

**HIGH-RESOLUTION METABOLOMICS FOR
PROFILING ENVIRONMENTAL EXPOSURES AND
BIOLOGICAL RESPONSE**

A dissertation submitted by

Douglas Ian Walker

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Civil and Environmental Engineering

Tufts University

August 2017

DISSERTATION COMMITTEE

Professor Kurt Pennell, Adviser, *Tufts University*
Professor Dean Jones, Co-Adviser, *Emory University*
Professor Doug Brugge, *Tufts University*
Professor Elena Naumova, *Tufts University*
Professor David Gute, *Tufts University*

Copyright © 2017
Douglas I. Walker
All Rights Reserved

“For the first time in the history of the world, every human being is now subjected to contact with dangerous chemicals, from the moment of conception until death.”

-Rachel Carson, *Silent Spring*

ABSTRACT

The environment is a critical determinant of human health. Environment and gene-environment interaction are suspected to contribute to over 85% of disease burden; however, only limited characterization is possible with current technologies. Thus, to develop a balanced view of human health and disease risk there is a critical need to establish analytical methodologies that provide enhanced population screening for identifying exposures and biological relevance.

The major objective of this dissertation was to establish high-resolution metabolomics (HRM) as a viable platform for environmental chemical surveillance and bioeffect monitoring in human populations. Recent advances in high-resolution mass spectrometry, liquid chromatography and data processing algorithms have positioned HRM as the preeminent analytical framework for precision medicine and exposome research. Incredible potential exists for expanding HRM methodology to provide the systematic measures required for population screening applications, yet there is a critical need to validate that environmental exposures lead to metabolic variations detectable in easily obtainable biological fluids. Here, the use of HRM for this purpose was demonstrated through application of metabolome-wide association study to five independent cohorts with well-characterized exposure measures. Metabolic variations were then evaluated for exposure biomarkers and alterations consistent with disease-related pathobiology.

The key outcome of this work is the establishment of HRM as a central platform linking exposures to internal dose and biological response. For each

study, the results demonstrate it is possible to detect metabolic variations due to acute or chronic environmental exposures using HRM. These metabolic alterations, which include biomarkers of environmental pollutants, biological response molecules and metabolic changes consistent with disease pathobiology, are the key functional inputs required to characterize environmental contributions to human disease. Overall, these findings demonstrate that HRM is a sufficiently robust analytical framework for population screening. Continued development is expected to greatly enhance measures of the occurrence, distribution and magnitude of exposures in humans, while advancing knowledge on mechanisms underlying environment-related diseases.

ACKNOWLEDGEMENTS

Successful completion of this dissertation would not have been possible without the guidance and mentorship provided by Dr. Kurt Pennell of Tufts University, Dr. Dean Jones of the Emory University School of Medicine and Dr. Gary Miller of the Emory University School of Public Health. From Dr. Pennell, I learned the importance of challenging myself by researching topics beyond my area of expertise. This proved invaluable as I transitioned from learning about the fate and transport of environmental pollutants in porous media to chemical biomonitoring in humans, and finally to understanding the biological effects of environmental exposures. In addition, without Dr. Pennell's patience, ability to see the larger goals of my research and willingness to let me pursue areas of research well beyond what I thought I was capable of none of this research would have been possible. Through my initial visit to his lab and subsequent residence, Dr. Jones mentorship and encouragement has been a crucial component of my development as a scientist. His wisdom, coaching and friendship has been a continued presence since arriving at Emory in 2013, and I owe much of what I know to his willingness to take time out of his busy schedule to discuss life, science, research and how to most effectively develop my career. I will never forget his patience and willingness to help me identify how I can improve as a scientist, which included strategies for successfully communicating my research, grantsmanship, managing projects, developing collaborations and writing effectively. I hope I can pass along the same level of guidance and mentorship to future students that I was able to receive from Dr. Jones. Without Dr. Miller, my

initial visit to Emory and subsequent dissertation research would not have occurred. He is an integral component of my involvement in human health research; his leadership and vision of the HERCULES Exposome Research Center has been an inspiration for developing new tools to identify environmental exposures in human populations and has provided me with unparalleled opportunities for professional development.

This dissertation was dependent upon samples collected through separately funded projects, and the contribution of researchers involved with the initial studies was critical. First, I would like to acknowledge Timothy Mallon and Patricia Rohrbeck from the Department of Defense, Phillip Hopke of Clarkson University, Mark Utell, Richard Phipps and Tom Thatcher of the University of Rochester for their contributions to **Chapters 4** and **9**. For their involvement with the trichloroethylene study in **Chapter 5**, I would like to acknowledge Nathaniel Rothman, Qing Lan, Wei Hu and Bryan Bassig of the National Cancer Institute, and Roel Vermeluen from Utrecht University. Because of their role in the original Trucker Study (**Chapter 6**) and enthusiastic contributions to our meetings on how to best incorporate metabolomics, I would like to acknowledge Francine Laden, Jaime Hart and Chirag Patel of Harvard University, and Ruthann Rudel of the Silent Spring Institute. I would also like to acknowledge Doug Brugge and John Durant of Tufts University, Kevin Lane of Boston University, Allison Patton of HEI and Christina Fuller of Georgia State University for their efforts on the CAFEH study. Without the careful study design, willingness to provide samples and excellent exposure assessment, **Chapter 7** would not have been possible.

Finally, I would like to acknowledge Michele Marcus, Elizabeth Marder and Melanie Jacobson of Emory University for their contributions to **Chapter 8**.

During the course of my PhD, I was fortunate to have worked with an excellent group of scientists. Their encouragement, friendship and assistance has been critical at all steps. Most importantly, I would like to thank the Clinical Biomarkers Laboratory at Emory University. The friendship and contributions of Vilinh Tran, Bill Liang, Michael Orr, Dr. Karan Uppal and Dr. James Roede were crucial for my success. I look forward to many future research collaborations and continued development of high-resolution metabolomics. I will also look back fondly on friendships formed in the CEE department and IMPES lab at Tufts University, which made those initial PhD years all the more enjoyable. I would also like to thank my committee members, Dr. Elena Naumova, Dr. David Gute and Dr. Doug Brugge, for agreeing to serve on my committee and providing such valuable guidance along the way.

Finally, I would like to thank my family for their patience and support over the years. None of this would have been possible without the love, understanding and encouragement from my wife Kayla, who provided support in many ways during those long days and late nights in the lab and while I was writing my thesis. I would also like to thank my in-laws, the Perry's, who have opened their home to me and always made me feel like a member of the family; and my parents, Matt and Sandra Walker, who have always provided the encouragement and support I needed. For teaching me to work hard and be the best possible, I would like to thank my grandparents, Stuart and Johanna Walker.

Research presented in this thesis was supported by the National Institutes of Health (Award numbers CP010121, ES016284, ES015462, ES012014, ES019776, ES023485, ES009047, IHL113451, ES025632, ES026071, ES016284, OD018006, ES026560, OD018006, MH107205), Department of Science and Technology of Guangdong Province, China (Award number 2007A050100004), Department of Defense (Contract 306889-1.00-64239), Tufts University graduate student travel fund and the Strategic Environmental Research and Development Program (SERDP) under Project ER-1612 (Contract W91HQ-08-C-0003). The content of this document has not been subject to agency review and does not necessarily represent the view of the sponsoring agency.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION	2
1.1. INTRODUCTION	2
1.2. RESEARCH OBJECTIVES	6
1.3. RELEVANCE TO THE FIELD OF ENVIRONMENTAL ENGINEERING	9
CHAPTER 2. LITERATURE REVIEW	12
2.1. THE METABOLIC PHENOTYPE	12
2.2. THE HUMAN EXPOSOME	15
2.3. POPULATION SCREENING	17
2.3.1. <i>Diurnal metabolic phenotype</i>	24
2.3.2. <i>Gender metabolic phenotype</i>	26
2.3.3. <i>Age-related metabolic phenotype</i>	28
2.3.4. <i>Disease metabolic phenotype</i>	31
2.3.5. <i>Dietary and nutritional metabolic phenotype</i>	35
2.3.6. <i>Geographic metabolic phenotype</i>	39
2.4. METABOLIC PHENOTYPING FOR ENVIRONMENTAL EXPOSURES	42
2.4.1. <i>Population screening for hazard identification</i>	45
2.4.2. <i>Metabolic phenotyping in environment wide association study of disease</i>	50
2.4.3. <i>Chemical exposure measurement in personalized medicine</i>	56
2.5. CONCLUSION.....	57
CHAPTER 3. RESEARCH METHODOLOGY	59

3.1.	STUDY SELECTION	59
3.1.1.	<i>Description of study cohorts</i>	60
3.1.1.1.	Feasibility study for post-deployment exposure surveillance system	60
3.1.1.2.	Molecular epidemiology studies of occupational and environmental exposures	62
3.1.1.3.	Biomarkers of exposure and effect in a traffic exposed population	62
3.1.1.4.	Community assessment of freeway pollution and health	64
3.1.1.5.	Michigan PBB cohort 30 years later: Endocrine disruption?	65
3.2.	HIGH-RESOLUTION METABOLOMICS METHODOLOGY	65
3.2.1.	<i>Plasma preparation</i>	67
3.2.2.	<i>Liquid chromatography and high-resolution mass spectrometry.</i>	68
3.2.2.1.	Dionex Ultimate 3000 with Q-Exactive HRMS (Chapters 4,6,8)	69
3.2.2.2.	Accela Surveyor with LTQ-Velos HRMS (Chapter 5).....	70
3.2.2.3.	Ultimate 3000 with Q-Exactive HF HRMS (Chapter 7)	71
3.2.3.	<i>Mass spectral feature extraction and alignment</i>	72
3.2.4.	<i>Metabolite annotation and characterization</i>	73
3.2.5.	<i>Metabolic pathway enrichment and bioeffect</i>	74
3.3.	HRM DATA ANALYSIS AND METABOLOME WIDE ASSOCIATION STUDY ..	75
3.3.1.	<i>Univariate statistical analyses for feature selection</i>	76
3.3.2.	<i>Multivariate statistical analyses for feature selection</i>	79
3.3.3.	<i>Network integration</i>	80

3.4.	REFERENCE STANDARDIZATION	82
CHAPTER 4. PILOT METABOLOME-WIDE ASSOCIATION STUDY		
OF BENZO(A)PYRENE IN SERUM FROM MILITARY PERSONNEL .. 84		
4.1.	ABSTRACT	86
4.2.	INTRODUCTION	87
4.3.	MATERIALS AND METHODS.....	90
4.3.1.	<i>Serum Samples</i>	90
4.3.2.	<i>BaP targeted quantification</i>	90
4.3.3.	<i>High-resolution metabolomics</i>	91
4.3.4.	<i>Feature selection</i>	92
4.3.5.	<i>Metabolite correlation and pathway enrichment</i>	92
4.4.	RESULTS	94
4.4.1.	<i>Targeted quantification of endogenous human metabolites</i>	94
4.4.2.	<i>BaP targeted profiling</i>	94
4.4.3.	<i>High-resolution metabolomics</i>	95
4.4.4.	<i>Serum BaP MWAS</i>	97
4.4.5.	<i>Metabolite correlation and pathway analysis</i>	100
4.5.	DISCUSSION	101
4.6.	CONCLUSIONS.....	108
CHAPTER 5. HIGH-RESOLUTION METABOLOMICS OF		
OCCUPATIONAL EXPOSURE TO TRICHLOROETHYLENE 109		
5.1.	ABSTRACT	111
5.2.	INTRODUCTION	113

5.3.	METHODS	116
5.3.1.	<i>Study design and exposure measurement</i>	116
5.3.2.	<i>High-resolution metabolomics</i>	117
5.3.3.	<i>Data analysis</i>	117
	122
5.4.	RESULTS	123
5.4.1.	<i>Study population</i>	123
5.4.2.	<i>TCE exposure MWAS</i>	123
5.4.3.	<i>Identification of TCE exposure products</i>	123
5.4.4.	<i>Biological response to TCE exposure</i>	125
5.4.5.	<i>Integration with immunological, renal and TCE markers</i>	127
5.5.	DISCUSSION	130
5.6.	CONCLUSIONS.....	134

CHAPTER 6. INTEGRATED MOLECULAR RESPONSE OF EXPOSURE TO TRAFFIC-RELATED POLLUTANTS IN THE US TRUCKING INDUSTRY..... 135

6.1.	ABSTRACT	136
6.2.	INTRODUCTION	138
6.3.	METHODS	142
6.3.1.	<i>Study population</i>	142
6.3.2.	<i>Microenvironment exposure measures</i>	142
6.3.3.	<i>Biomarker sampling</i>	143
6.3.4.	<i>High-resolution metabolomics</i>	144

6.3.5.	<i>Peripheral blood transcriptomics</i>	145
6.3.6.	<i>Metabolome wide-association study of traffic-related pollutants</i> <i>145</i>	
6.3.7.	<i>Metabolite annotation and pathway enrichment</i>	147
6.3.8.	<i>Network integration of the metabolome and transcriptome</i>	147
6.4.	RESULTS	149
6.4.1.	<i>Study population</i>	149
6.4.2.	<i>High-resolution metabolomics results</i>	150
6.4.3.	<i>Daily exposure MWAS</i>	151
6.4.4.	<i>Week-averaged exposure MWAS</i>	153
6.4.5.	<i>Network correlation of molecular response</i>	162
6.5.	DISCUSSION	167
6.6.	CONCLUSIONS.....	183

CHAPTER 7. METABOLOMIC ASSESSMENT OF YEARLONG

EXPOSURE TO NEAR-HIGHWAY ULTRAFINE PARTICLES 185

7.1.	ABSTRACT	186
7.2.	INTRODUCTION	188
7.3.	METHODS	192
7.3.1.	<i>Study population</i>	192
7.3.2.	<i>Exposure assessment</i>	193
7.3.3.	<i>High-resolution metabolomics</i>	193
7.3.4.	<i>Reference standardization</i>	196
7.3.5.	<i>Metabolome-wide association study of UFP exposure</i>	196

7.3.6.	<i>Metabolite annotation</i>	197
7.3.7.	<i>Metabolome correlation and pathway enrichment</i>	198
7.3.8.	<i>Correlation network with inflammatory and coagulation biomarkers</i>	199
7.4.	RESULTS	200
7.4.1.	<i>Study population</i>	200
7.4.2.	<i>High-resolution metabolomics</i>	200
7.4.3.	<i>Reference standardization</i>	203
7.4.4.	<i>Metabolome-wide association study of UFP exposure</i>	204
7.4.5.	<i>Metabolite correlation and pathway enrichment</i>	210
7.4.6.	<i>Metabolic correlation with inflammatory and endothelial biomarkers</i>	213
7.5.	DISCUSSION	215
7.6.	CONCLUSIONS.....	225

CHAPTER 8. METABOLOME WIDE ASSOCIATION STUDY OF POLYBROMINATED (PBB-153) AND POLYCHLORINATED (PCB-153) BIPHENYLS 227

8.1.	ABSTRACT	228
8.2.	INTRODUCTION	230
8.3.	METHODS	234
8.3.1.	<i>Study population</i>	234
8.3.2.	<i>Plasma PBB-153 and PCB-153 levels</i>	234
8.3.3.	<i>High-resolution metabolomics</i>	236

8.3.4.	<i>Data analyses</i>	237
8.3.5.	<i>Annotation and metabolic pathway enrichment</i>	238
8.3.6.	<i>Comparison to the National Health and Nutrition Examination Survey</i>	239
8.4.	RESULTS	240
8.4.1.	<i>Study population</i>	240
8.4.2.	<i>PBB-153 and PCB-153 plasma levels</i>	240
8.4.3.	<i>High-resolution metabolomic profiling</i>	243
8.4.4.	<i>F₀ PBB-153 and PCB-153 MWAS</i>	244
8.4.5.	<i>F₁ PBB-153 and PCB-153 MWAS</i>	247
8.4.6.	<i>Generational differences in exposure metabolic phenotype</i>	249
8.5.	DISCUSSION	252
8.6.	CONCLUSIONS.....	263

CHAPTER 9. DEPLOYMENT-ASSOCIATED EXPOSURE

SURVEILLANCE WITH HIGH-RESOLUTION METABOLOMICS 265

9.1.	ABSTRACT	267
9.2.	CRITICAL NEED FOR DEPLOYMENT-ASSOCIATED EXPOSURE ASSESSMENT	268
9.3.	HIGH-RESOLUTION METABOLOMICS (HRM): ADVANCED CLINICAL	
	CHEMISTRY	272
9.3.1.	<i>Mass spectrometry for metabolic profiling</i>	272
9.3.2.	<i>Quantification</i>	276
9.3.3.	<i>Applications for precision medicine.</i>	278

9.4.	HRM FOR ENVIRONMENTAL CHEMICAL ANALYSIS	280
9.4.1.	<i>Detection of low abundance chemicals</i>	280
9.4.2.	<i>Measurement of exposure and linking to body burden</i>	281
9.4.3.	<i>Characterizing internal dose response</i>	285
9.4.4.	<i>Exposure memory</i>	287
9.4.5.	<i>Addressing the dark matter of the exposome</i>	289
9.5.	HRM IN DEPLOYMENT EXPOSURE SURVEILLANCE	291
9.5.1.	<i>Retrospective analysis using the DoDSR</i>	291
1.1.1.	<i>Prospective use for exposure surveillance</i>	292
9.5.2.	<i>Integration with targeted phenotypic and other omics platforms</i> .	
	293	
9.6.	HRM: SUMMARY AND PERSPECTIVE	294
CHAPTER 10. CONCLUSIONS AND RECOMMENDATIONS FOR		
FUTURE RESEARCH..... 297		
10.1.	CONCLUSIONS.....	297
10.2.	RECOMMENDATIONS FOR FUTURE RESEARCH	310
10.3.	CONTRIBUTIONS	314
APPENDIX 1. LIST OF SUPPLEMENTARY FILES 319		
REFERENCES..... 320		

LIST OF FIGURES

Figure 1.1. <i>Source-to-outcome vision for exposure science.</i>	4
Figure 1.2. <i>Conceptual framework for use of HRM in environmental chemical surveillance and bioeffect monitoring.</i>	6
Figure 2.1. <i>The human metabolome represents expression of an individual's metabolic phenotype, and is a combination of chemicals from core biological processes, the microbiome, dietary exposures, drugs and exposure to xenobiotics from the environment and commercial products..</i>	13
Figure 2.2. <i>Advantages of high-resolution metabolomics for population screening and precision medicine research.</i>	19
Figure 2.3. <i>Comparison of untargeted and targeted chemical profiling techniques for population research.</i>	22
Figure 2.4. <i>Prevalence of multi-morbidities in different age groups from database of 1,751,841 patients from Barnett, et al (2012).</i>	30
Figure 2.5. <i>Number of articles listed in PubMed for years 2002-2014 after searching for the terms “disease”, “metabolomics” and “diagnostic”.</i>	33
Figure 2.6. <i>Categorization of metabolic phenotyping data obtained from seven mammalian species in Park et al. (2012) for bioeffect measurement and biomonitoring of environmental exposures.</i>	45
Figure 2.7. <i>Meet-in-the-middle approach to discover causal relationships between environmental exposures and disease outcome (Vineis et al. 2013).</i>	51

Figure 4.1. BaP serum concentrations in the 30 unidentified individuals. Quantified serum levels ranged from 0.13 to 37.2 ng/mL, with an average concentration of 3.39 ng/mL.	94
Figure 4.2. Metabolite coverage of metabolic pathways following database matching to KEGG. Matches are represented by black dots.	96
Figure 4.3. Unsupervised hierarchical clustering heatmap of the top 100 sPLS selected features associated with BaP. Individual grouping was found to be strongly associated with exposure level (colored boxes on left axis), reflecting exposure-metabolic signatures within the chemical profile.	97
Figure 4.4. Network association of the sPLS m/z features with raw HRM data. Red circles represent sPLS selected features; green triangles are correlating features with Spearman $ r \geq 0.6$ and $FDR \leq 20\%$	98
Figure 4.5. Top 15 metabolic pathways with enrichment p-value from Mummichog using the sPLS and network associated features.	100
Figure 4.6. Metabolite intensity differences from glycerophospholipid metabolism for individuals in the high and low BaP clusters.	103
Figure 4.7. Metabolite intensity differences from sulfur amino acid metabolism for individuals in the high and low BaP clusters.	104
Figure 5.1. Workflow of HRM data analysis and m/z feature characterization for TCE MWAS results	119
Figure 5.2. Manhattan plot for TCE MWAS in exposed and non-exposed workers. Linear regression identified m/z features that were associated with TCE exposure at $FDR \leq 20\%$ after correcting for age, sex and BMI.	124

Figure 5.3. Mummichog enriched pathways for metabolites associated with TCE exposure.....	126
Figure 5.4. Distribution of uric acid, glutamine and MTA in control, low exposed and high exposed workers.....	127
Figure 5.5. Network correlation analysis of molecular markers previously tested for association with TCE exposure and metabolic features detected in.....	129
Figure 6.1. Manhattan plot showing $-\log p$ as a function m/z feature retention time for daily exposure to $PM_{2.5}$, EC and OC.....	150
Figure 6.2. Manhattan plot showing $-\log p$ as a function m/z feature retention time for week averaged exposure to $PM_{2.5}$, EC and OC.....	154
Figure 6.3. Overlapping features associated with shift-length and week averaged exposure to the two pollutants at false discovery rate threshold 20%.	155
Figure 6.4. Metabolite and peripheral blood gene expression integration network. Discriminatory features and network correlation were completed using the R package xMWAS.	166
Figure 7.1. Metabolome-wide association study of UFP exposure using HILIC HRM data.....	206
Figure 7.2. Metabolome-wide association study of UFP exposure using C_{18} with negative ESI HRM data.	209
Figure 7.3. Correlation network structure of m/z features associated with UFP exposure from HILIC and C_{18} HRM data with independently measured biomarkers of inflammation and coagulation.....	214

Figure 8.1. PCB-153 and PBB-153 congener levels (log scale) measured in plasma obtained from individuals in generation F_0 and F_1	241
Figure 8.2. Results of PBB-153 and PCB-153 metabolome wide association study in F_0 . Metabolic changes associated with exposure were detected for both pollutants	242
Figure 8.3. Results of PBB-153 and PCB-153 metabolome wide association study in F_1 . As was observed in the older generation, metabolic changes associated with exposure were detected for both pollutants	246
Figure 9.1. High-resolution MS supports untargeted measurement of metabolic chemicals by reducing requirements for chemical separation and sample preparation while providing improved capability to measure low abundance chemicals in biological samples.	273
Figure 9.2. HRM profiling of plasma obtained from healthy individuals has indicated that metabolic intermediates from >80% of pathways present in the KEGG database can be detected (represented by black dots) using standardized sample preparation methods and dual-column chromatography.	274
Figure 9.3. Cross-platform comparison of targeted polycyclic aromatic hydrocarbon (PAH) measurements using gas chromatography MS and suspect screening for PAH metabolites by HRM.....	284
Figure 9.4. A targeted metabolome wide association study (MWAS) was completed by testing each metabolite for significant associations with triethylphosphate levels in plasma.....	287

Figure 9.5. HRM profiling of specimens collected from active duty individuals could readily be integrated into DoDSR protocols. In this example, HRM can evaluate the presence of exposure and effect biomarkers, identifying individuals at risk for exposure related diseases..... 292

Figure 9.6. Microbiome-MWAS of bronchoalveolar lavage fluid in HIV-1: **A** to **C** represent increasing stringency to identify network associations with greatest r meeting significance criteria. **D** to **F** contains the network subset associated with HIV-1, in order of increasing leniency. Adapted from Cribbs et al. (2016)..... 296

LIST OF TABLES

Table 2.1. <i>Representative dietary biomarkers for identifying food consumption and improving food recall in population studies from Scalbert et. al (2014). Used with permission from the American Society for Nutrition.</i>	37
Table 2.2. <i>Concentration of selected environmental chemicals in a cohort of 153 health individuals detected and quantified using high-resolution metabolomics and reference standardization (Go et al. 2015).</i>	48
Table 3.1. <i>Summary of environmental health studies selected for HRM profiling</i>	60
Table 4.1. <i>Serum concentrations of representative confirmed metabolites in the 30 unidentified samples.</i>	93
Table 5.1. <i>Demographic characteristics and TCE exposure level</i>	122
Table 5.2. <i>Top m/z features matching chlorinated isotope patterns</i>	122
Table 5.3. <i>Identified endogenous metabolites associated with TCE exposure and relevant to biological response to exposure</i>	125
Table 6.1. <i>Population characteristics of participants with plasma selected for HRM profiling</i>	149
Table 6.2. <i>Relevant metabolite annotations for m/z features associated with daily exposure to elemental carbon and organic carbon</i>	152
Table 6.3. <i>Relevant metabolite annotations for m/z features associated with week-averaged exposure to elemental carbon</i>	156
Table 6.4. <i>Steroid-related metabolites matching m/z features associated with week-averaged exposure to elemental carbon</i>	157
Table 6.5. <i>Mummichog enriched metabolic pathways for elemental carbon.</i>	157

Table 6.6. <i>Relevant metabolite annotations for m/z features associated with week-averaged exposure to organic carbon</i>	161
Table 6.7. <i>Mummichog enriched metabolic pathways for organic carbon</i>	162
Table 7.1. <i>Population characteristics and UFP exposure levels</i>	199
Table 7.2. <i>Measured amino acid, amino acid metabolites and clinical health marker levels</i>	201
Table 7.3. <i>Measured fatty acids, lipids, nucleotides, TCA cycle metabolites and co-factors</i>	202
Table 7.4. <i>Relevant metabolites identified by cross-sectional PLS-DA of HILIC with positive ESI data</i>	207
Table 7.5. <i>Relevant metabolites identified by cross-sectional PLS-DA of C₁₈ with negative ESI data</i>	208
Table 7.6. <i>Metabolic pathways associated with exposure using HILIC with positive ESI data</i>	211
Table 7.7. <i>Metabolic pathways associated with exposure using C₁₈ with negative ESI data</i>	212
Table 8.1. <i>Population characteristics of participants selected for HRM profiling</i>	239
Table 8.2. <i>Selected metabolites correlated with PBB-153 and PCB-153 in F₀</i> .	243
Table 8.3. <i>Metabolic pathway enrichment results for m/z features correlated with PBB-153 and PCB-153 in F₀</i>	244
Table 8.4. <i>Selected metabolites correlated with PBB-153 and PCB-153 in F₁</i> .	247

<i>Table 8.5. Metabolic pathway enrichment results for m/z features correlated with PBB-153 and PCB-153 in F₁</i>	248
<i>Table 8.6. Overlapping features correlated with PBB-153 or PCB-153 in both F₀ and F₁</i>	251
<i>Table 8.7. PBB-153 and PCB-153 correlation with circulating fatty acid levels in 2003-2004 NHANES</i>	261

**HIGH-RESOLUTION METABOLOMICS FOR
PROFILING ENVIRONMENTAL EXPOSURES AND
BIOLOGICAL RESPONSE**

Chapter 1. INTRODUCTION

1.1. Introduction

Estimates from genome wide association studies (GWAS) suggest that only 10-20% of diseases have a strong genetic component (Wild 2005), with the remaining 80-90% of unknown etiology (Rappaport and Smith 2010). Genotyping enables the identification of single gene mutations that are highly penetrant (Vineis et al. 2001); however, disease etiology is often multifactorial and driven by a combination of genetic, environment and lifestyle factors (Rappaport et al. 2014). A more complete understanding of how environment and lifestyle factors contribute to disease susceptibility and progression is required for mitigating risk, developing effective treatment strategies and identifying at risk populations. Currently, no unified method exists to characterize the sum involvement of environment and chemical exposures in disease.

Development of analytical strategies supporting environment-wide association studies (EWAS) (Patel et al. 2010) for risk assessment and identifying disease susceptibility has been identified as a critical need by the Institute of Medicine (Baird 2012), National Research Council (NRC 2009) and Department of Defense (Lindler 2015). In addition, it is well recognized within the environmental health, exposure science, epidemiology and medical research

communities that more comprehensive methods of characterizing environmental effects on human health are required (Wild 2005, Vineis and Perera 2007, Nicholson et al. 2008, Rappaport and Smith 2010, Wild 2012, Wild et al. 2013, Miller and Jones 2014, Rappaport et al. 2014, Vineis 2015). While the diverse nature of an individual's chemical experience combined with phenotypic heterogeneity will likely preclude development of methods for complete accounting of all exposures over a lifetime, identification of molecular markers providing tangible measurements for population research and risk assessment will greatly improve understanding of how environment influences human health.

Incorporating exposure information into population research has traditionally relied on monitoring approaches with varying levels of uncertainty. For example, heuristic models calibrated to chemical monitoring surveys have been employed to prioritize toxicity screening (Wambaugh et al. 2014), geospatial models have been used to predict respiratory exposures (Lane et al. 2015), recall surveys have been used to estimate dietary exposures (Chadeau-Hyam et al. 2011), ambient exposure measurements can be useful for estimating exposure to large groups, and breathing zone samplers have been developed to estimate exposure over short and long-term periods (Lan et al. 2010, O'Connell et al. 2014); however, these methods provide generalized estimates and cannot be used to assess internal exposure and biological relevance. Targeted biomonitoring utilizes measurement in biological specimens to estimate body burden of specific chemicals, providing information on internal dose and prevalence in a population. While biomonitoring has proven invaluable in population surveys, chemical

coverage is often limited. For example, the National Health and Nutritional Examination Survey (NHANES) applied targeted biomonitoring approaches to measure 212 chemicals in a cross-section of the US population; however, over

100,000 chemicals are registered for commercial use with the EPA, with recent estimates from the Toxic Substance Control Act suggesting 70,000 are commonly used. A recent survey by Dionisio et al. (2015) identified

approximately 20,000 used directly in consumer products. The ability to provide chemical monitoring on this magnitude far exceeds the capability of targeted platforms, therefore, advanced chemical profiling techniques are required to identify and monitor environmental exposures in humans.

High-resolution metabolomics (HRM) using liquid chromatography interfaced to ultra-high accuracy mass spectrometers has been identified as a promising analytical platform for broad environmental chemical surveillance and bioeffect monitoring. Due to increases in scan speed and data extraction

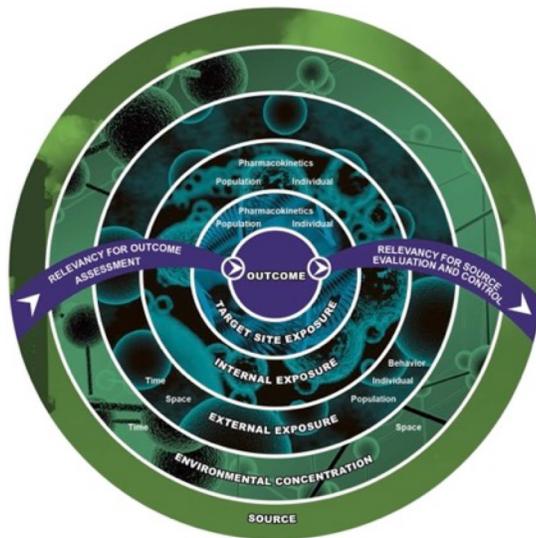


Figure 1.1. Source-to-outcome vision for exposure science. With the availability of advanced analytical and sensor technologies, exposure science has the capability to develop complementary measures over the continuum from exposure to internal dose and biological response. Application of HRM within this framework provides cost-effective measures of exposure biomarkers, internal dose, biological response and alternative endpoints. From (NRC 2012)

algorithms, modern instruments are now capable of detecting 10,000-75,000 unique chemical signals in a small volume (< 150 μ L) of blood. These chemicals include metabolites from core nutrient metabolism, lipids, the microbiome, diet-derived chemicals, pharmaceuticals and environmental contaminants (Jones et al. 2012, Wishart et al. 2013). The measures available using HRM therefore represents an integrated chemical profile derived from environment and endogenous processes. Measurement in human populations has the potential to provide quantitative assessment of exogenous chemicals. Changes in endogenous metabolites can be used to infer organ toxicity, key events in mode of action, toxin-targets and underlying mechanisms of chemical toxicity.

The advantages of HRM discussed above support the National Research Council (NRC) shift in risk assessment paradigm (NRC 2009) and recommendations in the 2012 NRC report, *Exposure Science in the 21st Century: A Vision and Strategy* (**Figure 1.1**) (NRC 2012, Lioy and Smith 2013). Measurement of the metabolic phenotype provides biomarkers of biologically relevant dose (reducing the need for apical endpoints and enables *in vivo* measurement of biochemical perturbations comparable to high-throughput *in vitro* assays for chemical hazard identification), dose-response relationships can be evaluated for individual chemicals and mixtures, chemical disposition, metabolism and products can be evaluated for differential effects, and chemical exposure surveillance can be completed in large populations at reasonable costs. Thus, the development of HRM as a tool for metabolic phenotyping of exposure and bioeffect will support the development of cost effective, advanced analytical

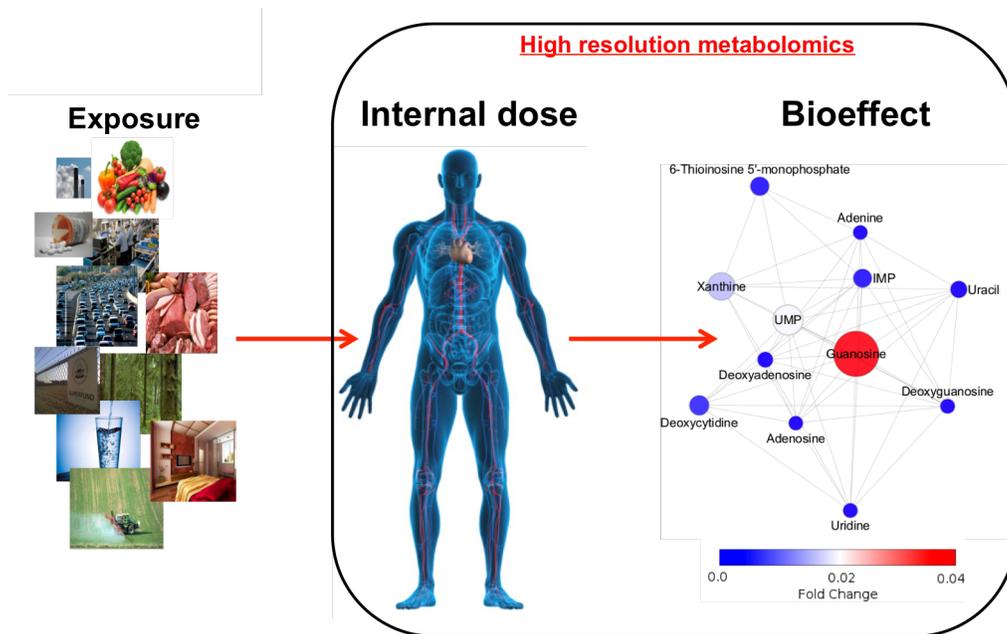


Figure 1.2. Conceptual framework for use of HRM in environmental chemical surveillance and bioeffect monitoring. Untargeted analytical techniques enable measurement of endogenous metabolism and plasma chemicals arising from environmental exposures, providing a means for chemical surveillance and biological response monitoring in general populations. The relationship between internal levels of exposure biomarkers and endogenous metabolism can then be related by testing for dose-response associations and changes consistent with early disease pathophysiology.

strategies providing the systematic information on exposures required for EWAS.

Many of these advantages have yet to be recognized, and there is a need to establish the feasibility of HRM to provide quantitative measurement of exposure biomarkers and biological response. The results of this dissertation provide a critical test for the suitability of HRM analytical strategies in exposure science and population screening.

1.2. *Research objectives*

The primary aim of this research was to develop HRM as a tool for monitoring chemical exposure and biochemical response in human populations.

The analytical approach is presented in **Figure 1.2**. Specific population cohorts with well-characterized exposures measured using traditional approaches were linked to alterations in the exogenous chemical profile (parent and exposure-related metabolites) and endogenous chemical profile (local and global changes in metabolic pathways) by HRM. This thesis begins with an overview of metabolic phenotyping (**Chapter 2**) and factors contributing to expression of the metabolic phenotype. **Chapter 3** provides justification for selecting each population in the following studies and description of research methodology. Results for specified research objectives are provided in **Chapters 4-8**. **Chapter 9** describes how HRM technologies can be developed as a public health surveillance tool, with particular focus on evaluating adverse exposures occurring to armed forces personnel during deployment. Conclusions and recommendations for future research are listed in **Chapter 10**. The objectives of this dissertation were to:

1. Address the suitability of biological samples stored in large bio-repositories for metabolomics profiling and detection of exposure related metabolic alterations. Thirty unidentifiable serum samples were obtained from the Department of Defense Serum Repository (DoDSR) and characterized using both targeted biomonitoring and untargeted HRM. Common endogenous metabolites were quantitated and compared to expected reference ranges. Metabolic changes associated with blood levels of benzo[a]pyrene (BaP), a ubiquitous environmental pollutant, were evaluated at the individual and pathway level. Identified biological

response was then compared to results from previously reported model system studies of BaP exposure (**Chapter 4**).

2. Evaluate the ability of HRM to detect dose-dependent metabolic alterations from personal exposure to trichloroethylene (TCE). Plasma collected from a Chinese population of 80 workers exposed to TCE and 95 frequency-matched, unexposed controls were profiled using HRM and tested for dose-dependent metabolic changes based upon shift-averaged TCE exposure levels. Mass-to-charge ratio (m/z) features associated with exposure were evaluated for the presence of halogenated compounds and endogenous metabolites. Relationship with independently measured biomarkers of immune function, renal damage and exposure were then tested (**Chapter 5**).
3. Identify the metabolic effects of shift- and week-averaged occupational exposure to pollutants from diesel exhaust. HRM profiling of repeat pre- and post-shift blood samples collected over the course of the workweek from non-smoking office workers, dockworkers and drivers employed in the trucking industry ($n=73$ participants) was completed and tested for association with micro-environment levels of traffic-related pollutants. Metabolic changes due to daily and week averaged $PM_{2.5}$, organic carbon (OC) and elemental organic carbon (EC) were determined and tested for correlation with the peripheral blood transcriptome to identify mechanisms underlying exposure-related metabolic changes (**Chapter 6**).

4. Investigate the feasibility of using HRM to evaluate plasma metabolite variations associated with yearlong time- and activity-adjusted exposure to ultrafine particulates (UFP). Metabolic changes related to high- and low-UFP exposure were determined in 59 individuals participating in the Community Assessment of Freeway Exposure and Health (CAFEH) study. Altered pathways were tested for changes consistent with air pollution-related disease pathology and for association with independently measured molecular markers of biological response (**Chapter 7**).
5. Evaluate HRM to detect metabolic changes from exposure to structurally similar persistent organic pollutants in two generations with different routes of exposure (dietary and maternal transfer). Blood levels of polybrominated biphenyl 153 and polychlorinated biphenyl 153 were measured in 80 blood plasma samples collected from individuals who ranged in age from 0-16 years during accidental contamination of livestock feed in Michigan and 76 who were born to parents residing in that area during that time period were tested for association with HRM data using a network-correlation approach. Similarities and differences for the two exposures and biological response in each generation were evaluated. Identified associations with fatty acid metabolism were also tested in an independent cohort selected from the National Health and Nutrition Examination Survey (**Chapter 8**).

