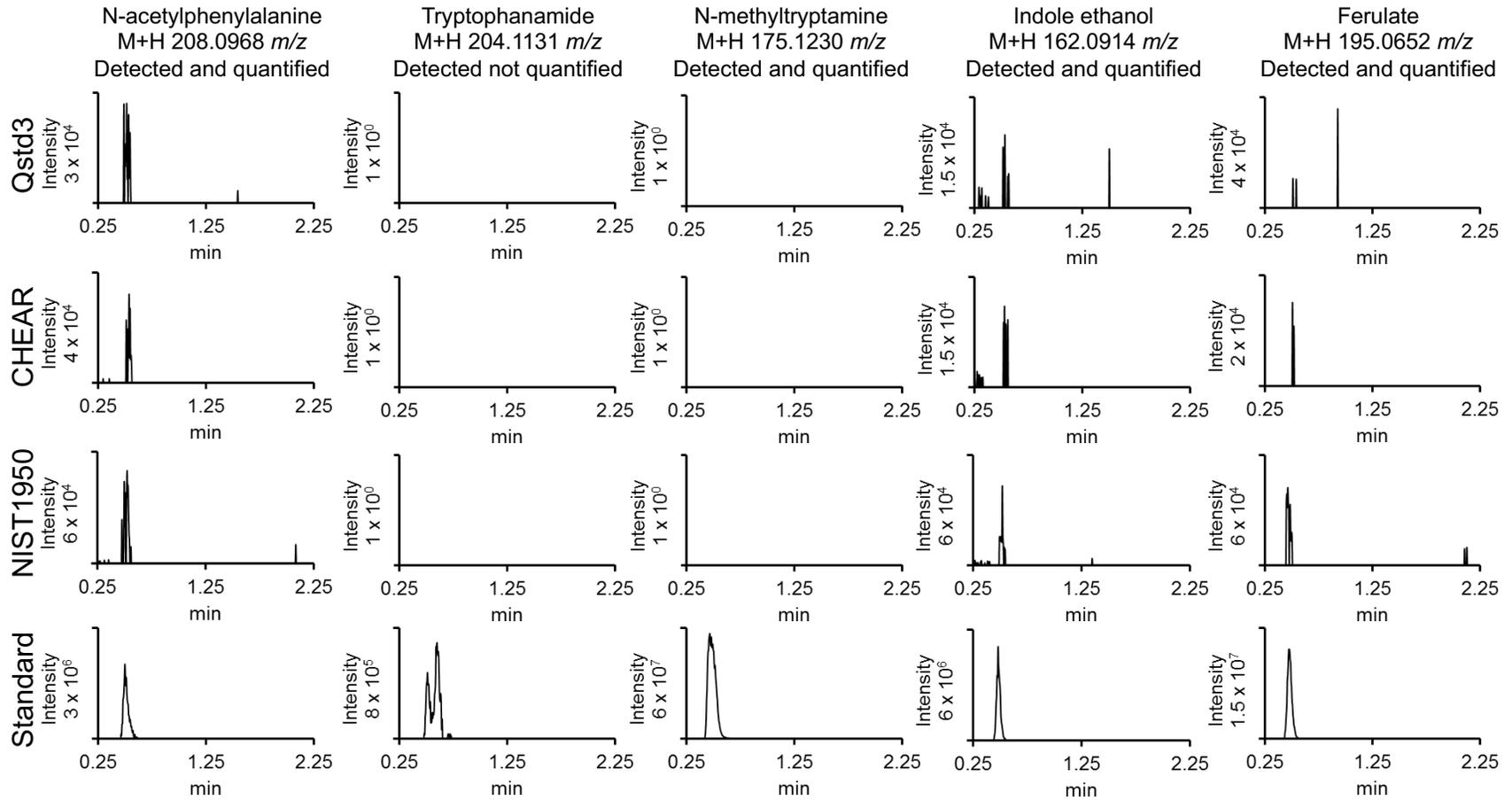


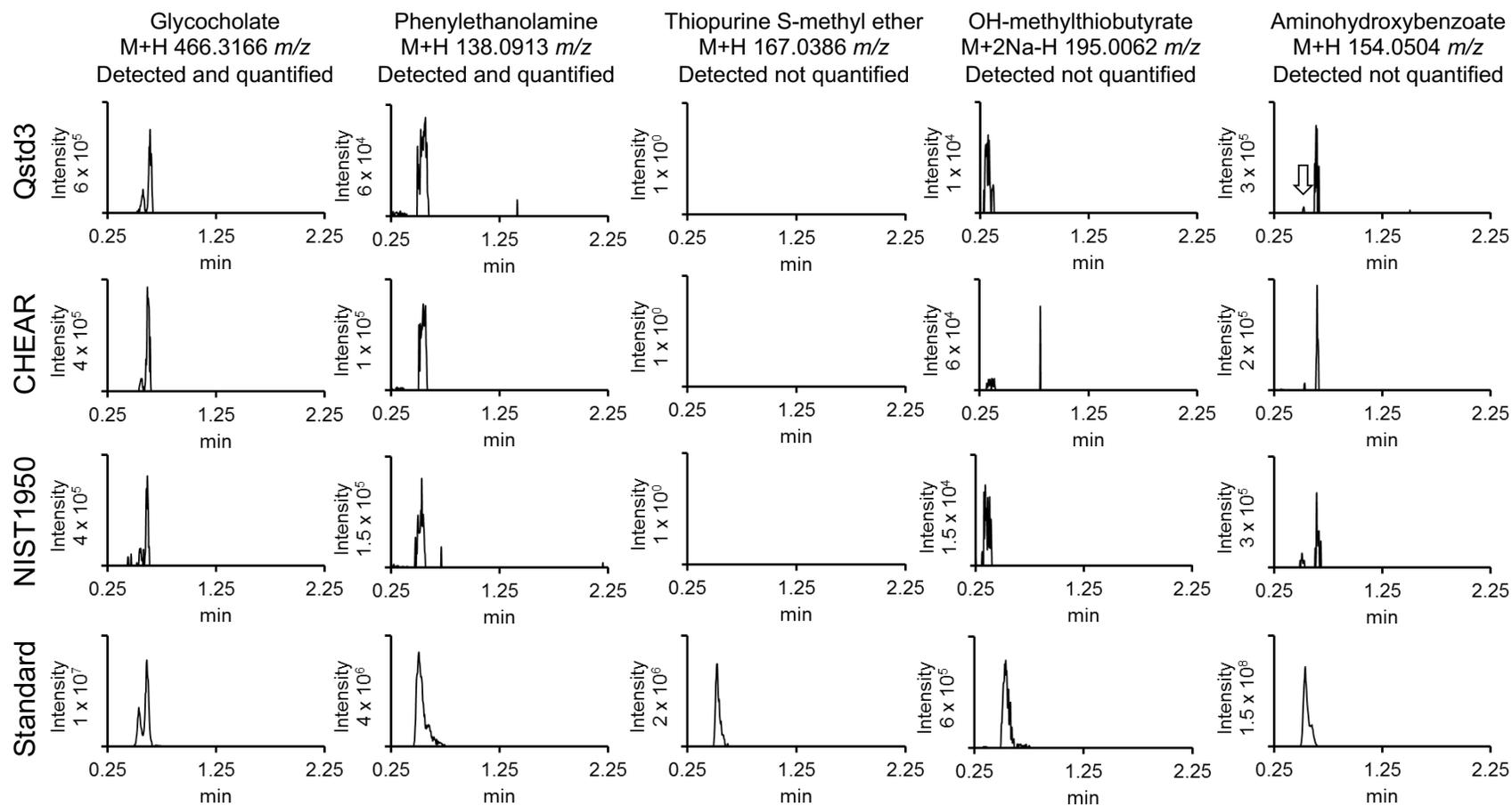