1.3. Relevance to the field of environmental engineering

The field of environmental engineering is focused on managing the interface between society and the environment, with the overall goal protection of human and ecological health. A critical component of environmental engineering centers on hazard identification; without the ability to properly identify health risks it is not possible to provide management solutions and reduce these risks to acceptable levels. In the United States, over 85,000 chemicals are registered with the EPA for manufacture, import and use in commercial products. This is in addition to 40,000 registered pesticide formulations, approximately 100,000 dietary phytochemicals, 5,000 chemicals approved for use as inert ingredients and 7,500 compounds registered as drugs or food additives with the FDA. An individual's history of these exposures defines our "chemical experience", which contributes directly to phenotype and health. This crystallizes a grand challenge for environmental engineering: How should risk be assessed and managed for hundreds of thousands of chemicals without understanding of their prevalence, levels in the environment and related health effects?

This thesis seeks to establish HRM as a platform for environmental exposure surveillance and biological response monitoring in human populations. The techniques demonstrated are directly relevant to environmental engineering in that it provides a universal platform for population screening to assess the occurrence and health effects of environmental exposures occurring within society. Since exposure and biological response measurements are completed using an untargeted methodology, the presence of environmental contaminants can be determined in a holistic manner and inform on unexpected and

uncharacterized chemical exposure risks. Thus, application of HRM in a public health surveillance framework has the potential to greatly expand the information that is currently available using traditional environmental monitoring and exposure science techniques. Using knowledge obtained with HRM, it will be possible for engineers to identify chemicals posing greater environmental exposure risk, develop new guidelines for appropriate chemical formulations and design interventions that decrease exposures through improved remediation efforts.

Chapter 2. LITERATURE REVIEW

Chapter 2 was published as a review in the edited book, *Metabolic Phenotyping in Personalized and Public Healthcare* (Eds. J. Nicholson, A. Darzi, E. Holmes and J. Lindon. 2016. Elsevier. ISBN 9780128003442). Co-authors and contributions are as follows. D. Walker: Conceived and wrote all sections, performed literature review, created figures and tables, approved final edits and submission; Y. Go: Read and edited final document; K. Liu: Read and edited final document; K. Pennell: Provided critical input for manuscript structure, edited intermediate drafts, edited final document; D. Jones: Provided critical input and guidance for manuscript structure, edited intermediate drafts, edited and approved final document

2.1. *The metabolic phenotype*

The human metabolome is an expression of an individual's metabolic phenotype, and includes all endogenous metabolites, chemicals from human-environment interaction and the reactants arising from interaction of these compounds with enzymatic and bacterial processes occurring within multiple body components (**Figure 2.1**). True extent of the human metabolic profile is unknown, with current estimates suggesting greater than 10^6 metabolites (Uppal et al. 2016). Because of individual differences, chemical profiles will vary in

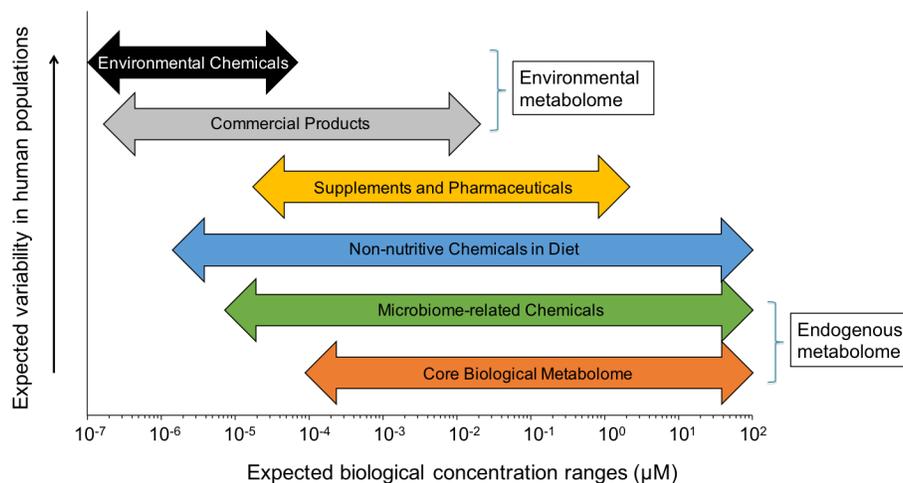


Figure 2.1. The human metabolome represents expression of an individual's metabolic phenotype, and is a combination of chemicals from core biological processes, the microbiome, dietary exposures, drugs and exposure to xenobiotics from the environment and commercial products. Within a population, individual differences will result in chemical measurements that vary in concentration and diversity due to unique chemical exposure experience and phenotypic characteristics for each person. While a core metabolome consisting of the metabolites required for life will be detected universally in population screening, environmental, drug and dietary chemicals in the metabolic phenotype are expected to vary greatly in concentration and presence. Furthermore, methods of analysis are required providing chemical detection over a wide range of concentration magnitudes.

concentration, properties and diversity due to the unique chemical exposure experience and phenotype for each person.

Recent efforts to curate a reference metabolome highlight the diversity present in metabolic phenotype. The Human Serum Metabolome (Psychogios et al. 2011) was developed to provide quantitative information on small molecular weight compounds, identifying 4,229 metabolites in human samples. When predicted metabolites, dietary and environmental chemicals are included, a much greater number of metabolites have been reported. For example, the Human Metabolome Database (Wishart et al. 2013) contains approximately 42,000

entries that are confirmed or expected in humans. The METLIN database (Smith et al. 2005) contains information on over 240,000 chemicals, including metabolites and exogenous compounds; however, there are also many synthetic chemicals present in the database not likely to be found in humans..

Untargeted metabolic profiling in human populations also detects a large number of chemicals that are not readily identifiable because of the absence of matches in chemical databases, lack of reference spectra and unavailability of analytical standards (da Silva et al. 2015). Environment and dietary sources are suspected to be the main contributors to unknowns. When considering transformation products, intermediates, byproducts and adducts, the number of possible chemical entities is enormous. For example, a search of the PubMed Compound database over the molecular weight range of 50-2000 dalton returned 54×10^6 chemicals. Although many of these chemicals are synthetic and not relevant to human exposures, upwards of 1 million chemicals are likely to contribute to the human metabolome. Plants are likely to contain 10^5 to 10^6 small molecules. Commercial use provides another source of chemical-human interaction contributing to metabolic phenotype. Chemicals registered with the United States EPA include 10^5 unique species, with recent estimates from the Toxic Substance Control Act indicating 70,000 are commonly used. To prioritize environmental chemicals for toxicity testing, a recent survey of 43,000 registered chemicals identified approximately 20,000 used directly in consumer products (Dionisio et al. 2015). When considering the multiplicative effect of transformation and detoxification products of dietary and environment derived

parent compounds, exogenous chemicals most likely exceed the endogenous contribution to metabolic phenotype, further supporting the role of external factors in the human metabolic profile and importance in population-based studies.

2.2. *The human exposome*

Establishing the role of environment in human health and disease will enable intervention with preventative measures and treatment therapies to reduce disease risk. Interaction of genetic and environment/lifestyle risk factors are suspected to be top contributors to the chronic disease burden; however, technologies available to measure genetic susceptibility far surpass those available for sequencing environmental exposures (Rappaport et al. 2014). The need to include environment in understanding human disease led Christopher Wild to introduce the concept of the exposome in 2005 (Wild 2005), which he defined as “encompassing life-course environmental exposures (including lifestyle factors), from the prenatal period onwards.” The exposome is envisioned as a complement to the genome, where life course of exposure and interaction with the genome defines risk for disease development. Unlike the genome, exposures are transient and change on both short and long-term time scales, making quantitative assessment challenging. A more tangible definition of the exposome was proposed by Miller and Jones (2014): “The cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behavior, and

endogenous processes.” Exposures in this framework are not only limited to external chemicals, but also include processes internal to the body (host factors) and wider socioeconomic influences (Rappaport and Smith 2010, Wild 2012). Importantly, by redefining the exposome in terms of cumulative measure of environmental influences and biological response, Miller and Jones acknowledge that the entirety of an individual’s exposure does not need to be quantified, with exposure effect and maladaptation to external influences linking environment to health and disease. This concept emphasizes the need to identify biological markers, such as plasma levels of environmental chemicals, transformation products, DNA adducts, epigenetic alterations and protein measurements in exposome research that can be used as surrogate markers of environment, providing tangible measurements that can be used in disease prevention and management (Miller and Jones 2014).

As discussed above, the metabolic phenotype represents an integrated chemical profile derived from environmental and endogenous processes; individual measurements provide a snapshot measure of the human exposome. Periodic chemical profiling of readily accessible biological fluids such as urine or blood has the potential to provide quantitative measurements of exogenous chemicals and biological response. In addition to chemical exposure, metabolic phenotyping also provides the ability to link external events such as stress and geographic location to biochemical perturbations in the metabolic profile. Measurement of metabolic phenotypes in a population-wide manner will therefore enable a means of evaluating environment–human interaction on a wide scale,

providing the systematic information on exposures required for environment-wide association studies (EWASs) (Patel et al. 2010) to complement GWAS.

2.3. Population Screening

To make large-scale population screening possible, cost-effective and high-throughput platforms are required to provide quantitative measurement of the metabolome. Because of the complexity of the human metabolome, no single platform will provide universal sequencing of all chemical species present in humans (Rappaport et al. 2014); however, current instrumentation and workflows provide sufficient chemical coverage for population screening of metabolic phenotypes. Techniques using nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) have been used for characterizing biological samples obtained from thousands of individuals to identify metabolic changes associated with aging, diet and the environment (Holmes et al. 2008, Holmes et al. 2008, Menni et al. 2013, Dunn et al. 2015), in addition to being identified as key technologies in exposome research (Wild 2012, Wild et al. 2013, Athersuch and Keun 2015, Go et al. 2015). Because of variations in instrument configuration, data processing approaches, and spectral stringency requirements, chemical coverage by these platforms varies widely; however, improved sensitivity obtained by MS-based metabolomics has led to its growing popularity. Increased use of MS for metabolomics has largely been driven by advances in data analytic tools (e.g. Smith et al. (2006), Yu et al. (2013), Uppal et al. (2013), Libiseller et al. (2015)), dedicated instrumentation (Park et al. 2012) and high resolution, high

scanning speed MS (Makarov et al. 2006, Kanu et al. 2008, Cribbs et al. 2014, Frediani et al. 2014, Go et al. 2014), making possible routine measure of 15,000-20,000 unique features in a 10-minute analysis time (Go et al. 2014, Go et al. 2015, Uppal et al. 2015).

The main features of high-resolution metabolomics (HRM) are summarized in **Figure 2.2**. The use of ultra-high resolution capabilities of Fourier-transform instruments offers advantages in providing coverage of low abundance environmental, dietary, microbiome and drug-related metabolites. Comparison of MS platforms and additional analytical platforms are discussed elsewhere (Scalbert et al. 2009, Dieterle et al. 2011). A rigorous analytic workflow using dedicated instruments is most advantageous. Triplicate analysis of each sample enhances reliability of detection and improves quantification by using a standard error of the mean instead of individual value. Simple extraction (addition of solvent with internal standards followed by removal of protein precipitate) instead of solid phase extraction or drying and reconstituting decreases error due to variation in efficiency of recovery or variability in detection of internal standard. Profile mode instead of centroid mode preserves the ability to discriminate chemicals with very similar mass-to-charge ratio (m/z), and use of a data re-extraction routine with options to optimize parameter settings improves detection and quantitative accuracy (Uppal et al. 2013, Libiseller et al. 2015).

High-resolution metabolomics for personalized medicine

Salient features of high-resolution metabolomics enabling large-scale population screening for personalized medicine

- **Replicate Injections:** By increasing technical replicates to three or more, the number of ions that can be reliably quantified is increased and effect sizes are reduced. For example, an analyte present in only one sample out of 100 can be reliably measured if present in three replicates, even though it is absent from the preceding 99 samples.
- **Ultra high-resolution mass detection:** Mass spectrometers with sufficient mass resolution, sensitivity and dynamic range to detect high concentration endogenous metabolites and low abundance environmental chemicals in a single instrumental run (e.g. Go, Walker (20)).
- **Advanced data extraction:** Adaptive data extraction algorithms (e.g. Yu, Park (22)) provide detection of 10,000 to 20,000 unique chemical signals post-noise removal.
- **Throughput:** Increased mass spectrometer scan speed allow “fast” chromatography methods of 5 minutes or less. With dual column configurations, 24-hour, six days per week operation, capacity for 30,000 samples per year per instrument.
- **Cost:** Cost is primarily driven by instrument lease, maintenance and personnel. Lease/y is ~ \$175,000; Personal cost/y ~\$325,000; Supplies disposable and maintenance is ~\$150,000. Including overhead of 50%, cost for 30,000 samples would approach ~\$1million, or 0.00001% of estimated United States healthcare expenditure.

Figure 2.2. Advantages of high-resolution metabolomics for population screening and precision medicine research.

Because of the advances made in metabolomics platforms, it is now possible to implement metabolic phenotyping for screening of human populations. Using standardized sample preparation and instrumental configurations, endogenous and xenobiotic chemical screening for 50,000 to 100,000 biologic samples could be completed each year for approximately \$5 million. For initial large-scale metabolic phenotyping, it is important to recognize a data driven approach is required. The gross influence of environment and phenotypic character on the metabolome is poorly understood, and in many instances evidence-based associations will have to be uncovered. This will necessitate an initial exploratory phase to develop metabolic characteristics associated with an individual. Approaches in how to successfully apply a discovery approach for population screening can be learned from large-scale GWAS and genome-sequencing studies currently underway. For example, the International HapMap Project was designed to identify common genetic variants associated with disease and their response to drugs and environmental factors (International HapMap 2003). Following completion of phases I and II, it became evident the presence of common variants alone could not explain inheritable disease risk and additional sequencing of low frequency single nucleotide polymorphisms (SNPs) was required (International HapMap et al. 2010). This highlights the need to apply metabolic phenotyping approaches in an untargeted framework (**Figure 2.3**). Untargeted profiling allows maximum chemical coverage and reduces bias by not limiting analytical targets to known metabolites and pathways since analytes are not selected *a priori*, and detected chemical signals are characterized after

categorizing their importance. While initial validation requires confirmation with co-elution and ion dissociation studies, development of cumulative detected metabolic feature databases and certified, well-characterized reference material (i.e. NIST SRM 1950 (Phinney et al. 2013)) will provide direct identification and quantification based on instrumentally acquired information (mass, retention time and intensity). This approach, termed “reference standardization” was recently used by Go et al. (2015) to demonstrate an application in human populations.

An additional benefit of adapting untargeted metabolomics platforms to metabolic phenotyping is the ability to measure chemicals arising from environmental influences, providing quantitative information for exposome research. High-resolution, MS based metabolomics platforms are now capable of detecting chemicals and their metabolites present at low nM ranges from environment, diet and lifestyle (Soltow et al. 2013, Go et al. 2015). The ability to provide measures of a wide range of chemicals arising from external exposures in a systematic and quantitative manner greatly enhances exposure assessment, which traditionally relied on external monitoring, lifestyle factors, modeling, observational data and targeted biomonitoring data. The simultaneous measurement of endogenous metabolites enables quantitative evaluation of the biological response through dose–response associations. A conceptual illustration of the exposure- bioeffect framework possible by metabolomics-based population screening is shown in **Figure 1.2**. While not directly establishing cause-effect relationships with distal health outcomes, it can provide information on exposure risk and whether it has occurred at a biologically relevant level.

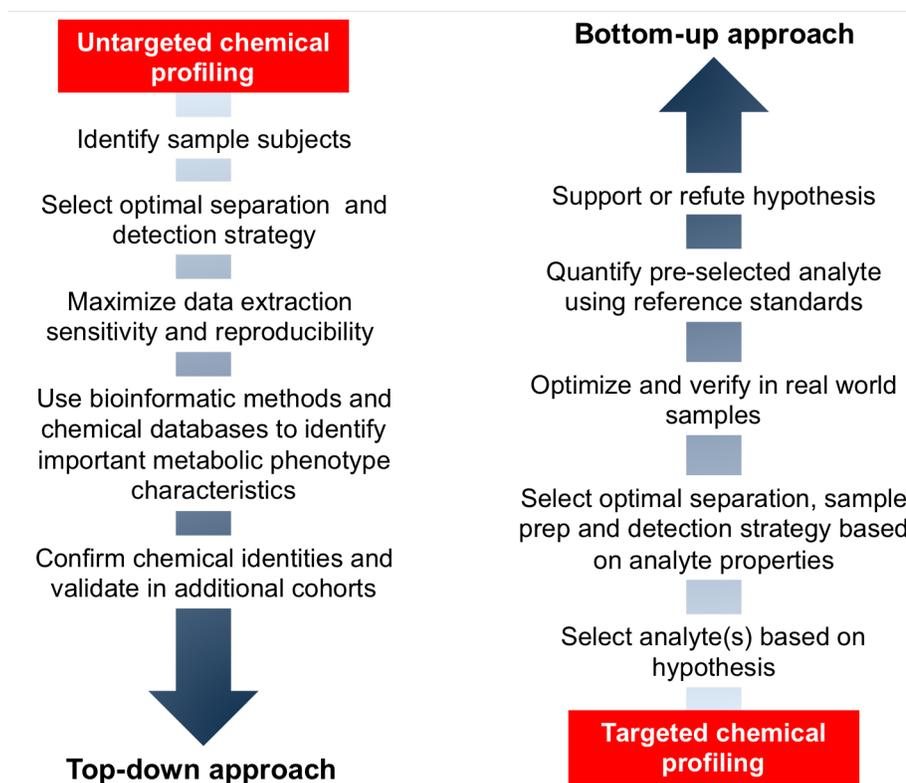


Figure 2.3. Comparison of untargeted and targeted chemical profiling techniques for population research. To derive the greatest benefit from population screening, untargeted chemical profiling is required. This is commonly referred to as a top-down analytical strategy, and is a data driven approach where the goal is to maximize detection of chemical species for biomarker discovery and metabolome wide association study. Candidate metabolic biomarkers are then confirmed and validated within complementary populations. Targeted approaches utilize pre-selected analytical targets for measurement within a group of subjects, and can be used to test hypothesis, estimate limited chemical exposures and test for specific biomarkers of disease.

Metabolic phenotyping using metabolomic platforms supports personalized medicine by providing comprehensive measure of tissue and/or biofluids chemistry for in-depth characterization of the metabolic profile. Translational applications include use as a diagnostic tool, identification of nutritional deficiencies, measurement of disease risk, tracking of therapeutic intervention efficacy, individualized exposure assessment, and metabolic health

forecasting. Adaptation into a health care paradigm and for public health surveillance could be accomplished by utilizing easily accessible biofluids such as blood or urine. Incorporating a metabolomic panel, i.e., “next generation” blood chemistry or urinalysis, into annual checkups will allow for periodic evaluation of endogenous, exogenous, and dietary health and would be complementary to additional omic measurements such as those obtained from the genome, epigenome, and proteome.

Before metabolic phenotyping can be applied in a personalized medicine context, a thorough understanding of how phenotypic characteristics, disease, and external factors influence the metabolome is required. Metabolome-wide association studies (MWAS) provide the framework for identifying specific metabolite associations with health endpoint or phenotypic traits (Holmes et al. 2008, Nicholson et al. 2008, Go et al. 2014). In MWAS, large-scale metabolic phenotyping is related to a single or series of variables to identify metabolite changes specific to the individuals expressing the variables of interest. MWAS can be employed for both discovery and diagnostic purposes and are analogous to GWAS, where genetic variants are tested as disease risk factors. Identification of metabolic phenotypes selected through MWAS provides insight into disease processes, dietary influences in the metabolome, exposure-related metabolic perturbations, and geospatial/cultural influences. Although these findings are initially correlative in nature, confirmation in validation cohorts and through elucidation of mechanistic processes is expected to identify metabolic indicators of health.

Recent studies have applied metabolomic profiling tools to evaluate how metabolic phenotype is influenced by numerous factors, including circadian rhythm, age, sex, disease, diet, climate and chemical exposures. To incorporate metabolomics into a personalized medicine framework, there is a need to understand how variation in the metabolic profile arises and is related to health and lifestyle factors. In the following section, I review how these factors contribute to expression of the metabolic phenotype and their relevance to population screening and MWAS. Emphasis is placed on use of metabolic phenotyping for exposome research and evaluating chemical exposure bioeffect in human populations.

2.3.1. Diurnal metabolic phenotype

Metabolism has long been known to exhibit diurnal cycles (Castro Cabezas et al. 2001, Levi and Schibler 2007). Time of day, amount of sleep, timing of meals and light/dark cycles are associated with metabolic fluctuations for control of energy expenditure and intake distribution during the waking period. Observational studies have indicated diurnal changes influence xenobiotic metabolism (Levi and Schibler 2007), while loss of circadian control has been associated with onset of cancer and diabetes, suggesting diurnal changes are an important measurement of overall metabolic health (Dallmann et al. 2012). Only a limited number of studies have applied metabolic phenotyping to evaluate diurnal changes in the metabolome. These have primarily examined the effect of fasting status, amount of sleep and the light/dark cycle in well-controlled environments.

Metabolites exhibiting circadian oscillations in humans are largely associated with gene expression under circadian control, such as fatty acid and lipid metabolism, in addition to cholesterol and bile acid biosynthesis (Park et al. 2009, Gooley and Chua 2014). For example, in the study by Gooley and Chua (2014), targeted profiling of 263 plasma lipids collected from 20 healthy individuals was completed over a span of 28 hours to assess inter-individual circadian variation. Rather than restricting their study to population-averaged diurnal changes, the authors examined individual changes in lipid levels over the sampling time course to evaluate how participant fluctuations differ from the group. Results showed lipids from all classes varied with the circadian pattern, and individual changes were greater than the group-rhythmic fluctuations; that is, the time of peak lipid levels varied greatly across the 20 individuals, even though the cohort was relatively homogeneous. Studies using metabolomic platforms aimed at metabolic intermediates have detected similar changes in lipid-related pathways (Dallmann et al. 2012, Gooley and Chua 2014) in addition to the amino acid glutamate, nutrients and hippurate (Park et al. 2009, Dallmann et al. 2012). Estimates from the previously discussed metabolomics studies suggest that up to 20% of metabolites could exhibit diurnal changes; however, the estimates were limited to characterized metabolites, and metabolic profiling was not completed in an untargeted manner.

Limited data exists on diurnal changes in the metabolome using untargeted approaches. In the study by Ang et al. (2012), untargeted metabolomics was used to test for diurnal variation in eight healthy adults based upon 1069 detected

features. Results were compared at selected time points, identifying 203 chemical signals exhibiting a significant time-of-day variation. Since data were collected within an untargeted framework, the authors used several techniques to characterize the metabolites; however, only 34 of the 203 metabolites could be positively identified, representing just 17% of the associated features. Similar to the previously discussed studies, the metabolites identified by Ang et al. (2012) included fatty acids and lipids, in addition to tyrosine, cortisol and methionine, and a large number of unidentified chemicals when searched against publicly available databases. Thus, diurnal variation within the metabolic phenotype has the potential to influence metabolite expression and MWAS results.

2.3.2. Gender metabolic phenotype

Gender is a strong determinant of metabolic phenotype due to fundamental biological differences occurring between the sexes (Kim et al. 2010) on both a genetic and metabolic level (Liu et al. 2012). Influences of environment, disease and treatment effectiveness also exhibit gender stratification, requiring consideration of sexual dimorphism during population screening. Recently, a number of statistically well-powered studies have applied metabolomics to highlight sexual dimorphisms in the metabolic phenotype. For example, in the study by Dunn et al. (2015), metabolic phenotyping of 1200 individuals from the United Kingdom was completed and evaluated for associations with gender, age, and health endpoints. Applying a two-way analysis of variance (ANOVA) using gender and age, the authors identified gender-dependent changes in both

endogenous and exogenous metabolites. The differentially expressed metabolites included metabolic intermediates previously associated with sex, including 4-hydroxyphenyllactic acid, creatinine, citrate, urate, and tyrosine, in addition to novel associations such as an increased level of methionine sulfoxide in females, which is a marker of oxidative stress. Caffeine was also increased in females and highlights the role of lifestyle contributions to the metabolic phenotype, since increased caffeine levels are mostly likely caused by a combination of different intakes and gender-specific metabolism.

Metabolic phenotyping in complementary populations has shown similar sexual dimorphisms, including differences in central metabolic intermediates. In the study by Mittelstrass et al. (2011), comparison of 131 serum metabolites measured in a population of 1649 males and 1732 females identified 103 metabolites exhibiting statistically significant differences between the sexes. Importantly, sex-associated metabolites were present in all detected metabolite classes, including acylcarnitines, amino acids, phosphatidylcholines, lysophosphatidylcholines, sphingomyelins, and hexoses. Even though this study was limited to common metabolic intermediates, broad differences were detected between the sexes. Application of untargeted approaches to evaluate gender dependent associations is expected to identify additional metabolites differentially expressed across the sexes. In the study by Krumsiek et al. (2015), untargeted profiling of serum obtained from 903 female and 853 male participants detected 180 metabolic features exhibiting significant differences between the sexes. The majority of these associations were identified endogenous metabolites, including

amino acids, vitamins/cofactors, carbohydrates, lipids, peptides, and nucleotides; however, 54 were unidentified metabolites, and 3 consisted of xenobiotics, which included 4-vinylphenol sulfate, piperine, and 2-hydroxyisobutyrate.

MWAS suggest that sex derived metabolic phenotype is not limited to differences in expected sex-dependent metabolites, such as steroid hormones, but also includes changes in core nutrition metabolism, chemicals from exogenous sources, and metabolic features with unknown function. In addition, the identified metabolic pathways influenced by gender are related to sex-specific disease susceptibilities, such as coronary heart disease and gout (Krumsiek et al. 2015).

2.3.3. Age-related metabolic phenotype

Over the lifespan of an individual, molecules, cells, and organs undergo a steady progression of damage occurring from the deterioration of homeostatic metabolic processes, resulting in loss of function, morbidity, and death (Fontana et al. 2010). Aging from these changes is considered the most universal contributor to metabolic decline and related diseases. Current estimates suggest only 25% of lifespan is attributed to heredity (Deelen et al. 2013), whereas lifestyle, environmental, and nutritional factors strongly influence the rate of aging. Estimating the contribution of age to the metabolic profile will inherently require deconvoluting the influence of diseases and the underlying processes contributing to disease onset. Confounding caused by the potential effect of age-related diseases and multimorbidity in populations representing young and old age groups is shown in **Figure 2.4**. As an advanced age is reached, the likelihood of

expressing clinical or preclinical manifestations of disease increases significantly (Barnett et al. 2012).

In addition to looking at metabolic phenotype differences based on age grouping, a number of studies have used age-related traits, including telomere length, forced expiratory volume, forced expiratory vital capacity, hip bone mineral density, blood pressure, cholesterol, and dehydroepiandrosterone sulfate to identify metabolic associations with physiologic age (Menni et al. 2013, Jove et al. 2015). For example, Menni et al. (2013) measured 280 metabolites in 6,605 individuals from the UK twin cohort study to identify metabolic correlations with age and clinical measurements known to be associated with aging. The authors identified 165 metabolites directly associated with age grouping when accounting for family relatedness, sex, and body mass index (BMI), suggesting that biochemical changes that occur during aging are expressed in the metabolic phenotype. Using the initially selected 165 metabolites, the authors were able to identify a subset of 22 measured chemicals that included endogenous and exogenous metabolites strongly correlating with age and with age-related clinical traits independently of age. The study by Menni et al. (2013) largely consisted of females (93% of individuals), but other studies have identified similar metabolic shifts with aging in men. In Dunn et al. (2015), two-way ANOVA of gender, age and the gender-age interaction was completed by comparing the metabolic phenotype of persons < 50 years old and > 64 years old. Age related changes were found to be both independent and dependent on gender, and similar to the results of Menni et al. (2013) included changes in tryptophan metabolism, serine, citrate

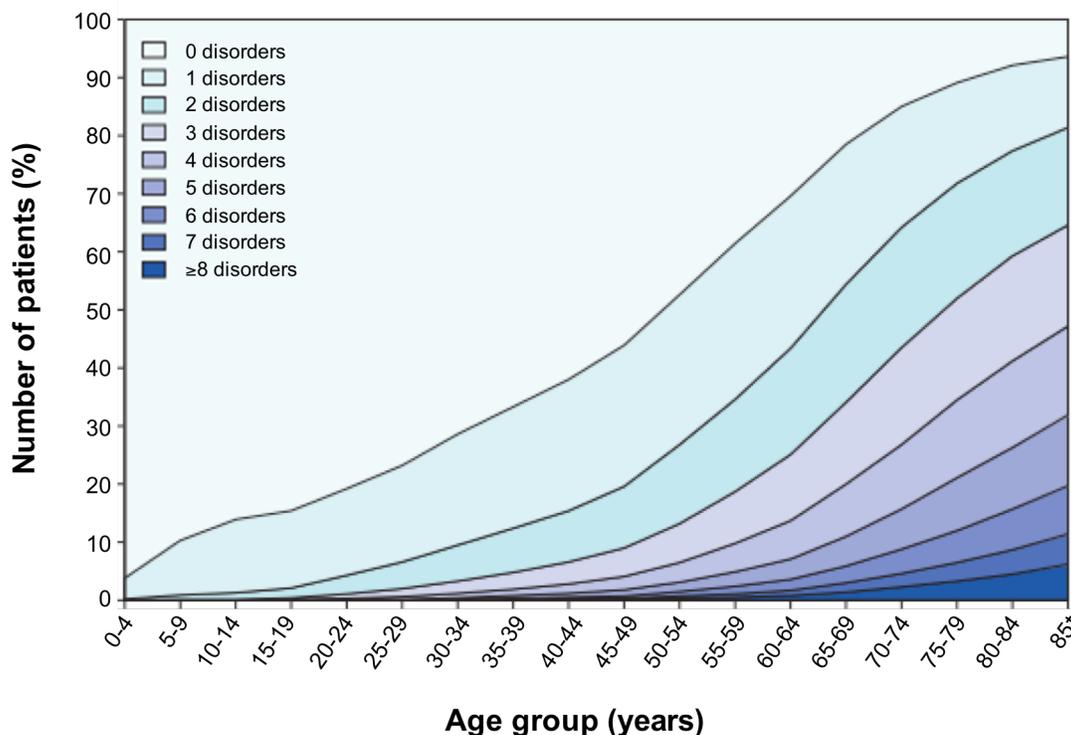


Figure 2.4. Prevalence of multi-morbidities in different age groups from database of 1,751,841 patients from Barnett, et al (2012). Used with permission from Elsevier.

and the food additive erythritol. Dunn et al. (2015) also detected significant changes in the metabolites known to be associated with aging. For example, cysteine levels were lower in the individuals >64 years old, which is consistent with evidence suggesting oxidative stress increases with age (Jones 2006).

Untargeted metabolomics have also been used to discover metabolic phenotype associations with physiological and chronological age. In the study by Zhao et al. (2014), metabolic profiles, measured by untargeted, high-resolution MS from 423 otherwise healthy Native Americans were tested for association with leukocyte telomere length, an indicator of biological aging. The authors were able to identify 19 metabolites that were associated with telomere length independent of chronologic age. The metabolites primarily consisted of lipids and

fatty acid products, including a positive association with anti-inflammatory fatty amides and glycerol phosphatidylethanolamines, suggesting a protective effect for telomere length. In this study, the authors limited the analysis to 1,364 metabolites that are also included in publicly available databases, representing only 18% of the chemical signals detected. Testing for associations with chronologic age in other studies has identified similar changes in metabolism, including fatty acid metabolites, phospholipids, and glycerides, in addition to unidentified chemicals providing no database matches (Jove et al. 2015).

In the previously described studies, both chronologic and physiologic aging were associated with changes in the metabolome. However, the natural aging process alone does not cause these changes, as disease status, in addition to the cumulative effects of environment, diet, and lifestyle factors, also contribute to expression of the metabolic phenotype over a lifetime. Observing changes in the metabolome of an individual over time has the potential support improved detection of disease risk and of poor health outcomes, providing increased opportunity for intervention and treatment.

2.3.4. Disease metabolic phenotype

The use of routine metabolic phenotyping in health care is expected to transform the understanding of disease, improve risk prediction, and increase efficacy of treatment. Therefore, population screening in healthy and diseased individuals has been one of the primary research focuses of metabolic phenotyping. For example, metabolic studies of neurodegenerative diseases

(Roede et al. 2013, Trushina and Mielke 2014), type II diabetes (Zhao et al. 2015), cancer (Kwon et al. 2015), HIV (Cribbs et al. 2014) and cardiovascular disease (Wurtz et al. 2015) have identified a number of metabolic markers related to disease progression, risk and the underlying derangement in metabolism characteristic of disease pathophysiology. The growing popularity of metabolomics in disease research and biomarker discovery can be demonstrated by searching PubMed for the terms “disease,” “metabolomics,” and “diagnostic,” with the number of publications in which these terms were used increasing from 3 in 2002 to 549 in 2014 (**Figure 2.5**). However, translating the findings from metabolomics to routine clinical use has proven challenging.

Metabolic markers have long been used to diagnose diseases and identify individuals requiring intervention. For example, elevated serum creatinine is used to evaluate and diagnose kidney disease, and hemoglobin A1C is an important diagnostic tool for metabolic syndrome and diabetes. Metabolic phenotyping will supplement traditional health indicators by enabling a more comprehensive measurement of blood chemistry. However, clinical diagnostic measurements often represent functional markers of disease pathophysiology, i.e. changes in physiologic function occur as a result of the disease itself. Due to the cost-effective nature of untargeted metabolomic approaches, it is possible to perform routine population screening for markers of metabolic changes that lead to eventual disease manifestation. By determining early biomarkers predictive of health outcome, lifestyle factors can be modified, progression can be more

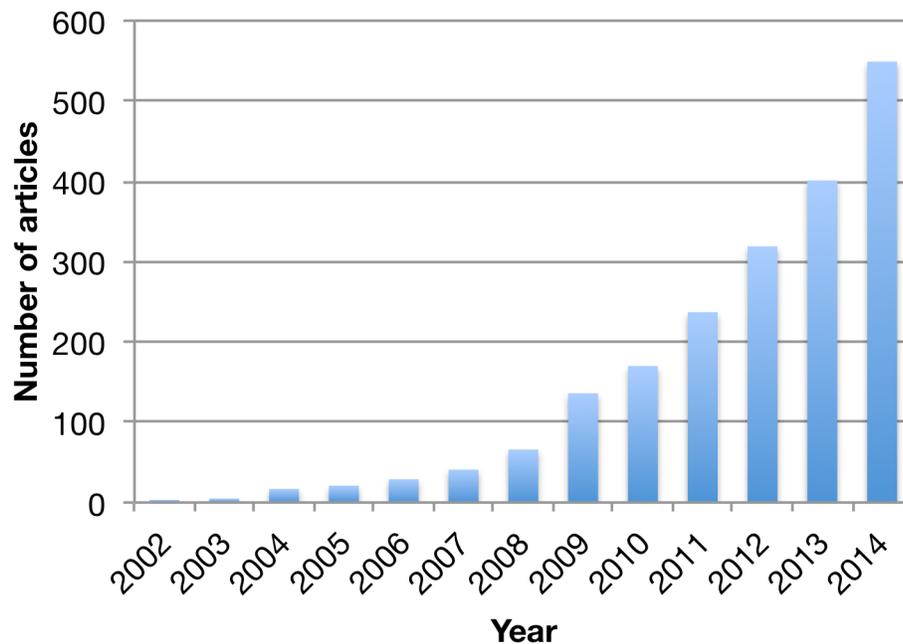


Figure 2.5. Number of articles listed in PubMed for years 2002-2014 after searching for the terms “disease”, “metabolomics” and “diagnostic”. The growing popularity of metabolomics for disease research is evident by the growing number of publications each year over that time span, which increased from three in 2002 to 549 in 2014.

accurately monitored, and intervention can be started for prevention rather than treatment.

In the application of population screening for identification of disease risk, one must acknowledge the heterogeneity in disease processes, which are combinations of genetic pre-disposition, environmental factors and lifestyle. Therefore, identification of predictive biomarkers will require an evidence-based approach, including large-scale population screening and animal models. The majority of studies completed to date are cross-sectional comparisons of healthy and diseased individuals, and only a limited number have applied metabolic phenotyping techniques to prospectively study metabolome associations with disease outcomes. Although cross-sectional studies provide insight into disease

mechanisms and possible therapeutic targets, there are complications due to treatment effects, reverse causality and an overall change in homeostasis due to altered metabolic functions. Prospective metabolic phenotyping studies overcome many of these limitations and improve ability to identify biomarkers predicting disease risk and onset. For example, in the prospective metabolomics study of cardiovascular disease (CVD) by Wurtz et al. (2015), NMR based chemical profiling was completed in a discovery population of 7,256 individuals (800 CV events during 15 yr follow up) and two different validation cohorts of 2,622 (573 CV events during 20-30 yr follow up) and 3,563 (368 CV events during 11-13 year follow up). Applying an MWAS framework to determine metabolic associations with CV event identified phenylalanine (Phe) and three fatty acids that acted as independent predictors of CVD onset 10-15 years later in both the discovery cohort and two validation cohorts. Dietary sources of Phe association were explored by evaluating the correlation with plasma aspartame; however, a significant correlation was not present, which suggested increased expression results from an underlying biological mechanism related to CVD. Alternative hypotheses can be inferred from an MWAS of Phe performed in marmosets (Go et al. 2014). In this study, Phe was strongly correlated with other amino acids, dietary chemicals and environmental chemicals containing an aromatic structure. Of interest, Phe also strongly associated with lipids, especially in females. The results of Wurtz et al. (2015) highlight the potential of population screening to identify novel disease biomarkers and use them in a predictive sense. Hazard ratio for future CVD were highest in the two youngest age groups (24-45 and 45-50),

which are considered at the lowest risk for CVD and have the most opportunity to decrease risk factors prior to a CV event.

Population screening using accessible biological fluids such as blood or urine has the capability to provide a wealth of chemical information that can be used to evaluate an individual's health status. Unlike genetic testing for disease associations, chemical measurement with broad coverage enables detection of endogenous metabolites and xenobiotics such as dietary chemicals, chemicals from environmental exposure, and pharmaceuticals. Incorporating this chemical information into understanding how a metabolic phenotype transitions from healthy to diseased for screening and for identifying the prognosis will be a major development in human health. Application of such an approach represents a paradigm shift in focus from disease treatment toward risk identification and prevention.

2.3.5. Dietary and nutritional metabolic phenotype

Metabolic profiling for nutritional and dietary metabolomics has received considerable attention recently (McNiven et al. 2011, Jones et al. 2012, Scalbert et al. 2014). The contributions of diet and nutrition to the metabolic phenotype are inherently complex because of external factors, including heterogeneity in intake, food preparation, diet composition, socioeconomic status, and availability, as well as host factors such as digestion, absorption and clearance. Furthermore, specific nutrients from dietary sources will impact hundreds of highly regulated molecular systems, resulting in a series of diet-related changes to the metabolic phenotype.

Therefore, diet is expected to contribute greatly to expression of the metabolic phenotype and heterogeneity present in humans.