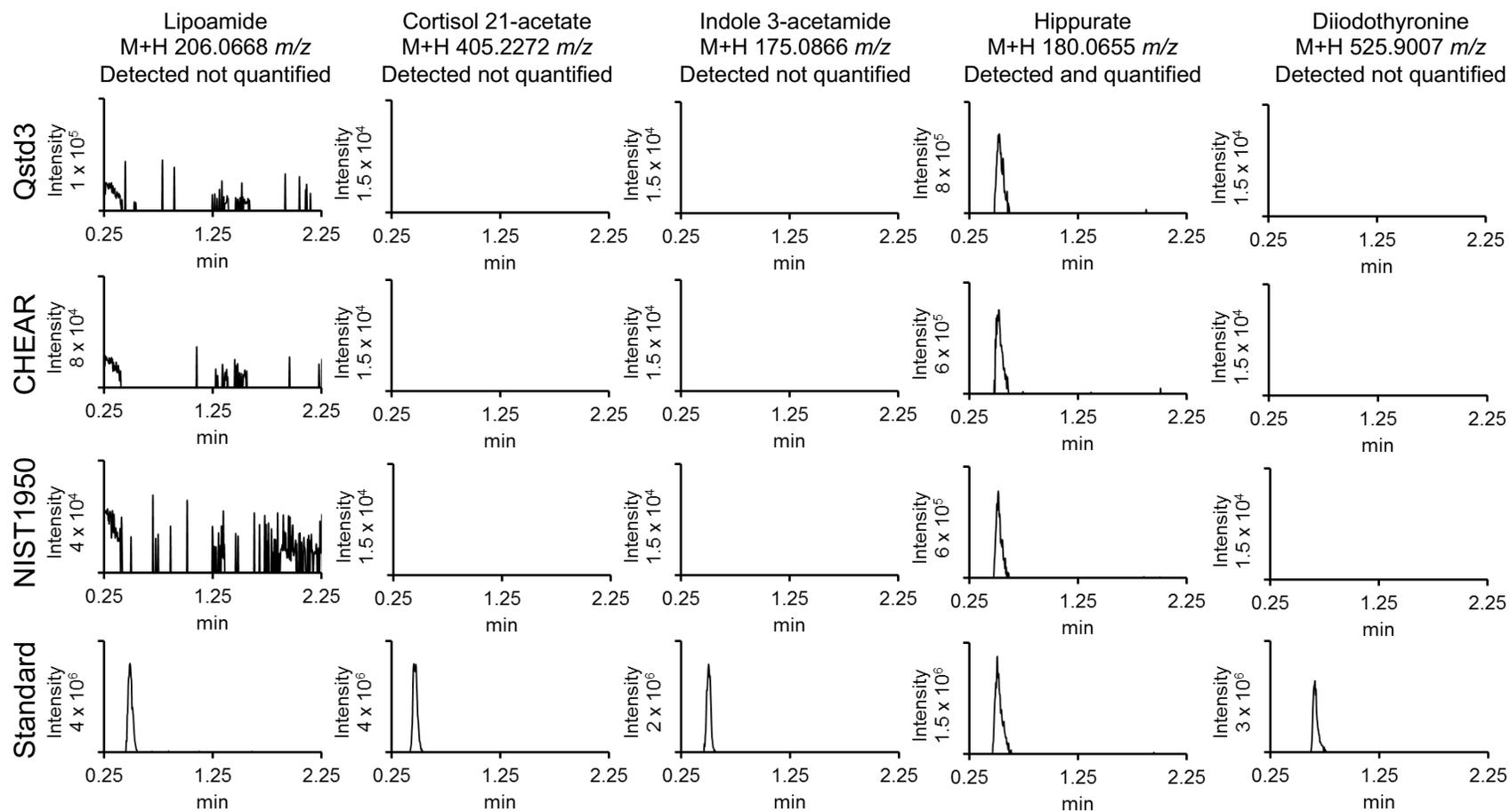


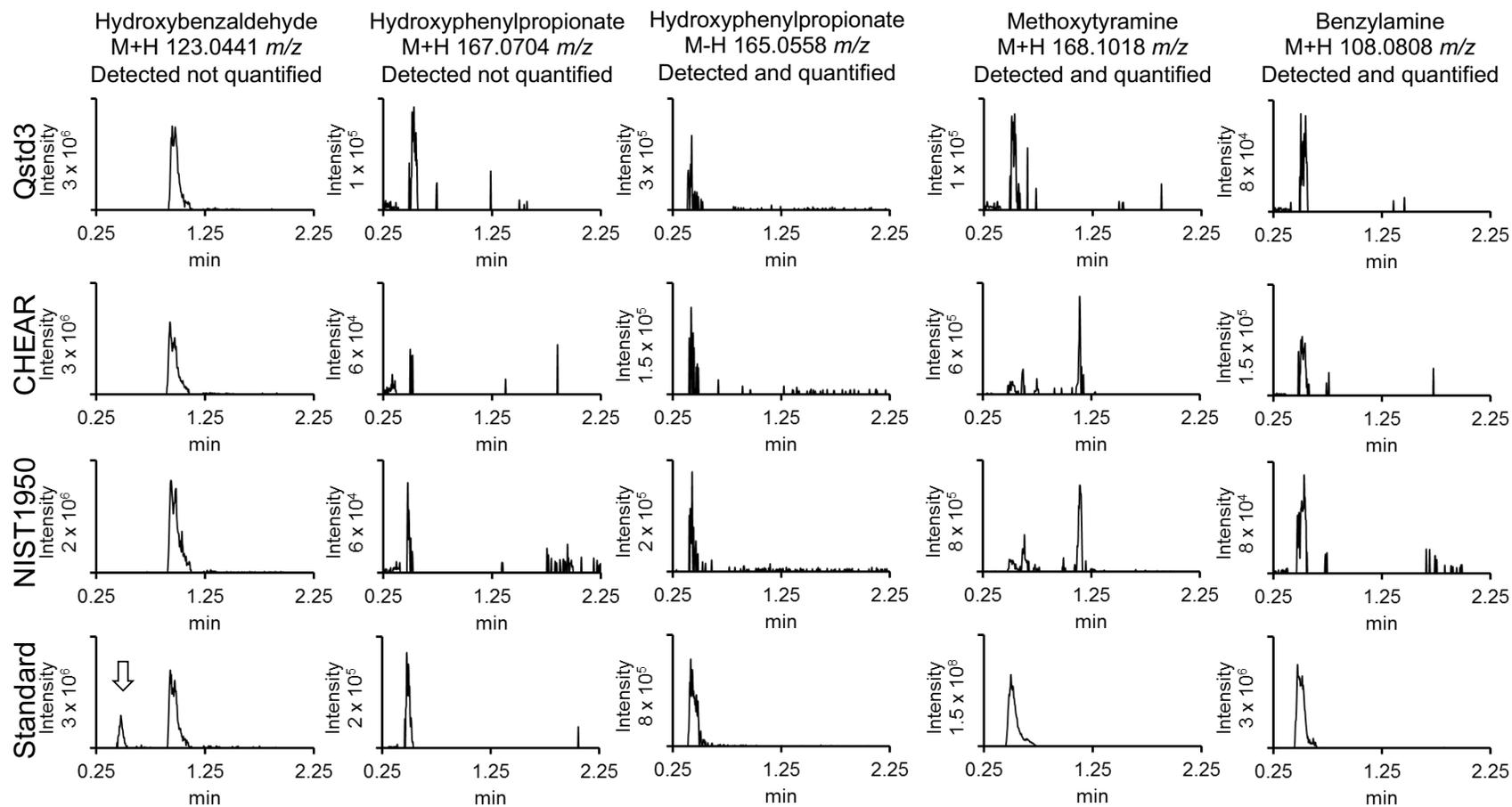