When considered as a source of chemical exposures, diet includes nutrients, non-nutritive chemicals, pesticides, preservatives, additives, and others. An average individual consumes approximately 40 of the required nutrients daily. Nutritive dietary organic chemicals can then be converted through intermediary metabolism to more than 1500 chemicals (Jones et al. 2012), however, most chemicals in food are non-nutritive in nature. For example, biologists estimate upwards of 200,000 metabolites are present within the plant kingdom (Goodacre et al. 2004), which contribute significantly to phytochemical body burden; 372 different polyphenols have been identified in different food types that are consumed regularly (Neveu et al. 2010), and 26,619 chemical constituents have been curated from 907 different types of food (Centre 2015). Application of untargeted analytical techniques in food and agricultural sciences to characterize animal, plants and derived products has further identified a large number of chemicals relevant to dietary exposures (Ibanez et al. 2013). Combining information obtained from population screening with dietary exposures is expected to provide a more complete understanding of the role of non-nutritive dietary chemicals in health and enable unbiased nutritional assessment for an individual. Metabolic phenotyping can improve identification of individuals requiring dietary interventions, provide a more mechanistic understanding of biological response to foods (e.g., food allergies), and improve food preparation and handling to avoid short-term and long-term health risks.

Table 2.1. Representative dietary biomarkers for identifying food consumption and improving food recall in population studies from Scalbert et. al (2014). Used with permission from the American Society for Nutrition.

Food category and food	Metabolic phenotype biomarkers of dietary intake
Fruit	
Apple	Kaempferol, isorhamnetin, m-coumaric acid, phloretin
Orange	Caffeic acid, hesperetin, proline betaine
Grapefruit	Naringenin
Citrus	Ascorbic acid, b-cryptoxanthin, hesperetin, naringenin, proline betaine, vitamin A, zeaxanthin
Fruit (total)	4-O-Methylgallic acid, b-cryptoxanthin, carotenoids (mix), flavonoids (mix), gallic acid, hesperetin, isorhamnetin, kaempferol, lutein, lycopene, naringenin, phloretin, vitamin A, vitamin C, zeaxanthin
Vegetables	
Carrot	α -Carotene
Tomato	Carotenoids (mix), lycopene, lutein
Vegetables, leafy	Ascorbic acid, b-carotene, carotenoids (mix)
Vegetables, root	Ascorbic acid, a-carotene, b-carotene
Vegetables (total)	Ascorbic acid, a-carotene, b-carotene, b-cryptoxanthin, carotenoids (mix), enterolactone, lutein, lycopene
Fruit and vegetables (total)	a-Carotene, apigenin, ascorbic acid, b-carotene, b-cryptoxanthin, carotenoids (mix), eriodictyol, flavonoids (mix), hesperetin, hippuric acid, lutein, lycopene, naringenin, phloretin, phytoene, zeaxanthin
Cereal products	
Whole-grain rye	5-Heptadecylresorcinol, 5-pentacosylresorcinol, 5-tricosylresorcinol
Whole-grain wheat	5-Heneicosylresorcinol, 5-tricosylresorcinol, alkylresorcinols (mix)
Whole-grain cereals (total)	5-Heneicosylresorcinol, 3,5-dihydroxybenzoic acid, 3-(3,5-dihydroxyphenyl)-1-propanoic acid, 5-pentacosylresorcinol, 5-tricosylresorcinol, alkylresorcinols (mix)
Seeds	
Soy products	Daidzein, genistein, isoflavones (mix), O-desmethylangolensin
Meats	
Meat	1-Hydroxypyrene glucuronide, 1-methylhistidine
Meat, beef	Pentadecylic acid
Animal products (total)	1-Methylhistidine, 3-methylhistidine, margaric acid, pentadecylic acid, phytanic acid
Dairy products	
Milk, dairy products	Iodine, margaric acid, pentadecylic acid, phytanic acid
Fish	
Fatty	DHA, EPA, long-chain ω -3 PUFAs, PCB toxic equivalents, pentachlorodibenzofuran, PCB 126, PCB 153, ω -3 PUFAs
Lean	Long-chain ω -3 PUFAs
Beverage (non-alcoholic)	
Tea	4-O-Methylgallic acid, gallic acid, kaempferol
Coffee	Chlorogenic acid
Beverage (alcoholic)	
Wine	4-O-Methylgallic acid, caffeic acid, gallic acid, resveratrol metabolites
Alcoholic beverages (total)	5-Hydroxytryptophol/5-hydroxyindole-3-acetic acid, carbohydrate-deficient transferrin, ethyl glucuronide, g-glutamyltransferase, aspartate aminotransferase, alanine aminotransferase

Current application of metabolic phenotyping to nutritional research has included studies to correlate specific dietary biomarkers with food intake, with the primary goal of characterizing diet and assessing nutritional status in a way that provides a less biased estimate than can be made from patient recall.

Representative dietary biomarkers for different food classes are provided in **Table 2.1**. Although the results presented here are limited to previously published findings, it is expected that further characterization of the diet-metabolic phenotype interaction will identify additional biomarkers correlated with diet. For example, a number of studies have indicated that habitual diet (e.g. high-fat, high-protein vs. grain-based diets) influences the metabolic phenotype, resulting in specific compounds present in the metabolome that are capable of identifying primary food consumption. In the study by O'Sullivan et al. (2011), urine was used to assess dietary patterns in 125 subjects. Following a 3-day dietary assessment and categorization based on established food groups, urinary profiles were compared. The results showed that the differences in the metabolic phenotype within different dietary groups were driven by food and nutrient intakes, including protein, sugar, starch, salt, magnesium, and alcohol. O'Sullivan et al. (2011) used the metabolites differentiating the groups to identify healthy, unhealthy and traditional diets, which was based upon food contributing to total energy intake. The group consuming the unhealthiest diet, which included high intake of meat products, white bread, butter and preserves (and corresponding lowest intake of fruit, vegetables and whole-meal bread) had elevated trimethylamine N-oxide (TMAO), which has been associated with increased risk of CVD (Wang et al. 2011). The results from this study highlight the ability to assess diet-related contributions to the metabolic phenotype. Additional metabolomics studies have extended the relationship between diet and the metabolic phenotype to include dietary glycemic load (Barton et al. 2015),

phytochemical dietary load (Walsh et al. 2007) and Western vs. Eastern diets (Holmes et al. 2008).

Diet strongly contributes to metabolic phenotype, health and exposure burden. Therefore, the ability to provide a quantitative measure of dietary factors is an important component in population screening using untargeted metabolomic techniques. While specific biomarkers of food intake have been identified (Scalbert et al. 2014), more general diet-metabolome interactions are still poorly understood. Application of untargeted metabolomics approaches in population screening has the potential to greatly expand characterization of how dietary exposures are related to the metabolic phenotype. Understanding of diet-metabolome interaction will support identification of dietary risk factors in disease and allow interventions to reduce risk.

2.3.6. Geographic metabolic phenotype

Primary physical location will influence the metabolic phenotype by changing the composition of external factors contributing to the metabolome. For example, the type and the quantity of food intake are strongly influenced by culture and availability; climate, altitude, and length of light-dark cycles change total energy expenditure and metabolic rate; spatial-temporal variations in sources of environmental chemicals, chemical use, and availability of commercial products result in different exposures. Geographic variations among population groups linked to ethnicity and lifestyle are also closely associated with disease risk (Holmes et al. 2008).

Comparison of metabolic phenotypes from diverse locations has primarily identified diet-related chemicals as discriminatory metabolites. In the study by Holmes et al. (2008), 24-h urinary metabolic profiles of 832, 1,138, 496, and 2,164 samples were obtained from individuals in China, Japan, the United Kingdom, and the United States, respectively. Metabolomic profiles obtained from the four populations identified well-defined Eastern and Western metabolic phenotypes, with the geographic differences more pronounced than those caused by gender. Regional differences were also identified, including the presence of varied metabolic patterns in Northern and Southern China. Metabolites discriminating between the different locations predominantly included metabolites of dietary origin and were correlated with energy intake, micronutrients, BMI, alcohol consumption, and dietary cholesterol. Similar results were obtained in Walsh et al. (2014), which tested the influence of biofluid collection location on the plasma and urinary metabolome of 219 adults from seven European countries. Regional differences were detected in diet-related metabolites, including TMAO, creatinine, N-methylnicotinate, and hippurate, which was attributed to variations in meat, fish, and polyphenol-rich food consumption.

In addition to direct effects of geography, environmental conditions such as temperature and altitude, contribute to changes in metabolism in response to environmental stressors (O'Brien et al. 2015). Increases in altitude and decreased atmospheric pressure have been associated with changes in oxygen-dependent metabolism and have been studied extensively for more than a century. More

recent research has coupled metabolomic analyses of individuals in an induced high-altitude state to show changes of lactic and succinic acid with altered activity of hypoxia-inducible factor (Tissot van Patot et al. 2009). Environmental stresses and hot/cold response in extreme climates can exacerbate disease (McMichael et al. 2006), alter basal metabolic rate (Snodgrass et al. 2005) and induce heat dissipation processes. Although these are important considerations, the associated metabolic changes have not been well characterized in human populations (O'Brien et al. 2015). With recent advances in personal monitoring equipment, including digital sensors in smartphones and passive personal exposure monitors (O'Connell et al. 2014), it is now possible to track and collect a range of data on micro-environment and metabolic status. The combination of sensor data and metabolic phenotyping data will support improved identification of environmental stressors and health maintenance. Using currently available technology, heart rate, body temperature, cholesterol, glucose, blood pressure and steps taken can now be routinely measured in real-time. Combined with geographical information systems (GIS), weather and temperature information can provide daily, weekly, monthly and yearly-integrated measurements of environment available for integration with population screening data.

The varied nature of chemical exposures, spatial-temporal factors, disposition, changes in detoxification based on age, gender, weight, and health status result in a unique exposure history for each individual. Routine application of metabolic phenotyping in a population-wide framework will provide systematic information on exposures required for EWAS (Patel et al. 2010). In the

following sections, the use of metabolic phenotyping for assessing chemical exposures and contribution to disease burden is discussed. Strategies for incorporating chemical surveillance and bioeffect monitoring into a personalized medicine framework are also presented.

2.4. *Metabolic phenotyping for environmental exposures*

Contribution of low-dose, chronic environmental exposures to disease burden is poorly understood. To support detailed measurement of the environment for linking exposure events to health outcomes, analytical frameworks that provide quantitative assessment of chemical exposures in a high-throughput manner are required. Traditionally, the study of environmental risk factors in disease and health outcomes have relied on characteristics such as age, location, profession, and lifestyle risk factors determined through questionnaires to identify populations with a high probability of exposure. For example, residence near farmland, occupational use of solvents, and a self-reported smoking habit have been used to test associations between chemical exposure and health outcome (Brown et al. 2005). Although these approaches enable identification of at-risk populations on a large scale, considerable uncertainty exists in estimating exposure, making it challenging to link disease outcomes with exposure in a dose-dependent manner. To overcome this limitation, targeted molecular epidemiology approaches that use molecular measurements of environmental risk, including biomarkers of exposure, effect, and biological response, have been developed. Molecular measurement can be used to evaluate chemical internal dose,

biologically effective dose, early biological response and susceptibility, which can then be linked to disease or alternative biological endpoints. The use of biomonitoring, which estimates body burden of toxic chemicals through direct measurement in biological fluids, has especially contributed to understanding potential association between exposure and diseases, such as Parkinson's disease (PD), Alzheimer's disease (AD) and type II diabetes (Patel et al. 2010, Hatcher-Martin et al. 2012, Richardson et al. 2014). However, biomonitoring is typically completed using targeted approaches, which are time consuming and often limited to specific classes of chemicals (e.g. polychlorinated biphenyls (PCBs), brominated flame retardants), with targeted biomonitoring assays rarely exceeding 100 chemicals. Considering the large number of chemicals present in the environment, diet and in commercial products with potential for direct human exposure, it becomes apparent the coverage by targeted biomonitoring alone is not sufficient for EWAS.

Metabolic phenotyping using high-throughput, untargeted metabolomics has been identified as one of the most promising analytical frameworks for providing detailed estimates of the exposome (Park et al. 2012, Soltow et al. 2013, Vineis et al. 2013, Wild et al. 2013, Go et al. 2014, Miller and Jones 2014, Rappaport et al. 2014, Athersuch and Keun 2015). Because of the ability to detect chemicals from both endogenous and exogenous sources, a diverse series of exposure biomarkers, internal dose and the biological response are obtained; each is an important functional input for exposome research. While only limited examples of metabolic phenotyping for exposure surveillance and bioeffect

monitoring exist, application in animal models and human populations provide insight into how it can be incorporated into population screening and public health framework. In the study by Park et al. (2012), metabolic profiles obtained using untargeted metabolomics from seven mammalian species were compared to test whether humans would exhibit greater variation in plasma chemical content compared with animals housed in well-controlled research facilities. The resulting analysis identified 1485 metabolites present in all species, and an additional 2335 variably detected in the different species. To differentiate between endogenous metabolites and environmental chemicals, plasma metabolomic profiles were evaluated for significance of within-species and between-species correlation to test for metabolite variations across the species. The resulting analysis identified two separate modules discriminating chemicals with small and large variations among the species, with the first module enriched in endogenous metabolites (644), and the second module including environmental chemicals and endogenous metabolites (841) involved in xenobiotic metabolism. The results from Park et al. (2012) support the use of metabolic phenotyping for providing relevant biomarkers of exposure and response in population screening. Using an untargeted assay, the authors were able to simultaneously measure endogenous metabolism and environmental chemicals in a high-throughput manner, which can then be used for biomonitoring and identify biological response to support EWAS (**Figure 2.6**) results were obtained by Soltow et al. (2013), Go et al. (2014) and Go et al. (2014), supporting the use of metabolic phenotyping for environmental chemical biomonitoring and exposome research.

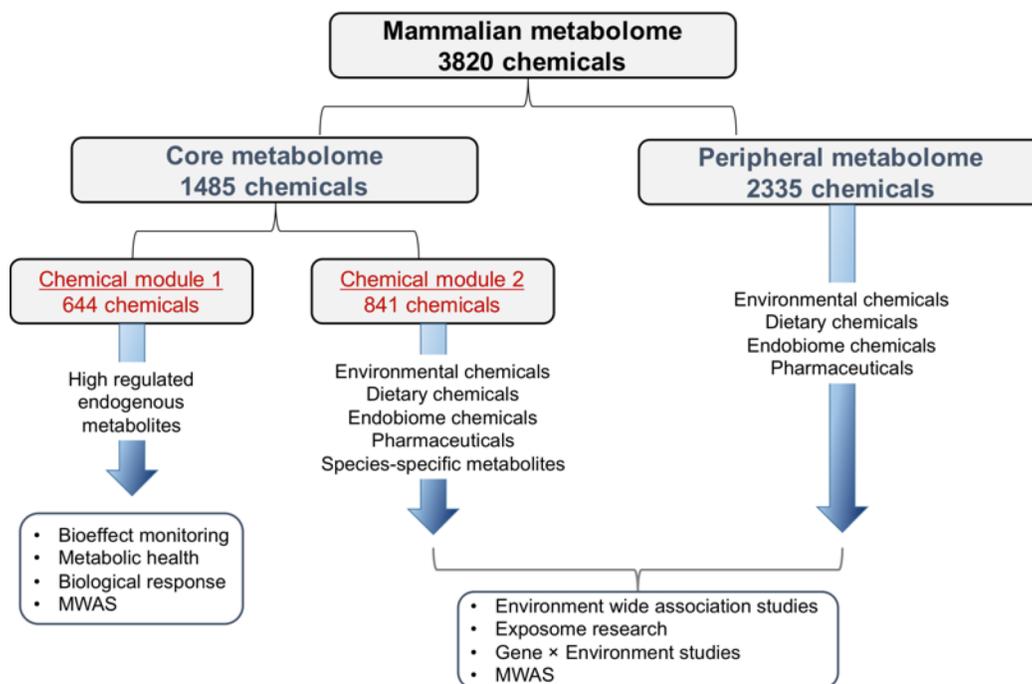


Figure 2.6. Categorization of metabolic phenotyping data obtained from seven mammalian species in Park et al. (2012) for bioeffect measurement and biomonitoring of environmental exposures. Comparison of metabolic profiles from the seven mammalian species identified a core metabolome consisting of chemicals common to the different species and a peripheral metabolome, which supported the core metabolome. Evaluating variation of metabolite levels and detection within the core metabolome identified two separate chemical modules. Chemical module 1 included endogenous metabolites having characteristics suitable for studies of biologic responses to toxic exposure, evaluating metabolic health and MWAS. Module 2 included environmental chemicals, variable endogenous metabolites, dietary chemicals, metabolites derived from the microbiome, and pharmaceuticals. Metabolites in module 2 and the peripheral metabolome can be used to support environment-wide association studies and gene-environment studies in exposome research. Used with permission by Elsevier.

2.4.1. Population screening for hazard identification

In addition to providing wider chemical coverage than available from targeted biomonitoring, population screening using untargeted HRM will assist in hazard identification for exposure to commercial chemicals. As discussed above, current estimates suggest that upward of 100,000 chemicals are used in industry,

commercial products, food preparation/storage, and pharmaceutical formulations. Although tremendous advances have been made in high-throughput toxicity screening (HTS) to assess chemical toxic effects (Tice et al. 2013), a large number in commercial use have not been thoroughly assessed for short and long-term toxicity to human populations. Furthermore, translating *in vitro* findings to quantitative risk assessment for human populations has proven challenging. For example, endpoints are typically assessed after exposure to a single chemical while not accounting for synergistic effects from mixtures; dose-response represents acute exposures and is not characteristic of the low dose, chronic exposure typically observed in human populations; it is not possible to measure maladaptation following a history of exposure leading to mutagenesis and cancer; *in utero* toxicity resulting in developmental outcomes cannot be determined; and chemical availability, disposition, biotransformation and contribution to health status cannot be evaluated *in vitro*. In many instances, animal model experiments are required to address the limitations of HTS and validate toxic mechanisms; however, it is not possible to complete these tests for all chemicals (and mixtures) in current use. Metabolic phenotyping can assist in hazard identification by surveying populations for the occurrence of single chemicals and mixtures to which humans are routinely exposed. Prevalence can then be used to prioritize for in-depth toxicity testing in model systems.

Although the use of untargeted approaches will increase the ability to detect known and unknown chemical exposures, it must be acknowledged that complete coverage of all chemical exposures will not be possible. Many

exogenous chemicals are rapidly transformed and excreted, and can only be detected immediately following exposure. Metabolomic profiling sensitivity, methodology, sample processing and biological fluid will also influence the ability to detect chemical occurrence. Uncertainty will be present when providing absolute concentrations of chemicals detected by untargeted metabolomics, which is influenced by the availability of authentic reference standards, use of surrogate standards and increasing analyte measurement number at the expense of sensitivity. However, the uncertainty in metabolic phenotyping is less than present in current methods used to estimate exposure in the general population for chemicals without detailed measurement (Wambaugh et al. 2013, Wambaugh et al. 2014). For example, in Wambaugh et al. (2014), chemicals with biomonitoring data available in the National Health and Nutrition Examination Survey were utilized to develop a model to predict exposure based on use, persistence, production, and physiochemical properties. Following parameter estimation using the training data set, exposure estimates for 7986 chemicals within the ACToR database was completed and then ranked on the basis of whether exposure occurs at concentrations consistent with the toxic response observed in HTS. The resulting confidence intervals for predicted concentrations were as high as four orders of magnitude, making accurate estimates in human population challenging.

To adequately provide chemical exposure surveys in humans by population screening, measured absolute concentrations are required. Therefore, strategies enabling quantification of untargeted metabolomic profiling for environmental chemicals will be needed to identify which exposure related

Table 2.2. Concentration of selected environmental chemicals in a cohort of 153 health individuals detected and quantified using high-resolution metabolomics and reference standardization (Go et al. 2015). Used with permission from Oxford University Press.

Metabolite	Source	Detected median concentration	Previously reported concentration ranges
Caffeine (μM)	Dietary	22.0	26-129 ^a
Chlorobenzoic acid (nM)	Insecticide	38.0	Not available
Chlorophenylacetic acid (nM)	Insecticide	28.0	Not available
Chlorsulfuron (nM)	Herbicide	0.6	Not available
Cotinine (nM)	Nicotine metabolite	3.9	6.7-13.5 ^a
Dibutylphthalate (nM)	Plasticizer	20.5	15-989 ^b
Dipropylphthalate (nM)	Plasticizer	57.0	Not available
Hippuric acid (μM)	Dietary	4.7	6-28 ^a
Octylphenol (nM)	Commercial products	34.0	0.14-2.2 ^c
Pirimicarb (nM)	Insecticide	0.8	Not available
Styrene (nM)	Industrial processes	5.3	0.4-5.3 ^a
Tetraethylene glycol (nM)	Industrial processes	2.9	Not available
Triethylphosphate (nM)	Plasticizer/Flame retardant	6.9	Not available
Triphenylphosphate (nM)	Plasticizer/Flame retardant	22.0	Not available
Tris(2-chloropropyl)phosphate (nM)	Flame retardant	45.0	Not available
Xylcarb (nM)	Insecticide	4.6	Not available

a. HMDB, Wishart et al. (2013)

b. Toxnet, <http://toxnet.nlm.nih.gov>,

c. Qin et al. (2013), and represent urinary measurements, expressed as nM/L urine

chemicals pose a risk to health. One approach could include the use of metabolic phenotyping to identify chemicals of interest, upon which a targeted analytical technique (e.g. Barr and Needham (2002)) using authentic chemical standards could then be used to re-analyze all samples. Targeted analytical techniques are labor and time intensive, making it challenging and costly to analyze sample sets comprised of 10^4 to 10^5 subjects. A second approach could include a series of analytical standards containing specified targets or use ^{13}C -labeled internal standards spiked into samples. While providing absolute concentrations, this method would result in loss of advantages gained using untargeted profiling, since chemical standards would have to be selected *a priori*. Recently, a quantification strategy relying on reference standardization was developed to provide absolute concentrations for a wide range of endogenous and exogenous chemicals detected

using untargeted high-resolution mass spectrometry (Go et al. 2015). Using reference standardization, quantification post-data acquisition is possible by referencing a pooled sample analyzed with each series of samples. The pooled reference, which is representative of the analytical samples, can be characterized and analytes quantified using traditional analytical chemistry techniques, such as quantification by methods of addition (e.g. Niessen et al. (2006)) or standardization of a certified reference material, such as NIST SRM 1950 (McGaw et al. 2010). Known concentrations within the reference standard can be used to determine a chemical response factor and calculate analytical sample concentrations based on single-point calibration. The benefit of this approach is that targeted quantification is only required in the reference sample, chemicals do not need to be selected *a priori* for quantification and population wide estimates of plasma chemical concentrations can be determined without having to re-analyze samples using a targeted approach.

By supporting quantitation of large numbers of chemicals detected in human samples, reference standardization can provide the systematic biological and environmental chemical measurements required for exposome research. For example, in the study completed by Go et al. (2015) plasma from a healthy population of 153 individuals was analyzed using untargeted metabolic profiling. Reference standardization was then applied to quantify a select number of metabolites and exogenous chemicals. **Table 2.2** contains a representative list of chemicals derived from exogenous sources that were detected, confirmed and quantitated within the plasma samples. A thorough search of the literature

indicated less than half had reported reference levels, with the remaining previously undetected in normal human populations. Similar results were obtained by Roca et al. (2014) and Jamin et al. (2014), where untargeted chemical measurements were able to detect both common exposure biomarkers and environmental chemical metabolites previously uncharacterized in human populations. Thus, analytical platforms used in metabolic phenotyping now enable a more comprehensive measurement of internal chemical exposure in humans than available from targeted biomonitoring techniques. The resulting survey of chemicals in human populations can be used to determine potential for exposure, identify chemicals to prioritize for toxicity screening and test for links between environment, the genome and health.

2.4.2. Metabolic phenotyping in environment wide association study of disease

Because of the ability to routinely measure a large number of environmental chemicals in the metabolic phenotype, EWAS can be completed with untargeted metabolomic profiling data to identify whether chemical exposures are related to health outcome and disease. This can be accomplished through analytical strategies seeking to link environment to alternative endpoints (e.g. biological response detected in metabolomics, complementary measured molecular markers) or identifying exposure-related associations with disease state. A number of strategies have been proposed for the use of metabolic phenotyping to identify environment-disease links (Vineis and Perera 2007, Chadeau-Hyam et al. 2011, Vineis et al. 2013, Athersuch and Keun 2015). For

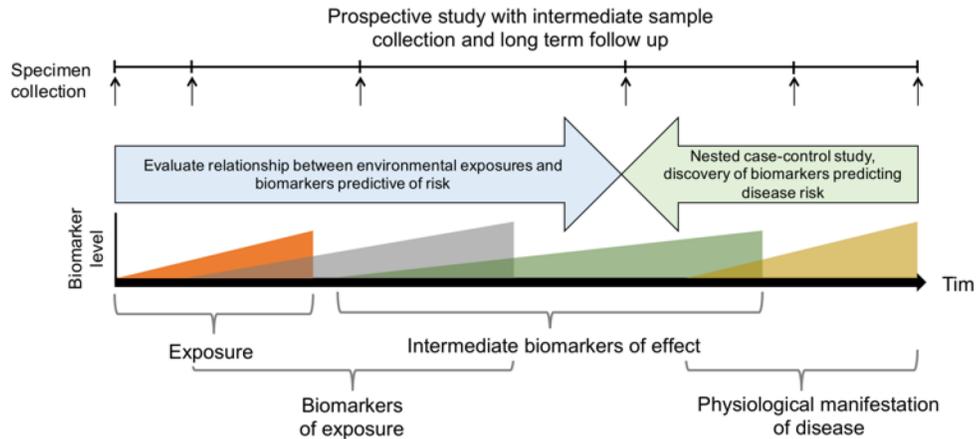


Figure 2.7. *Meet-in-the-middle* approach to discover causal relationships between environmental exposures and disease outcome (Vineis et al. 2013). Application of MITM to prospective studies can be used to link disease outcomes to intermediate biomarkers. In human populations, this is completed through three stages. The first requires epidemiological evidence linking exposure to disease outcome. Within a prospective study, this would require linking exposure and/or biomarkers of exposure to disease outcome through retrospective analysis of nested case-control studies. The second stage requires identifying associations between exposure and an intermediate biomarker. Intermediate biomarkers can be measured through metabolic phenotyping and MWAS, such as altered endogenous metabolite levels associated with exposure, or complementarily measured molecular markers, such as methylation levels or inflammatory response. The third stage requires identifying the relationship between intermediate biomarkers associated with environmental exposure and disease outcome. Within this framework, if association between disease and exposure biomarkers is present in all three stages, a casual relationship between exposure and disease outcome is reinforced. Metabolic phenotyping applied to prospective studies supports the MITM principle by enhancing coverage of endogenous metabolites and xenobiotics, allowing more in-depth testing for exposure and disease associations.

example, to avoid complication by factors related to reverse causality and identify exposures from the environment contributing to development of disease, a *meet-in-the-middle* (MITM) approach has been proposed (Vineis and Perera 2007, Chadeau-Hyam et al. 2011, Vineis et al. 2013). With application of the MITM study design, the causal relationship between disease outcome and environment is

assessed through a prospective search for intermediate biomarkers related to past exposure and associated with eventual disease development (**Figure 2.7**). In human populations, this is completed through three stages, which includes identification of 1) exposure and disease associations, 2) relationship between exposure biomarkers and intermediate biomarkers (such as alterations in the metabolic phenotype) and 3) relationship between intermediate biomarkers and disease outcome (Vineis and Perera 2007, Vineis et al. 2013). If associations between disease and exposure biomarkers are identified in all three stages, it reinforces the casual nature of the association. Measurement used in the MITM framework can also be determined through animal exposure and disease models, which assists in validating biological plausibility.

To provide a proof-of-concept implementation of the MITM approach, Chadeau-Hyam et al. (2011) used metabolic phenotyping, dietary and lifestyle information to test for environment-related biomarkers of colon and breast cancer outcomes in a nested case-control study design. Using metabolomics data, the authors first tested for metabolic features distinguishing patients with breast cancer and those with colon cancer from healthy controls, followed by testing for dietary and lifestyle exposure associations. Although breast cancer exhibited no discernable differences between healthy and diseased individuals, colon cancer resulted in eight chemical regions associated with outcome. Comparison of chemical regions significant in colon cancer to those exhibiting associations with dietary exposure identified well-recognized risk factors (smoking) and protective factors (fiber intake), supporting the use of MITM for identifying putative

markers linked to both exposure and disease outcomes. Metabolic phenotype markers determined through application of MITM and validated in model systems represent prime candidates for incorporation into population-based environmental epidemiologic studies and for use to identify disease risk factors.

In addition to measurement for chemical surveillance and disease-environment biomarker identification, metabolic phenotyping can be applied to elucidate biological response to chemical exposure. This provides insight into toxicant targets, mechanisms, whether exposure occurs at a biologically relevant dose and chemical biological fate, including metabolism, distribution and excretion. In the study by Jeanneret et al. (2014), metabolic profiling of an individual who received an extremely high dose of dioxin was completed to identify changes in endogenous metabolites correlated with exposure. The authors identified a series of steroid-related metabolites associated with high dioxin exposure, which were then tested as discriminatory metabolic phenotype biomarkers in 11 workers exposed to dioxin residues in a pesticide production plant and matched, healthy controls. A series of 24 metabolites distinguished exposed workers from controls, which indicated changes in expression of cytochrome P450s, hepatotoxicity, bile acid biosynthesis and oxidative stress; these results are consistent with the known aryl hydrocarbon receptor (AhR) binding of dioxins. Importantly, the data showed that changes in the metabolome occurring due to a recent, acute dioxin exposure were still detected in the metabolic phenotype of individuals exposed 40 years prior. Thus, in addition to

biomonitoring, population screening by untargeted metabolomics provides the capability to identify biological markers indicating a history of exposure.

Metabolic phenotyping in tandem with traditional exposure assessment, such as personal monitoring or targeted biomonitoring for exposure biomarkers, can also be used to identify biological response to environmental stressors. Wang et al. (2015) used targeted measures of urinary polycyclic aromatic hydrocarbons (PAH) metabolites to stratify individuals exposed to low- and high-levels of air pollution from industrial coking operations. The detected PAH biomarkers provided a means to evaluate dose-dependent changes in the urinary metabolome based on internal PAH exposure, which indicated association with muscular breakdown products, increased lipid peroxidation and depletion of antioxidants. The presence of metabolites related to oxidative stress is consistent with PAH related diseases, supporting the ability to detect early metabolic events in the exposure-disease continuum. By using individual biomarkers of exposure, the authors were also able to identify biological associations changing in a dose dependent manner, which avoids confounding due to unknown/uncharacterized exposures and provides testable hypothesis for mechanistic study in experimental systems.

In addition to the two studies discussed above, metabolic profiling of human exposure to cadmium (Ellis et al. 2012, Gao et al. 2014), pesticides (Bonvallot et al. 2013), welding fumes (Wei et al. 2013) and arsenic (Zhang et al. 2014) support the use of metabolomic profiling for exposure related biological changes, measurement of alternative toxicant endpoints and insight into toxic

mechanisms in human populations. Metabolic phenotyping can also be applied in an environment wide association study framework to systematically examine chemical associations with disease. For example, in a HRM study of PD progression (Roede et al. 2013), groups of subjects (80 from 146 PD and 20 from controls) were matched by age, gender, smoking status and pesticide exposure to identify metabolic phenotype related to PD and PD progression (slow vs. rapid). Significant differences between 80 patients and 20 controls included matches to a polybrominated diphenyl ether (PBDE), tetrabromobisphenol A, octachlorostyrene and pentachloroethane. The chemical feature corresponding to PBDE had a mean intensity 50% above controls, while the match to 2-amino-1,2-bis(p-chlorophenyl)ethanol, was more than 50% higher in individuals with rapid-disease progression. Metabolic phenotyping of other aging-associated diseases have identified positive relationships between chemical exposures and disease. In a pilot study of neovascular age-related macular degeneration (NVAMD), MWAS of 26 patients and 19 controls selected 94 metabolic features associated with NVAMD (Osborn et al. 2013). The identified features included a match to β -2,3,4,5,6-pentachlorocyclohexanol (β -PCCH), a hydroxylated metabolite of β -lindane originating from commercial grade insecticide formulations. To evaluate if the identity of β -PCCH was correct, the intensity was correlated against the metabolome to test for associations with additional exposure related chemicals. Both the ^{37}Cl form of β -PCCH and an ambiguous match to 7-hydroxy-1,2,3,6,8-pentachlorodibenzofuran and 1,2,3,7,8-pentachlorodibenzodioxin were found to correlate with β -PCCH, suggesting a possible association for chemical exposures

in NVAMD. Although the populations from these two studies are too small to make firm conclusions and the identification of the chemicals associated with PD and NVAMD were not confirmed by comparison to authentic reference standards, the presence of disease-exposure associations provide possible targets for follow-up to determine if low-level environmental chemicals could be causally related to these diseases. These two studies indicate the potential of metabolomic profiling to identify the role of exogenous chemicals in disease.

2.4.3. Chemical exposure measurement in personalized medicine

With the exception of occupational medicine, chronic environmental exposures are often not considered in the prevention and diagnosis of diseases. In many instances, chemical exposures represent a modifiable disease risk, and primary prevention can be achieved by reducing exposure risk in an individual. Before this can be applied in a personalized medicine context, a framework for providing exposure assessment is required, in addition to evidence supporting a causal relationship between exposure and health outcome. Metabolic phenotyping by untargeted metabolomic profiling meets the former requirement by providing a platform for universal chemical surveillance and bioeffect monitoring in human populations. Routine application is now possible at reasonable costs, making large-scale use of population screening for chemical surveillance feasible. This will enable a more thorough accounting of chemical exposures for an individual, but the relationship to diseases, particularly those with long latency periods, is currently unknown. This does not limit future use of metabolic phenotyping to

predict disease risk, but it must be acknowledged that a discovery period will be required prior to deriving the full benefit of applying population screening techniques for chemical surveillance. Once the link between disease and environment is more fully elucidated, the advanced clinical and environmental measurements provided by untargeted metabolic profiling techniques, in addition to information provided by the genome (Schwartz et al. 2004, Sun 2014), adductome (Rappaport et al. 2012) and proteome (Go and Jones 2014) will improve the ability of healthcare practitioners to include exposome-based measurements health outcome predictions and incorporate phenotypic information into designing preventative measures and treatment regimens for chronic diseases.

2.5. Conclusion

Sequencing of the human genome has failed to identify a strong genetic basis for many chronic diseases, with a combination of genetic, environment and lifestyle factors suspected to contribute to disease onset and progression. Unlike currently available techniques enabling population screening for genetic risk factors of disease, no platform currently exists permitting comprehensive measurement of environment and lifestyle factors in humans. To adequately incorporate environment into a personalized medicine framework, there is a need to develop technologies that can provide quantitative measures of the influence of external factors on health. Untargeted measurement of metabolic phenotype using high-throughput methods provides a platform for universal chemical surveillance platform and bioeffect monitoring, which vastly improves the chemical coverage

available from targeted analytical techniques. Application of untargeted metabolic phenotyping in human populations enables a more thorough accounting of dietary, environmental and pharmaceutical chemicals in biological fluids, in addition to providing advanced measurement of endogenous metabolism. When used as a population-screening tool and in predictive health, metabolic phenotyping is expected to provide considerable insight into molecular markers of health and disease risk. Phenotypic traits, including time-of-day, sex, age and disease status are associated with a number of changes to metabolite levels, and will be important considerations when drawing conclusions for personalized medicine. External factors, such as dietary patterns, external environment, location and chemical exposures are also strong contributors to the presence and distribution of metabolites. Providing an objective measurement of both phenotypic and environment factors contribute to the usefulness of metabolic phenotyping in support of personalized medicine. The ability to provide sensitive measurements capable of detecting small changes in individuals chemical profile will have important implications for disease prediction, tracking and preventative medicine. Therefore, the adoption and application of metabolic phenotyping in a population-screening framework is expected to enhance personalized health forecasting and provide valuable insight into the processes driving the development of diseases related to lifestyle and environmental factors.

Chapter 3. RESEARCH METHODOLOGY

3.1. Study selection

To support the objectives of this thesis, blood samples previously collected as part of five independent, federally funded environmental health studies were profiled using HRM (**Table 3.1**). Project summaries are provided below. Studies were selected for HRM profiling based upon availability of exposure data, alternative endpoint measurements, study design, exposure duration and route. Initial efforts were focused on testing the quality of biorepository samples for HRM analysis and testing exposure metabolic response (**Chapter 4**). A critical evaluation of HRM as a platform for exposure science and public health use is the ability to measure relevant changes associated with a range of exposure periods. Additional populations were selected to be representative of different exposure timescales. In **Chapter 5**, exposure was acute, with monitoring completed over an 8-hour shift period and followed by blood collection immediately at shift end. Exposure assessment for **Chapter 6** included ambient monitoring of diesel exhaust during days 1 and 5, with biological specimens collected pre- and post-shift on both days. The study selected for **Chapter 7** has measures on personalized, time- and activity-adjusted exposure predictions for a one-year period that is based upon residence, work and

Table 3.1. Summary of environmental health studies selected for HRM profiling

Chapter	Project title	Project number	Study size	PI
4	Feasibility study for post-deployment exposure surveillance system	306889-1.00-64239	30	Dr. Timothy Mallon, DOD
5	Molecular epidemiology studies of occupational and environmental exposures	Z01-CP010121	175	Dr. Nathaniel Rothman, NCI
6	Biomarkers of exposure and effect in a traffic exposed population	R01-ES016284	95	Dr. Francine Laden, Harvard University
7	Community assessment of freeway pollution and health	R01-ES015462	59	Dr. Doug Brugge, Tufts University
8	Michigan PBB cohort 30 years later: Endocrine disruption?	R01-ES012014	177	Dr. Michele Marcus, Emory University

school location. The population selected from the Michigan PBB Registry (**Chapter 8**) is representative of long-term exposure, with plasma levels of lipophilic PBBs available in a multi-generational cohort from exposures that occurred during the 1970s. For each of these studies, independent measurements for a variety of biomarkers are available, allowing additional assessment of alterations in metabolic phenotype. A short description of each study is provided below. The specific demographics, details on exposure measures, additional biological response markers and statistical analysis for MWAS details are provided in each chapter.

3.1.1. Description of study cohorts

3.1.1.1. Feasibility study for post-deployment exposure surveillance system

In 2011, the United States Institute of Medicine (IOM) recommended that the Department of Defense (DoD) collect individual breathing zone samples and conduct long-term studies of troop health outcomes to address concerns about perceived health risks from exposure during deployment (Tollerud et al. 2011,

Baird 2012, Lindler 2015). Realistically, there are inherent limits to exposure assessment in deployed settings. For example, the use of personal monitoring equipment limits mobility in active combat situations, logistics of sampler collection is challenging with large-scale troop movements, and assessment for biologically relevant dose requires additional molecular measurements. Furthermore, the post-exposure window of opportunity for measuring exposures or immediate consequences may range from hours to days for some agents. Therefore, valid and reliable measures are needed to characterize exposures that do not disrupt effective operation during deployment. Retrospective profiling of biological specimens collected pre- and post-deployment for biomarkers of exposure, effect, relevant dose and susceptibility provide a means of assessing the occurrence of chemical exposure related to poor health outcomes. Through the DoD Serum Repository (DoDSR), an extensive system exists for collection, cataloguing and storing of serum samples collected pre- and post-deployment from active duty armed forces personnel (Rubertone and Brundage 2002, Perdue et al. 2015). To establish the feasibility of using DoDSR samples for identifying the occurrence of exposures during deployment, a multi-university team was assembled to measure of different exposure biomarkers, markers of biological response and microRNA (miRNA) expression in serum samples obtained from the repository. This was completed in four phases: 1) Pilot study testing to establish the quality of samples for chemical characterization using unidentifiable samples; 2) Profiling of 400 serum samples which included pre- and post-deployment samples from troops who deployed to Iraq and Afghanistan and 400

controls matched on time in service; 3) Linking measures completed in the 200 deployed and 200 control individuals to demographics, health outcomes and additional measures; 4) Validation of findings in liver and cell lines to establish biomarker feasibility. Additional information about the study design can be found in Mallon et al. (2016).

3.1.1.2. Molecular epidemiology studies of occupational and environmental exposures

Ongoing and planned studies include evaluation of populations exposed to benzene, chlorinated solvents, organochlorines, disinfection byproducts, polycyclic aromatic hydrocarbons, arsenic, formaldehyde, diesel exhaust, coal combustion products, and selected pesticides. On-going case-control studies of stomach, esophagus, brain, bladder, renal cell, thyroid, prostate and breast cancer, non-Hodgkin's lymphoma (NHL), and benzene-induced hematotoxicity and hematologic malignancies and cohort studies of women in China and agricultural workers in the United States are evaluating a range of potential genetic risk factors and their interaction with occupational and environmental exposures. Many of these studies are also collecting tumor samples for future molecular analyses. (Adapted from <https://projectreporter.nih.gov/>, accessed February 2, 2015)

3.1.1.3. Biomarkers of exposure and effect in a traffic exposed population

Particulate matter (PM) attributable to traffic and other combustion sources have been associated with cardiac-related health outcomes, respiratory disease, lung cancer, and overall mortality. Experimental evidence in animals and humans suggests that exposure to traffic-related PM results in systemic inflammation, oxidative stress and cell cycle dysregulation, as well as direct genotoxicity. However, the precise mechanisms are poorly defined and have not been well characterized in large "real-world" populations. This research project was developed to prospectively characterize personal exposures to vehicle exhaust while concurrently collecting repeated blood and urine samples in a subpopulation of the Trucking Industry Particle Study, an ongoing national epidemiologic and exposure assessment study. Relationships between measured exposures and biomarkers of exposure and oxidative and inflammatory effects were assessed. Specifically, serum levels of 1, 2 naphthoquinone and 1, 4-naphthoquinone albumin adducts and urinary naphthols (biomarkers of traffic-related exposure) and serum levels of C-reactive protein, interleukin-6 and soluble intercellular adhesion molecule-1 and urine levels of 8-hydroxy-2'-deoxyguanosine (biomarkers of effect) were measured. External exposure measurements include elemental carbon, naphthalene, and bound polycyclic aromatic hydrocarbons (PAH), a novel specific indicator of PAHs attached to elemental carbon particles that can concentrate and transport them into the target areas of the lungs and internal organs. In addition, a genome-wide gene expression micro array analysis using whole blood RNA was completed. (Adapted from <https://projectreporter.nih.gov/>, accessed February 2, 2015)

3.1.1.4. Community assessment of freeway pollution and health

This project was developed as a community-based participatory research study of the relationship between air pollution gradients and health effects in individuals living next to major highways. The core study involves measuring 5 traffic-related pollutants (i.e., UFP, PM_{2.5}, NO_x, CO, black carbon, particulate PAH) in 6 neighborhoods within 400 meters of highways in the Boston area. A background site >1000m from highways will also be monitored. Time spent within near highway zones was evaluated, with rigorously measured highway pollution gradients in the neighborhoods. The investigators documented exposures at work, school and while commuting. For a subset of non-smoking households, blood samples were collected and analyzed for CRP and fibrinogen. This study will be (1) the first to test associations between highway pollution gradients and biological markers of health, (2) the first CBPR study of highway pollution, and (3) the most comprehensive collection of data on time spent in the exposure zone and many confounders and effect modifiers. Bivariate and regression analyses will be completed, as well as mathematical models analyze the large set of data. This proposal was developed to actively involve communities living very near to major highways in scientific research on the health effects of pollutant gradients coming off the highways (to about 200 meters). The study emphasizes the role of ultrafine particles on development of asthma in children, reduction of lung function in non-asthmatic children and increases in markers of inflammation and

heart disease risk (C-reactive protein) in adults. (Adapted from <https://projectreporter.nih.gov/>, accessed February 4, 2015)

3.1.1.5. Michigan PBB cohort 30 years later: Endocrine disruption?

This project includes a targeted investigation of endocrine-related outcomes in individuals in the Michigan PBB Registry and their offspring to follow up intriguing findings of a study recently completed. In 1973, inadvertent substitution of a livestock feed supplement with fire retardant led to widespread contamination of meat and dairy products in Michigan with polybrominated biphenyls (PBBs), a class of chemicals toxicologically similar to PCBs, PBDE, DDE and other halogenated organics suspected to disrupt endocrine function. Over 4,000 individuals with high likelihood of exposure were subsequently enrolled into the Michigan PBB Registry, and serum samples were analyzed for PBB. The current project will pursue key findings as follows. The proposed investigations allows targeted, efficient follow up of the important findings observed in the previous study, and will provide further insight into possible impacts of this class of chemicals on pubertal development, reproductive health and ovarian function. (Adapted from <https://projectreporter.nih.gov/>, accessed February 4, 2015)

3.2. *High-resolution metabolomics methodology*

Profiling of all samples was completed using the HRM methodology developed in the Clinical Biomarkers Laboratory at Emory University. This platform, which is

based upon liquid chromatography with high-resolution mass spectrometry (LC-HRMS), has demonstrated excellent performance in metabolic phenotyping for research on nutritional metabolomics, the exposome, animal and in-vitro exposure models and human health and disease (Johnson et al. 2008, Johnson et al. 2010, Jones et al. 2012, Park et al. 2012, Roede et al. 2012, Li et al. 2013, Roede et al. 2013, Soltow et al. 2013, Yu et al. 2013, Cribbs et al. 2014, Fitzpatrick et al. 2014, Go et al. 2014, Go et al. 2014, Go et al. 2014, Go et al. 2014, Lin et al. 2014, Neujahr et al. 2014, Patel et al. 2014, Roback et al. 2014, Yu and Jones 2014, Zhao et al. 2014, Burgess et al. 2015, Uppal et al. 2015, Zhao et al. 2015, Jones 2016). In addition, development of HRM in the Clinical Biomarker Laboratory has resulted in innovations vastly improving chemical detection and data analysis when compared to contemporary metabolomic platforms, including development of adaptive data extraction and alignment algorithms for improved detection of low abundance chemical signals (Yu et al. 2009, Yu et al. 2013, Yu and Jones 2014), software routines for improved feature reliability and detection (Uppal et al. 2013), high-throughput sample preparation and instrumental analysis approaches providing triplicate injections, complementary chromatography stationary phases and ionization polarity (Johnson et al. 2010, Soltow et al. 2013, Liu et al. 2016), implementation of metabolic and rule based annotation strategies for improved matching confidence and bioeffect assessment (Li et al. 2013, Uppal et al. 2016), reference standardization and related strategies for quantitation of untargeted HRM data (Go et al. 2015, Accardi et al. 2016), network based approaches for simplifying multi-metabolite MWAS (Uppal et al. 2015) and

algorithms for integration of complex datasets (Roede et al. 2014, Cribbs et al. 2016). Importantly, the HRM framework in the Clinical Biomarkers Laboratory has been optimized for reliable detection of low abundance environmental chemicals while also providing suitable measure of high abundance metabolic intermediates, resulting in routine detection of 10,000-75,000 unique mass spectral features. Thus, in combination with the advantages described in **Figure 2.2**, use of this platform provides the optimal HRM strategy to evaluate feasibility as an exposure science and public health tool. Description of HRM methodology, instrument analysis, data extraction, alignment and reduction are provided below.

3.2.1. Plasma preparation

Plasma or serum samples were prepared and analyzed daily in batches of twenty. Preparation of samples for HRM is based on the approach described by Park et al. (2012), which includes protein precipitation and extraction of metabolites with acetonitrile. Numerous sample preparation schemes have been developed for untargeted metabolomics, including solid phase extraction, ultrafiltration and solvent precipitation/extraction with various solvent/denaturant combinations (Vuckovic 2012). While there are benefits and disadvantages for each, this workflow has been developed to maximize throughput, chemical space coverage and minimize introduction of preparation artifacts while maintaining a sample extract consistent with the liquid-chromatography mobile phase.

Prior to analysis, sample aliquots were removed from storage at -80°C and thawed on ice. Each cryotube was then vortexed briefly to ensure homogeneity,

and 65 μL was transferred to a clean microfuge tube. Immediately after, each aliquot was treated with 130 μL of ice-cold LC-MS grade acetonitrile (Sigma Aldrich) containing 3.25 μL of internal standard solution with eight stable isotopic chemicals selected to cover a range of chemical properties. Internal standards included [$^{13}\text{C}_6$]-D- glucose, [^{15}N]-indole, [2- ^{15}N]-L-lysine dihydrochloride, [$^{13}\text{C}_5$]-L-glutamic acid, [^{15}N]-L-tyrosine, [trimethyl- $^{13}\text{C}_3$]-caffeine, [$^{15}\text{N}_2$]-uracil, [3,3- $^{13}\text{C}_2$]-cystine, and [^{15}N , $^{13}\text{C}_5$]-L-methionine and were prepared as described by Soltow et al. (2013). Following addition of acetonitrile, the sample was equilibrated for 30 min on ice; upon which precipitated proteins were removed by centrifuge (16.1 $\times g$ at 4°C for 10 min). The resulting supernatant (130 μL) was removed, added to a low volume autosampler vial and maintained at 4°C until analysis (<22 h).

A quality control (QC) pooled reference sample and the NIST SRM 1950 certified plasma reference standard (Simon-Manso et al. 2013) were included at the beginning and end of the analytical batch for reference standardization and quality control. Analytical reproducibility was confirmed via the stable isotope internal standards added prior to sample extraction. Only features exhibiting a median coefficient of variation (CV) $\leq 100\%$ were used in data analysis.

3.2.2. Liquid chromatography and high-resolution mass spectrometry

Due to advances in chromatography and instrument technology that occurred during the completion of this dissertation, different platform configurations were used to analyze the study samples. While these instruments were based on the

same Orbitrap HRMS technology, optimization of each instrument resulted in slightly different operational parameters. During acquisition, data files were saved in profile mode using the .RAW format and converted to .CDF with the Thermo Xcalibur file converter or .mzXML using MSConvert (Chambers et al. 2012). Specific details for each platform are provided below.

3.2.2.1. Dionex Ultimate 3000 with Q-Exactive HRMS (Chapters 4,6,8)

Sample extracts were analyzed in triplicate by liquid chromatography with Fourier transform mass spectrometry (Dionex Ultimate 3000, Q-Exactive, Thermo Fisher, m/z range 85-1275), as described by Soltow et al. (2013). Following a 10 μ L sample injection, analyte separation was accomplished using a 2.1 cm x 10 cm x 4.6 μ m C₁₈ column (Higgins Analytical) and acetonitrile gradient (A= 2% formic acid, B= water, C= acetonitrile) consisting of an initial 2 min period of 80% A, 5% B, 15% C, followed by linear increase to 0% A, 5% B, 95% C and then held for an additional 4 min, resulting in a total runtime of 10 min per injection. Mobile phase flow rate was held at 0.35 mL/min for 6 min, and then increased to 0.5 mL/min. Following analytical separation, the column was flushed with a wash solution consisting of 2% formic acid in acetonitrile. The mass spectrometer was operated in positive electrospray ionization (ESI) mode at a resolution of 70,000 (FWHM). Source tune settings were optimized using a standard mixture, resulting in capillary temperature, sheath gas, auxiliary gas and spray voltage settings of 275°C, 45 (arbitrary units), 5 (arbitrary units) and 4.5 kV, respectively. High-resolution detection of m/z features was accomplished by

maximum injection time of 10 milliseconds and automatic gain control target of 1×10^6 . During untargeted data acquisition, no exclusion or inclusion masses were selected, and data was acquired in MS1 mode only.

3.2.2.2. Accela Surveyor with LTQ-Velos HRMS (Chapter 5)

Sample extracts were analyzed in triplicate by dual column (anion exchange and C₁₈) liquid chromatography-Fourier transform mass spectrometry (Accela, LTQ Velos Orbitrap, Thermo Fisher, m/z range 85-850), as described by Soltow et al. (2013). For the purpose of this study, only data acquired from the C₁₈ based reverse phase separation was utilized. Following a 10 μ L sample injection, separation was accomplished using a 2.1 cm x 10 cm x 4.6 μ m C₁₈ column (Higgins Analytical) and acetonitrile gradient (A= 2% formic acid, B= water, C= acetonitrile) consisting of an initial 2 minute period of 5% A, 60% B, 35% C, followed by linear increase to 5%A, 0% B , 95% C and then held for an additional 4 minutes, resulting in a total runtime of 10 minutes per injection. Mobile phase flow rate was held at 0.35 mL/min for 6 minutes, then increased to 0.5 mL/min. Following analytical separation, the column was then flushed with a wash solution consisting of 2% formic acid in acetonitrile to remove retained lipids. The mass spectrometer was operated in positive ESI mode at a resolution of 60,000. Source tune settings were optimized using a standard mixture, resulting in capillary temperature, sheath gas, auxiliary gas and spray voltage settings of 275°C, 45 (arbitrary units), 5 (arbitrary units) and 4.5 kV, respectively. High-resolution detection of m/z features was accomplished by maximum injection time

of 500 milliseconds and AGC target of 5×10^5 . No exclusion or inclusion masses were selected, and data was acquired in MS1 mode.

3.2.2.3. *Ultimate 3000 with Q-Exactive HF HRMS (Chapter 7)*

Sample extracts were analyzed in triplicate using a dual column, dual polarity approach that includes hydrophilic interaction (HILIC) chromatography with positive ESI and C_{18} chromatography with negative ESI (Ultimate 3000, Q-Exactive HF, Thermo Fisher, m/z range 85-1275), as described by Liu et al. (2016). Both positive and negative mode data were used in the subsequent data analysis. Following a 10 μ L sample injection, HILIC separation was accomplished using a 2.1 cm x 10 cm x 2.6 μ m HILIC column (Thermo Accucore) and acetonitrile gradient (A= 2% formic acid, B= water, C= acetonitrile) consisting of an initial 1.5 minute period of 10% A, 10% B, 80% C, followed by linear increase to 10% A, 80% B, 10% C at 6 min and then held for an additional 4 minutes. Separation by C_{18} was with 2.1 cm x 10 cm x 2.6 μ m column (Thermo Accucore) and the following gradient: initial 2 minute period of 94.5% A, 5% B, 0.5% C, followed by linear increase to 4.5% A, 95% B, 0.5% C at 6 min and then held for an additional 4 minutes. Mobile phase flow rate was held at 0.35 mL/min for 6 minutes, and then increased to 0.5 mL/min. Following analytical separation, the column was then flushed with a wash solution to remove retained compounds. The mass spectrometer was operated using ESI mode at a resolution of 120,000. Source tune settings included capillary temperature, sheath gas, auxiliary gas, sweep gas and spray voltage settings of 300°C, 45 (arbitrary

units), 25 (arbitrary units), 1 (arbitrary units) and +3.5 kV, respectively for positive mode, and 200°C, 30 (arbitrary units), 5 (arbitrary units), 1 (arbitrary units) and +3.0 kV for negative mode. S-Lens RF level was maintained at 45. High-resolution detection of m/z features was accomplished by maximum injection time of 10 milliseconds and AGC target of 1×10^6 . In addition to collecting MS1 data for all samples, data-dependent MS² spectra were collected for two randomly selected samples.

3.2.3. Mass spectral feature extraction and alignment

Following analysis of each batch, data files were sorted by chromatography or polarity and extracted separately. Feature detection and peak alignment was completed using apLCMS (Yu et al. 2009) and xMSanalyzer (Uppal et al. 2013). Both packages are available as public domain software for use in the R programming environment (R Core Team 2013). Feature tables were generated using the *cdf.to.ftr()* function in apLCMS, providing ion m/z , aligned retention time and intensity for each analytical replicate using up to six different combinations of extraction parameters. The apLCMS data extraction parameters have been optimized for each analytical configuration using the methodology described by Uppal et al. (2013). To improve feature reliability and improve detection, the *xMSwrapper.apLCMS()* function in xMSanalyzer was used to merge apLCMS feature tables collected at multiple parameter settings. xMSanalyzer feature tables are generated by selecting the optimal peak detection and integration from different parameter settings and include the ion m/z ,

retention time, total number of samples that feature was detected in, the number of samples that have a feature detected in two or more of their technical replicates, median technical replicate coefficient of variation for a given feature, the number of samples with the feature present in 50% or greater of the technical replicates, maximum intensity measured, peak score and ion abundance. Non-detectable signals were represented as zeroes, and data were transformed to reduce data heteroscedasticity (van den Berg et al. 2006).

3.2.4. Metabolite annotation and characterization

The identity of detected m/z features was first evaluated by comparison to a database of chemicals with confirmed m/z , retention time and mass spectral patterns prepared by the Clinical Biomarkers Laboratory. Following this initial characterization, xMSannotator (Uppal et al. 2016) was used to provide matches to both endogenous and exogenous chemicals present in the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Human Metabolome Database (HMDB). Tentative identifications were completed using common adducts corresponding to a given ionization polarity, with mass accuracy ≤ 10 ppm (defined as relative m/z error $\times 10^6$). The combination of confidence scoring based upon multiple adducts, isotopic grouping and mass defect has shown to provide greater accuracy when annotating high-resolution mass spectrometry data (Uppal et al. 2016). While many of the metabolic pathways have been characterized in humans using different approaches, often m/z features detected using untargeted techniques provide no matches in chemical databases. Origins of unknown, detected

metabolic features include uncharacterized metabolites, conjugates, previously unknown chemical exposures and products arising from promiscuous enzymes (Little et al. 2011, da Silva et al. 2015, Walker et al. 2016). When possible, unidentifiable significant features were evaluated using a number of strategies, including mass defect filtering, isotopic distribution grouping and structural elucidation with ion disassociation; however, confirmation of identity is challenging due to the limited availability of analytical standards and low abundance in samples (Uppal et al. 2016).

3.2.5. *Metabolic pathway enrichment and bioeffect*

Metabolite pathway enrichment was completed using Mummichog (Li et al. 2013), which enables pathway identification without *a priori* feature identification. Mummichog was developed to address the challenge in annotating untargeted metabolomics data, and connects genome-scale metabolic models to metabolomics data prior to annotation. This is accomplished by examining metabolite matches within the context of metabolic networks; when used with *m/z* features that have been selected based upon association with a given outcome, correct identification will be distributed in a meaningful way, while incorrect identifications are distributed randomly. Therefore, this not only protects against errors in annotation, but also protects against false positives during the feature selection stage. Mummichog, metabolic models will include MetaFishNet (Li et al. 2010), and pathways exhibiting a *p*-value ≤ 0.05 will be considered significantly associated with exposure. Correlation clusters generated using

MetabNet will also be used to evaluate pathway-level associations. The results from these analyses were investigated for biological response to exposure.

3.3. *HRM data analysis and metabolome wide association study*

HRM was evaluated as an analytical platform to detect exposure related changes using as MWAS framework to identify associations with the exposure measures specific to each study. HRM provides advanced chemical profiling, enabling comprehensive measurement of metabolites arising from exposure and the majority of human biochemical pathways. Exposure to environmental chemicals has the potential to influence local and global changes in gene transcription and enzyme activity, resulting in variation to metabolite levels associated with exposure dose. Testing for dose-related associations can therefore identify metabolic changes related to chemical exposure.

Traditional molecular epidemiology studies of exposed populations have relied on a small number of targeted measures to evaluate biomarkers of exposure and response (Schulte and Perera 1993). Relation between measured marker, endpoint, covariates and outcome can be assessed using a finite number of statistical tests with discrete values. Data obtained from HRM profiling is inherently more complex. Analytical (formation of multiple ions are possible for a single chemical; ion suppression due to co-eluting high-abundance metabolites), chemical (isotopic forms) and biological (metabolites from the same pathway/chemical) correlations exist within the data; large number of features relative to sample size requires corrections for multiple hypothesis testing; ion

specific differences in intensity distribution and interpretation of non-detected values is complicated by inability to determine detection limit for large number of features easily. Thus, data normalization, dimensionality reduction, false discovery rate correction and additional informatics approaches were required prior to data inference. Due to the differences in exposure measurement, type and duration, analytical approaches to identify exposure related changes to the metabolic phenotype were specific to each population. Univariate, multivariate and network based data analysis approaches used for MWAS are described below.

3.3.1. Univariate statistical analyses for feature selection

Univariate statistical techniques were used to identify metabolic features associated with exposures in **Chapters 5, 6 and 8**. For **Chapter 5**, a linear regression framework was implemented in the statistical software R version 3.1.2 that incorporated m/z feature \log_2 intensity as the dependent variable and exposure classification as the independent variable while correcting for age, sex and body mass index. Regression analysis was completed separately for each detected m/z feature meeting the missing value and CV thresholds using the $lm()$ function; p for the exposure variable meeting the false discovery rate (FDR) < 20% were then selected for additional characterization.

Due to the availability of repeat measures for the study in **Chapter 6**, a linear mixed-effects regression model was applied to determine m/z features associated with the two timescales of exposure (shift- and week-averaged). To

identify m/z features associated with shift-length exposure to traffic-related pollutants, the following model was implemented in R:

$$\gamma_{ijk} = \beta_0 + \beta_1 C_{exp,yjk} + \beta_2 Age_j + \beta_3 Day_k + \beta_4 BMI_j + U_j + \varepsilon_{ijk}$$

where γ_{ijk} is the \log_2 transformed intensity profile for m/z feature i corresponding to individual j on day k using post-shift samples only, $C_{exp,yjk}$ is the interquartile-range normalized exposure y , which includes shift-averaged ambient exposure to $PM_{2.5}$, OC and EC for individual j on day k , the random intercept for each individual U_i and within subject error ε_{ijk} . To estimate effects due to week-averaged exposure, the model was modified to:

$$\begin{aligned} \gamma_{ijkm} = & \beta_0 + \beta_1 C_{wexp,yj} + \beta_2 Age_j + \beta_3 Day_k + \beta_4 BMI_j + \beta_5 Time_m + U_j \\ & + \varepsilon_{ijk} \end{aligned}$$

where γ_{ijkm} is the \log_2 transformed intensity profile for m/z feature i corresponding to individual j on day k at time m (pre- or post-shift), $C_{wexp,yj}$ is the interquartile-range normalized, week-averaged exposure y , which includes week-averaged ambient exposure to $PM_{2.5}$, OC and EC for individual j , the random intercept for each individual U_i and within subject error ε_{ijk} . Regression analysis was performed using the *lmer()* function available in the R package *lme4* (Bates et al. 2015) and completed separately for each exposure and time-scale combination. The likelihood ratio test of the complete model against the null

model, which excluded the corresponding exposure measures as the independent variable, was completed using the R *anova()* function to obtain *p* for each *m/z* feature associated with exposure. Model fits were evaluated by calculating the conditional and marginal r^2 with the *r.squaredGLMM()* function from the R package MuMIn (Bartoń 2013) and reported for all *m/z* features meeting the significance threshold. To account for multiple hypothesis testing, a Benjamini-Hochberg false discovery rate (FDR) (Benjamini and Hochberg 1995) threshold of 20% was applied to identify *m/z* features associated with each pollutant. The resulting *m/z* features with *p* below the FDR threshold were selected for annotation and comparison of metabolic changes associated with different types and lengths of exposure.

In **Chapter 8**, a network based correlation approach was used to detect *m/z* features associated with the flame retardant PBB-153 and persistent environmental pollutant PCB-153. Spearman correlation coefficients and corresponding *p* values were determined individually for each *m/z* feature and pollutant measure using the R package MetabNet (Uppal et al. 2015), which calculated pairwise correlation coefficients for complete observations (non-zero intensities only) using the *cor()* function. *P* values were then estimated with the Student's t-test approximation with *n* based upon the number of complete observations. Correlations with Spearman $|r| \geq 0.3$ and $p < 0.01$ were selected for further characterization. For these *m/z* features, effect size variability was estimated by 95% confidence intervals determined for 1000 iterations using the

function *spearman.ci()* available from the R package RVAideMemoire (Herve 2017).

3.3.2. *Multivariate statistical analyses for feature selection*

In **Chapters 4** and **7**, *m/z* features associated with each environmental pollutant were identified using multivariate approaches based upon partial least squares (PLS) variable selection techniques. This approach provided insight into the metabolic variations discriminating exposure and systemic patterns present due to different exposures rather than evaluating dose-dependent effects. In **Chapter 4**, sparse partial least squares (sPLS) regression analysis (Le Cao et al. 2008) was utilized to identify features associated with benzo[a]pyrene serum levels. The sPLS regression analysis performs a supervised dimensionality reduction approach to maximize covariance between the continuous response variable (benzo[a]pyrene serum level) and predictor variables (*m/z* feature intensity profile) (Patel et al. 2014). Features discriminating exposure levels were detected using the *sPLS()* function available from the R package mixOmics (Dejean et al. 2014), with final selection by Lasso penalization of the PLS loading vectors for feature selection and ranked based on model contribution. The top 100 features were selected as a manageable number for subsequent analyses.

Rather than using a regression framework, in **Chapter 7** a PLS discriminant analysis (PLS-DA) (Wold et al. 2001) was employed to identify metabolic variations contributing the greatest to differences between low and high average, annual ultra-fine particle (UFP) exposure. The optimal number of latent

components was determined using the *plsgenomics* package in R. *MixOmics* was used to build the PLS-DA model and select discriminatory features based on variable importance for projection (VIP) score ≥ 2.0 (Le Cao et al. 2009). Model accuracy was evaluated by R^2/Q^2 for PLS components 1 and 2, 10-fold cross-validation accuracy and permuted accuracy based on data shuffling (random data predictive accuracy). Association of m/z features meeting the VIP threshold were then evaluated using Cohen's d .

3.3.3. *Network integration*

In **Chapters 5, 6 and 7**, integration of metabolite features with independently measured molecular endpoints was completed to test how metabolic variations related to exposure were consistent with biological changes associated with disease processes and provide insight into mechanisms underlying biological response. In **Chapter 5**, network integration with immunological response measures, renal damage biomarkers and urinary TCE metabolites was completed using a Spearman correlation threshold of $|r| \geq 0.3$ and FDR $< 20\%$. Correlations were determined using the *metabnet()* function available from the R package MetabNet (Uppal et al. 2015). For this analysis, correlation coefficients and corresponding p values were determined individually for each m/z feature associated with TCE exposure and molecular measure. Relationships were calculated for all observations (including zero intensities) using the *cor()* function. P values were then estimated with the Student's t-test approximation with n based upon the number of all observations.

Due to the large number of m/z features associated with exposures, network integration was combined with upfront PLS variable selection to combine protein biomarkers and peripheral blood gene expression with exposure associated metabolic variations in **Chapters 6** and **7**, respectively. These analyses were completed using the function `run_xmwas()`, which is available in the R package xMWAS (Uppal et al. 2017), and combines feature selection with network association scoring, cluster identification and centrality scoring for integration of complex omic datasets. In **Chapter 7**, xMWAS was used to combine m/z features identified as systemically associated with UFP exposure from positive and negative ESI with plasma markers of inflammation and coagulation. Protein markers were used as central nodes (X matrix), with HILIC positive and C_{18} negative m/z features input as the Y and Z matrices, respectively. The top 100 features from HILIC and C_{18} were then selected based-upon VIP rank order, which was determined using the PLS methodology described above. For the VIP ranked m/z features, the integrated network structure was determined using the mixOmics `network()` function to identify correlations with $|r| \geq 0.3$ and $p < 0.05$. Clusters were then identified as communities of nodes tightly connected with each other but sparsely connected with the rest of the network using multilevel community detection (Vincent et al. 2008).

In **Chapter 7**, a network correlation approach including initial variable selection based upon multilevel sPLS was used to identify the relationship between peripheral blood gene expression and metabolites associated with diesel exhaust exposures. To address diurnal effects, only post-shift feature expression

from both days was considered. Initial feature selection for both datasets was completed using multilevel sPLS regression analysis (Liquet et al. 2012), which utilizes a supervised multivariate dimensionality reduction method to perform simultaneous discriminatory analysis and m/z feature selection while accounting for the dependency structure of both datasets and individual repeated measurements (Patel et al. 2014). The top 1500 discriminatory molecular probes from the first three latent variables were then combined with the metabolite levels using the *network()* function for determining network structure and clustering. Correlations with $|r| \geq 0.4$ and $p < 0.05$ were then selected for visualization in Cytoscape, with gene and metabolite pathway enrichment determined using MetaCore (Thomson Reuters) and *Mummichog*, respectively. Previously described interactions between network associated genes and environmental chemicals were also evaluated using the Chemical Toxicogenomics Database (CTD) (Davis et al. 2013).

3.4. Reference standardization

In **Chapters 4** and **7**, a reference standardization approach was used to provide absolute concentrations of a select number of endogenous metabolites that were then compared to expected concentration ranges. This was accomplished by identifying detected m/z features that have been previously confirmed by comparison of adduct m/z , retention time and MS² analysis to authentic reference standards and database spectra (Go et al. 2015, Go et al. 2015). Using this approach, metabolite levels are first determined in Q-Std3 by methods of addition

(Boyd et al. 2008) or comparison to NIST standard reference material (SRM 1950; Metabolites in Frozen Human Plasma) (McGaw et al. 2010, Phinney et al. 2013), providing a reference standard for each analyte that was subsequently analyzed at the beginning and end of each batch. Average metabolite response was then determined based upon the most reliable adduct formed, and concentrations were calculated using single point calibration. Response factors (calculated as the ratio between the known concentration of the compound being quantified and ion intensity in Q-Std3) were determined using the study-averaged Q-Std3 intensities.

Chapter 4. PILOT METABOLOME-WIDE ASSOCIATION STUDY OF BENZO(A)PYRENE IN SERUM FROM MILITARY PERSONNEL

This chapter is taken from a manuscript that was published in the *Journal of Occupational and Environmental Medicine* (PMID: 27501104). This manuscript represents the first MWAS of blood-levels of benzo(a)pyrene and provided initial HRM characterization of serum samples stored in the Department of Defense Serum Repository (DoDSR). The results provide evidence human samples stored in the DoDSR are of sufficient quality for HRM profiling and identification of exposure-related metabolic associations. This study established the feasibility for using metabolomic techniques to characterize metabolite associations with environmental pollutants using an MWAS framework. Co-authors and contributions are as follows. D. Walker: Conceived and executed statistical analysis, interpreted results, performed literature review, wrote all sections, created figures and tables, approved final edits and submission; K. Pennell: Supervised D. Walker, provided critical input for manuscript structure, edited intermediate drafts, edited final document; K. Uppal: Provided software package for sPLS regression analysis; X. Xia: Quantitated serum PAH levels; P. Hopke: Oversaw PAH analysis, read and edited final document; M. Utell: Provided input and guidance for manuscript structure, edited final document; R. Phipps: Read

and edited final document; P. Sime: Read and edited final document; P.

Rohrbeck: Obtained DoDSR samples, read and edited final document; T. Mallon:

Project PI, provided critical input for manuscript structure, edited intermediate

drafts, edited final document; D. Jones: Supervised D. Walker, oversaw HRM

analysis, provided critical input and guidance for manuscript structure, edited

intermediate drafts, edited final document.

4.1. Abstract

Objective: A pilot study was conducted to test the feasibility of using Department of Defense Serum Repository (DoDSR) samples to study health and exposure-related effects.

Methods: Thirty unidentified human serum samples were obtained from the DoDSR and analyzed for normal serum metabolites with high-resolution mass spectrometry and serum levels of free benzo(a)pyrene (BaP) by gas chromatography-mass spectrometry. Metabolic associations with BaP were determined using a metabolome-wide association study (MWAS) and metabolic pathway enrichment.

Results: The serum analysis detected normal ranges of glucose, selected amino acids, fatty acids, and creatinine. Free BaP was detected in a broad concentration range. MWAS of BaP showed associations with lipids, fatty acids, and sulfur amino acid metabolic pathways.

Conclusion: The results show the DoDSR samples are of sufficient quality for chemical profiling of DoD personnel.

4.2. Introduction

The Department of Defense Serum Repository (DoDSR) contains serum samples from service personnel originally collected as part of mandatory HIV testing and later extended to include pre- and post-deployment samples (Rubertone and Brundage 2002). While not currently utilized for routine biomonitoring, serum samples stored in the repository have potential as a valuable resource for assessing deployment-related exposures, which has been recognized as an important need in the National Academy of Sciences 2000 report “Strategies to Protect the Health of Deployed U.S. Forces: Detecting, Characterizing, and Documenting Exposures” (NRC 2000). The recent availability of high-throughput chemical profiling platforms with expanded coverage, such as high-resolution metabolomics (HRM), now makes possible cost-effective environmental chemical surveillance and bioeffect monitoring using DoDSR samples (Jones 2016, Walker et al. 2016). HRM, which utilizes liquid chromatography and high-resolution mass spectrometry, provides quantitative measure of a large number (10,000-20,000) of endogenous and exogenous chemicals in biological samples that can be used to profile biomarkers of exposure, effect, and alterations in metabolism consistent with disease (Jones et al. 2012, Jones 2016, Walker et al. 2016).

Before the DoDSR can be appropriately used as a resource for evaluating exposure, there is a need to test sample suitability for chemical profiling. Since HRM is a new technology, there is no way to assess sample alterations occurring during storage; however, quality can still be evaluated by testing for expected metabolic patterns and biological associations with targeted measurement of

specific exposure biomarkers. Thus, combining the metabolic measurements obtained from HRM and levels of exposure biomarkers measured using either HRM or other platforms, it is possible to test for exposure-related metabolic changes. Comparison of *in vivo* metabolic associations measured in the serum can then be interpreted in light of known toxicological targets and toxicodynamics, allowing assessment of sample suitability for measuring biologically relevant effects.

Deployment-related exposures have the potential to occur from a number of sources. Of particular interest is identifying respiratory exposure to combustion products generated from burn pits, which is suspected as a possible risk factor for new onset asthma and respiratory symptoms (Szema et al. 2010, Szema et al. 2011). Serum levels of BaP, a carcinogenic polycyclic aromatic hydrocarbon arising from the combustion of organic material, was selected as an exposure biomarker to test the suitability of DoDSR samples for elucidating biological response to environmental exposures. Previous biomonitoring studies have shown that BaP was measureable in serum because of respiratory exposure to smoke, exhaust and incomplete combustion (Nemoto and Takayama 1983, Neal et al. 2008, Pleil et al. 2010, Al-Daghri et al. 2014). During the course of deployment, service personnel are known to potentially experience a wide range of BaP exposures from different sources (Baird 2012, Smith et al. 2012), although due to the unknown time of collection for serum samples in this study it is not possible to link measured BaP to exposures during deployment.

The present pilot study was designed to evaluate the feasibility of using DoDSR samples for measuring biomarkers of exposure and associated biological response. HRM chemical profiling of thirty unidentified DoDSR samples was completed to evaluate 1) whether metabolites could be reliably detected in the serum samples 2) serum levels of parent BaP and 3) metabolic pathway alterations associated with serum BaP detected using a metabolome-wide association study (MWAS) framework (Holmes et al. 2008, Osborn et al. 2013, Go et al. 2014). It is important to recognize that although serum levels of BaP represent exposure biomarkers relevant to deployed troops, it is not possible to directly associate serum BaP levels with deployment due to the unidentified nature of the samples. By testing the suitability of DoDSR for metabolic studies, we hope to define sample quality for metabolic studies and use as a resource for chemical profiling.

4.3. *Materials And Methods*

4.3.1. *Serum Samples*

Thirty unidentified serum samples (i.e., information on source, date of collection and meta-data is unknown) were obtained from the DoDSR for BaP targeted quantification and HRM. The repository consists of approximately 50 million serum samples originally collected for mandatory HIV testing in armed forces personnel (Perdue et al. 2015). Following collection, samples were stored at -30°C. Prior to analysis, specimens were thawed at 4°C, 500 µL was aliquoted into separate microfuge tubes after vortexing, refrozen and shipped on dry ice.

4.3.2. *BaP targeted quantification*

BaP was measured using sample preparation and GC-MS methodologies optimized for analysis of DoDSR samples (Xia et al. 2016) and based upon previously established methods. (Sirimanne et al. 1996). Following addition of an internal standard solution consisting of deuterated anthracene (d_{10}), serum polycyclic aromatic hydrocarbons (PAHs) were first extracted using Triton X-100, followed by hexane. Sample extracts were analyzed by a Thermo Fisher Trace gas chromatograph interfaced to a quadrupole mass spectrometer (Thermo Fisher DSQ) operated in selective ion monitoring mode (SIM). Concentration of BaP was determined by comparing the peak intensities to a five-point calibration curve. Method fitness of purpose and detection limit was evaluated by spiking serum samples with BaP at two concentration ranges, with BaP recovery exceeding >80%. Method limit of detection was 0.0028 ng/mL.

4.3.3. *High-resolution metabolomics*

Samples were prepared for metabolomics analysis using established methods described elsewhere (Park et al. 2012, Soltow et al. 2013, Go et al. 2015). Briefly, serum aliquots were removed from storage at -80°C, and thawed on ice, upon which 65 µL of serum was added to 130 µL of acetonitrile containing a mixture of stable isotopic standards, vortexed and allowed to equilibrate for 30 min. Following protein precipitation, triplicate 10 µL aliquots were analyzed by reverse-phase C₁₈ liquid chromatography with Fourier transform mass spectrometry (Dionex Ultimate 3000, Q-Exactive, Thermo) operated in positive electrospray ionization mode and resolution of 70,000 (FWHM) over mass-to-charge (*m/z*) range of 85-1,250. Samples were analyzed in batches of 15, in addition to a quality control (QC) pooled reference sample included at the beginning and end of the analytical batch for quantification and standardization. Raw data files were extracted using hybrid apLCMS (Yu et al. 2013) with modifications by xMSanalyzer (Uppal et al. 2013), with each unique ion defined by *m/z*, retention time and ion abundance.

Tentative chemical identifications for all detected features was completed by matching the accurate mass *m/z* features to commonly observed adducts at a ±5 ppm mass error threshold using the *feat.batch.annotation.KEGG()* function in xMSanalyzer (Uppal et al. 2013) and the Kyoto Encyclopedia for Genes and Genomes (KEGG) database. Since no information was known about the origin of the serum samples, 15 common human endogenous metabolites were quantified

within the 30 samples by reference standardization (Go et al. 2015). Serum concentrations from the National Health and Nutrition Examination Survey (NHANES) and Human Metabolome Database (HMDB) were included for comparison.