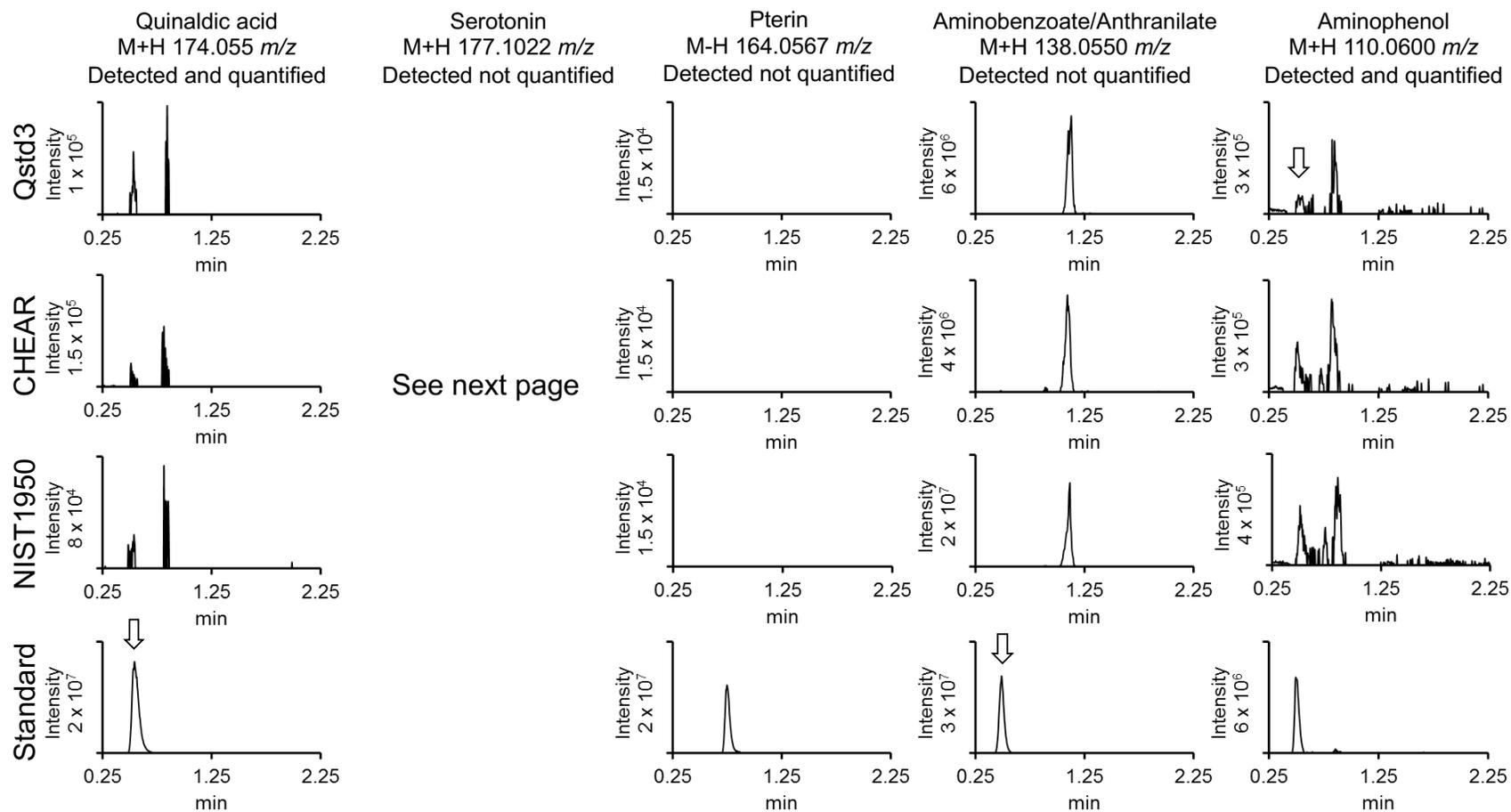




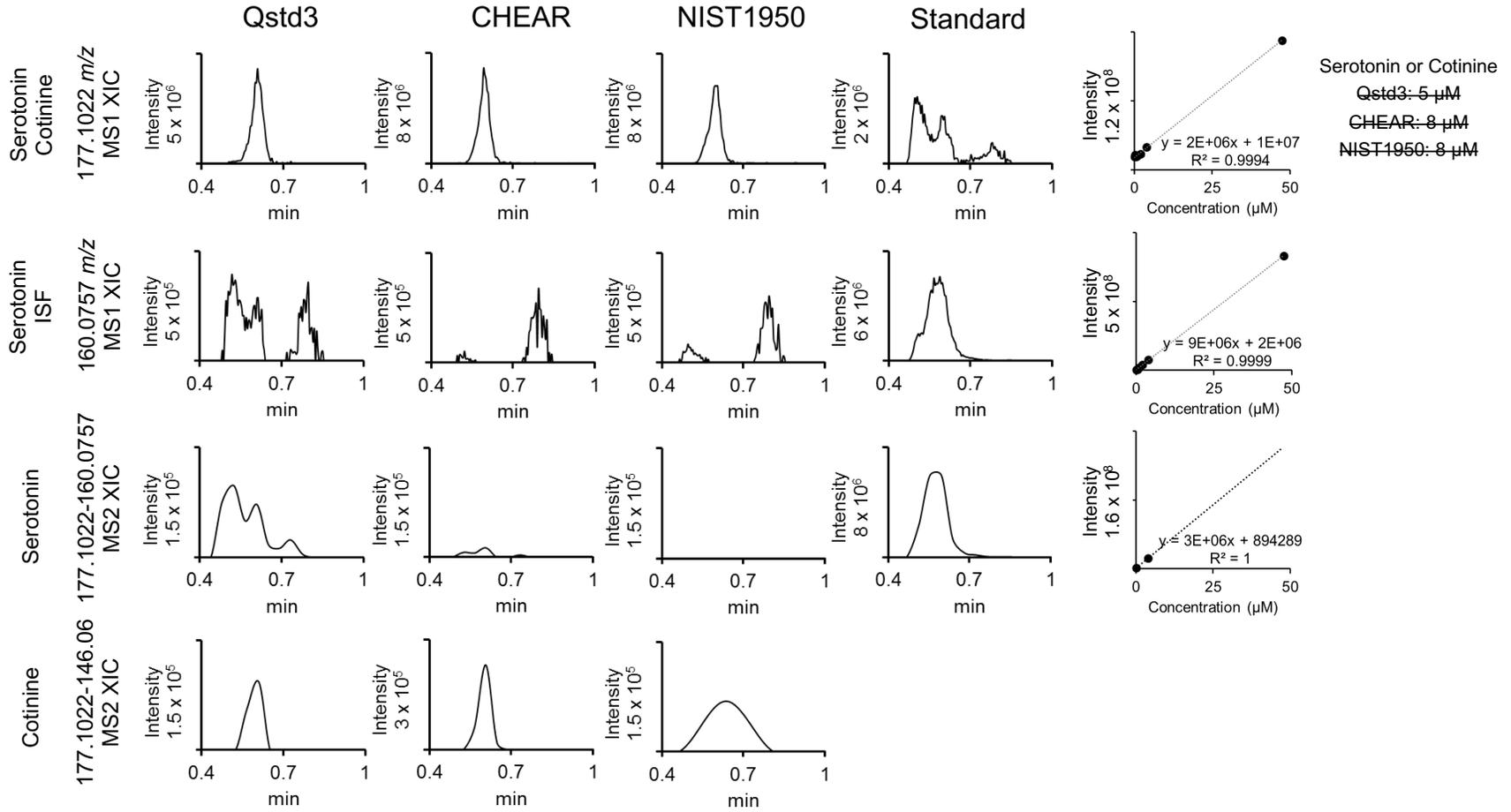