4.3.4. Feature selection

Sparse partial least squares (sPLS) regression analysis (Le Cao et al. 2008) was utilized to identify features associated with BaP serum levels. The sPLS regression analysis performs a supervised dimensionality reduction approach to maximize covariance between the continuous response variable (BaP serum level) and predictor variables (m/z feature intensity profile) (Patel et al. 2014). The top 100 features were selected as a manageable number for subsequent analyses. Feature selection was completed using the *sPLS()* function within the R package mixOmics (Dejean et al. 2014). Visualization of the sPLS significant features and grouping based on similarity in expression were evaluated with unsupervised hierarchical clustering. The corresponding heat map was plotted using dendrogram clustering and z -score normalized intensity profiles. BaP levels were then categorized based on sample clustering.

4.3.5. Metabolite correlation and pathway enrichment

To evaluate pathway correlations of metabolites associated with measured BaP levels, a metabolome-wide Spearman correlation analysis was applied to the 100 m/z features selected with sPLS to capture associations within the raw data.

Table 4.1. Serum concentrations of representative confirmed metabolites in the 30 unidentified samples.

Metabolite	Detected <i>m/z</i>	Adduct	Avg. concentration \pm SD (μ M)	Serum concentration range (μ M) ¹
Arginine	175.1194	+H	148 \pm 39	130 \pm 30 ^a
Glycine	120.0035	+2Na-H	280 \pm 62	330 \pm 106 ^a
Histidine	156.0771	+H	100 \pm 12	143 \pm 27 ^a
Ornithine	133.0974	+H	83 \pm 28	93.8 \pm 41 ^a
Phenylalanine	166.0867	+H	131 \pm 18	78 \pm 21 ^a
Threonine	120.0659	+H	136 \pm 22	128 \pm 41 ^a
Tryptophan	205.0974	+H	56 \pm 7	55 \pm 9.7 ^a
Tyrosine	182.0814	+H	84 \pm 23	55 \pm 9.7 ^a
Linoleic acid	281.2481	+H	2309 \pm 157	2310-5190 ^b
α -Linolenic acid	279.2324	+H	54 \pm 10	26-151 ^b
Glucose	203.0529	+Na	4310 \pm 1153	4971 \pm 373 ^a
Kynurenine	209.0925	+H	2 \pm 0.4	2.4 \pm 0.5 ^a
Carnitine	162.1128	+H	52 \pm 9	30 \pm 8 ^a
Creatinine	114.0667	+H	93 \pm 13	87 \pm 19 ^a
Creatine	132.0770	+H	16 \pm 8	36.7 \pm 28.3 ^a

a. Wishart, D. S., et al. (2009). HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res* **37**(Database issue): D603-610.

b. U.S. Centers for Disease Control and Prevention. Second National Report on Biochemical Indicators of Diet and Nutrition in the U.S. Population 2012. Atlanta (GA): National Center for Environmental Health; April 2012.
http://www.cdc.gov/nutritionreport/pdf/Nutrition_Book_complete508_final.pdf

Spearman rank correlation coefficient and Benjamini and Hochberg (1995) false discovery rate were calculated pair-wise for the sPLS selected features using the R package MetabNet (Uppal et al. 2015). Correlations with $|r| \geq 0.6$ and FDR $\leq 20\%$ were selected to test for metabolic pathway enrichment in Mummichog (Li et al. 2013). Enriched pathways were selected at significance score ≤ 0.05 . A select number of metabolite matches were confirmed by MS/MS, and compared to reference standards (when available) or database spectra (Smith et al. 2005, Wishart et al. 2009, Horai et al. 2010). Confirmed identifications by MS/MS included the metabolites selected for absolute quantification, in addition to linoleate, sphinganine, sphingosine, cholesterol, acetyl-carnitine, threonine, dihydrobiopterin, histamine, 4-chlorophenylacetate and methionine.

4.4. Results

4.4.1. Targeted quantification of endogenous human metabolites

Table 4.1 shows the average concentration of 15 metabolites selected for quantification by reference standardization, which includes essential and non-essential amino acids, glucose, fatty acids and metabolites known to be involved in various disease processes. Creatinine and creatine were detected at average concentrations of 93 ± 13 and $16 \pm 8 \mu\text{M}$, respectively. Due to the tightly regulated range of glucose in healthy individuals, serum glucose levels represent an excellent indicator of overall

health. The average glucose level in these serum samples was $4312 \pm 822 \mu\text{M}$, which corresponds to a glucose level of 77 mg/dL and is within the range typically observed during a fasting glucose test (70-99 mg/dL). One sample had an extremely low glucose level of 4 mg/dL.

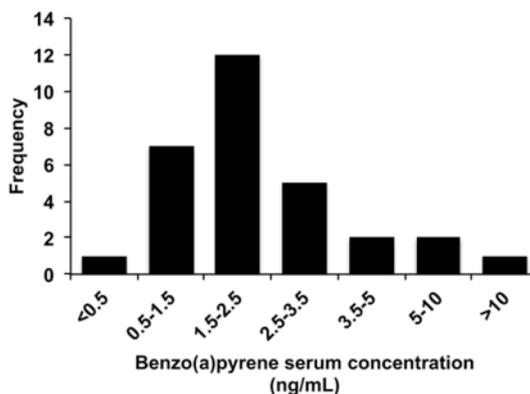


Figure 4.1. BaP serum concentrations in the 30 unidentified individuals. Quantified serum levels ranged from 0.13 to 37.2 ng/mL, with an average concentration of 3.39 ng/mL.

4.4.2. BaP targeted profiling

BaP was detected over a wide concentration range in the serum (**Figure 4.1**), with a minimum, 25th percentile, median, 75th percentile and maximum concentration (ng/mL) of 0.13, 1.45, 2.09, 3.18 and 37.2, respectively, corresponding to an

average serum BaP concentration of 3.39 ng/mL \pm 1.19 SE. The presence of outlier BaP levels was evaluated by subtracting or adding 2.0 \times IQR (interquartile range) from the 1st and 3rd BaP quartiles. For the thirty individuals, one serum level corresponding to 37.2 ng/mL exceeded the IQR limits. While biologically plausible, this sample was removed prior to sPLS analysis to avoid biasing results due to the study size.

4.4.3. High-resolution metabolomics

Following data extraction and alignment, 7,810 unique m/z features were detected with a median replicate relative standard deviation (RSD) of 19.6%. Technical replicates were averaged and filtered based on the requirement that each feature must be present in greater than 10% of the samples, resulting in 7,584 remaining features. Median relative standard deviation for triplicate measurements following filtering was 19.8%, providing a median standard error of the mean value of 11.4%. Annotation using m/z values included 6,010 unique KEGG IDs, which represented 70 metabolic pathways with ten or greater matches to metabolic intermediates present. These included 625 within the overall human metabolic map (**Figure 4.2**). It should be noted that these categories are estimates based upon accurate mass matches because confirmation of structures was not practical for this number of chemicals with the amount of sample available.

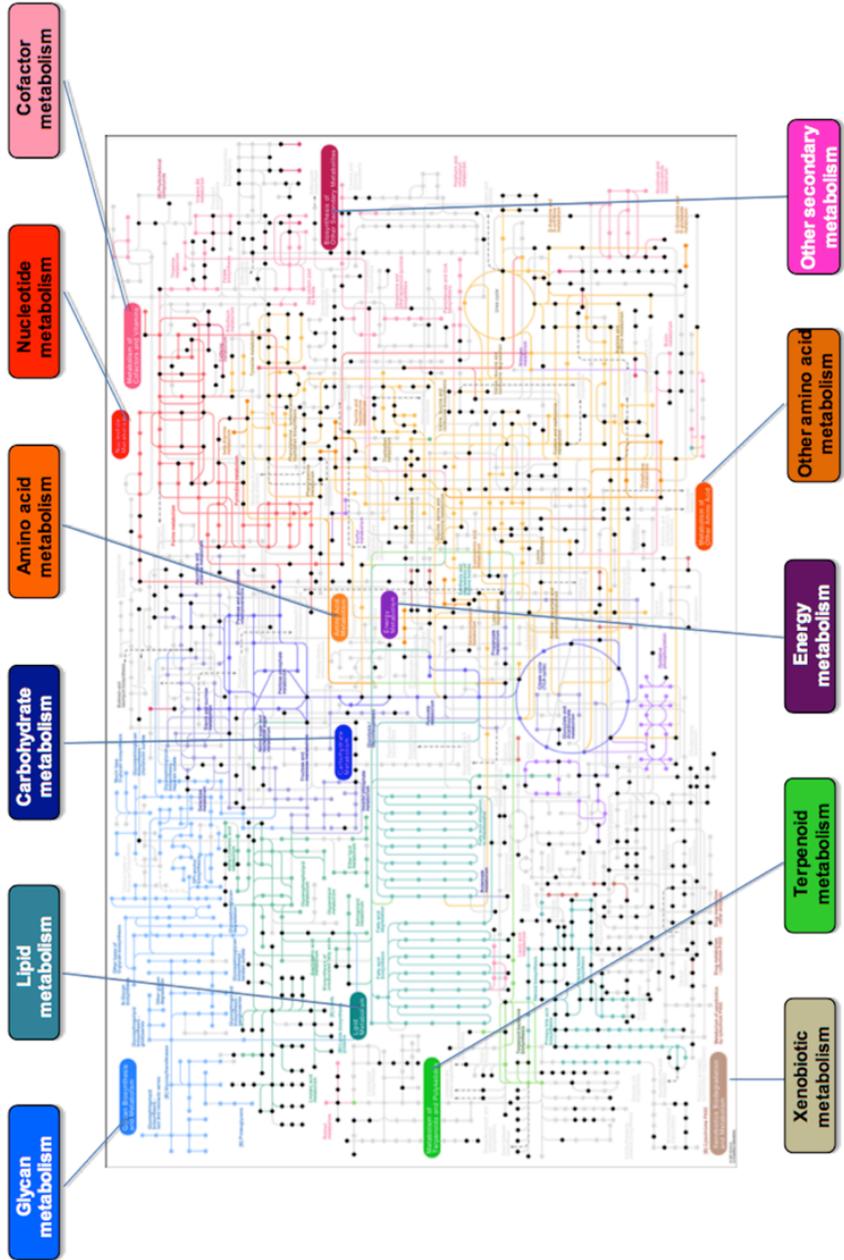


Figure 4.2. Metabolite coverage of metabolic pathways following database matching to KEGG. Matches are represented by black dots. A total of 625m/z features matched KEGG metabolites present within the human metabolic pathways, which included amino acid, lipid, steroid, and xenobiotic metabolism.

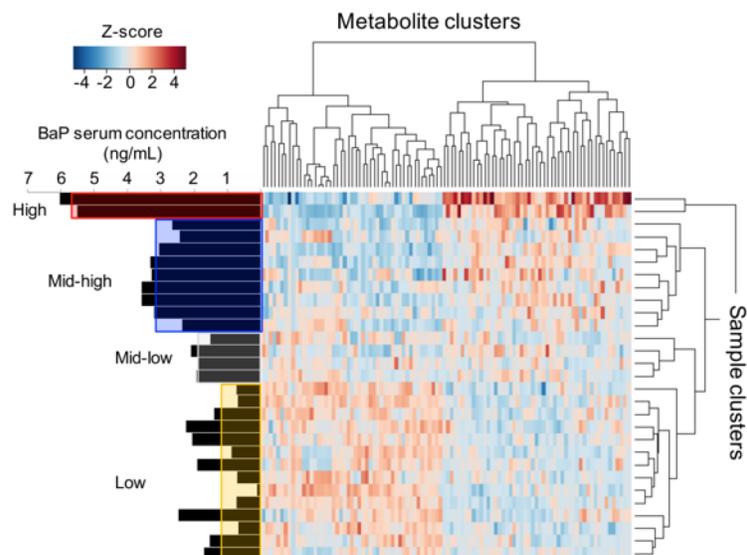


Figure 4.3. Unsupervised hierarchical clustering heatmap of the top 100 sPLS selected features associated with BaP. Individual grouping was found to be strongly associated with exposure level (colored boxes on left axis), reflecting exposure-metabolic signatures within the chemical profile.

4.4.4. Serum BaP MWAS

To test for associations with BaP, an sPLS based MWAS was completed using as BaP serum levels as the continuous dependent variable and detected features as predictor variables. We selected the top 100 features from sPLS for further analysis, representing a manageable number for subsequent characterization. Average intensity of the selected features ranged over 500-fold, with 81% of the features corresponding to compounds with $m/z \leq 600$. Retention time distribution was skewed towards early eluting compounds, with 62% detected in 3 minutes or less. To determine if groups of metabolites are associated with BaP serum levels, we performed two-way unsupervised hierarchical clustering (**Figure 4.3**). The top 100 features grouped into 15 clusters (**Supplementary File 1**), which represented both increased and decreased expression with BaP serum levels (left

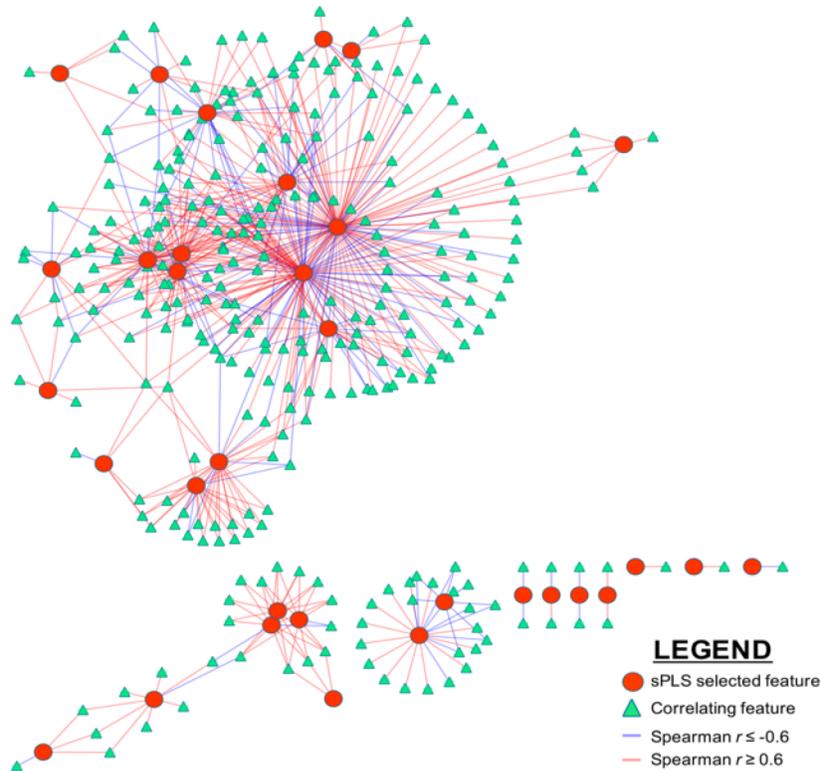


Figure 4.4. Network association of the sPLS m/z features with raw HRM data. Red circles represent sPLS selected features; green triangles are correlating features with Spearman $|r| \geq 0.6$ and $FDR \leq 20\%$.

axis). Clustering of samples based upon sPLS selected metabolic profiles (horizontal axis) resulted in 4 separate groups based upon BaP levels, including a high-exposure group ($n=2$, average BaP serum concentration 5.77 ng/mL), mid-high exposure ($n=9$, average BaP serum concentration 3.06 ng/mL), mid-low exposure ($n=4$, average BaP serum concentration 1.87 ng/mL) and low exposure ($n=14$, average BaP serum concentration 1.28 ng/mL). Clusters were used to define low (1.28 ng/mL; $n=11$) and high (high + mid-high; average= 3.55 ng/mL; $n=14$) BaP groups for comparing relative expression of metabolites in significant metabolic pathways.

To provide identification of sPLS features based upon m/z , the KEGG database was searched for matches to common adducts at a mass error threshold of 5 ppm (listed in **Appendix 1.1**). Feature grouping by BRITE categorization (34), included lipids (n=34, 21.0%), phytochemicals (n=22, 13.6%), pesticides (n=7, 4.3%), compounds with biological roles (n=5, 3.1%), carcinogens (n=5, 3.1%), pharmaceuticals (n=5, 3.1%) and endocrine disrupting compounds (n=1, 0.6%), with the remaining 51.2% (n= 83) non-classified. The assigned identities of the annotated features include important biological metabolites and compounds originating from exogenous sources. Centrally acting, biological molecules important in overall metabolism and signaling include threonine (+H-H₂O), dihydrobiopterin (+H), histamine (+H), hexadecanol (+ACN+H), sphinganine (+H-H₂O), hydroxybutanoic acid (+H), linoleate (+H-H₂O), and 2-hydroperoxy-octadecatrienoic acid (+Na). Previous studies for this platform have confirmed identities for threonine, dihydrobiopterin, histamine, sphinganine and linoleate. Within the annotated features, exogenous chemicals were detected as well, including included chlordecone (+Na), cartap (+ACN+H), bendiocarb/dioxacarb (+ACN+Na), cymoxanil (+ACN+H), prometryn/terbutyn (+Na), imazethapyr (+ACN+Na), 4-chlorophenylacetate (+H-H₂O) and trimethoxytoluene (+H). Methoxypyrene (+ACN+H), a product of PAH metabolism, was also identified as one of the significant features. For the remaining 57 sPLS significant features, no matches were present in the KEGG database.

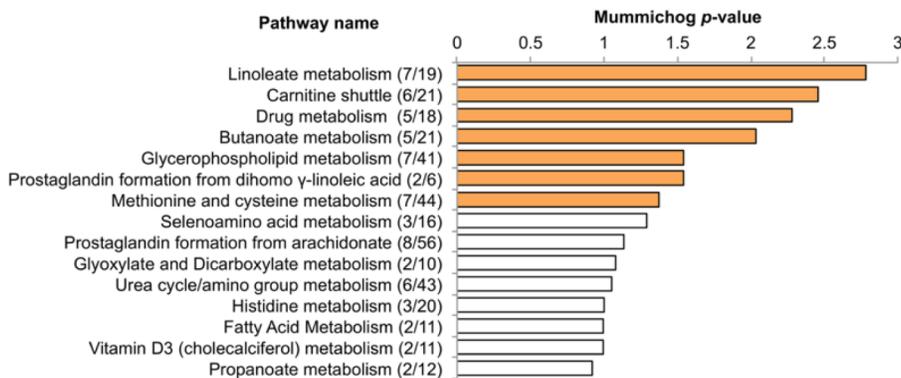


Figure 4.5. Top 15 metabolic pathways with enrichment p -value from Mummi- chog using the sPLS and network associated features. Seven pathways were enriched ($p \geq 0.05$, filled bars) within the MWAS selected features.

4.4.5. Metabolite correlation and pathway analysis

Metabolome-wide Spearman correlation analysis was applied to the m/z features selected with sPLS to test for correlations within the raw HRM data. An additional 388 m/z 's were identified as associated with the sPLS features (**Figure 4.4**). The full list was utilized for pathway enrichment. One or more of the MWAS features were present in 58 of the 114 metabolic pathways. The top 15 identified pathways are shown in **Figure 4.5**. Seven were found to be enriched based on a significance score ≤ 0.05 (filled bars). Enriched pathways included linoleate metabolism ($p= 0.0017$), carnitine shuttle ($p= 0.0035$), drug metabolism ($p= 0.005$), butanoate metabolism ($p= 0.009$), glycerphospholipid metabolism ($p= 0.029$), prostaglandin formation ($p= 0.029$) and methionine/cysteine metabolism ($p= 0.043$). Box and whisker plots of metabolites from these pathways are shown in **Figure 4.6** and **Figure 4.7**.

4.5. Discussion

This pilot study analyzed 30 unidentified serum samples obtained from the DoDSR and measured free BaP, a PAH commonly found in diesel exhaust and linked to cancer in a number of organ systems, including liver and lung. BaP measured in serum was correlated with HRM profiles of metabolic pathway intermediates. MWAS of BaP showed associations with lipids, fatty acids, methionine and cysteine metabolic pathways and demonstrated associations with metabolic pathways that have previously been connected with BaP exposure *in vivo* and *in vitro*.

Serum sample quality was assessed by comparing levels of representative human metabolites to published values for the general population. Quantitated serum chemicals including glucose, creatinine, amino acids and carnitine were found in concentrations that are in the normal range for healthy individuals (**Table 4.1**) based on the average expected blood concentrations in HMDB (Wishart et al. 2013) and the National Health and Examination Survey (NHANES). While the measured value of carnitine in this study was higher than the average value listed in HMDB, carnitine levels up to 79 μM have been reported (Shihabi et al. 1992). The average glucose level was comparable to healthy individuals (70-99 mg/dL). One sample was observed to have a serum glucose level of 4 mg/dL, suggesting a sample collection artifact. When collecting blood serum, coagulant and storage conditions strongly influence inhibition of glycolysis. This has been observed to result in an 85% decrease in

serum glucose levels following storage for 24 hours (Fobker 2014), and may be useful as a reliable marker of serum handling.

Biomonitoring of BaP exposure is most commonly accomplished by measuring hydroxylated urinary metabolites or formation of DNA adducts (Haugen et al. 1986, Strickland and Kang 1999, Aquilina et al. 2010); however, serum concentrations of parent BaP can be utilized as a general indicator of respiratory and dietary exposures. BaP was one of the most commonly detected PAHs in post-mortem blood from male and female Caucasians and African Americans (Ramesh et al. 2014), which ranged in concentration from 0.6 ng/mL to 10.5 ng/mL. A similar concentration range of 0-3.7 ng/mL was reported by Al-Daghri et al. (2014); however, Pleil et al. (2010), reported median and maximum BaP levels of 0.019 and 0.195 ng/mL, respectively, in plasma obtained from a healthy, non-smoking, adult population with no recent history of remarkable exposures. Thus, reported measures of BaP vary widely, and further characterization of BaP distribution in general and exposed populations is required.

Detected features within the metabolic profile were observed to associate with BaP serum levels and included alterations in immune and neurotransmitter related pathways. Histamine, which acts as a localized mediator of inflammation through regulation of four receptors and the corresponding intracellular signals (Jutel et al. 2006), was decreased with increasing BaP. Exposure to diesel fuel particles has been shown to increase allergic inflammation in both animal models and humans through increased local antigen-specific immunoglobulin E and

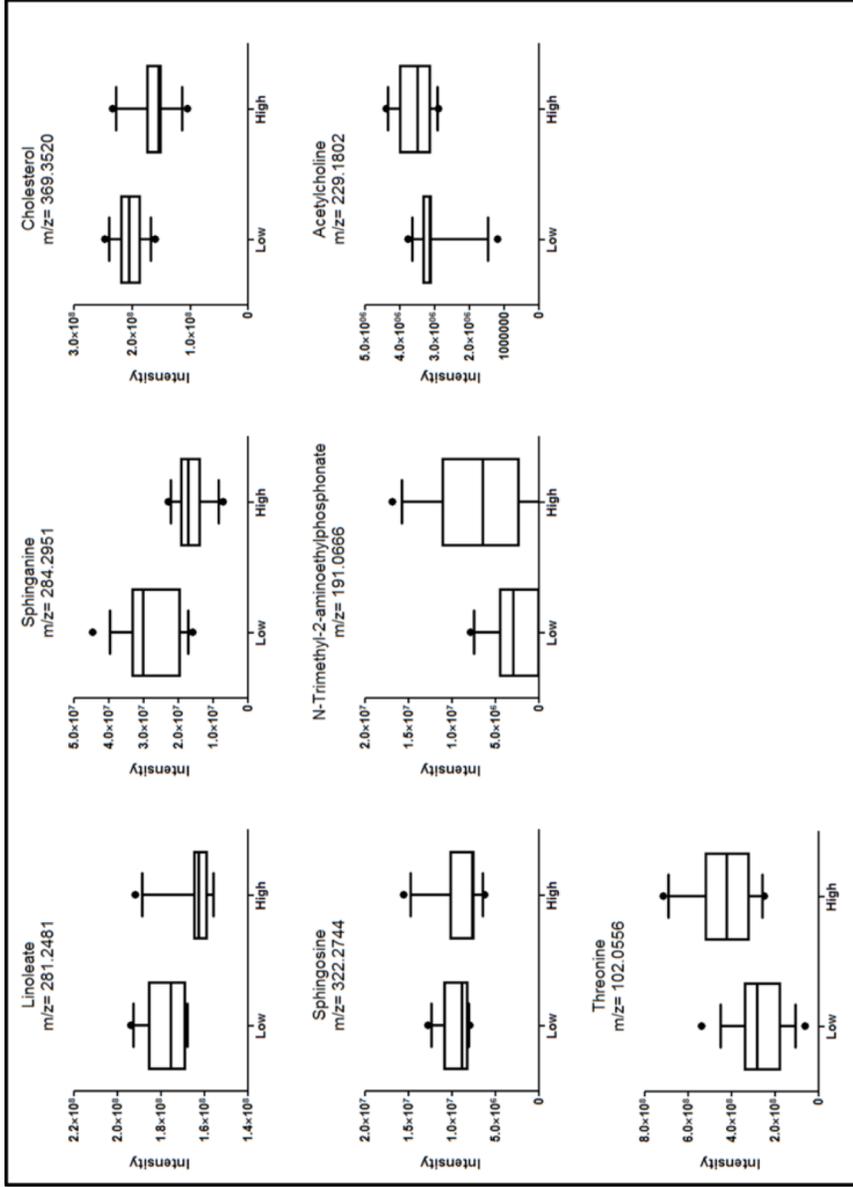


Figure 4.6. Metabolite intensity differences from glycerophospholipid metabolism for individuals in the high and low BaP clusters.

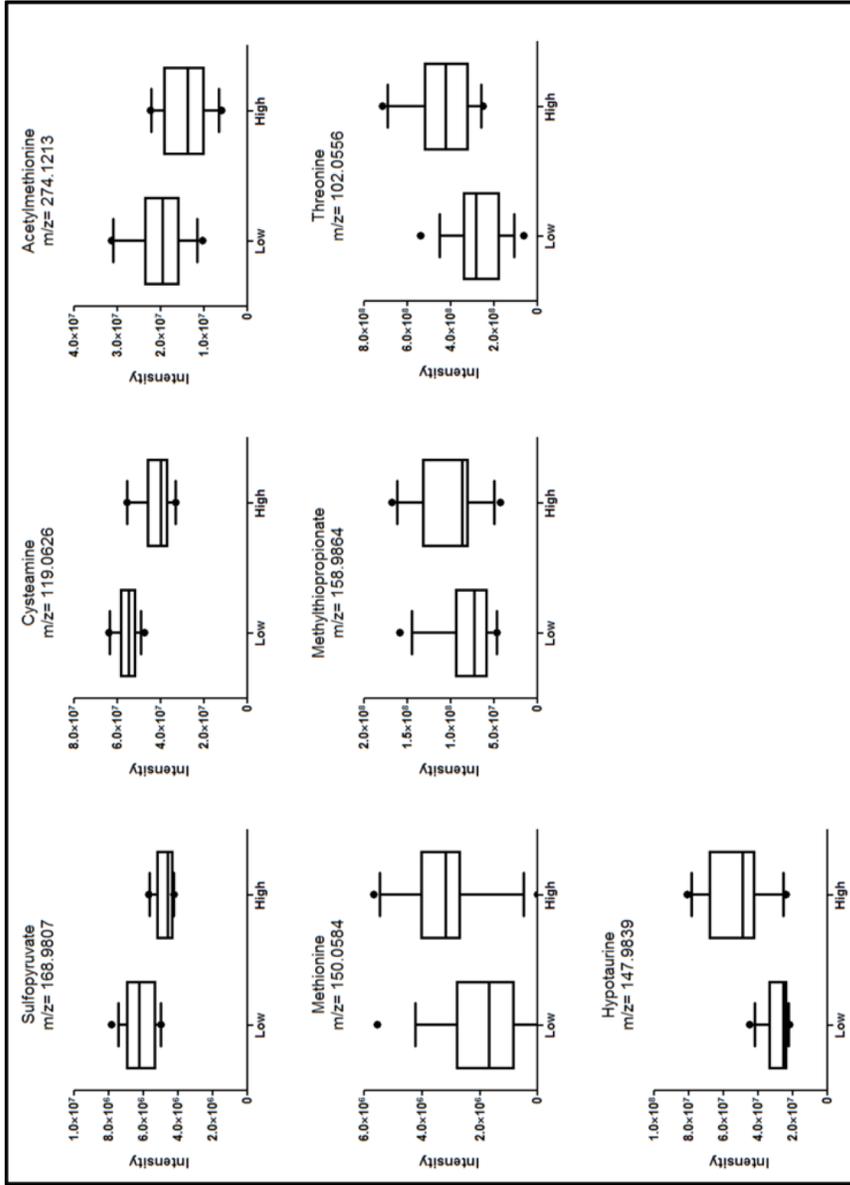


Figure 4.7. Metabolite intensity differences from sulfur amino acid metabolism for individuals in the high and low BaP clusters.

cytokine release (Diaz-Sanchez 1997, Nel et al. 1998, Ohtoshi et al. 1998). In addition, animal and *in vitro* models of BaP exposure have indicated increases in histamine production following exposure periods (Burchiel and Luster 2001, Park et al. 2011). While histamine was decreased in our study, association with BaP level nonetheless suggests changes in metabolic pathways important to immune response. Threonine, which is utilized for mucin synthesis and immunoglobulin A production, was increased with higher serum levels of BaP. Dihydrobiopterin, which is produced during synthesis of tyrosine, dopamine and serotonin, was also increased within the higher exposed group. Animal models indicate that elevated serum levels of dihydrobiopterin led to endothelial nitric oxide synthase dysfunction (Noguchi et al. 2011). Association of these metabolites with BaP suggests a possible relationship to both immune and cardiovascular function, consistent with known pathophysiology of respiratory exposure to combustion products.

Pathway enrichment identified associations with both lipid/fatty acid and oxidative metabolic pathway intermediates. For example, seven significant metabolic intermediates from glycerophospholipid metabolism (**Figure 4.6**) were identified in the enrichment analysis. Linoleate, a polyunsaturated, essential fatty acid utilized for biosynthesis of prostaglandins and cell membranes was decreased in the higher BaP groups, as well as sphinganine and sphingosine, which provide critical roles in ceramide lipid formation, and free cholesterol. Glycerophospholipid metabolism has previously been identified as an altered pathway following metabolic profiling of a BaP rat exposure model (Wang et al.

2014), and included differential expression of sphinganine and phytosphingosine. The authors attributed this to possible disruption to cell membranes based on the decreased levels of serum ceramide (d18:0/14:0) and increased sphinganine and phytosphingosine by sphingomyelinases alterations. This trend was not observed in the present study. Lipid profiling of human mammary cells dosed with BaP also identified alterations in sphingomyelins and glycerophospholipids (Jungnickel et al. 2014), which was determined to be receptor independent and metabolically associated. BaP was also shown to alter cholesterol-rich micro domains, resulting in changed fatty acid composition and reduced cholesterol in rat liver epithelial cells. The data strongly suggests BaP exposures alter blood lipid composition, which was observed in APoE +/- knockout mice (Curfs et al. 2003) and suggests the relationship is not due to binding and co-transport on lipoproteins alone.

Oxidative metabolic pathway associations with BaP included Cys and Met metabolism (**Figure 4.7**), an important component of central amino acid metabolism and antioxidant protection. Cys related intermediates include sulfopyruvate, cysteamine and hypotaurine. Cysteamine is a degradation product of cysteine, and acts as a precursor to hypotaurine, which in turn is oxidized to taurine. Hypotaurine and cysteamine are expected to function as protective antioxidants through oxidation of the sulfinic group (Aruoma et al. 1988, Fontana et al. 2004) through scavenging of OH, HOCl and H₂O₂. Relative expression levels of cysteamine and hypotaurine are in agreement with the two metabolites acting as protective antioxidants. Cysteamine was reduced in the high exposure

group and negatively correlated with hypotaurine. Metabolism of BaP causes formation of a carbon-6 localized radical cation (Cavalieri and Rogan 1995), which significantly contributes to DNA binding and adduct formation. It is also possible that the oxidation of cysteamine to hypotaurine occurred during storage; however, the high degree of enrichment within this pathway and dose dependent change with BaP suggests this behavior is due to biological response and not a sample storage artifact.

Met, a biochemical precursor to Cys, was elevated within the higher exposed group. Met is an essential amino acid, and a key metabolic precursor. Met is active in immune regulation, polyamine metabolism, DNA methylation (Cavuoto and Fenech 2012) and glutathione (GSH) formation, which detoxifies BaP via GSH S-transferase catalyzed conjugation of BaP (Srivastava et al. 1999). Met has also been evaluated for a chemoprotective effect against BaP genotoxicity using HepG2 cells (Roh et al. 2012). Increased Met levels were found to reduce uptake of BaP while increasing discharge of BaP DNA adducts, up-regulating intracellular GSH and possibly altering protein expression contributing to formation. In our study, acetyl-Met, which acts as a bioavailable form of Met (Boggs 1978), was down-regulated within the higher serum level group. Previous studies have shown that acetyl-Met treatment can prevent liver toxicity by protecting against GSH depletion and stabilizing cellular methylation (Skoglund et al. 1986, Lertratanangkoon and Scimeca 1993, Smith et al. 2011). Acetyl-Met is present and detectable within mammalian tissues and body fluids

(Smith et al. 2011), and the decreased level could arise due to a protective liver effect, a target organ for BaP toxicity.

4.6. Conclusions

The present study shows that DoDSR samples are of sufficient quality for chemical profiling of DoD personnel and identification of the BaP-associated metabolic perturbations. Insight into biological perturbations enables a basis for understanding possible health risks of environmental exposures and changes in metabolism that can result in manifestation of exposure-related disease pathophysiology. The ability of archived DoDSR samples to provide relevant biomarkers of exposure and effect make possible the assessment of exposures in military personnel while avoiding the additional burden and risk from the use of personal monitoring equipment.

While the lack of exposure data and small sample size limit the ability to generalize conclusions, the results show that the approach enables detection of metabolic alterations in humans that have either been observed experimentally in animal models or are physiologically plausible based upon existing knowledge. Thus, the data establish feasibility to use DoDSR serum samples for retrospective exposure surveillance. Future integration of this approach with well characterized, larger sample sizes will enable sequencing for both environmental surveillance and bioeffect monitoring.

Chapter 5. HIGH-RESOLUTION METABOLOMICS OF OCCUPATIONAL EXPOSURE TO TRICHLOROETHYLENE

This chapter is taken from a manuscript that was published in the *International Journal of Epidemiology* (PMID: 27707868). This manuscript represents the first untargeted HRM characterization of shift-length exposure to trichloroethylene (TCE), a degreasing agent and the most commonly detected contaminant in ground water. The results establish HRM as a central platform linking personal exposure to internal dose and biological response. Using untargeted, HRM analysis of blood obtained from exposed workers, personal TCE exposure levels were linked to internal dose and perturbations in endogenous metabolism. A systemic metabolic response to TCE was observed in exposed workers, which included a large number of unidentified chlorinated chemicals and alterations in endogenous metabolism consistent with known toxic targets, including renal, liver and the immune system. This MWAS of TCE represents the first example of using HRM to characterize dose-dependent changes in the metabolome from chemical exposures. The population was selected to represent an acute exposure, with TCE levels assessed over the length of an 8 to 10-hr work shift followed by immediate collection of biofluids. Co-authors who directly and contributed to this work and efforts are as follows. D. Walker: Designed and completed statistical

analysis, interpreted results, performed literature review, wrote all sections, created figures and tables, approved final edits and submission; K. Uppal: Provided software for network integration and annotation; L. Zhang: Read and edited final document; R. Vermeulen: Provided input and guidance for manuscript structure, edited final document; W. Hu: Provided population data; K. Pennell: Supervised D. Walker, provided critical input for manuscript structure, edited intermediate drafts, edited final document; D. Jones: Supervised D. Walker, oversaw HRM measures and efforts, provided critical input and guidance for manuscript structure, edited intermediate drafts, edited final document; N. Rothman: Designed and participated in sample collection for original occupational exposure study, provided critical input and guidance for manuscript structure and statistical analysis, edited intermediate drafts, edited final document; Q. Lan: Designed and participated in sample collection for original occupational exposure study, provided critical input and guidance for manuscript structure and statistical analysis, edited intermediate drafts, edited final document.

5.1. Abstract

Introduction: Occupational exposure to trichloroethylene (TCE) has been linked to adverse health outcomes including non-Hodgkin's lymphoma, kidney and liver cancer; however, TCE mode of action for development of these diseases in humans is not well understood.

Methods: Untargeted metabolomics analysis of plasma obtained from 80 TCE exposed workers (full shift exposure range of 0.4 to 230 ppm_a) and 95 matched controls were completed by ultra-high resolution mass spectrometry. Biological response to TCE exposure was determined using a metabolome wide association study (MWAS) framework, with metabolic changes and plasma TCE metabolites evaluated by dose-response and pathway enrichment. Biological perturbations were then linked to immunologic, renal and exposure molecular markers measured in the same population.

Results: Metabolic features associated with TCE exposure included known TCE metabolites, unidentifiable chlorinated compounds and endogenous metabolites. Exposure resulted in a systemic response in endogenous metabolism, including disruption in purine catabolism and decreases in sulfur amino acid and bile acid biosynthesis pathways. Metabolite associations with TCE exposure included uric acid ($\beta= 0.13, p= 3.6 \times 10^{-5}$), glutamine ($\beta= 0.08, p= 0.0013$), cystine ($\beta= 0.75, p= 0.0022$), methylthioadenosine ($\beta= -1.6, p= 0.0043$), taurine ($\beta= -2.4, p= 0.0011$) and chenodeoxycholic acid ($\beta= -1.3, p= 0.0039$), which are consistent with known

toxic effects of TCE, including immunosuppression, hepatotoxicity and nephrotoxicity. Correlation with additional exposure markers and physiologic endpoints supported known disease associations.

Conclusions: High-resolution metabolomics correlates measured occupational exposure to internal dose and metabolic response, providing insight into molecular mechanisms of exposure-related disease etiology.

5.2. *Introduction*

Trichloroethylene (TCE) is a widely used industrial solvent and common organic contaminant in groundwater (Moran et al. 2007). TCE is classified as carcinogenic to humans by the IARC (Guha et al. 2012) and both a carcinogenic and non-carcinogenic health hazard by the US EPA (Chiu et al. 2013). TCE exposure has been linked to increased risk for kidney cancer (Charbotel et al. 2006, Moore et al. 2010), non-Hodgkin's lymphoma (NHL) (Raaschou-Nielsen et al. 2003, Purdue et al. 2011) and liver cancer (Scott and Jinot 2011, Hansen et al. 2013). Human and animal studies have identified numerous non-cancer adverse effects of TCE, including immune dysfunction (Iavicoli et al. 2005, Cooper et al. 2009), nervous system (Kilburn 2002), renal (Lash et al. 1995, Bruning and Bolt 2000) and liver (Brautbar and Williams 2002) toxicity.

The biological plausibility of TCE toxicity has been verified in animal and *in vitro* studies (Chiu et al. 2013), but limited data exist establishing biochemical changes in humans. Recently, cross-sectional molecular epidemiology studies evaluating TCE exposure on immune function (Lan et al. 2010, Bassig et al. 2013, Zhang et al. 2013) and nephrotoxicity (Vermeulen et al. 2012) have observed effects that occurred at levels of exposure below the current US OSHA permissible limits of 100 parts-per-million of air (ppm_a). Participants were selected using extensive screening procedures to avoid confounding exposures, past occupational use of volatile organic compounds (VOC) and previous history of cancer. The results indicated immunosuppression consistent with increased NHL risk (Lan et al. 2010), elevated urinary nephrotoxicity markers (Vermeulen

et al. 2012) and immunotoxic effects including decreased serum immunoglobulin G (IgG), immunoglobulin M (IgM) and IL-10 (Bassig et al. 2013, Zhang et al. 2013). These findings suggest a complex biochemical response to TCE exposure with multiple targets of toxicity.