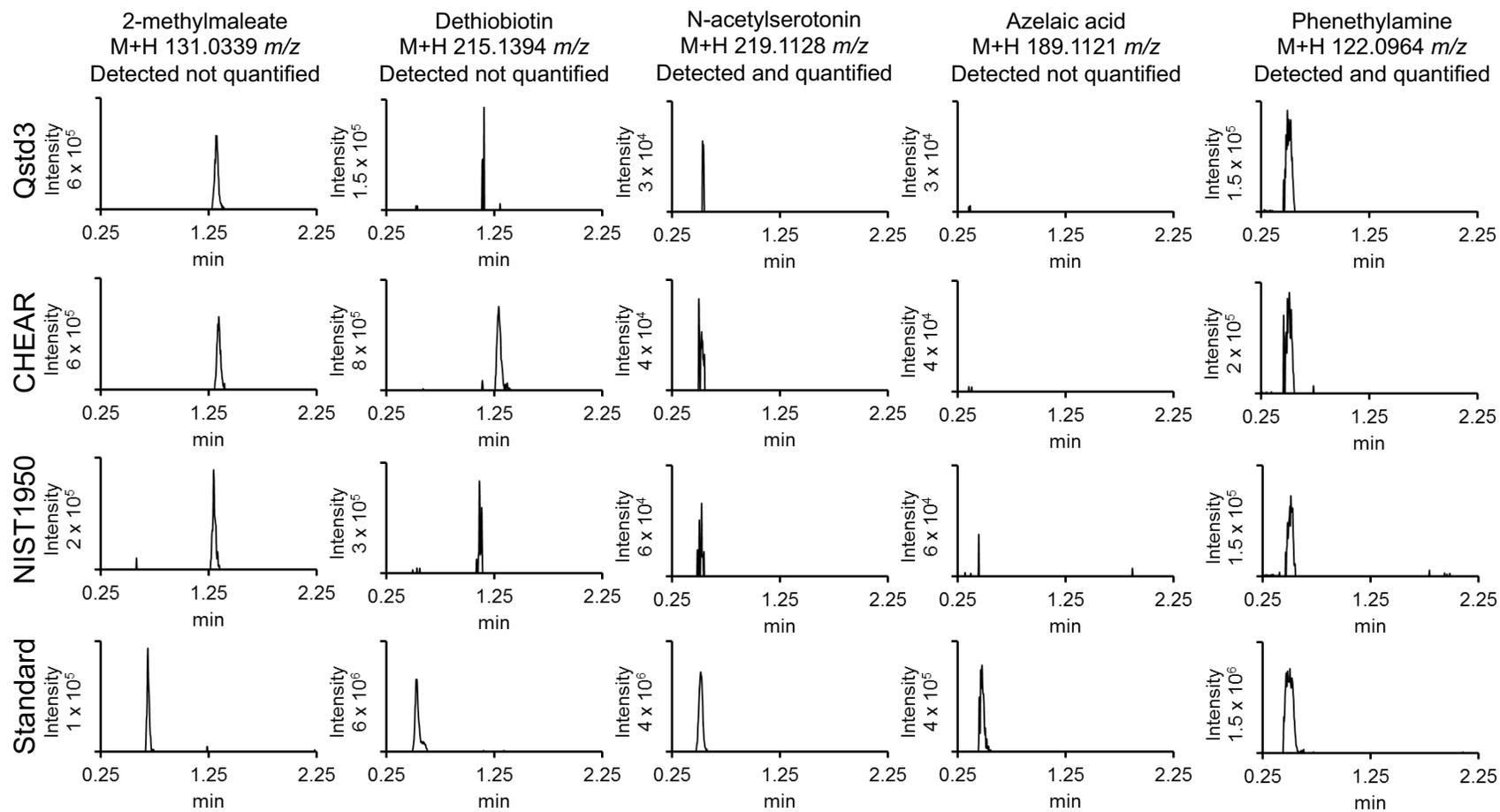


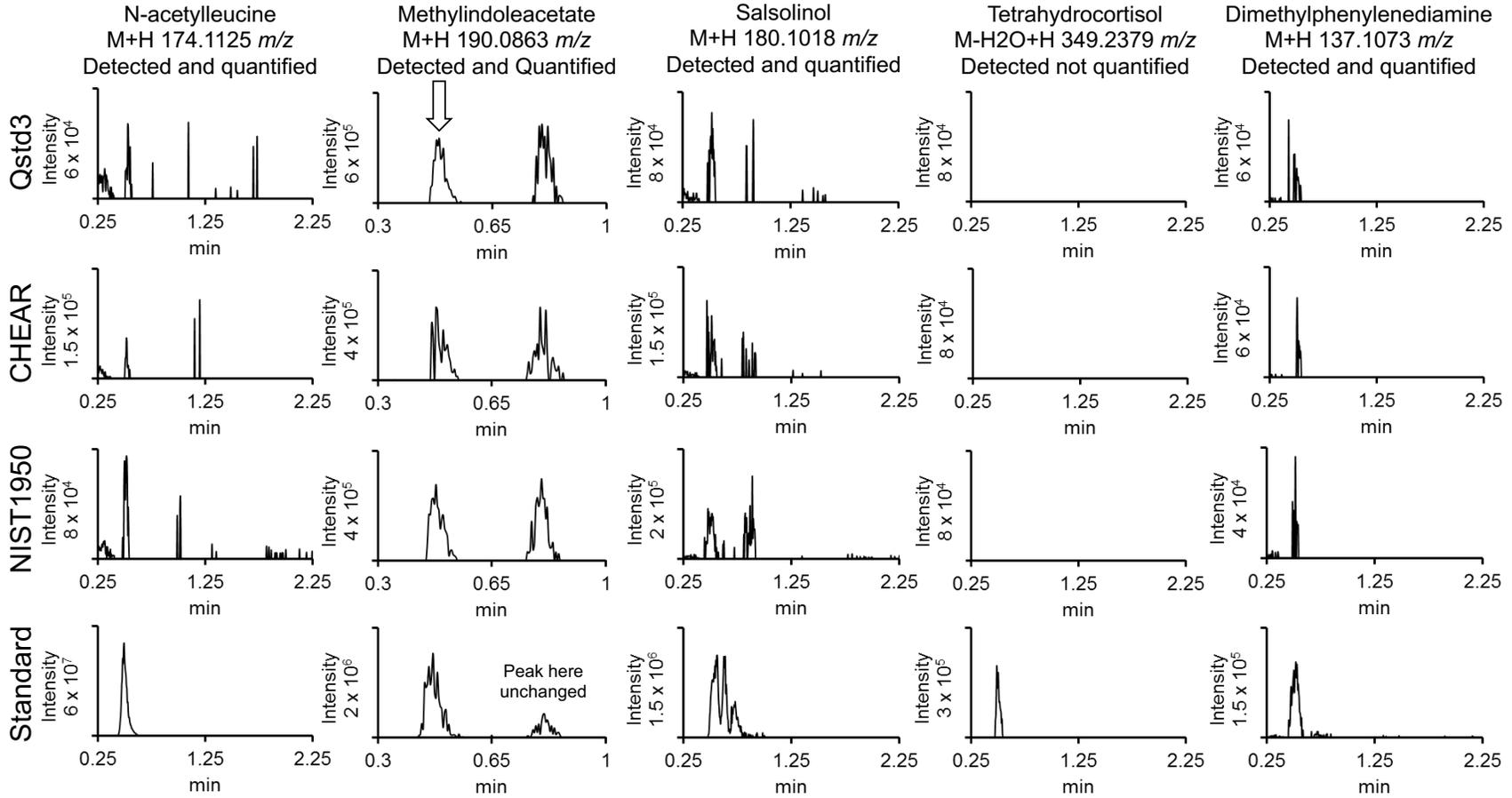