Comprehensive profiling of the metabolic phenotype associated with TCE exposure can provide insight into the biochemical changes occurring within exposed workers. High-resolution metabolomics (HRM) using ultra-high resolution mass spectrometry (Makarov et al. 2006, Scigelova et al. 2011) with data extraction algorithms (Uppal et al. 2013, Yu et al. 2013, Yu and Jones 2014, Libiseller et al. 2015) now enables measurement of >10,000 chemicals in biological samples with quantitative reproducibility (Go et al. 2015). With mass spectrometry based HRM platforms, chemicals are detected as ions in a gas phase and measured as a mass-to-charge ratio (m/z). The m/z signals arise from interaction of a neutral molecule with cations present in solution, often resulting in multiple signals corresponding to a single chemical species. Thus, initial measures are most accurately described as m/z feature, which is defined by the accurate m/z , retention time and associated intensity and a non-targeted analytic structure is used, with chemometric approaches for selection and identification of metabolic species present in a sample. The resulting chemical profile provides in-depth coverage of the metabolic phenotype, including detection of metabolic intermediates, dietary chemicals, xenobiotics and microbiome-related metabolites (Holmes et al. 2008, Holmes et al. 2008, Nicholson et al. 2008, Psychogios et al. 2011, Jones et al. 2012, Park et al. 2012, Scalbert et al. 2014, Athersuch and Keun

2015). A previous metabolomic study of TCE exposure in mice linked liver-related effects to increased expression of peroxisome proliferator-activated receptor α (PPR α) target genes, which influences both inflammatory response, cell proliferation, lipid metabolism and glucose metabolism (Fang et al. 2013).

In this study, we used a metabolome wide association study (MWAS) (Nicholson et al. 2008, Go et al. 2014, Burgess et al. 2015) to identify dose-dependent metabolic changes in TCE-exposed workers and unexposed controls using a well characterized, worker population (Lan et al. 2010, Vermeulen et al. 2012). Direct TCE and endogenous metabolites were characterized, and pathway analysis was used to identify biologic response to TCE exposure. We further examined metabolite association with separately measured TCE exposure, immunological and nephrotoxicity markers.

5.3. Methods

5.3.1. Study design and exposure measurement

Samples were collected in 2006 as part of a cross-sectional study conducted in Guangdong, China to assess the early biological effects of occupational exposure to TCE; a full description of factory and subject selection is described in Lan et al. (2010). Prior to selection of workplaces, monitoring was performed to assure there were no exposures to benzene, styrene, ethylene oxide, formaldehyde or epichlorohydrin and negligible levels of other chlorinated solvents. Factories selected for study included metal (n= 4), optical lens (n= 1) and circuit board (n= 1) cleaning processes. Control workers were enrolled from two clothes manufacturing factories, one food production factory and a hospital within the same geographic area, and were frequency matched based on age (\pm 5 years) and sex. Exclusion criteria were history of cancer, chemotherapy and radiotherapy, in addition to previous occupations with notable exposure to benzene, butadiene, styrene and/or ionizing radiation. The study was approved by Institutional Review Boards at the US National Cancer Institute and the Guangdong National Poison Control Center, China.

In June and July of 2006, replicate full-shift personal air exposure measurements were taken for 80 exposed workers and 95 controls using 3M organic vapor-monitoring badges. All samples were analyzed for TCE, with a subset (48 from TCE exposed workers) selected for additional VOCs, including benzene, methylene chloride, perchloroethylene and epichlorohydrin. In two of the factory locations, workers intermittently used respirators; only one subject

wore gloves. Each study subject was asked to provide a 29 mL peripheral blood sample, post-shift urine sample, overnight urine sample, buccal cell mouth rinse sample and undergo a brief physical exam evaluating blood pressure, height, weight, temperature and signs of current upper or lower respiratory infection. A questionnaire requesting information about occupational history, demographic characteristics and lifestyle was also completed by each subject. Organic vapor monitoring samples were obtained on a subgroup of control workers in the food and clothes production factories

5.3.2. High-resolution metabolomics

Plasma was prepared and analyzed daily in batches of 20 by HRM with C₁₈ liquid chromatography using the methodology of Go et al. (2015). Data was extracted with apLCMS (Yu et al. 2009) and xMSanalyzer (Uppal et al. 2013). Full details are provided in **Chapter 3.2**. From the 10,017 ions detected we removed all features that were not detected in greater than 50% of the individuals from at least one group. The remaining 7,830 *m/z* features were log₂ transformed and used for defining a TCE exposure metabolic phenotype.

5.3.3. Data analysis

Statistical analysis was carried out using R version 3.1.2 (R Core Team 2014). A flow diagram detailing all data analysis steps is provided in **Figure 5.1**.

Step 1: TCE exposure MWAS

The TCE exposure MWAS was completed using a linear regression framework. For each m/z feature, the \log_2 transformed intensity was used to test for dose-response across the categories of exposure, which included controls, low exposed ($<12 \text{ ppm}_a$) and high exposed ($\geq 12 \text{ ppm}_a$), defined as a continuous variable (control=0; $< 12 \text{ ppm}_a= 1$; $\geq 12 \text{ ppm}_a= 2$). Equal spacing between the groups was used to bias feature selection towards those showing dose-response while providing representative metabolic associations with TCE exposure across the population and reducing false positives. The statistical model including adjustments for age (continuous), sex (factor) and body mass index (BMI; continuous), which are known to influence chemical disposition and toxicokinetics (Ernstgard et al. 2003, Gochfeld 2007). To evaluate if worker smoking and alcohol use should be considered, the data was re-analyzed separately with smoking and alcohol use status included as a covariate. The resulting analysis increased the number of m/z features with $\text{FDR} \leq 20\%$ by one and eight when accounting for smoking and alcohol use, respectively. Thus, neither was included as covariates in the regression model. To account for multiple comparisons, we applied a Benjamini and Hochberg (1995) false discovery rate (FDR) threshold of 20% (raw p -value= 4.4×10^{-3}), which controls the rate of false findings rather than falsely rejected null hypotheses (Vinaixa et al. 2012, Patel 2016).

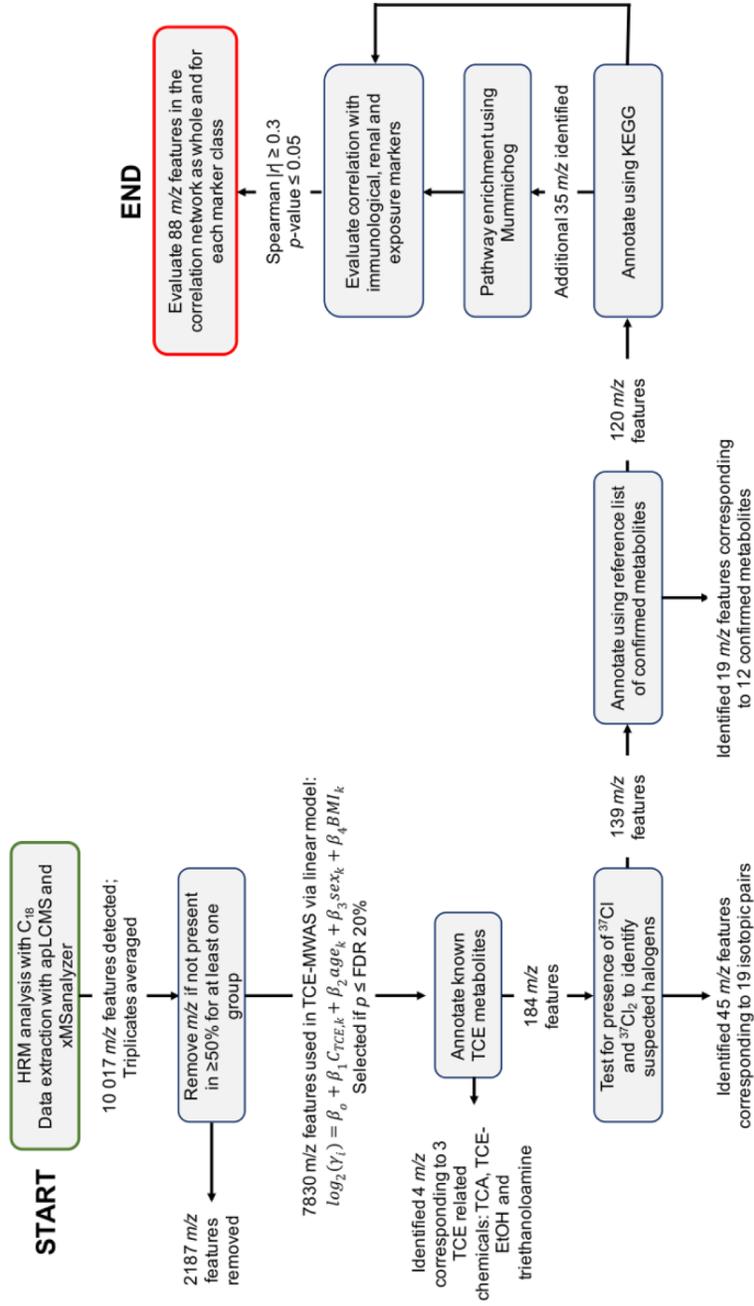


Figure 5.1. Workflow of HRM data analysis and m/z feature characterization for TCE MWAS results

Step 2: Identification of TCE exposure products

We first characterized the m/z features associated with TCE exposure to identify direct exposure products by matching the mass for common positive electrospray ionization adducts of TCE additives and metabolites (Lash et al. 2000) at match accuracy of ± 10 ppm ($\pm 10e-6 \times$ theoretical m/z mass) to TCE associated m/z features. Unidentified masses were tested for the presence of ^{37}Cl and $^{37}\text{Cl}_2$ isotopes with the *pattern.search()* function in the R package nontarget (Loos 2015) with retention time tolerance of ± 10 s and mass accuracy of ± 5 ppm.

Step 3: Biological response to TCE exposure

The remaining m/z features were matched to a reference database of 75 metabolites previously confirmed with MS² and co-elution studies (Go et al. 2015). Additional features not matching these metabolites were annotated with the KEGG database (Kanehisa et al. 2012, Tanabe and Kanehisa 2012), which provides information on 487 pathways containing 17,620 unique metabolites and is a common metabolic reference for metabolomic studies. Identities were assigned using evidence scoring provided in *Mummichog* (Li et al. 2013) for matching to +H, +Na, +K, -H₂O+H, -2H₂O+H, +ACN+H, +ACN+Na, +2ACN+H, +2Na-H and +2H adducts at ± 10 ppm mass tolerance. Enriched metabolic pathways were selected using *Mummichog* scoring threshold of 0.05.

Step 4: Integration with TCE exposure and physiological markers

Correlation for each of the 188 m/z features with biomarkers of TCE exposure, immune function and renal damage were calculated using MetabNet (Uppal et al. 2015) in exposed workers only. These measures have been described previously (Kim et al. 2009, Lan et al. 2010, Vermeulen et al. 2012, Zhang et al. 2013) and include white blood cell count (WBC), lymphocytes (LY), CD4+ T-cells (CD4), soluble CD27, soluble CD30, mitochondrial DNA copy number (mtDNA), IgG, IgM, urinary isoenzymes of glutathione-s-transferase (α GST, piGST), kidney injury molecule 1, (KIM-1), N-acetyl-beta-glucosaminidase (NAG), trichloroacetic acid (TCA), total, free and conjugated trichloroethanol (TCE-EtOH). All urinary measurements were normalized by creatinine to account for differences in urinary output (Cone et al. 2009). Marker-metabolite correlations were selected for consideration and plotting in Cytoscape (Su et al. 2014) based upon effect size (Spearman $|r| \geq 0.3$) and p -value threshold ≤ 0.05 . Due to the use of effect size for prioritizing network connectivity, correction for multiple hypothesis testing was not applied.

Table 5.1. Demographic characteristics and TCE exposure level

Subjects	Controls Controls (n= 95)	Exposed		
		Total (n= 80)	< 12 p.p.m (n= 39)	≥ 12 p.p.m (n= 41)
<i>Demographic characteristics</i>				
Age, mean (SD) ^a	27 (7)	25 (7)	24 (5)	27 (8)
BMI ^b , mean (SD) ^a	22 (3)	21 (3)	21 (2)	22 (3)
Current smoking, n (%) ^c	37 (39%)	34 (43%)	17 (44%)	17 (41%)
Alcohol use, n (%) ^c	40 (42%)	26 (33%)	13 (33%)	13 (32%)
Sex, n (%) ^c				
Female	23 (24%)	23 (29%)	15 (38%)	8 (20%)
Male	72 (76%)	57 (71%)	24 (62%)	33 (80%)
<i>TCE exposure</i>				
TCE air level (ppm _a), mean (SD) ^d	< 0.03	22.19 (35.9)	5.19 (3.5)	38.36 (44.6)
Minimum exposure	NA	0.4	0.4	12.0
Maximum exposure	NA	229	11.7	229

^aMean ± standard deviation; ^bBMI, body mass index; ^cNumber

^dMean ± standard deviation of TCE exposure levels, expressed as parts-per-million of air.

Table 5.2. Top *m/z* features matching chlorinated isotope patterns

<i>m/z</i> ^a	Time (s)	Identity	Isotopes ^b	Adjusted <i>r</i> ²	<i>β</i> coefficient	<i>P</i>
206.8750	81.3	Trichloroacetic acid ^c	M+2, M+4	0.76	9.87	3.1E-55
224.8456	88.4	Trichloroethanol ^d	M+2	0.68	8.37	1.0E-43
264.8338	78.7	Unknown	M+2, M+4	0.55	7.16	9.8E-32
332.8201	83.3	Unknown	M+2, M+4	0.43	5.69	3.9E-23
280.8067	88.5	Unknown	M+2, M+4	0.41	5.08	1.2E-21
324.7895	77.2	Unknown	M+2	0.39	5.55	7.2E-21
340.7621	77.1	Unknown	M+2, M+4	0.32	4.11	7.1E-16
274.8629	84.8	Unknown	M+2	0.29	4.01	2.4E-15
392.7587	86.1	Unknown	M+2	0.20	4.14	3.2E-10
380.7691	52.6	Unknown	M+2, M+4	0.16	0.08	2.6E-04

Regression parameters for the monoisotopic mass only are listed. Regression results for all features with FDR ≤ 20% are provided in **Supplementary Text II**.

^amass-to-charge ratio

^bM+2= ³⁷Cl (+1.9971); M+4= ³⁷Cl₂ (+3.9941)

^cDetected as +2Na-H adduct form

^dDetected as +2K-H adduct form

5.4. Results

5.4.1. Study population

Demographics, including age, sex, BMI, current smoking and alcohol use status were comparable among the exposed and non-exposed workers (**Table 5.1**). The median exposure level of 12.0 ppm_a was used to define a low and high exposure threshold. Measured TCE levels for exposed factory worker levels ranged from 0.4 ppm_a to 229 ppm_a (**Table 5.1**); 96% of the workers exposed to TCE levels were below the OSHA 8-h limit.

5.4.2. TCE exposure MWAS

At FDR threshold $\leq 20\%$, MWAS identified 188 *m/z* features associated with TCE exposure (**Figure 5.2**). These features were selected for characterization and identification of the metabolic response to TCE exposure; 61 exhibited *p*-values $\leq 10^{-4}$ with adjusted *r*² ranging from 0.08 to 0.76. Regression coefficient for exposure categorization ranged from -5.4 to 9.9. The list of *m/z* features and regression results are provided in **Supplementary File 2**.

5.4.3. Identification of TCE exposure products

We tested for the presence of exposure products by searching the 188 *m/z* features for TCE metabolites from cytochrome P450, glutathione (GSH) conjugation by GSH S-transferase (Lash et al. 1995, Lash et al. 2000, Chiu et al. 2007) and nine common additives used in commercial formulations of TCE (IARC 2014). Three *m/z* features were consistent with TCE metabolites. A match

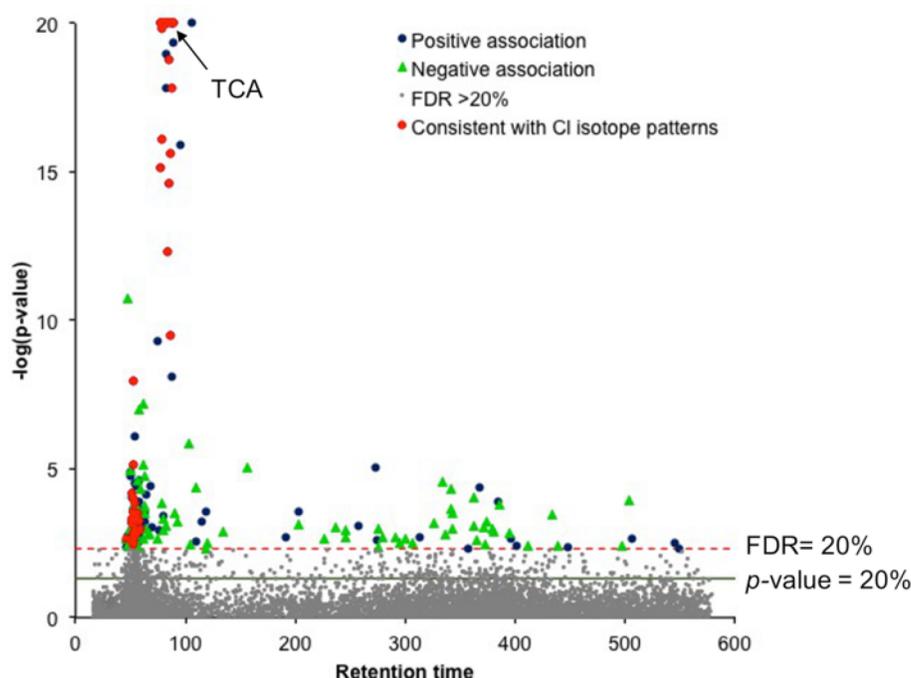


Figure 5.2. Manhattan plot for TCE MWAS in exposed and non-exposed workers. Linear regression identified m/z features that were associated with TCE exposure at $FDR \leq 20\%$ after correcting for age, sex and BMI. Each m/z features is colored by association ($0 \leq \beta \leq 0$). Those with isotopic patterns matching halogenated compounds (based upon the presence of ^{37}Cl and $^{37}\text{Cl}_2$ masses within a retention time window of ± 10 s) are indicated by a red triangle.

to the +2Na-H adduct of TCA was detected, and exhibited the highest association with exposure (**Table 5.2**). Additional adduct masses of TCA included the +Na+K form. The +2K-H adduct of TCE-EtOH was also present, as was the stabilizer triethanolamine.

TCA and TCE-EtOH alone did not explain m/z features associated with exposure, and the presence of features with a negative mass defect suggests additional chlorinated chemicals (Zhu et al. 2006, Jobst et al. 2013). To test for additional chlorinated metabolites, we applied isotopic pattern searching for m/z

Table 5.3. Identified endogenous metabolites associated with TCE exposure and relevant to biological response to exposure

m/z^a	Identity	ID level ^b	Adjusted r^2	β	p
439.3008	7 α -Hydroxycholest-4-en-3-one	3	0.12	-0.15	1.1E-04
394.3031	Chenodeoxycholic acid	4	0.03	-1.31	3.9E-03
132.0764	Creatine	1	0.21	-0.22	4.9E-05
241.0310	Cystine	1	0.05	-0.75	2.2E-03
191.0397	Glutamine	1	0.14	0.07	1.3E-03
307.0197	Homocystine	3	0.07	-0.06	2.4E-03
188.0705	Indolelactic acid	1	0.07	-0.16	2.9E-04
336.0527	Methylthioadenosine	3	0.09	-1.00	4.4E-03
400.3405	Palmitoylcarnitine	1	0.20	-0.18	9.3E-05
119.0488	Phenylacetic acid	1	0.05	0.11	8.0E-04
163.9776	Taurine	1	0.08	-2.39	1.1E-03
249.0609	Tryptophan	1	0.15	-0.15	6.5E-08
146.0596	Tyrosine	1	0.05	-0.09	1.5E-03
212.9993	Uric acid	1	0.34	0.13	3.6E-05
279.2310	α -Linolenic acid	1	0.03	-0.09	3.9E-03

a. When multiple adducts or isotopes were detected, the highest ranked m/z was used for regression parameters. All TCE exposure association m/z features are listed in **Supplementary File 2**

b. ID level indicates degree of annotation with 1: m/z and retention time matched to authentic standards confirmed with MS², 2: Multiple/isotopes present; 3: m/z matched single adduct mass within 10 ppm mass error, 4: m/z matched adduct mass of multiple isobaric species, most probable identification used.

ions exhibiting M+2 and M+4 mass spacing, which corresponds to the presence ³⁷Cl (+1.9971) and ³⁷Cl₂ (+3.9941). We identified 45 m/z features corresponding to 19 unique isotopic pairs (**Figure 5.2**). The top 10 monoisotopic m/z masses ranked according to p -values are provided in **Table 5.2**. Two included the C₂HCl₂O₂³⁷Cl and C₂HClO₂³⁷Cl₂ forms of TCA and the C₂H₃Cl₂O³⁷Cl isotope of TCE-EtOH; however, the remaining showed no matches to known TCE related metabolites.

5.4.4. *Biological response to TCE exposure*

After removal of the 49 m/z features identified as probable chlorinated chemicals, we annotated the remaining 135 m/z features. Nineteen m/z features

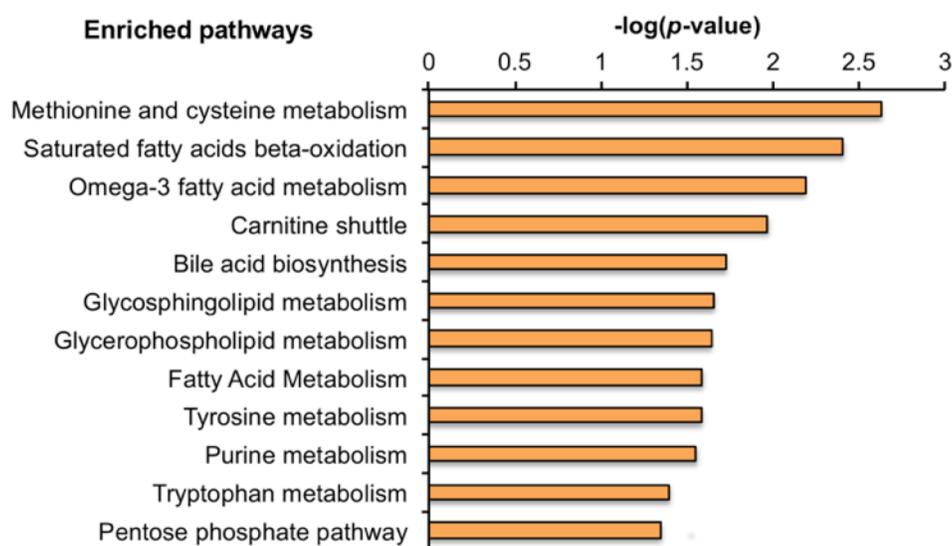


Figure 5.3. *Mummichog* enriched pathways for metabolites associated with TCE exposure.

corresponded to 12 confirmed metabolites (**Table 5.3**). The remaining had identities assigned based on systematic evidence scoring by isotopic and adduct pairing, which identified an additional 35 *m/z* features (**Supplementary File 2**). In total, 54 *m/z* features matching 46 unique metabolites were identified. The remaining had no probable matches.

We used *Mummichog* (Li et al. 2013) to evaluate metabolic pathway enrichment, which identified 12 pathways at *p* threshold ≤ 0.05 (**Figure 5.3**). Detected metabolites in each pathway are provided in **Supplementary File 2**. Metabolites from sulfur amino acid metabolism were decreased with exposure, and included cystine and homocystine (**Table 5.3**). Disruption in bile acid biosynthesis was also detected, with decreased levels of taurine, chenodeoxycholic acid and 7α -hydroxy-cholestene-3-one in exposed workers (**Table 5.3**). Changes in fatty acid metabolism were observed, including reduced levels of palmitoylcarnitine. Metabolites from purine metabolism, including uric

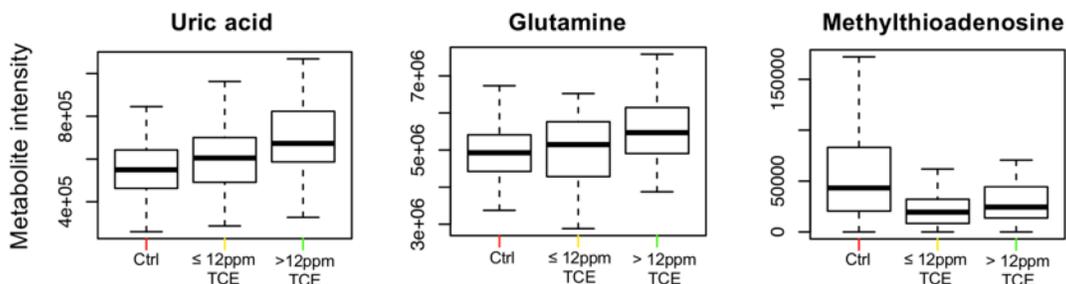


Figure 5.4. Distribution of uric acid, glutamine and MTA in control, low exposed and high exposed workers. Uric acid and glutamine increased in an exposure dependent manner, while MTA was decreased in all exposed workers relative to controls.

acid and glutamine were increased in a dose-dependent manner in association with TCE exposure, while methylthioadenosine (MTA), a metabolite of methionine and purine salvage, was decreased in TCE-exposed workers (**Figure 5.4**). Metabolic changes consistent with other physiological effects of TCE were also detected, including tryptophan and tyrosine (**Table 5.3**).

5.4.5. *Integration with immunological, renal and TCE markers*

To visualize systemic effects, we tested for metabolic associations with molecular markers of immune function, nephrotoxicity and TCE exposure. For all 188 *m/z* features selected through MWAS, 88 were correlated with at least one marker and included 48 identified as metabolites or halogenated chemical species (**Figure 5.5**). Excluding sCD27 and mtDNA, a high degree of connectivity was present. Chlorinated metabolites were negatively correlated with IgG, IgM and CD4+ T-cells, none of which matched known TCE biotransformation products. KIM-1 and additional nephrotoxicity markers were positively associated with chlorinated isotopes. A negative association of KIM-1 with uric acid was present, while MTA was negatively associated with the urinary isoenzymes. Correlation

with urinary markers of TCE exposure showed positive associations with unidentified and identified chlorinated metabolites, including TCA, TCE-EtOH and unidentified chlorinated chemicals.

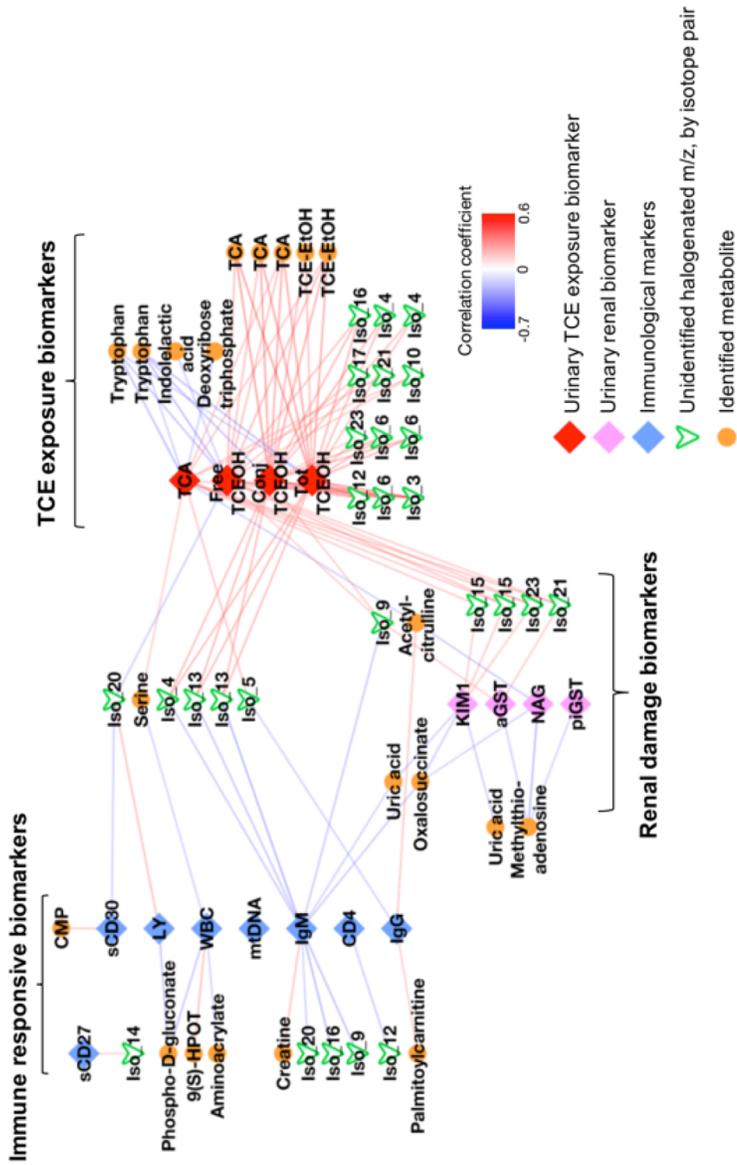


Figure 5.5. Network correlation analysis of molecular markers previously tested for association with TCE exposure and metabolic features detected in this study. Only correlations with Spearman $|r| \geq 0.3$ and p -value ≤ 0.05 corresponding to identified metabolites or probable halogens were included. Immune markers: CD4= CD4⁺ T-cell count, IgG= immunoglobulin G, IgM= immunoglobulin M, LY= lymphocyte cell count, mtDNA= mitochondrial DNA turnover rate, sCD27= soluble CD27, sCD30= soluble CD30, WBC= white blood cell count. Nephrotoxicity markers: aGST= urinary α glutathione-s-transferase, piGST= urinary pi glutathione-s-transferase, KIM1= urinary kidney injury marker 1, NAG= urinary n-acetyl-beta-glucosaminidase. TCE exposure markers: Conj. TCEOH: urinary trichloroethanol glucuronide, Free TCEOH: urinary trichloroethanol, Tot TCEOH: sum of urinary trichloroethanol and trichloroethanol glucuronide, TCA: urinary trichloroacetic acid.

5.5. *Discussion*

Metabolic phenotyping with HRM enables an unbiased approach to investigate biologic changes and mode of action for environmental and occupational exposures. The present study provides an application of HRM to occupational exposure effects in humans and demonstrates the utility of HRM to link VOC exposure to metabolic pathways associated with relevant disease. MWAS of TCE exposure showed alterations to metabolic features characterized as exposure byproducts and endogenous metabolites. In addition to detecting known metabolites of TCE, isotopic spacing analysis and strong positive association with exposure categorization suggests the presence of additional chlorinated metabolites, likely derived from TCE. While the chemical identity of these metabolites was not determined, it suggests the existence of additional, uncharacterized TCE related metabolic products.

Pathway enrichment of metabolites associated with exposure was consistent with TCE detoxification. The presence of cysteine and methionine metabolites is in agreement with metabolism of TCE by GSH conjugation. Association with bile acid biosynthesis is likely related to detoxification of TCE by glucuronidation, with TCE rapidly metabolized to TCE-EtOH in the liver, glucuronidated by glucuronyltransferase and secreted in bile. TCE-EtOH glucuronide undergoes enterohepatic recirculation (Chiu et al. 2013) followed by metabolism to TCA (Stenner et al. 1997). Blood bile acid alterations have been identified as a marker of liver damage and hepatotoxicity (Li and Chiang 2012, Lake et al. 2013). Exposure related alterations in fatty acid metabolism and

reduced palmitoylcarnitine suggest alteration in peroxisome proliferator-activated receptor α , which was also observed in animal TCE exposure models (Fang et al. 2013).

Enriched pathways provided additional evidence of toxicity, with the identified endogenous metabolites relevant to diseases associated with TCE exposure. Elevated levels of blood uric acid have been identified as both an independent risk factor and marker of kidney disease (Mohandas and Johnson 2008, Sonoda et al. 2011). TCE related alterations in purine catabolism resulting from changes in expression of adenosine deaminase or purine nucleoside phosphorylase could account for the immunosuppressive effects observed by Lan et al. (2010). Both enzymes are essential for developing and maintaining immune function (Moriwaki et al. 1999, Cristalli et al. 2001), with abnormalities reported in immunodeficiency, leukemia diseases and neoplastic transformations (Smyth et al. 1978, Segura et al. 1989, Koizumi and Ohkawara 1992, Kopff et al. 1993). The detected TCE associated changes in plasma glutamine further suggest alterations in immune-related purine catabolism rather than hepatic cells, since glutamine is the primary transporter of ammonia to the liver following deamination of guanine and adenine (Mathews et al. 2012). MTA can act as a secondary precursor to adenine biosynthesis through 5'-methylthioadenosine phosphorylase (MAP) (Avila et al. 2004). Disruption to MAP has been observed in tumor cells, which rely on *de novo* purine biosynthesis for adenine and DNA synthesis (Nobori et al. 1996). MTA has been observed to be hepatoprotective following treatment with carbon-tetrachloride (CCl₄) (Simile et al. 2001). Interestingly, MTA enhances

production of IL-10 (Avila et al. 2004), and MTA alterations in exposed workers were consistent with TCE related decreases in IL-10 (Bassig et al. 2013) . Finally, TCE is known to have central nervous system (CNS) depressant effects, with *in vitro* and long-term exposure associated with dopaminergic neurodegeneration (Liu et al. 2010, Zaheer and Slevin 2011, Caudle et al. 2012). Tryptophan and tyrosine both act as precursors for the synthesis of neurotransmitters, and were decreased with exposure

TCE metabolic products are transported through multiple tissues and organ systems (Chiu et al. 2013). The presence of metabolic features consistent with unidentified chlorinated chemicals indicates a complex detoxification response to TCE exposure. A different correlation network for each class of markers suggests the possibility of uncharacterized TCE metabolites being toxic to specific organ systems. Association with immune function markers was consistent with the immunosuppressive effects observed by Lan et al. (2010) and Zhang et al. (2013). Chlorinated metabolites were negatively correlated with IgG, IgM and CD4+ T-cells, suggesting unidentified metabolites could be directly involved in mediating immunotoxicity. KIM-1 and additional nephrotoxicity markers were positively associated with chlorinated isotopes, consistent with the findings of Vermeulen et al. (2012), while plasma uric acid and MTA showed negative correlation. The negative association of uric acid with KIM-1 was unexpected since urinary expression of KIM-1 is increased with tubular necrosis and various renal diseases, suggesting perturbed purine metabolism as a potential causal mechanism. A similar trend was observed for MTA and the urinary

isoenzymes. α GST and piGST are shed from the proximal and distal tubular cells, respectively, and are upregulated following kidney damage (Branten et al. 2000). A large number of suspected chlorinated chemicals endogenous metabolites, were correlated with urinary TCA and TCE-EtOH measures. While a correlation for identified TCE metabolites (TCA and TCE-EtOH) with immune and renal markers was not detected, many unidentified chlorinated chemicals were, suggesting greater toxicity to immune and renal systems.

We acknowledge some limitations in this work. First, there is a predominance of males in our study population; thus, while sex was included as a covariate in the MWAS model, sex based differences were not evaluated. Second, a limited sample population of 175 individuals from a specific geographical location was used, and an independent cohort was not available to replicate the findings. Third, this study was focused on known TCE and endogenous metabolites because the limited sample availability precluded in-depth structural characterization of unidentified metabolites. Lack of reference standards and low abundance will make structural elucidation of these features challenging. Finally, the results of this study are correlative in nature. We could not account for unknown confounders, nor identify the exact mechanism through which these metabolic associations occurred. Despite these limitations, MWAS identification of pathways known to participate in detoxification of TCE and pathways consistent with toxicological targets of TCE suggests biological relevance. In addition, molecular markers previously shown to be associated with TCE exposure were correlated with TCE metabolites and exposure associated

endogenous metabolites. The results therefore show that HRM provides a useful approach to link occupational exposures to metabolic perturbations and obtain leads to underlying mechanisms of exposure-related disease.

5.6. Conclusions

Environmental and workplace exposures are a significant contribution to chronic disease burden. This study of a population with rigorous selection criteria and TCE exposure assessment shows that HRM is able to link internal TCE metabolites and perturbations of endogenous metabolism with isolated occupational TCE exposure. The results show associations with unidentified chlorinated products as well as a multiple mechanistic and disease markers of hepatic function, kidney damage and immune dysfunction. While our study was limited to a high, occupational exposure, the metabolic associations represent an important first step in identifying how TCE alters metabolism and leads to disease risk. More generally, the results establish the feasibility to use HRM as an occupational surveillance tool to assist in epidemiologic studies of specific exposure risks, underlying toxic mechanisms and exposure related disease susceptibility.

**Chapter 6. INTEGRATED MOLECULAR RESPONSE OF EXPOSURE
TO TRAFFIC-RELATED POLLUTANTS IN THE US TRUCKING
INDUSTRY**

Collaborators who contributed to this work and efforts are as follows. D. Walker: Designed and completed all presented data analysis, interpreted results, performed literature review, wrote manuscript, created figures and tables; J. Hart: Co-I of original study, assisted in exposure monitoring and biofluid collection, organized and provided population data; B. E. Garshick: Co-I of original study, assisted in design of exposure monitoring and biofluid collection; F. Laden: PI, responsible for all aspects of original study design; K. Pennell: Supervised efforts of D. Walker; D. Jones: Supervised D. Walker, oversaw HRM measures and efforts

6.1. Abstract

Introduction: Exposure to traffic-related pollutants is associated with increased risk of cardiopulmonary disease and mortality. Suspected mechanisms include endothelial dysfunction, oxidative stress and inflammation; however, the precise biochemical pathways from exposure to effect and the how different constituents of air pollution contribute to these adverse health outcomes are not known. In the present study, we used an integrated molecular response approach that included high-resolution metabolomics and peripheral blood gene expression to identify biological response to daily and week-averaged diesel exhaust exposure.

Methods: Blood samples collected over the course of a workweek from 73 non-smoking males employed in the US trucking industry were analyzed using untargeted, high-resolution mass spectrometry and tested for association with daily and week-averaged microenvironment levels of elemental carbon (EC), organic carbon (OC) and particulates with diameter $\leq 2.5 \mu\text{m}$ (PM_{2.5}). Metabolic associations with each pollutant were evaluated for the presence of environmental chemicals and changes to biochemical processes associated with disease risk. Identified metabolites were then tested for relationship with the peripheral blood transcriptome using multivariate variable selection and network correlation.

Results: No detected metabolic associations with PM_{2.5} were present at the false discovery rate threshold of 20%. Week averaged EC and OC resulted in the greatest exposure-associated metabolic alterations, with changes detected in

plasma levels of chemicals present in exhaust emissions and metabolites from central biochemical processes. EC exposure included dose-related changes for 845 metabolic features consistent with increased lipid peroxidation, biomarkers of oxidative stress, thrombotic signaling lipids, and metabolites associated with disruption to nitric oxide production. OC exposure resulted in a different pattern of biological response, with 650 metabolic features suggesting alterations to antioxidant levels, oxidative stress biomarkers and intermediates in nitric oxide production. Network correlation with post-shift gene expression results provided additional evidence of changes in processes related to endothelial function, immune response, inflammation and oxidative stress.