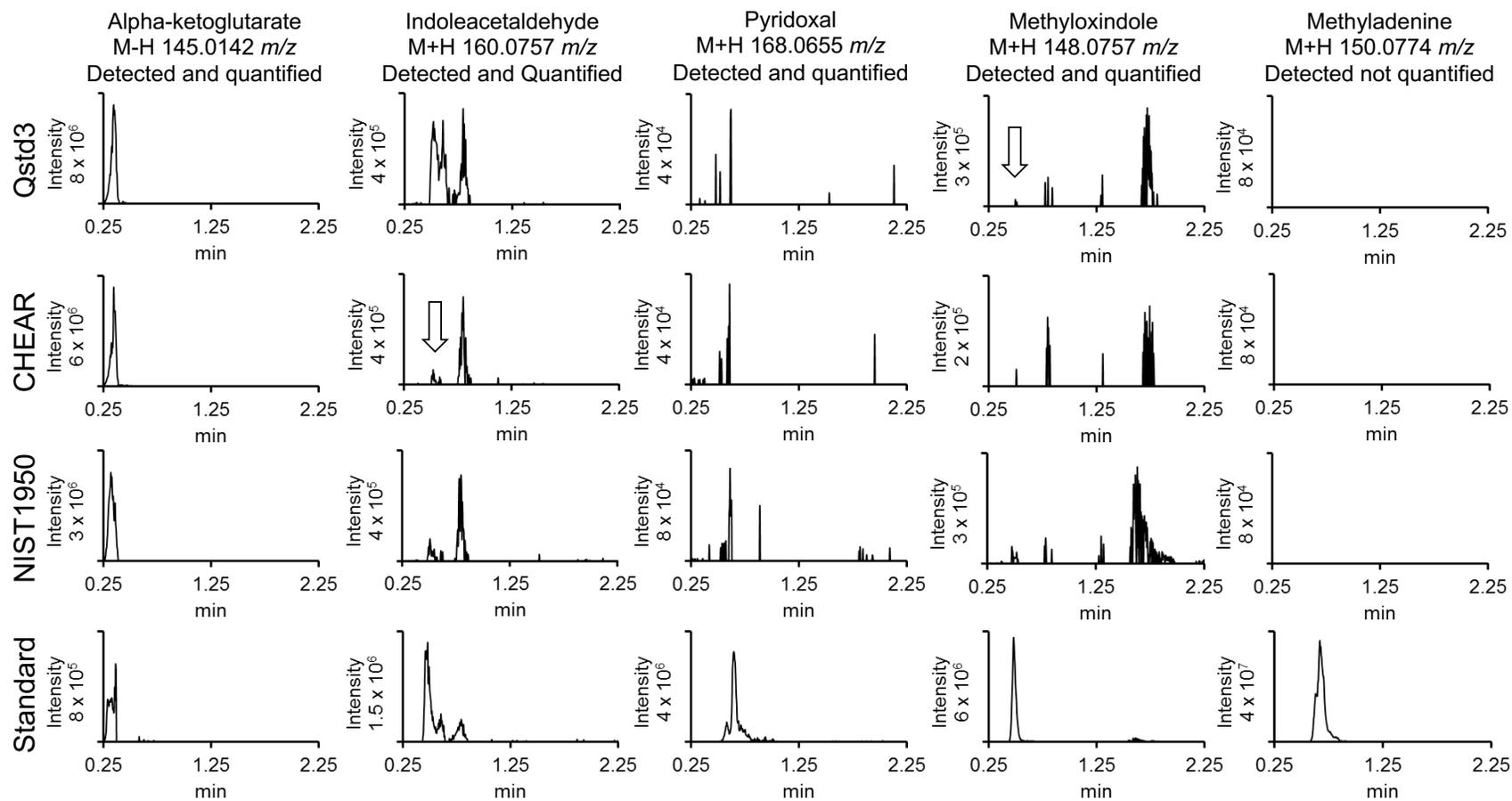
See next page

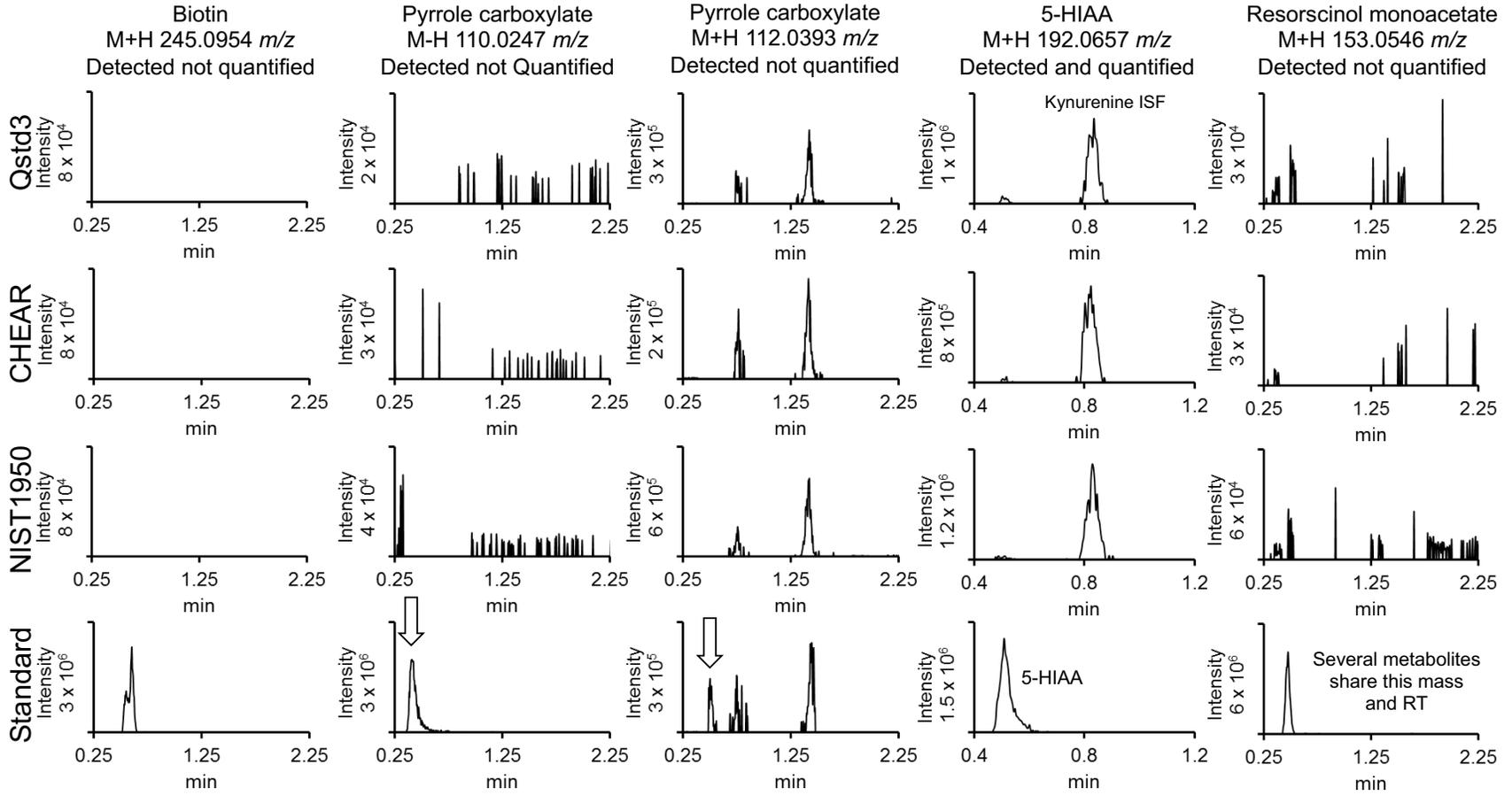


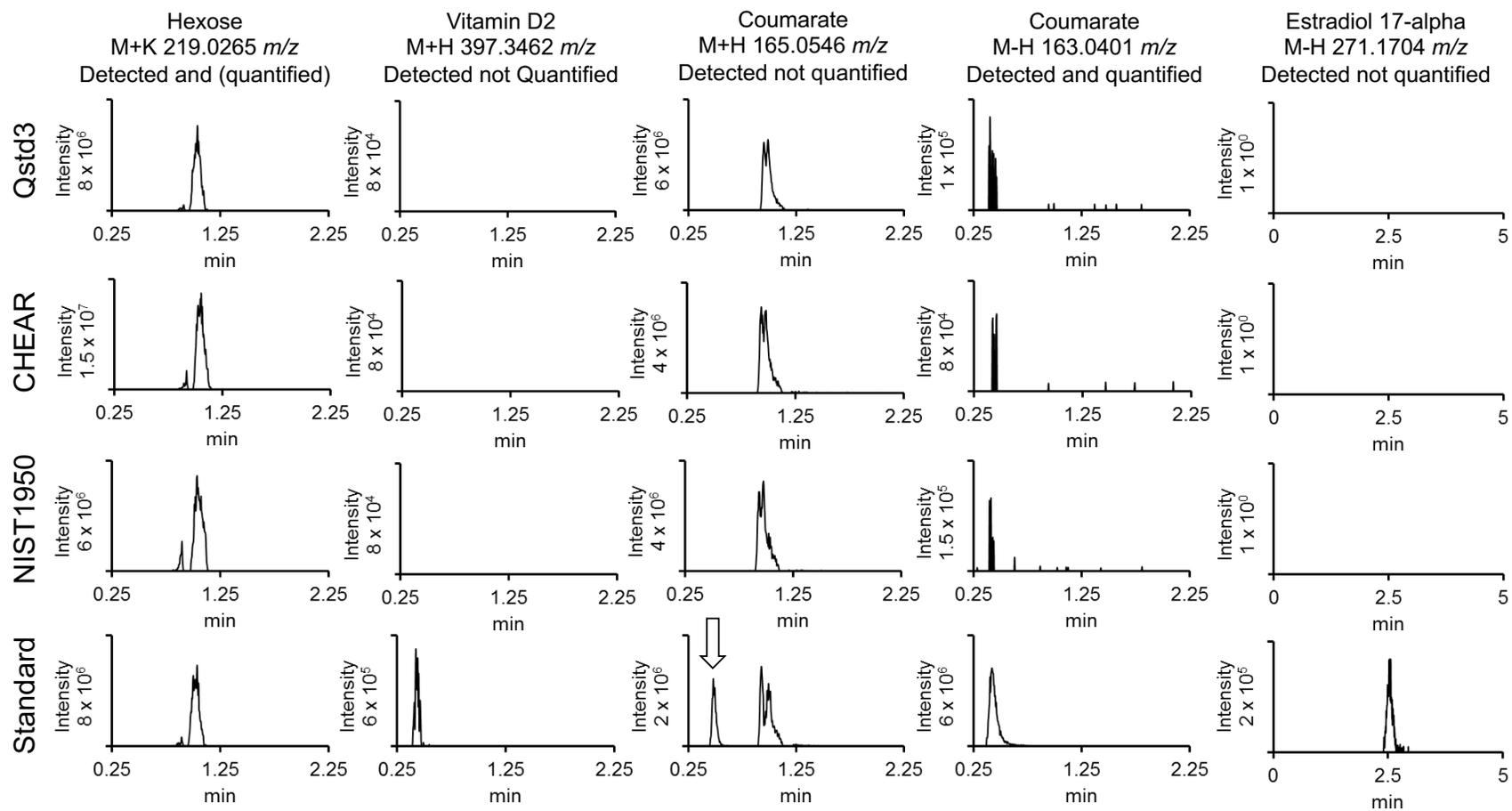
Serotonin or Cotinine
Qstd3: 5 µM
CHEAR: 8 µM
NIST1950: 8 µM