Conclusions: This study provides the first integrated assessment of molecular response to environmental exposure in humans. Identified metabolic and transcriptomic changes associated with exposure to traffic-related pollutants are consistent with increased risk of cardiovascular diseases and the adverse health effects of air pollution.

6.2. Introduction

Exposure to traffic-related pollution has been linked to increased risk of all-cause mortality, cardiovascular disease (CVD), cardiopulmonary outcomes and lung cancer mortality in epidemiological studies (Pope et al. 2002, Pope et al. 2009, Hart 2016), with 2010 estimates suggesting 3.1 million deaths were attributable to ambient particulate exposure alone (Lim et al. 2012). Suspected underlying mechanisms that contribute to disease risk include increased oxidative stress, endothelial dysfunction and inflammation; however, study in human populations has shown conflicting results (Donaldson et al. 2001, Sorensen et al. 2003, Chuang et al. 2007, Brugge et al. 2013, Bates et al. 2015). One of the primary challenges in addressing the role of these mechanisms is the difficulty inherent in estimating exposure to vehicle emissions and linking these levels to biomarkers of disease processes. Observational studies with long-term data are complicated by low-resolution exposure estimates and availability of biological samples. Assessment of short-term exposure may not be representative of real-life timescales and exposure, resulting in an acute response that provides a measure of immediate biological change only. Diseases are often initiated by chronic disruption to systems under homeostatic control, therefore changes from long-term exposures are most likely relevant to human health. Thus, the use of refined exposure measures combined with study designs that allow assessment at different time-scales is critical for assessing how biological response contributes to air pollution related disease pathobiology.

Exhaust emissions consist of a complex mixture of particle bound and free chemicals; however, exposure is typically assessed using aggregate measures, such as number of particles below a defined aerodynamic diameter threshold. Source, fuel type, combustion efficiency and distance from the source contribute to both the individual constituents and composition of particulate matter present in traffic-related pollution. Growing evidence suggests the composition of particulate matter, specifically elemental carbon (EC), is a possible mediator for increased disease risk. EC, which often includes measure of black carbon/soot, is produced through pyrolysis of hydrocarbons from incomplete combustion and is a major component of carbonaceous particulates arising from diesel exhaust (Schauer 2003). Although limited, population-based studies that include measures of EC either as a single exposure or within a multi-pollutant model have shown EC can be attributed to increased emergency department visits for congestive heart failure, mortality and cardiovascular deaths (Lipfert et al. 2006, Ostro et al. 2007, Beelen et al. 2008, Sarnat et al. 2015). Cumulative EC exposure based upon 5- and 1-year lag periods is associated with lung cancer mortality in trucking industry workers in the US (Garshick et al. 2008), while the International Agency for Research on Cancer has classified diesel exhaust as a Group 1 carcinogen (Benbrahim-Tallaa et al. 2012). Although health risks of EC have been documented, the contribution of EC and additional traffic-related pollutants to metabolic response has not been well characterized.

The human metabolome is a functional measure of the interaction between the genome, diet, environment and the biochemical processes required for life

(Wishart et al. 2013). Identifying the metabolic changes associated with environmental exposures has the potential to improve understanding of response to exposures, providing new insight into the biological changes underlying exposure-related diseases. Untargeted chemical profiling techniques based upon ultra-high resolution mass spectrometry now make possible measure of over 20,000 chemical features. This analytical methodology provides sufficient metabolic characterization for precision medicine and exposome research, and has been identified as a central platform linking external exposure to internal dose, biological response and disease pathobiology (Walker et al. 2016). When combined with complementary response measures, such as gene expression, protein levels or epigenetic changes, it is possible to develop an *in vivo*, systems-biology level understanding of how environmental exposures influence critical biochemical processes in humans. Development of aggregated biological response patterns represents a new paradigm for combining toxicology with molecular and environmental epidemiology; however, no examples exist of application in humans with well-defined exposures.

In the present study, we used untargeted metabolomics and peripheral blood gene expression to characterize the integrated molecular response of US trucking industry workers following occupational exposure to three traffic related pollutants. Using microenvironment exposure levels and blood samples collected at multiple time points throughout the workweek, we evaluated metabolic alterations associated with daily and week-averaged EC, organic carbon (OC) and particulate matter (diameter $\leq 2.5\mu\text{m}$) for changes in metabolites indicating

internal dose to chemical exposures and alterations in endogenous biochemical processes. Using metabolic changes associated with week-averaged exposures, we applied a data-driven network approach to identify correlating gene expression patterns and pathway enrichment. These results represent the first integrated assessment of human molecular response to environmentally relevant traffic-related pollutants and support exposure associated changes to pathways implicated in increased risk for CVD.

6.3. Methods

6.3.1. Study population

Details on the study population, recruitment and sample collection have been reported previously (Neophytou et al. 2013, Neophytou et al. 2014). Briefly, 95 participants were recruited from 10 trucking terminals located throughout the northeastern US (CT, MA, MD, NJ, NY and PA) and classified based upon job duties with different patterns of exposure, including pick-up and delivery drivers, freight dock workers and office clerks. Participants were selected for enrollment in the study if they had at least two days off prior to the start of the study.

Measurements and sample collection were completed between February 2009 and October 2010, with terminals sampled one at a time for up to 8 continuous days. To reduce confounding due to smoking and sex differences, the present analysis was restricted to 73 non-smoking, male workers. All participants provided informed consent, and the study protocol was approved by the Institutional Review Board of the Brigham and Women's Hospital and the Human Subject Committee of the Harvard School of Public Health.

6.3.2. Microenvironment exposure measures

Microenvironment samples, including indoor office spaces, terminal docks and within truck cabs, was collected at all 10 terminal locations for 24 hr/day over the full workweek (6-9 days). Exposure measures of traffic-related pollutants included EC and OC in PM_{1.0} (particulate matter with diameter $\leq 1 \mu\text{m}$) and total PM_{2.5} (particulate matter with diameter $\leq 2.5 \mu\text{m}$). Collection protocols and

detailed methodology on assigning exposures are described elsewhere (Davis et al. 2006). Briefly, EC and OC were measured by collecting PM_{1.0} on a 22-mm quartz tissue filter preceded by precision machined cyclone separator (SCC1.062 Triplex, BGI Inc., Waltham, MA USA), which was then analyzed by thermal-optical carbon analyzer using the NIOSH 5040 method (NIOSH 2003). PM_{2.5} was collected on a 37-mm Teflon filter (Gelman/Pall, Port Washington, NY, USA) after passing through a precision-machined cyclone pre-selector (GK2-5-SH, BGI, Inc., Waltham, MA USA) to remove particles greater than 2.5 µm in aerodynamic diameter. The method was consistent with the EPA PQ200 Federal Reference Method (Yanosky and MacIntosh 2001, Tainio et al. 2005). Individualized, personal exposures were then estimated for each participant using a weighted average of the reported time spent at each work location.

6.3.3. Biomarker sampling

For each truck terminal, initial blood samples were collected from each participant prior to the day's work shift on their first day back after at least two days off. At the end of the first work shift, a second blood sample was collected, and then pre- and post-shift samples were collected again on the last day of the same workweek. Following each blood draw, blood tubes were drawn and stored at 4°C until processing. Plasma samples for HRM were centrifuged, aliquoted and stored in the vapor phase of liquid nitrogen freezers at < -130°C. RNA was immediately extracted from samples selected for gene expression analysis using the Qiagen RNeasy kit and stored at -80°C.

6.3.4. High-resolution metabolomics

Untargeted high-resolution metabolomic (HRM) profiling was completed following protocols developed in prior studies (Go et al. 2015, Walker et al. 2016). Briefly, plasma aliquots were removed from storage at -80°C and thawed on ice. A 65 µL of plasma was then added to 130 µL of acetonitrile containing a mixture of stable isotopic standards, vortexed, and allowed to equilibrate for 30 minutes. Triplicate 10 µL aliquots were analyzed by reverse-phase C₁₈ liquid chromatography (Targa 100 mm x 2.1mm x 2.6 µm, Higgins Analytical Inc) with detection by high-resolution mass spectrometry (Q-Exactive, Thermo Scientific, San Jose, CA). Analyte separation was accomplished using water, acetonitrile and 2% [v/v] formic acid in water (solution A) mobile phases operating under the following gradient: initial 2 min period of 80% A, 5% B, 15% C, followed by linear increase to 0% A, 5% B, 95% C at 6 min and then held for an additional 4 min. Mobile phase flow rate was held at 0.35 mL/min for 6 min, and then increased to 0.5 mL/min. The high-resolution mass spectrometer was equipped with an electrospray ionization source operated in positive ion mode with spray voltage of 4.5 kV, probe, capillary temperature 275°C, sheath gas flow 45 (arbitrary units), auxiliary gas flow 5 (arbitrary units) and S-lens RF level of 69. Resolution was set at 70,000 (FWHM) and mass-to-charge (*m/z*) scan range 85-1275. Samples were analyzed in batches of 20, in addition to a quality control (QC) pooled reference sample included at the beginning and end for quantification and standardization.

Upon injection of all study and quality control samples, mass spectral features with replicate coefficient of variation (CV) $\leq 100\%$ were extracted and aligned using apLCMS (Yu et al. 2013) with modifications by xMSanalyzer (Uppal et al. 2013) and batch effect correction by ComBat (Johnson et al. 2007). Detected chemical signals were defined by accurate mass-to-charge ratio (m/z), retention time and intensity, referred to as m/z features throughout. Prior to statistical analysis, replicate injections were averaged and m/z features not detected in $>50\%$ of the participants were removed.

6.3.5. *Peripheral blood transcriptomics*

The transcriptomic results for this cohort have been described previously (Chu et al. 2016). Gene expression profiling was conducted using the Illumina HumanHT-12 v4 Expression BeadChip, with RNA labeling and array hybridization performed according to protocol. Image capture was performed using the Illumina BeadArray Reader. Standard QC and pre-processing procedures were applied, with background correction and quantile normalization procedures completed using the R package *lumi* (Du et al. 2008). The final data set was \log_2 transformed prior to network integration and included information for 47,295 probes, with only post-shift samples from both days used.

6.3.6. *Metabolome wide-association study of traffic-related pollutants*

Due to the availability of repeat HRM and exposure measures in this study, we completed metabolome wide association study (MWAS) of the three pollutants

separately for daily and week-averaged workplace exposure. Association for each m/z feature with the three pollutants was determined using linear mixed effects regression models, which included a random intercept for each subject to account for baseline inter-individual differences. Fixed effects for both models included the interquartile range (IQR) normalized exposure measure, age, day and BMI. We tested for daily-exposure associated metabolic changes by using m/z feature intensity for post-shift blood samples collected on both days and daily exposure for each participant. Week-averaged effects were evaluated using the first and last day averaged exposure for each participant and HRM results from all time points. To account for diurnal effects, we included sampling time (pre- or post-shift) as an additional covariate to allow m/z intensity to vary with time irrespective of exposure measures. Regression analysis was performed using the R package *lme4* (Bates et al. 2015) and completed separately for each exposure and time-scale combination. The likelihood ratio test of the complete model against the null model, which excluded the corresponding exposure measures as the independent variable, was used to obtain p for each m/z feature associated with exposure. Model fits were evaluated by calculating the conditional and marginal r^2 and reported for all m/z features meeting the significance threshold. To account for multiple hypothesis testing, a Benjamini-Hochberg false discovery rate (FDR) (Benjamini and Hochberg 1995) threshold of 20% was applied to identify m/z features associated with each pollutant. The resulting m/z features with p below the FDR threshold were selected for annotation and comparison of metabolic changes associated with different types and lengths of exposure.

6.3.7. Metabolite annotation and pathway enrichment

High-resolution mass spectrometry provides accurate mass measures of ion m/z , which is related to chemical monoisotopic mass, an intrinsic molecular property. The m/z features correlated with each exposure were first matched to a reference database of 120 metabolites previously confirmed with MS² and co-elution studies (Go et al. 2015). Additional m/z features not matching these metabolites were annotated based upon positive electrospray ionization adducts using the KEGG database (Kanehisa et al. 2012) and METLIN (Smith et al. 2005). Identities were assigned using evidence scoring (Uppal et al. 2016) and ± 5 parts-per-million (ppm) mass tolerance ($\Delta m_{\text{error}}/m_{\text{theoretical}} \times 10^6$). Enriched metabolic pathways were selected using a Mummichog (Li et al. 2013) scoring threshold ≤ 0.05 .

6.3.8. Network integration of the metabolome and transcriptome

To evaluate integrated metabolic and gene expression response to exposure, we used the network correlation approach developed for the R package xMWAS (Uppal et al. 2017) to link identified metabolites associated with EC and/or OC to peripheral blood transcriptomic results. To avoid diurnal effects, only post-shift feature expression from both days was considered. Initial feature selection for both datasets was completed using a multilevel sparse partial least squares (msPLS) regression analysis (Liquet et al. 2012), which utilizes a supervised multivariate dimensionality reduction method to perform simultaneous

discriminatory analysis and m/z feature selection while accounting for the dependency structure of both datasets and individual repeated measurements. The top 1500 discriminatory molecular probes from the first three latent variables were then combined with the metabolite levels using the *network()* function available in the R package mixOmics for network structure and clustering. Correlations with $|r| \geq 0.4$ and $p < 0.05$ were then selected for visualization in Cytoscape, with gene and metabolite pathway enrichment determined using MetaCore (Thomson Reuters) and Mummichog, respectively. An enrichment score of 0.1 was used to identify metabolic pathways and $p < 0.05$ for gene expression pathways. Previously described interactions between network associated genes and environmental chemicals were also evaluated using the Chemical Toxicogenomics Database (CTD) (Davis et al. 2013).

6.4. Results

6.4.1. Study population

A summary of the participant characteristics is provided in **Table 6.1**. HRM profiling was limited to 73 individuals who were otherwise healthy, non-smoking males from all three-job duties with blood samples available on both sampling days.

Table 6.1. Population characteristics of participants with plasma selected for HRM profiling

Characteristic	Total
Number of individuals	73
Age (years, mean±SD)	49.8 ± 8.3
BMI (mean±SD)	29.8 ± 4.5
Race (<i>n</i> , (%))	
White	68 (93%)
Hispanic	5 (7%)
Primary job title (<i>n</i> , (%))	
Pick-up and delivery driver	38 (52%)
Dockworker	14 (19%)
Officeworker	21 (29%)
Average day 1 exposures (mean±SD)	
PM _{2.5} (µg/m ³)	10.0 ± 5.7
EC (µg/m ³)	0.6 ± 0.5
OC (µg/m ³)	8.8 ± 2.9
Average day 5 exposures (mean±SD)	
PM _{2.5} (µg/m ³)	9.7 ± 4.8
EC (µg/m ³)	0.6 ± 0.4
OC (µg/m ³)	8.6 ± 4.3
Average work-week exposures (mean±SD)	
PM _{2.5} (µg/m ³)	10.0 ± 4.5
EC (µg/m ³)	0.6 ± 0.4
OC (µg/m ³)	8.7 ± 3.1

Comparison of exposure levels at the beginning and end of the workweek showed no differences between the two workdays ($p > 0.7$). Week-averaged exposures showed a weak correlation between PM_{2.5} and EC (Pearson $r = 0.39$, $p = 0.0007$) and between PM_{2.5} and OC (Pearson $r = 0.27$, $p = 0.021$), while no correlation was present between EC and OC (Pearson $r = 0.03$, $p = 0.78$). Although this population primarily consisted of white workers, no effort was made to stratify based upon race.

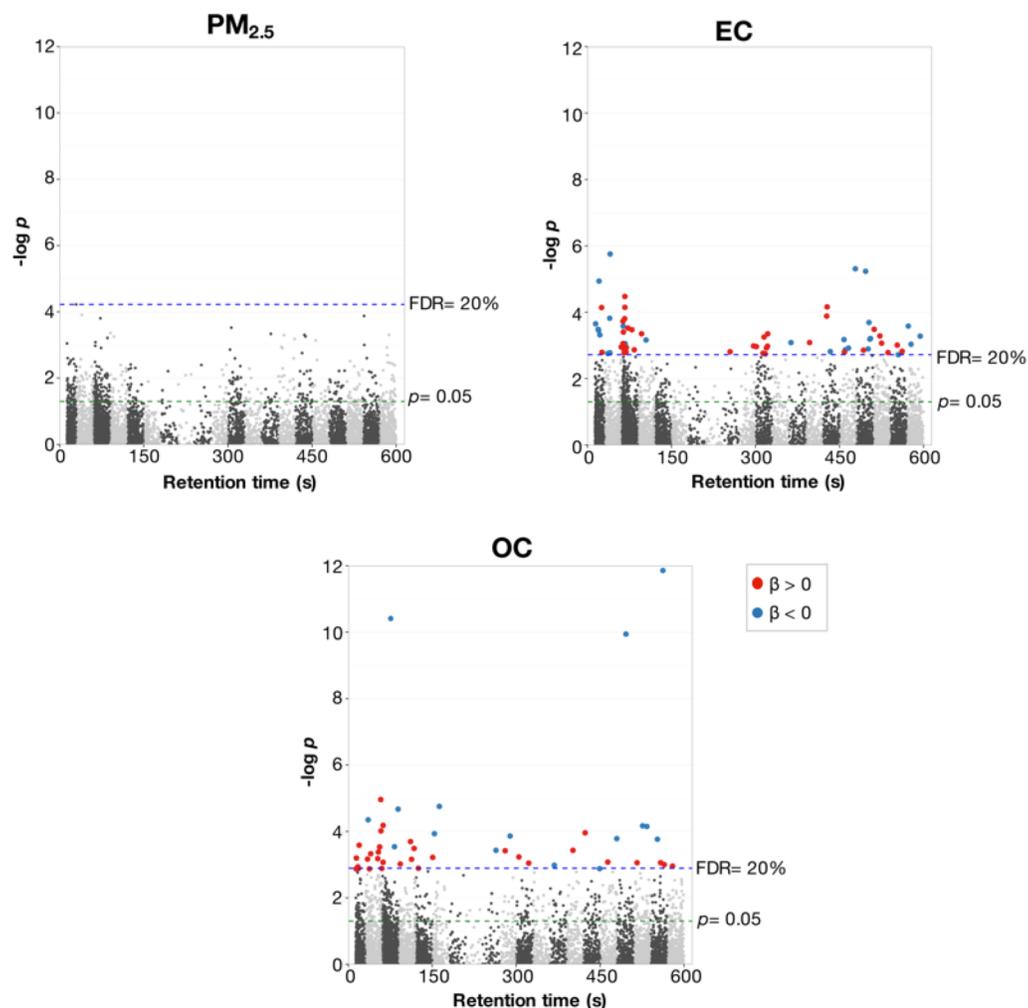


Figure 6.1. Manhattan plot showing $-\log p$ as a function m/z feature retention time for daily exposure to $\text{PM}_{2.5}$, EC and OC. MWAS was completed using linear mixed effects models and post-shift metabolite intensity only. Fixed effects included daily shift length exposure, age, day and BMI.

6.4.2. High-resolution metabolomics results

HRM profiling using C_{18} chromatography with positive ESI detected 19,380 unique m/z features with mean replicate CV of 37.6%. Prior to MWAS and integration with gene expression data, m/z features detected in less than 50% of individuals were removed. This resulted in 7,042 and 7,035 m/z features remaining for the daily and week-averaged analyses, respectively.

6.4.3. Daily exposure MWAS

To evaluate acute effects of exposure to traffic-related pollutants, we used an MWAS framework to test the relationship between shift-averaged exposures and m/z feature intensity for post-shift samples obtained on both workdays (**Figure 6.1**). Using an FDR threshold of 20%, 68 and 48 unique m/z features were associated with EC and OC, respectively. No associations with shift-averaged measures of $PM_{2.5}$ were detected. For both EC and OC, interquartile range normalized regression coefficients varied from -4.2 to 4.5, respectively; 55% were positively associated with EC and 67% positively associated with OC.

Primary metabolic effects of daily exposure to EC were first evaluated by pathway enrichment. Only butanoate metabolism ($p= 0.01$) and urea cycle/amino group metabolism ($p= 0.04$) were associated with EC exposure, with only two metabolites from these pathways identified. Matching of accurate mass m/z showed changes in metabolites related to inflammation, endothelial function, lipid peroxidation, environmental chemicals consistent with exhaust emissions and co-factors (**Table 1**). OC included changes in selenoamino acid metabolism ($p= 0.003$), glycerophospholipids ($p= 0.003$), pyruvate metabolism ($p= 0.009$), serine metabolism ($p= 0.01$), glycolysis/gluconeogenesis ($p= 0.03$) and arginine/proline metabolism ($p= 0.03$). Similar to the EC results, only a small number of metabolites were identified from each pathway, with only selenoamino acid, glycerophospholipid and serine metabolism having more than 2 matches.

Table 6.2. Relevant metabolite annotations for m/z features associated with daily exposure to elemental carbon and organic carbon

m/z	Exposure	Identity	Mass error (ppm)	Identification confidence ^a	β (95% CI)	p
149.1325	EC	pentyl-Benzene	0.5	4	-0.73 (-1.1, -0.4)	1.1E-05
155.1067	EC	4-oxo-2-Nonenal	0.5	4	-0.9 (-1.3, -0.5)	4.9E-06
389.1350	EC	16-Hydroxy-4-carboxyretinoic acid	3.0	4	-2.4 (-3.7, -1.0)	1.1E-03
105.0341	EC	Benzoic acid	5.7	1	-3.7 (-5.8, -1.5)	8.2E-04
167.9929	EC	2-Mercaptobenzothiazole	4.0	4	2.3 (0.9, 3.7)	1.6E-03
422.2357	EC	Leukotriene E4	-0.5	4	3.2 (1.4, 5.1)	8.7E-04
223.1054	EC	Dihydrobiopterin	-4.0	4	3.4 (1.5, 5.5)	1.2E-03
854.373	EC	Heme A	8.7	4	3.9 (2.0, 5.8)	7.2E-05
311.1286	OC	Porphyrin	1.0	4	-1.8 (-2.7, -1.0)	7.1E-05
245.0633	OC	Ethyl glucuronide	0.3	4	-1.3 (-1.9, -0.6)	1.4E-04
192.1014	OC	3,4-Dimethylbenzoic acid	2.0	4	-1.6 (-2.2, -0.9)	1.8E-05
136.0622	OC	Adenine	3.0	4	-2.4 (-3.8, -1.0)	1.1E-03
157.1224	OC	4-hydroxynonanal	0.6	4	-2.8 (-4.2, -1.4)	1.7E-04
589.2770	OC	Mesoporphyrin	2.0	4	2.7 (1.1, 4.3)	1.4E-03
402.0927	OC	Lactoylgutathione	-3.7	4	3.3 (1.6, 5.1)	2.5E-04
107.0537	OC	Serine	3.7	1	3.5 (1.7, 5.3)	2.0E-04
132.0657	OC	L-Glutamate-5-semialdehyde	1.5	4	3.9 (1.8, 6.1)	3.7E-04

Metabolites were consistent with products of chemicals found in diesel fuel and exhaust, nucleotides and related intermediates, porphyrin metabolites, the lipid peroxidation product 4-hydroxynonenal (4-HNE) and lactoylglutathione (LGSH) (**Table 6.2**). The majority of environmental chemical metabolites and 4-HNE were decreased with OC exposure, suggesting daily OC levels may be representative of different emission sources.

6.4.4. *Week-averaged exposure MWAS*

Previous results in this cohort have shown changes in gene expression, inflammatory markers, oxidative stress and urinary exposure biomarkers are largely dependent on week-averaged, rather than daily exposures (Neophytou et al. 2013, Neophytou et al. 2014, Chu et al. 2016). To test how week-averaged exposure influences the metabolic phenotype, we performed a second series of MWAS for the three traffic-related pollutants using samples at all time-points and week-averaged exposures (**Figure 6.2**). At the FDR threshold of 20%, 845 and 650 *m/z* features were associated with EC and OC, respectively; none met the FDR criteria for PM_{2.5}. Comparison of the significant *m/z* features for both MWAS showed 168 overlapped between OC and EC (**Figure 6.3**), while 49 (72%) and 28 (58%) were associated with either pollutant at both time points. For both EC and OC, IQR normalized regression coefficients varied from -4.0 to 5.3 and -5.2 to 6.4, respectively, which included 75% that were positively associated with EC and 59% positively associated with OC.

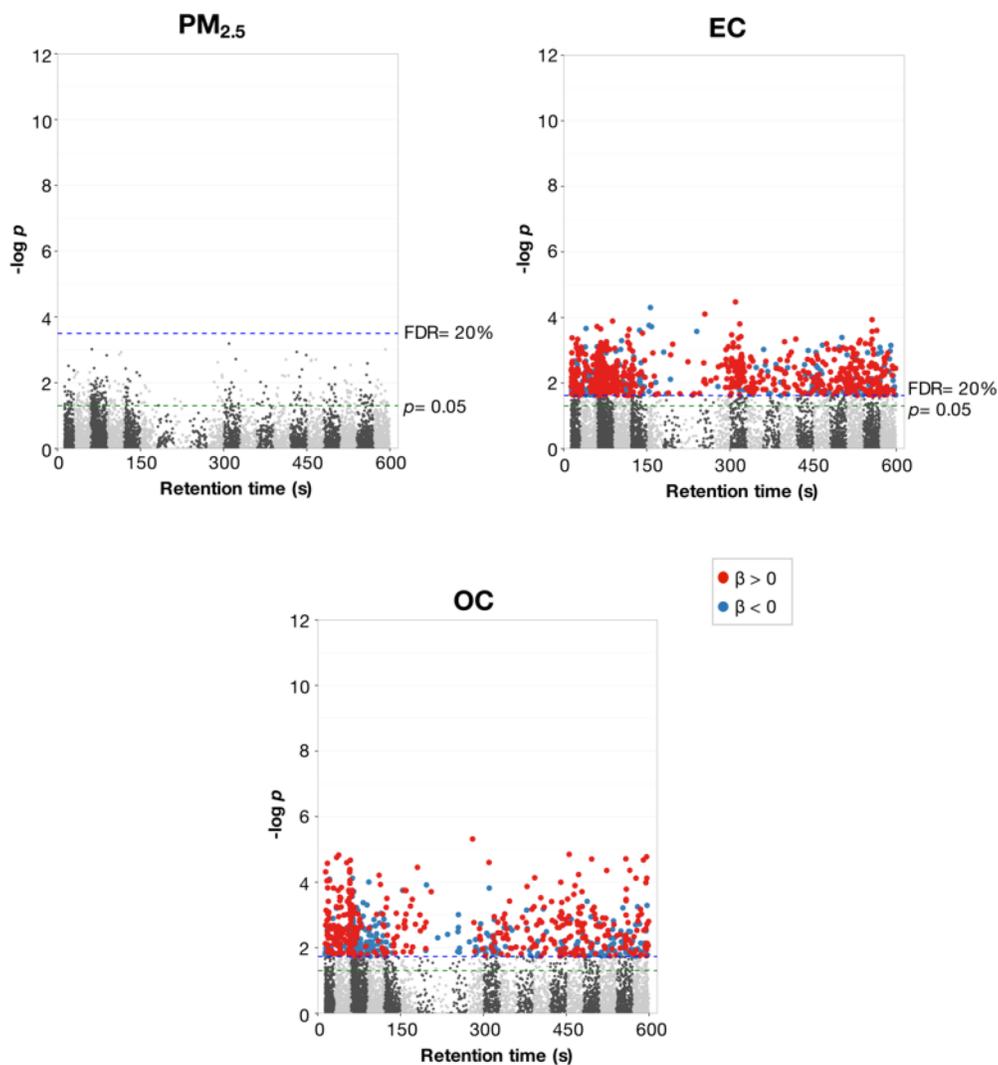


Figure 6.2. Manhattan plot showing $-\log p$ as a function m/z feature retention time for week averaged exposure to $PM_{2.5}$, EC and OC. Exposure levels were determined as the average of both monitoring days and MWAS was completed using linear mixed effects models with metabolite intensities from all time points. Fixed effects included week-averaged exposure, age, day, sampling time (pre- or post-shift) and BMI.

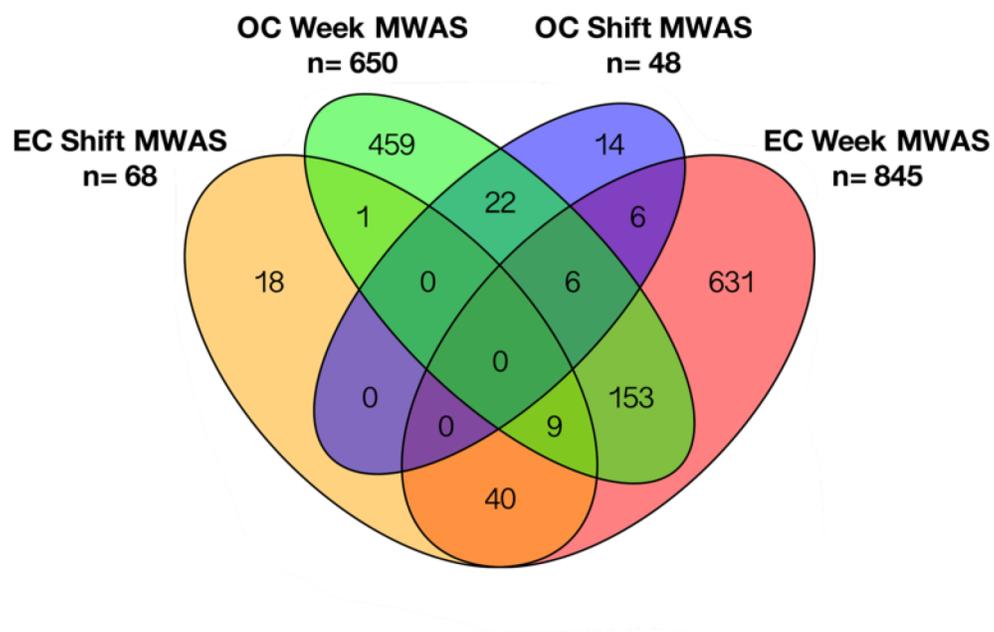


Figure 6.3. Overlapping features associated with shift-length and week averaged exposure to the two pollutants at false discovery rate threshold 20%.

Table 6.3. Relevant metabolite annotations for m/z features associated with week-averaged exposure to elemental carbon

m/z	Identity	Mass error (ppm)	Identification confidence ^a	Exposure β	p
371.1013	Hydroxyl-N-acetyl-cysteinyldihydronaphthalene	-6.2	4	2.3 (0.6, 4.0)	9.2E-03
339.1572	Nicotine glucuronide	6.0	4	2.3 (0.8, 3.8)	3.0E-03
143.0341	Muconic acid	1.0	4	2.6 (0.7, 4.4)	7.8E-03
93.0703	Toluene	4.0	4	2.8 (0.8, 4.8)	6.7E-03
109.0638	Cresol	9.0	4	2.8 (1.2, 4.5)	1.0E-03
129.0701	Naphthalene	1.5	4	2.9 (1.1, 4.7)	2.0E-03
196.1335	4-oxo-2-nonenal	1.0	4	0.5 (0.1, 0.9)	1.5E-02
263.2366	Linoleic acid	-1.1	1	1.6 (0.5, 2.8)	6.9E-03
297.2410	12-oxo-10E-octadecenoic acid	4.0	4	1.8 (0.6, 2.9)	3.2E-03
371.1007	Glutathione	3.0	4	2.4 (0.7, 4.1)	6.0E-03
302.1426	4-hydroxynonenal mercapturic acid	0.3	4	2.4 (0.8, 4.0)	3.6E-03
158.1260	4-hydroxynonenal	1.9	4	2.6 (0.7, 4.6)	8.9E-03
177.0400	Ascorbic acid	3.4	1	3.9 (1.9, 5.9)	2.3E-04
351.2144	Prostaglandin I2	6.3	4	1.7 (0.3, 3.1)	1.7E-02
335.2214	Thromboxane B2	-0.6	4	2.6 (.5, 4.8)	1.5E-01
375.2141	Thromboxane A2	-0.5	4	2.7 (0.8, 4.5)	5.8E-03
624.2953	Hepoxilin A3-C	0.6	4	2.9 (1.0, 4.8)	3.6E-03
854.3730	Heme A	8.7	4	3.5 (1.7, 5.2)	2.2E-04
227.1097	Hydroxy-L-homoarginine	7.0	4	2.6 (0.8, 4.4)	6.3E-03
169.0592	Glutamine	4.7	1	-0.5 (-1.0, -0.1)	1.7E-02
156.0278	Aspartic acid	6.4	1	2.7 (0.6, 4.8)	1.1E-02
445.1649	Dihydrofolate	-2.5	1	-2.2 (-3.7, -0.7)	5.4E-03
459.1986	Riboflavin	-0.2	4	-3.8 (-6.0, -1.6)	8.2E-04

a. Identification confidence based upon Schymanski et al. (2014)

Table 6.4. Steroid-related metabolites matching m/z features associated with week-averaged exposure to elemental carbon

m/z	Suspected Identity	Mass error (ppm)	Number of steroid matches	Exposure β	p
142.1228	4-Methylpentanal	1.4	1	-0.9 (-1.6, -0.2)	7.2E-03
255.2108	Androsterone	0.4	5	2.9 (1.2, 4.6)	1.2E-03
271.2061	Testosterone	1.8	8	3.0 (0.9, 5.0)	4.9E-03
289.1791	2-OH-Estradiol	-2.4	4	2.6 (1.0, 4.3)	1.9E-03
291.2294	Androsterone	-8.6	5	2.7 (0.8, 4.6)	5.0E-03
307.1309	Estrone-quinone	1.3	4	2.1 (0.7, 3.6)	3.6E-03
313.2171	17alpha-Hydroxyprogesterone	2.9	9	2.8 (0.9, 4.7)	3.8E-03
367.1885	Dehydrocorticosterone	1.1	2	2.4 (0.5, 4.3)	1.6E-02
367.2455	Urocortisol	-6.5	1	2.2 (0.7, 3.8)	4.7E-03
377.2065	7alpha-Hydroxypregnenolone	0.5	4	-1.8 (-3.4, -0.2)	2.4E-02
390.2614	17alpha,21-Dihydroxypregnenolone	-6.4	3	2.6 (0.8, 4.4)	5.3E-03
449.2513	Androsterone glucuronide	-4.7	3	2.6 (0.7, 4.5)	8.7E-03

Table 6.5. Mummichog enriched metabolic pathways for elemental carbon.

Pathway	Number of metabolites associated with EC exposure	Number of metabolites detected from pathway	Mummichog P
Linoleate metabolism	11	19	0.005
Prostaglandin formation from arachidonate	18	50	0.011
Fatty acid metabolism	6	12	0.011
Glyoxylate and dicarboxylate metabolism	5	10	0.014
Aspartate and asparagine metabolism	21	65	0.020
Leukotriene metabolism	13	38	0.020
Vitamin B9 (folate) metabolism	5	12	0.027
Lysine metabolism	8	23	0.031
Glycosphingolipid metabolism	8	23	0.031
Urea cycle/amino group metabolism	13	42	0.040
Drug metabolism - cytochrome P450	12	39	0.044
Alanine and aspartate Metabolism	7	21	0.045
De novo fatty acid biosynthesis	5	14	0.049

Metabolite identification for both exposures was first completed by comparison to a reference database of confirmed metabolites (Level 1 identification), which was then complemented with annotation by accurate mass matching and confidence scoring. A subset of relevant metabolite matches associated with EC is provided in **Table 6.3**. Tentative identifications included metabolites originating from exhaust emissions, lipid peroxidation products, oxidative stress biomarkers, endothelial function-related metabolites and co-factors. Lipid peroxidation products and oxidative stress metabolites included 4-oxo-2-nonenal, 4-HNE, 4-HNE mercapturic acid, glutathione and 12-oxo-10E-octadecenoic acid. An additional 14 *m/z* matching 40 different steroid-related metabolites were also present (**Table 6.4**). Due to multiple isobaric species matching a given accurate mass *m/z*, it is not possible to definitely identify each match; however, the results suggest possible alterations in endocrine function. Pathway enrichment identified 13 metabolic pathways associated with EC exposure, including changes to fatty acid metabolism, pro-inflammatory lipid signaling, co-factor metabolism and nitrogen catabolism (**Table 6.5**). Of these, only urea cycle/amino group metabolism was also associated with shift-length EC exposure.

Annotation of *m/z* features associated with week-averaged OC also showed exposure resulted in changes to critical biochemical processes. A subset of relevant metabolite matches is provided in **Table 6.6**. Similar to EC, metabolites from exposure to exhaust emissions were present, including PAH and VOC metabolites. However, many of these exhibited a negative association with increasing OC levels. Additional identified *m/z* features were consistent with

changes in co-factor metabolism, increased oxidative stress and disruption to nitric oxide (NO) production and endothelial function. Oxidative stress-related changes included increased methionine sulfoxide and LGSN; LGSN was also associated with daily OC exposure. Inconsistent with the elevated markers of oxidative stress, the lipid peroxidation product 4-hydroxyhexenal showed a negative relationship with OC exposure. This was also observed for 4-HNE in the OC shift-averaged exposures. Metabolites critical to NO production and endothelial function included dihydrobiopterin, citrulline, homocysteine and agmatine, while metabolic products of NO, including ornithine, acetylarginine and N-(omega)-hydroxyarginine were elevated. In addition, thromboxane A₂ (TxA₂), which is a metabolite of platelet activation and aggregation factors, was decreased. Enrichment in steroid metabolites was not observed, with only 6 *m/z* matching 16 different sterol species. Of these, only 7 α -hydroxypregnenolone was associated with both pollutants. Pathway enrichment identified 23 pathways, which were consistent with processes related to endothelial function, co-factor metabolism, inflammatory signaling, nitrogen catabolism, mitochondrial bioenergetics and amino acid metabolism (**Table 6.7**).

Although overall metabolic response to the two pollutants differed, 168 (11%) of the *m/z* features were associated with week-averaged exposures to both pollutants. Comparison of regression coefficients showed all but one unidentifiable *m/z* feature (*m/z*= 741.1941) exhibited opposite direction of change with exposure, i.e. those having a positive association with one pollutant were negatively associated with the other. Identified metabolites included endogenous

metabolic intermediates and chemical products of environmental exposures. Endogenous metabolites were related to amino acid metabolism (aspartic acid, lysine, serine), co-factor metabolism (riboflavin, pantothenic acid) and endothelial function (gamma-hydroxy-L-homoarginine, TxA₂). Environmental chemicals were largely consistent with exposure to volatile organic chemicals, and included dichloroethylene, 1-nitro-5,6-dihydroxy-dihydronaphthalene, muconic acid, benzoic acid and 2,6-dichloroindophenol. Systemic metabolic changes were also consistent across the two exposures, with six pathways related to inflammation, fatty acid metabolism, amino acid metabolism and nitrogen catabolism associated with exposure to both pollutants.