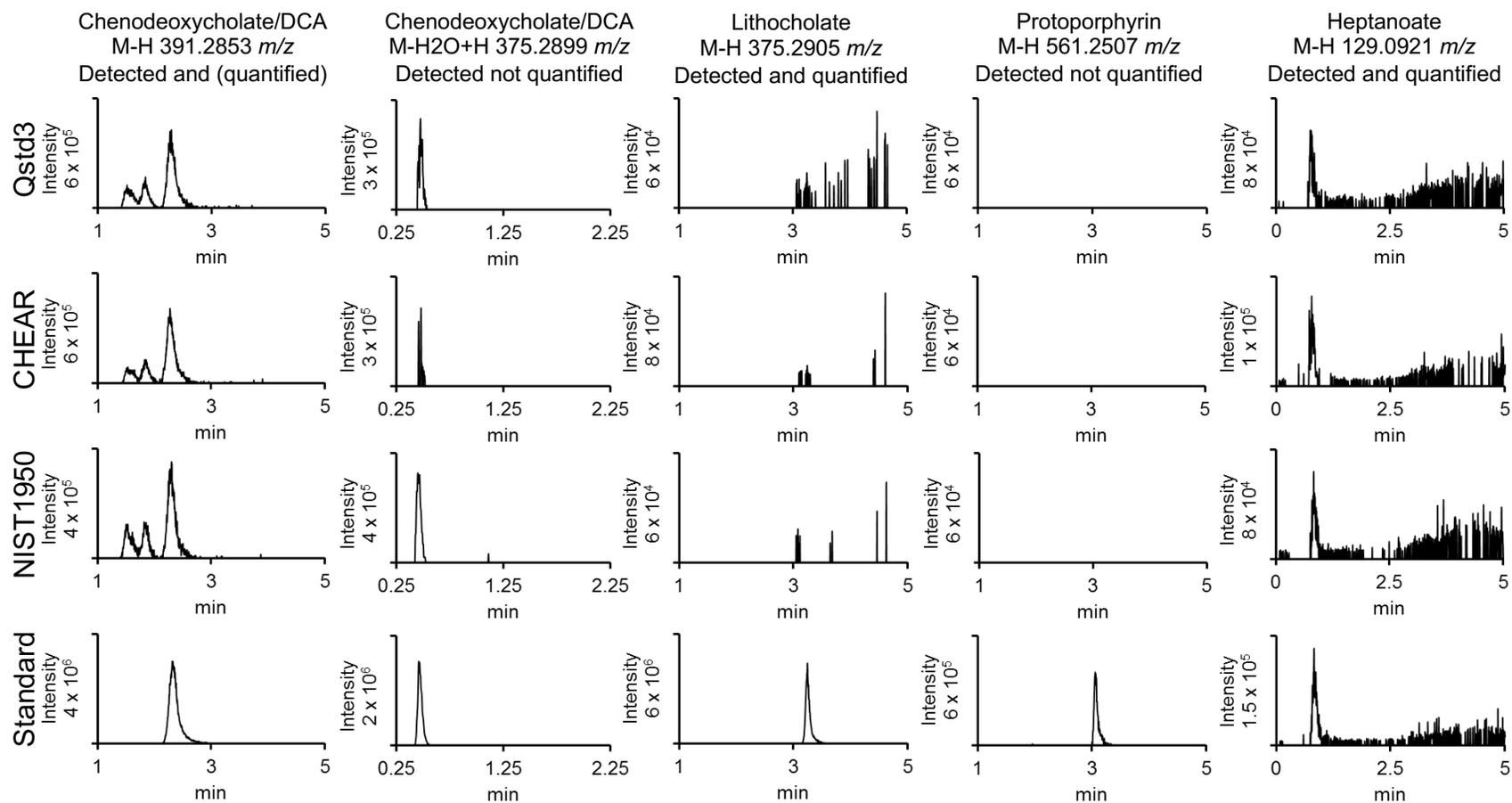


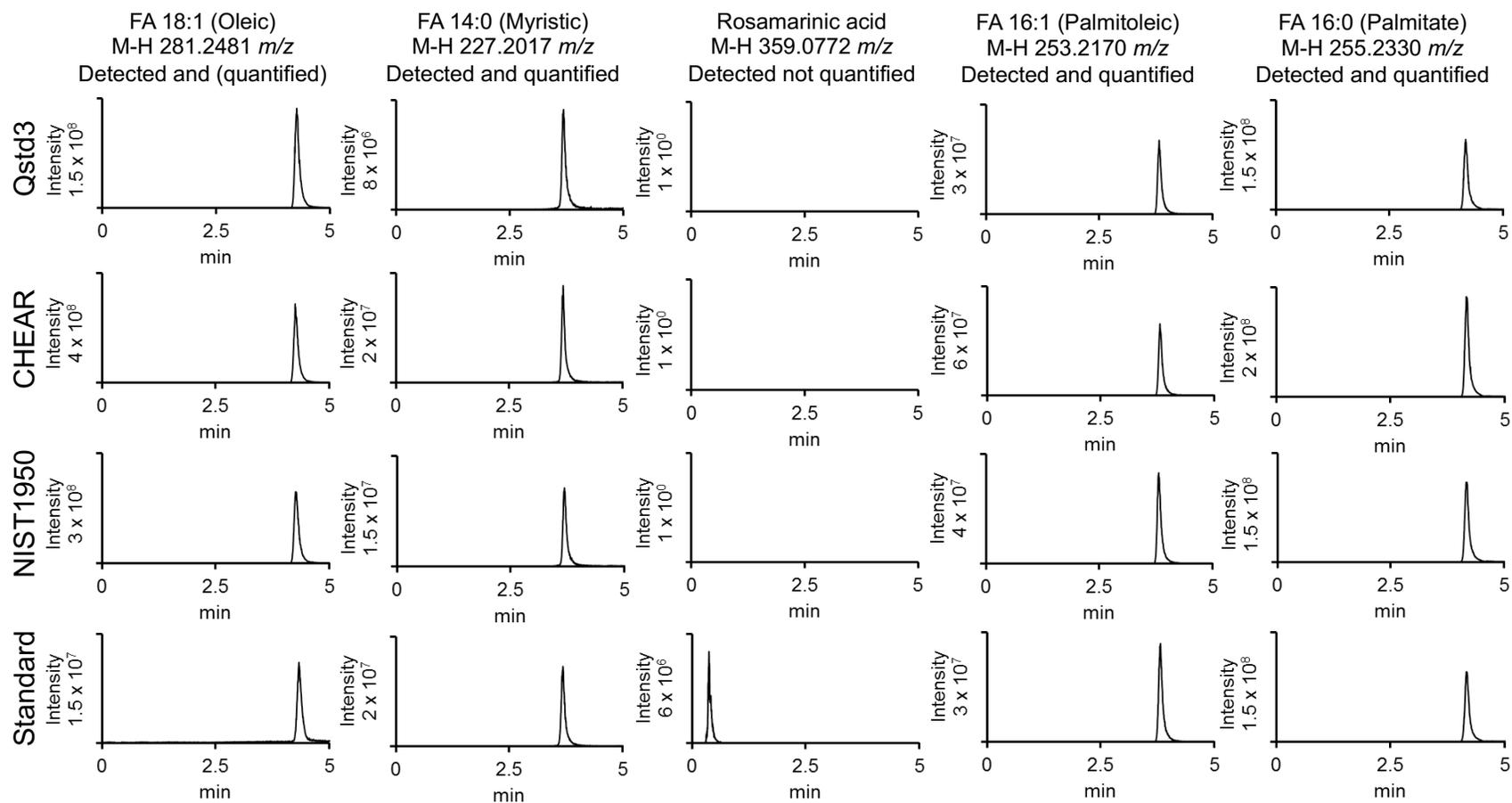


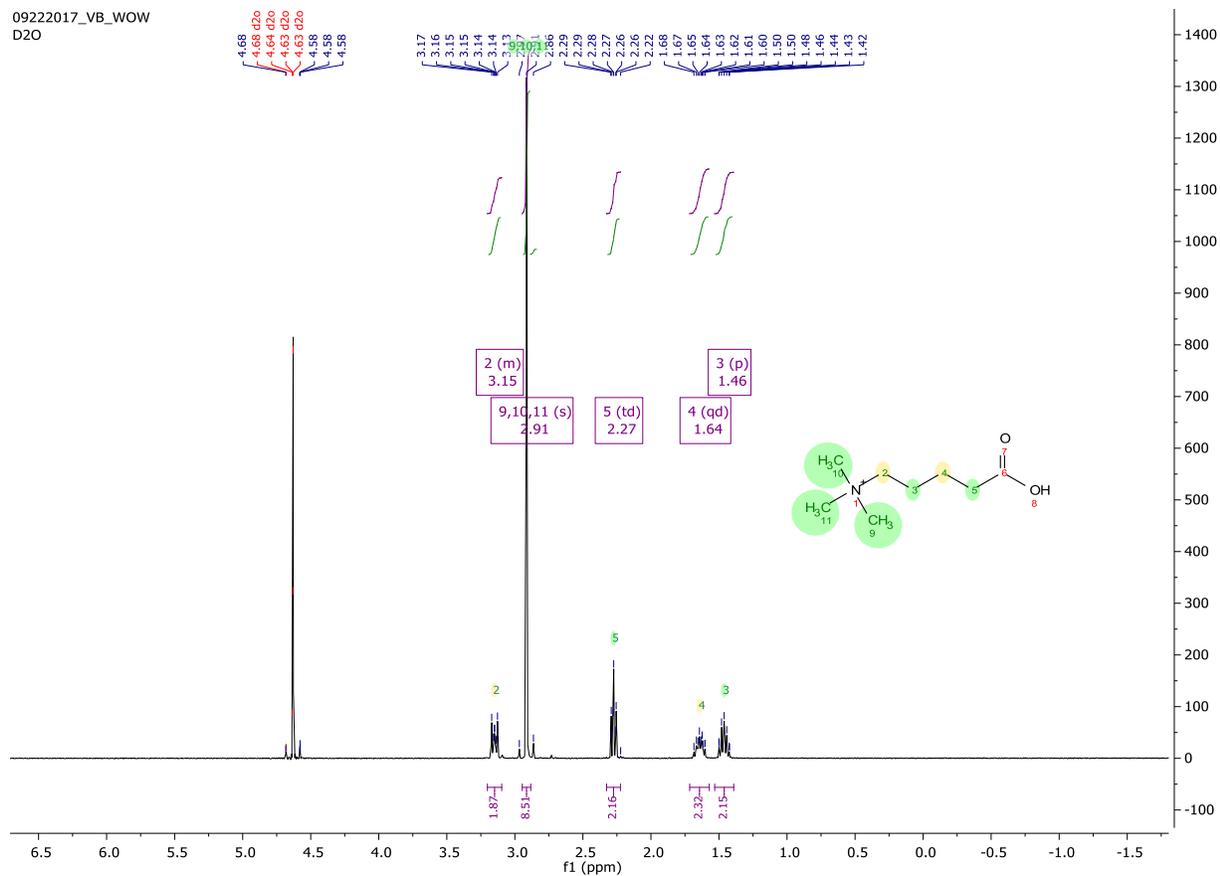




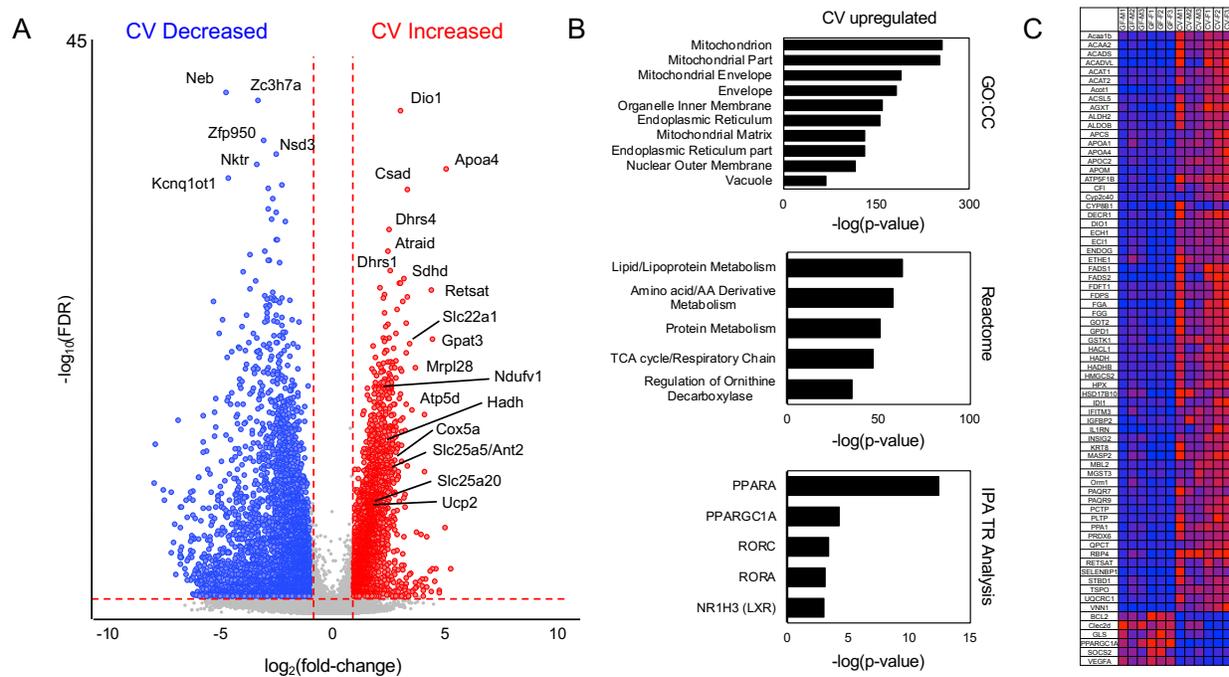








Appendix 3 (Chapter 4). ^1H NMR of synthesized δ -valerobetaine in D_2O . The protons are assigned as followed: δ 3.15 (multiplet, 2H), δ 2.91 (singlet, 9H), δ 2.27 (triplet, 2H), δ 1.64 (quintuplet, 2H), δ 1.46 (quintuplet, 2H).



Appendix 4 (Chapter 4). RNAseq analysis shows the influence of the microbiome on hepatic gene expression. A) Differential expression analysis using DeSeq2 revealed 2,711 transcripts increased in CV liver and 3,458 transcripts increased in GF liver (\log_2 fold-change > 1, FDR < 0.05). Of the 2,711 transcripts increased in CV liver, 2,508 were mapped. Of the 3,458 transcripts decreased in GF liver, 2,491 were mapped (MSigDB). B) Overrepresentation analysis of differentially expressed genes (DeSeq2, FDR < 0.05) shows the microbiome influences transcriptional regulation of hepatic energy metabolism (mitochondria GO:CC and Reactome lipid/lipoprotein metabolism, Amino acid metabolism, TCA cycle metabolism). Upstream analysis with IPA software reveals PPAR-alpha as the top transcriptional regulator targeted by the microbiota. C) PPAR-alpha target gene expression in mammalian liver.