Table 6.6. Relevant metabolite annotations for m/z features associated with week-averaged exposure to organic carbon

m/z	Identity	Mass error (ppm)	Identification confidence ^a	Exposure β	p
285.0939	Benzo[a]pyrene-7,8-dihydrodiol-9,10-oxide	8.0	4	4.3 (1.9, 6.8)	6.7E-04
195.0199	Naphthaldehyde	3.0	4	-3.0 (5.0, -1.1)	3.1E-03
143.0341	Muonic acid	1.0	4	-3.3 (-5.4, -1.1)	3.2E-03
221.0940	Hippuric acid	8.0	1	-2.5 (-4.2, -0.8)	5.2E-03
150.0915	Cresol	0.0	4	-1.3 (-2.3, -0.4)	7.4E-03
265.9608	1-Nitronaphthalene-7,8-oxide	-3.0	4	4.1 (2, 6.2)	2.2E-04
204.0086	Methionine sulfoxide	2.0	4	0.8 (0.3, 1.4)	3.5E-03
402.0921	Lactoylglutathione	5.0	4	3.4 (1.0, 5.7)	5.1E-03
177.0691	Homocysteine	-0.6	1	-2.4 (-4.3, -0.4)	1.7E-02
156.1021	4-hydroxyhexenal	0.0	4	-2.6 (-4.5, -0.6)	9.3E-03
214.0578	Citrulline	-3.3	1	-0.7 (-1.2, -0.1)	1.6E-02
303.1179	Dihydrobiopterin	1.0	4	-3.8 (-5.6, -2)	9.8E-05
335.2214	Thromboxane B2	-0.6	4	-3.8 (-6.2, -1.4)	2.4E-03
227.1097	Hydroxy-L-homoarginine	7.0	4	-3.9 (-5.9, 1.8)	4.2E-04
169.0860	Agmatine	8.3	4	-4.2 (-6.4, -1.9)	3.8E-04
209.0081	Ornithine	-3.8	4	1.3 (0.3, 2.2)	1.1E-02
217.1045	Acetylglutamine	4.0	4	1.9 (0.4, 3.3)	1.1E-02
229.0682	N-(omega)-Hydroxyarginine	-5.2	4	2.7 (0.9, 4.4)	2.9E-03
132.0657	Lactate	1.5	4	3.1 (1.2, 4.9)	1.7E-03
195.9779	Threonine	3.6	1	-3.7 (-5.6, -1.9)	1.5E-04
156.0278	Aspartic acid	6.4	1	-3.6 (-6, -1.2)	4.2E-03
216.1960	alpha-Tocopherol	-8.3	4	1.9 (0.6, 3.2)	6.1E-03
134.0635	Asparagine	-5.2	1	2.6 (1, 4.2)	1.7E-03
107.0537	Serine	3.7	1	4.1 (2.2, 6)	6.1E-05
202.1074	Pantothenic acid	0.0	4	3.9 (1.6, 6.3)	1.4E-03
171.0142	Thiamine	0.6	4	2.7 (0.6, 4.8)	1.4E-02

a. Identification confidence based upon Schymanski et al. (2014)

Table 6.7. Mummichog enriched metabolic pathways for organic carbon.

Pathway	Number of metabolites associated with OC exposure	Number of metabolites detected from pathway	Mummichog <i>P</i>
Urea cycle/amino group metabolism	17	42	0.0002
TCA cycle	7	12	0.0002
Alanine and Aspartate Metabolism	9	21	0.0003
Selenoamino acid metabolism	8	18	0.0003
Glycine, serine, alanine and threonine metabolism	16	49	0.0005
Beta-Alanine metabolism	6	14	0.0007
Ascorbate (Vitamin C) and Aldarate Metabolism	7	19	0.0010
Arginine and Proline Metabolism	11	35	0.0012
Butanoate metabolism	7	21	0.0018
Aminosugars metabolism	8	25	0.0018
Lysine metabolism	7	23	0.0034
Fatty acid activation	5	15	0.0037
Pyruvate Metabolism	5	16	0.0054
Tyrosine metabolism	20	80	0.0057
Vitamin E metabolism	9	35	0.0086
Histidine metabolism	5	18	0.011
Linoleate metabolism	5	19	0.015
Glycerophospholipid metabolism	10	42	0.015
Valine, leucine and isoleucine degradation	6	24	0.016
Sialic acid metabolism	6	27	0.034
Aspartate and asparagine metabolism	14	65	0.036
Leukotriene metabolism	8	38	0.043
Tryptophan metabolism	12	58	0.050

6.4.5. Network correlation of molecular response

Integration of HRM results with additional biological response measures has the potential to provide enhanced understanding of mechanisms underlying effects of environmental exposures in humans. Thus, we used msPLS-regression for variable selection followed by association scoring to integrate peripheral blood gene expression with identified metabolites that were significantly associated with EC and/or OC. The molecular response network is provided in **Figure 6.4. A**

high-degree of connectivity was observed for the two datasets, resulting in 1,234 molecular probes and 146 metabolites with $|r| \geq 0.4$ and $p \leq 0.05$. Multi-level community detection identified four clusters in the molecular response network. Excluding cluster 2, the remaining were largely gene expression dominated, with clusters 1, 3 and 4 including 120, 370 and 699 genes, respectively, associated with 65, 2 and 22 metabolites. Cluster 2 contained 57 metabolites and 45 genes. To assess the biological functions associated with each of the clusters, we tested for metabolite and gene expression pathway enrichment for each cluster. The complete results are provided in **Supplementary File 3**. Cluster 1 showed association with 15 metabolic pathways and 23 gene expression pathways, which included pathways related to inflammation and immune response, DNA damage, cell adhesion and vascular function. Metabolites present in cluster 2 suggested a different pattern of response, including 26 metabolic pathways and 13 gene expression pathways. These were consistent with developmental processes, endothelial NO synthases, nucleotide metabolism and oxidative stress processes. Oxidative stress pathways included changes to ascorbate/aldarate metabolism, sulfuramino acid metabolism and TNF-alpha-induced ROS-dependent Caspase-3 signaling, with the metabolites ascorbic acid, glutathione and cysteamine belonging to this cluster. Metabolites from xenobiotic metabolism were also present in cluster 2, including naphthalene and hydroxy-N-acetyl-L-cysteinyl-dihydronaphthalene. Cluster 3 consists almost entirely of gene expression nodes, with the vitamin E metabolite 7'-carboxy-alpha-chromanol and pyrroline the only two identified metabolites present. Due to this, 60 gene expression pathways were

enriched from a diverse range of processes, including adipocyte differentiation, apoptosis, immune response, oxidative and reticulum stress response, inflammation and red blood cell adhesion. Similarly, cluster 4 only showed enrichment in the Vitamin E metabolism pathway and contained the metabolites tocopherol (vitamin E) and ubiquinol-10. The gene expression results showed 47 pathways, relevant processes included NO synthase signaling in muscle tissue, platelet aggregation, TNF signal transduction, mitochondrial dysfunction, immune response and regulation of lung epithelial progenitor cell differentiation.

Gene expression results were also evaluated by searching CTD for reported gene-chemical interactions using the top two most connected gene nodes from each cluster. For cluster 1, MATR3 and PPM1A showed interaction with benzo[a]pyrene, volatile organic chemicals and vehicle emissions. In addition, BPDE, dinitrotoluene and nanoparticle exposures had reported effects on expression of MATR3. Cluster 2 nodes included one gene providing no matches in CTD (LOC100131165). The second gene, SNORA14A, only showed interactions with 1-(4-methoxyphenyl)-2-(3,4,5-trimethoxyphenyl)ethane, a methylated analog of resveratrol that has been shown to reduce levels of induced NO synthase. Chemical interaction with the top gene expression nodes in cluster 3 consisted of a large number of environmental pollutants originating from exhaust emissions. For KIDINS220, traffic-related pollutants included dinitrobenzene, benzene and PAH dipole epoxides, while dinitrotoluenes, PAHs and dipole epoxide metabolites, carbon nanotubes, tobacco smoke and vehicle emissions have documented interaction with expression of IVNS1ABP. Cluster 4 resulted in

similar interactions, including dinitrotoluenes, dioxin and PAH related chemicals for the top gene, while interactions with dinitrotoluene, dinitrobenzene, a PAH dipole epoxide, dioxins and nitroarenes were present for OTUB1.

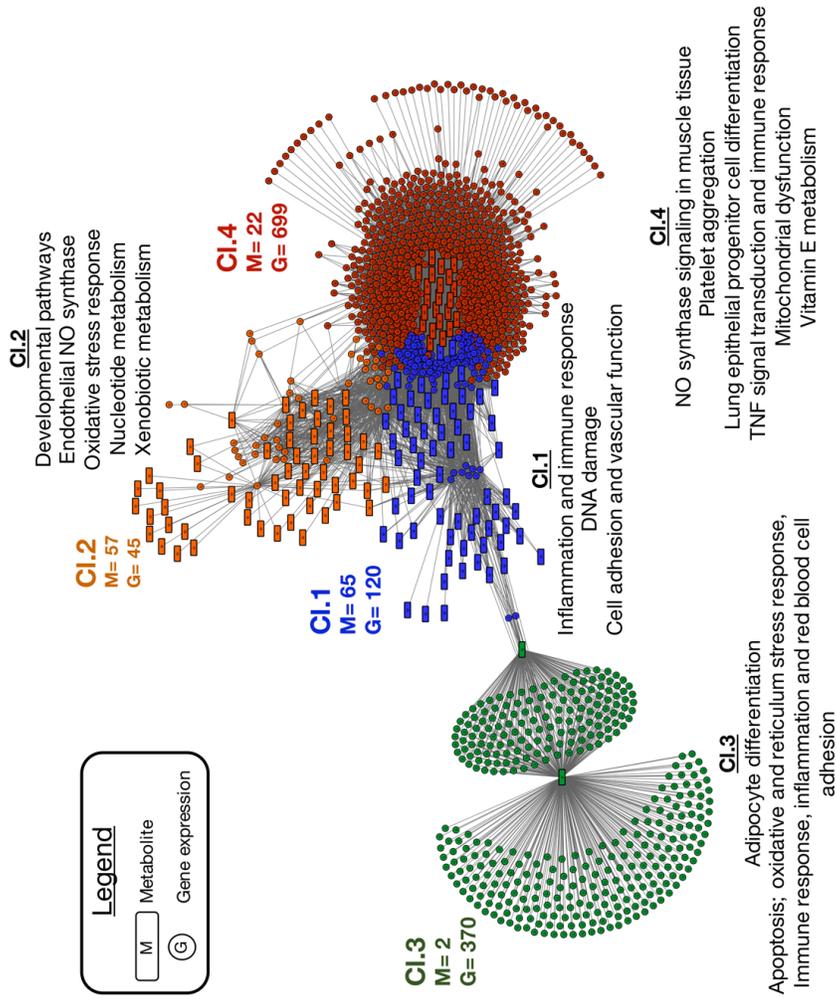


Figure 6.4. Metabolite and peripheral blood gene expression integration network. Discriminatory features and network correlation were completed using the R package xMWAS. The top discriminatory genes and identified metabolites were first using multilevel sparse partial least squares regression analysis and network connections determined based upon $|r| \geq 0.4$ and $p < 0.05$. Clusters were identified using multi-level community detection and evaluated for both metabolite and gene pathway enrichment. Cl= Cluster; M= metabolite node, G= gene expression node.

6.5. Discussion

The present study provides the first integrated molecular characterization of biological response to different constituents of traffic-related pollution. Using a well-characterized occupational exposure cohort, untargeted metabolomics was used to identify metabolic changes associated with PM_{2.5}, EC and OC from daily and week-averaged exposures. The results showed week-averaged EC and OC resulted in the greatest metabolic alterations. These included biomarkers of exposure to exhaust emissions and contrasting metabolic changes consistent with processes related to oxidative stress, endothelial function and inflammation. Combined with the post-shift gene expression data, the results support disruption to pathways associated with increased risk for adverse cardiovascular outcomes and identify novel biological changes with a plausible role in exposure-related disease etiology.

Although exposure to traffic-related pollutants is typically modeled using aggregate measures such as total PM₁₀ or PM_{2.5}, particulate composition will vary depending on combustion efficiency, fuel type and distance from emission source. MWAS of the three pollutants identified different metabolic response based upon exposure timescale and type. For both daily and week-averaged exposures, PM_{2.5} levels showed no metabolic associations at the FDR threshold. Only one other study has examined the short-term changes to the plasma metabolome based upon environmental measures of PM_{2.5}. In the study by Breitner et al. (2016), prior day increases in PM_{2.5} levels were associated with changes in arginine, glycine, ornithine and acetyl-carnitines. Since prior-day exposures were driving dose-

dependent metabolic changes, it suggests there is a lag-time between exposure and biological response. This is consistent with previous studies in this cohort, where week-averaged but not shift length PM_{2.5} exhibited association with inflammatory markers, oxidative stress, exposure biomarkers and peripheral blood gene expression (Neophytou et al. 2013, Neophytou et al. 2014, Chu et al. 2016). The use of long-term measures of particulate exposure (≥ 1 year) has shown to provide better detection of metabolic associations with PM_{2.5}. In Menni et al. (2015), estimates of ambient exposure to PM₁₀ and PM_{2.5} were determined using dispersion models and participant residence zip code. The results showed changes to amino acids, co-factor metabolism and benzoic acid, in addition to a negative association with pulmonary function. MWAS of yearlong ultrafine particulate exposures has also shown metabolites discriminating low and high ultrafine particulate exposure, including changes consistent with endothelial dysfunction, oxidative stress and inflammation (**Chapter 7**). Thus, these results suggest longer timescales are needed to elucidate the metabolic effects of exposure to PM_{2.5}. While we attempted to account for chronic exposure using week-averaged levels, PM_{2.5} is a ubiquitous environmental pollutant with exposure occurring outside of the workplace. Combined with the potential for lag effect, this makes untangling the metabolic response to particulate matter following relatively short monitoring periods challenging. Furthermore, these studies were primarily focused on traffic-related pollutants from automotive emissions and were not specific to diesel exhaust.

While PM_{2.5} showed no association with the metabolome, EC and OC exhibited dose-associated changes to both environmental chemical exposure markers and alterations to central metabolic pathways. Due to the complexity and varied composition of EC and OC in particulate matter, findings on the health effects of these two pollutants have largely been inconsistent across different population studies and are complicated by measurement techniques and heterogeneity (Schauer 2003). However, both toxicological and epidemiological studies suggest EC is an important contributor to adverse health effects linked to traffic related exposures. Human population studies show including EC in models results in increased effect sizes for disease and biomarker outcomes (Janssen et al. 2011). To date, the metabolic response to EC and OC has not been well characterized, and only limited data is available on metabolic effects of diesel exhaust exposure. To measure acute response to biodiesel fuel exhaust exposure, Surowiec et al. (2016) used a multi-platform metabolomics assay to evaluate human lung lavage fluids from different lung regions following a 1-hr controlled exposure. Metabolite expression in lavage fluids showed exposure resulted in appreciable differences for 46 metabolites, including the amino acids aspartic acid, glutamic acid, glycine, ornithine, glutamine and alanine, fatty acids, benzoic acid metabolites and co-factor metabolism. Additional studies in humans have shown diesel exhaust to result in altered DNA methylation profiles (Jiang et al. 2014), peripheral gene expression (Peretz et al. 2007) and changes to bioactive lipids in lung lavage fluids (Gouveia-Figueira et al. 2017). Model systems provide additional evidence of metabolic changes associated with exposure to diesel

emissions. To evaluate systemic response to diesel fuel particulates, Oeder et al. (2015) applied a multi-omic approach that included proteomic, transcriptomic and metabolomic assessment of exposure to emissions from diesel and heavy fuel oil using human epithelial lung cells. The results showed diesel fuel, which is considered the cleaner burning alternative, resulted in a systemic biological response that was not observed for the heavy fuel oil particulates. These included changes to energy and protein metabolism, mRNA processing and chromatin modification, which was attributed to the greater elemental carbon levels. Heavy fuel oil also showed changes to gene and protein expression consistent with oxidative stress and inflammation. Metabolomic profiling of mouse models using mixed vehicle emissions that included both gasoline and diesel exhaust identified similar metabolic changes, including alterations to oxidative stress-related metabolites, inflammation, lipid peroxidation and metabolites related to NO production (Brower et al. 2016).

Identified metabolic alterations associated with EC exposure were largely consistent with risk factors for CVD. While not present following shift-length exposure, week-averaged EC levels were associated with metabolites with important antioxidant functions and biomarkers of increased oxidative stress, including four lipid peroxidation products: 4-oxo-2-nonenal, 12-oxo-10E-octadecenoic acid, 4-HNE mercapturic acid and 4-HNE. The elevated levels suggest increased formation of reactive oxygen species (ROS). Lipid peroxidation products are relatively stable, cytotoxic aldehyde products that react with a wide range of different biomolecules, including proteins, lipids and nucleic acids

(Weber et al. 2013) and have been implicated in the pathogenesis of multiple chronic diseases (Selley 1998, Akude et al. 2010). Increased oxidative stress and formation of *in vivo* ROS has been well documented in exposure to particulate matter (Lodovici and Bigagli 2011, Bates et al. 2015); however, limited data is available on the exact mechanisms through which the ROS-generating potential of the particulates results in a more oxidized state *in vivo*. In **Chapter 7**, HRM identified similar changes in metabolites associated with yearlong ultrafine particulates from near highway exposures, suggesting similar biological response to the two pollutants. Linoleic acid, which is a polyunsaturated omega-6 fatty acid, was also positively associated with EC exposure. Polyunsaturated fatty acids are particularly susceptible to lipid peroxidation, which occurs through oxidant attack at carbon-carbon bonds with subsequent hydrogen abstraction and oxygen insertion (Yin et al. 2011). In addition, the two antioxidants GSH and ascorbic acid were identified. GSH and its corresponding oxidized form, glutathione disulfide (GSSG) are one of the primary extracellular thiol/disulfide redox couples and provide multiple critical functions, which include protecting against oxidative stress, chemical detoxification and arginine turnover during NO synthase (Hofmann and Schmidt 1995). Ascorbic acid is a protective antioxidant with functions similar to GSH, has also been linked to endothelial processes, including chemical stabilization of tetrahydrobiopterin during NO synthesis (Heller et al. 2001), vascular endothelial function (Brown and Hu 2001) and has beneficial effects on vascular dilation (Jackson et al. 1998), among other protective effects (Padayatty et al. 2003). Increased oxidation, particularly of thiol

couples, and subsequent shift towards a more positive steady state redox potential has been shown to result in activation of pro-inflammatory cytokines (Iyer et al. 2009), regulate early events in atherosclerosis (Go and Jones 2005) and is associated with CVD and endothelial function (Ashfaq et al. 2008), while a high burden of oxidative stress has been associated with mortality in patients with coronary heart disease (Patel et al. 2016).

Additional associations with EC exposure consistent with CVD risk included alterations to metabolites related to thrombotic signaling, coagulation and endothelial function. These include the eicosanoid metabolites TxA_2 and thromboxane B₂ (TxB_2), which were positively associated with exposure. TxA_2 provides multiple physiological functions, which include signaling for activation of new platelets and increased aggregation, vasoconstriction, endothelial adhesion molecule expression, cell proliferation and acts as a hypertensive agent (Ishizuka et al. 1998, Ding and Murray 2005, Smyth 2010); TxB_2 is an inactive form of TxA_2 with minimal biological function. Elevated biosynthesis of TxA_2 and increased receptor expression are elevated in CVD (Smyth 2010), with irreversible inhibition of COX-1, which is the platelet enzyme that biosynthesizes TxA_2 from arachidonic acid, being one of the primary anti-platelet therapies for prevention of thrombotic events (Sanmuganathan et al. 2001). Additional inducers of TxA_2 production and receptor expression include inflammation, oxidative stress and NO, with TxA_2 and other prostaglandin formation almost entirely driven by NO following inflammatory events (Mollace et al. 2005). This is consistent with the EC-associated increase in hydroxy-L-homoarginine, which is a

minor metabolic byproduct of NO production by oxidation of the arginine (Arg) analogue, homoarginine, by NO synthase (Moali et al. 1998). Increased prostaglandin I₂ (PGI₂) and hepoxilin A₃-C (HxA₃-C) provides additional evidence of disruption to this pathway. PGI₂ is biosynthesized by the isoenzyme COX-2 from arachidonic acid, with physiological functions including mediation of vasodilation and platelet inhibition (Whittle et al. 1978, Cheng et al. 2002) and acts as a homeostatic control for the vasoactive effects of TxA₂. HxA₃-C is the glutathione adduct of hepoxilin A₃, which is also an arachidonic acid metabolite with pro-inflammatory and vasodilation signaling properties. Metabolic pathways associated with EC exposure included changes to the bioactive eicosanoid pathways linoleate metabolism, prostaglandin formation from arachidonate and leukotriene metabolism. Taken together, these results provide strong evidence that exposure to EC results in changes to physiologically active signaling pathways that are implicated in a wide range of processes related to CVD and endothelial dysfunction.

Week-averaged EC exposure was also associated with changes in amino acids, sterols and co-factor metabolism. Decreased glutamine (Gln) and increased aspartic acid provide additional evidence of disruption to NO production and endothelial function. Gln is a precursor for *de novo* synthesis of Arg through intestinal uptake of Gln and release of citrulline, which is then converted to Arg at the kidney (Ligthart-Melis et al. 2008). Arg is a direct precursor of NO, which is generated in the endothelium by endothelial NO synthase (Gornik and Creager 2004). While Arg was not detected in the MWAS, the hydroxy-L-homoarginine

was elevated and suggests utilization of Arg analogues. Both Gln and Asp were also associated with long-term exposure to UFP, suggesting a non-specific metabolic response to traffic-related pollutants. The presence of both male and female sex hormones associated with EC was a surprising result since the cohort was all male; however, this may be due to incorrect identification because of the large number of isobaric sterol species possible and the inability to provide unambiguous identification of the different species. Both population and mechanistic studies have shown exposure to traffic-related pollutants can lead to endocrine disruption effects. In the study by Guven et al. (2008), toll-booth collectors with elevated traffic emissions exposure exhibited significantly lower sperm counts and greater abnormal sperm morphology when compared to office-workers. Both mice and cell models have also shown exposure to diesel exhaust particulates can lead to decreased estrogen receptor mRNA expression and antiandrogenic effects (Kizu et al. 2003, Takeda et al. 2004), which was primarily attributed to binding of the aryl-hydrocarbon receptor.

Changes in co-factor metabolites, which included dihydrofolic acid and riboflavin for the week-averaged exposure and dihydrobiopterin (BH₂) following daily exposure are consistent with documented biological effects of air pollution. Both dihydrofolic acid (Vitamin B9 derivative) and riboflavin (Vitamin B2) were decreased with exposure to EC. Active forms of folic acid are critical methyl-donors at the cellular level for regulation of homocysteine, synthesis of *S*-adenosylmethionine, methylation and DNA synthesis (Crider et al. 2012). Exposure to black carbon and other traffic-related exposures has shown to result

in increases in methylation patterns, which was also observed to influence biological and vascular response (Breton and Marutani 2014). Vitamin B supplementation has been shown to mediate epigenetic changes following controlled exposure to PM_{2.5}, suggesting a protective role for these vitamins (Zhong et al. 2017). BH₂ is the oxidative metabolite of tetrahydrobiopterin, an essential cofactor for NO synthase-catalyzed formation of NO from Arg. Elevated levels of circulating BH₂ have shown to be associated with endothelial NO synthase dysfunction (Noguchi et al. 2011), while oxidative stress can result in oxidation of tetrahydrobiopterin to BH₂, decreasing the amount available for NO synthase (Leiva et al. 2015). While consistent with the disruptions to endothelial function observed for the week-averaged exposure, BH₂ was associated with daily exposures only, suggesting this could be a transient effect.

Although many of the biochemical processes associated with OC were similar to those observed for EC, specific metabolite changes differed and overlapping metabolites showed opposing direction of change following exposure to the two pollutants. A possible explanation for this trend is different patterns of exposure and particulate origin. Prior characterization of this cohort has shown exposure patterns differ based upon job title and pollutant, with pickup/delivery drivers and dockworkers exhibiting highest exposure levels of EC, while office workers and pickup/delivery drivers showed the highest OC levels (Neophytou et al. 2013). The OC fraction of particulate matter consists of a complex mixture of polar and non-polar chemicals that arise from a number of processes after initial combustion, including organic matter condensation, binding of semi-volatile

organics on the particulate surface and interaction with aerosols, in addition to atmospheric changes occurring during transport. Thus, while EC represents exposures from diesel fuel exhaust, the measured OC could arise from a variety of sources and may not be due to emissions alone. This is consistent with the detected changes in chemicals that originate in exhaust emissions, which showed positive associations of volatile PAHs and other organic compounds with EC and negative associations with OC.

Both daily and week-averaged OC exposures resulted in changes to metabolites providing antioxidant effects and markers of oxidative stress. Unlike EC, which showed a positive association with lipid peroxidation products, OC exposure indicated a negative relationship with 4-HNE following shift-length exposures. As discussed above, 4-HNE is formed by oxidant attack on carbon-carbon bonds, suggesting ROS forming potential of traffic related pollutants enriched in OC is decreased. However, this is inconsistent with increased methionine sulfoxide (MetSO), LGSB and decreased homocysteine levels. MetSO is the oxidation product of methionine with reactive oxygen species. Non-enzymatic oxidation of Met residues is a common post-translational protein modification showing oxidative damage and is a recognized marker of oxidative stress (Moskovitz et al. 1997, Mashima et al. 2003). LGSB is formed by GSH dependent degradation of methylglyoxal (MG) by glyoxalase I, which is subsequently split by glyoxalase II to regenerate GSH and form lactate (Chang and Wu 2006), another metabolite associated with OC exposure. MG has been implicated in a number of disease processes, including hypertension, diabetes and

formation of atherosclerotic lesions (Chang and Wu 2006). In addition, MG has been shown to induce formation of ROS and to irreversibly bind to Arg, cysteine and lysine residues, resulting in formation of advanced glycation end products (Chang et al. 2005, Vasdev et al. 2007). Hyperhomocysteinemia, which includes abnormal levels of homocysteine, has been identified as a tentative biomarker of cardiovascular disease and thrombosis (Jacobsen 1998). While levels were decreased with exposure in this study, homocysteine is a critical intermediate in the conversion of methionine to cysteine, which is required for production of other sulfur amino acids, including taurine and glutathione. As discussed, above, high burden of oxidative stress and decreased antioxidants has been identified in the pathology of CVD and mortality. Pathways consistent with antioxidant processes were also detected, including vitamin E metabolism and ascorbate/aldarate metabolism. In addition, alpha-tocopherol was positively associated with OC, which has previously been observed to protect against decreasing lung function following PM_{2.5} exposure (Menni et al. 2015). Alterations to the TCA cycle provide additional evidence of oxidative stress from OC exposure. Mitochondrial energy production is the result of close coordination between the TCA cycle and electron transport chain, with disruption to either resulting in increased oxidative stress and mitochondrial dysfunction (Pieczenik and Neustadt 2007).

OC exposure showed association with the greatest number of metabolites related to endothelial function. This included alterations to Arg and proline metabolism, and changes to metabolites with functions that include NO

production from Arg, required co-factors, decreased thrombotic signaling molecules and endogenous inhibitors of NO. These changes indicate OC results in disruption to NO production itself, while many of the EC associations suggested increased platelet formation and aggregation. Interestingly, many of the possible negative effects of NO metabolism associated with EC showed an opposite relationship with OC, including decreased BH₂, TxB₂ and hydroxy-L-homoarginine. While elevated levels of these metabolites are associated with a range of disease pathobiology, decreases suggest that exposures to emissions with high OC content may have less detrimental effects on endothelial function. OC was also associated with decreased citrulline and agmatine, while ornithine, acetylarginine and n-(omega)-hydroxyarginine were increased. Citrulline is formed from ornithine by ornithine carbamoyltransferase, a critical step in the urea cycle and nitrogen catabolism, or through the NO synthase dependent oxidation of Arg during NO production. Significant metabolites and pathways suggest disruption to both processes could be occurring, with urea cycle/amino group metabolism the top pathway associated with OC. Changes consistent with possible disruption to NO included agmatine, acetylarginine and n-(omega)-hydroxyarginine. Agmatine, which is an endogenously produced decarboxylation product of Arg, inhibits inducible NO synthase and acts as a modulator of NO production (Raghavan and Dikshit 2004). Decreased levels suggest elevated rates of NO production, which is consistent with the changes in n-(omega)-hydroxyarginine but not the decreased levels of citrulline. N-(omega)-hydroxyarginine is the first oxidative product of Arg by NO synthase and can act

both as an inhibitor of Mn dependent arginase enzymes and catalytic intermediate in NO production (Cox et al. 2001). Taken together, these results suggest OC exposure contributes to metabolic alterations that have important implications for nitrogen catabolism and endothelial function. While changes in the urea cycle may not directly influence risk factors for CVD and other exposure related diseases, it could influence bioavailability of NO intermediates and disrupt endothelial function. For example, studies have shown decreased citrulline can lead to NO synthase decoupling and increase severity of acute respiratory distress syndrome symptoms and sepsis (Luiking et al. 2009, Neill et al. 2009, Ware et al. 2013). Thus, these results provide additional evidence exposure to traffic related pollution has implications for endothelial function and CVD risk.

An important indicator of environmental exposure is the ability to evaluate internal dose relative to external exposures. MWAS of both EC and OC identified chemical exposure biomarkers that are consistent with pollutants present in exhaust emissions. For EC exposure, positive associations with biomarkers of PAH exposure, volatile organic compounds and tobacco exposure were present. These include toluene and naphthalene, which are ubiquitously present in diesel exhaust emissions (Tsai et al. 2012, Alam et al. 2016) and the detoxification products hydroxyl-N-acetyl-cysteinyl-dihydronaphthalene (phase II metabolite of naphthalene), muconic acid (biomarker of benzene exposure) and cresol, which is a metabolite of toluene and a combustion product present in exhaust emissions (Sperlingova et al. 2007, Zhang et al. 2011). Plasma levels of cresol have also shown to be associated with yearlong UFP from near-highway exposures,

suggesting its use as a non-specific biomarker of exhaust emission exposure
(Chapter 7).

Although participants selected for HRM profiling were purposely biased towards non-smokers, nicotine glucuronide was also associated with EC. The lack of additional tobacco use biomarkers suggests this is due to passive exposure, possibly in outdoor areas where EC levels were highest. OC exposures resulted in different patterns of association. Environmental chemical levels including naphthaldehyde, muconic acid, hippuric acid and cresol were negatively associated with OC. Naphthaldehyde is a degradation product of methylnaphthalene, which is present in small amounts in exhaust and other combustion emissions. Hippuric acid is a non-specific biomarker of toluene and benzene, which can be confounded by consumption of benzoic acid and dietary phenols (Amorim and Alvarez-Leite 1997). OC exposure was associated with elevated levels of two reactive PAH diol epoxides: benzo[α]pyrene-7,8-dihydrodiol-9,10-oxide (BPDE) and 1-nitronaphthalene-7,8-oxide (NNDE). BPDE is a chemical carcinogen formed through oxidation of benzo[α]pyrene via a series of cytochrome P450 enzymes, resulting in a metabolically active species with the ability to bind to a wide range of biomolecules (Delaney and Essigmann 2008, Chung et al. 2010). Benzo[α]pyrene is an incomplete combustion product that is found in cigarette smoke, grilled foods and automobile exhaust fumes. The presence of NNDE suggests exposure to nitronaphthalene, a mutagenic nitroaromatic that has been detected in ambient air, diesel exhaust and photocopier toner emissions. While the biological activity of NNDE in humans

has not been well characterized, formation of metabolically active species and protein adduction is suspected in the cytotoxic effects of nitronaphthalene, which includes lesion formation in Clara cells and pulmonary toxicity in neonatal rats (Kovacic and Somanathan 2014).

Following characterization of the metabolome, we tested for additional associations with peripheral blood gene expression using a data-driven network approach. To our knowledge, this is the first study to integrate exposure-associated metabolic alterations with peripheral blood gene expression in a human cohort. Exposure to environmental chemicals can influence local and global changes in gene transcription and enzyme activity, resulting in a multitude of biological changes that can influence metabolic changes and contribute to underlying toxicological mechanisms. Integrating gene expression results with MWAS therefore has the potential to enhance understanding of how functional molecular mechanisms interact due to environmental exposures. Enriched gene pathways were consistent with the identified metabolic changes associated with EC and OC exposure, providing additional evidence that exposure to traffic-related pollution leads to changes in endothelial function, inflammation and oxidative stress. Most importantly, characterization of enriched metabolic and gene expression pathways within each cluster consisted of different biochemical processes, suggesting biological response to air pollution is systemic and has important implications for a diverse range processes critical for homeostatic control and health. Comparison to curated data in CTD showed many of the top genes for each cluster have previously been demonstrated to interact with

pollutants present in exhaust emissions, including PAHs and dipole epoxide metabolites, vehicle emissions, volatile organic compounds and nitroaromatics. Thus, not only were the *in vivo* identified biological changes from gene expression and the metabolome consistent, identified genes have been independently shown to interact with traffic-related pollutants *in vitro*.

We acknowledge several limitations. First, this study was limited to an all-male, non-smoking, pre-dominantly Caucasian cohort employed in the US trucking industry. Due to the sample population size, we could not adequately stratify by sex, age or race. Although samples were collected over the course of a workweek, only exposure occurring at the workplace were considered, and we could not account for those that occurred during off hours. Second, results showed week-averaged exposures resulted in the greatest molecular changes; however, these were estimated based upon one week only, and due to the length of the study we could not evaluate the effects of longer exposure periods and how levels varied over time. Third, the results from this study are correlative in nature. We could not account for unknown confounders or confirm the results in an independent cohort. In addition, we evaluated correlation of peripheral blood gene expression with the metabolome. Previous results in humans have shown signals in whole blood generally reflect organism wide processes (Bartel et al. 2015); however, although the metabolome represents an integrated profile of multiple body compartments and processes, it is not possible to assess cell specific metabolite or gene expression. Fourth, while confidence-based methods were used for annotation of metabolites, unidentifiable signals were not extensively

characterized. Lack of reference standards and low abundance of many of the m/z features render characterization and identification of these features challenging. Finally, exposure was determined using aggregate measures that included PM_{2.5}, EC and OC. Specific levels of volatile and semi-volatile levels that arise from different fuel types and sources were not determined. Despite these limitations, we identified EC and OC exposure-associated changes consistent with increased oxidative stress, endothelial dysfunction and inflammation, which are common risk factors for diseases related to air pollution exposure. In addition, HRM detected exogenous chemicals associated with each pollutant, providing novel biomarkers for assessing exposures. The results demonstrate the use of a combined HRM and gene expression approach to characterize molecular mechanisms underlying environmental exposures in human populations. Continued application of these approaches in human population will provide improved understanding of how environmental exposures contribute to adverse health outcomes.

6.6. Conclusions

Exposure to diesel exhaust emissions and other traffic-related pollutants has been linked to cancer and cardiopulmonary disease; however, the mechanisms linking exposure to disease are unknown. To elucidate the molecular response of occupational exposure to diesel exhaust in the US trucking industry, we used an MWAS framework to identify plasma metabolic changes in workers following daily and week-averaged exposure to PM_{2.5}, EC and OC. The results showed

PM_{2.5} did not result in any significant metabolic alterations. MWAS of EC and OC identified contrasting metabolic changes to pathways consistent with endothelial function, inflammation and oxidative stress. These include metabolites that function as thrombotic signaling molecules, antioxidants, biomarkers of oxidative stress, lipid peroxidation products and intermediate metabolites in NO production. Volatile organic chemicals, their metabolites, PAHs and dipole epoxide metabolites were also associated with the two exposures and link external exposure to internal dose of compounds present in exhaust emissions. Week-averaged levels of both pollutants resulted in the greatest degree of metabolic alterations, suggesting persistent workplace exposure influences long-term biological changes that are not mediated by time off. Data-driven integration of peripheral blood gene expression levels with exposure-associated metabolites identified additional functional changes related to endothelial function, NO synthase, inflammation and immune response. Taken together, the results show exposure to diesel exhaust emissions influences metabolite and gene expression pathways implicated in CVD and atherosclerosis risk, providing additional evidence of the adverse effects of traffic-related pollution.

**Chapter 7. METABOLOMIC ASSESSMENT OF YEARLONG
EXPOSURE TO NEAR-HIGHWAY ULTRAFINE PARTICLES**

Contributions to this work are as follows. D. Walker: Designed and completed HRM data analysis, interpreted results, performed literature review, wrote manuscript, created figures and tables; K. Lane: Contributed to data preparation and analysis in original study, performed time and activity adjustments; assembled demographic data, selected samples for HRM profiling; K. Uppal: Developed algorithms for network analysis; K. Liu: Quantified metabolite levels in reference materials for reference standardization; A. Patton: Collected data developed models in original study for predicting UFP exposures; J. Durant: Co-I of original study; oversaw ambient exposure measures and model development; D. Jones: Supervised D. Walker, oversaw HRM measures and efforts; read and edited chapter; D. Brugge: PI of original CAFEH study, designed sample collection, measurements and participant enrollment, read and edited chapter; K. Pennell: Supervised efforts of D. Walker, read and edited chapter.

7.1. *Abstract*

Introduction: Exposure to traffic-related particulates from vehicle exhaust has been associated with increased risk of adverse cardiopulmonary outcomes. Underlying mechanisms appear to include systemic inflammation, endothelial dysfunction and oxidative stress; however, the precise biochemical pathways impacted by traffic-related particulate exposures are not known. In this study, we employed high-resolution metabolomic profiling of plasma from participants enrolled in the Community Assessment of Freeway Exposure and Health Study (CAFEH) to assess metabolic alterations associated with year-long, time-activity adjusted, near-highway ultrafine particles (UFP).

Methods: For this pilot study, blood samples from 59 non-smoking participants were selected based on high (avg 24,000 particles/cm³) and low (avg 16,000 particles/cm³) annual average UFP exposures and analyzed using high-resolution mass spectrometry. Metabolic associations with exposure were first evaluated using quantified levels of 79 metabolites, followed by untargeted identification of discriminatory metabolic features using partial least squares discriminatory analysis (PLS-DA). Systemic response was assessed through correlation networks and pathway enrichment, which was then tested for corresponding relationships with measured markers of inflammation and coagulation.

Results: Comparison of metabolite concentrations identified five metabolites differential expressed between low- and high UFP exposure consistent with increased oxidative stress and endothelial dysfunction, including arginine (Cohen's $d= 0.57$), glutamine (Cohen's $d= 0.58$), cystine (Cohen's $d= 0.68$) and methionine sulfoxide (Cohen's $d= 0.72$). Expanding to the entire metabolomic dataset, feature selection using PLS-DA detected an additional 316 m/z features. Identified metabolites were related to lipid peroxidation, endogenous inhibitors of nitric oxide and environmental chemicals, including volatile organic chemicals and bioactive polycyclic aromatic hydrocarbon metabolites. Network correlation analysis showed 38 metabolic pathways were associated with UFP exposure, including those related to inflammation, oxidative stress, endothelial function and mitochondrial bioenergetics. Metabolites were then tested for correlation with independently measured biomarkers of inflammation and endothelial function, linking metabolic associations with exposure to disease pathobiology.

Conclusions: The results suggest long-term exposure to UFP is associated with antioxidant pathways, *in vivo* generation of reactive oxygen species and processes critical to endothelial functions. However, the presence of a systemic biological response indicates that metabolic alterations associated with exposure to traffic related pollutants are complex. Application of these techniques in larger populations and model systems is needed to confirm mechanisms underlying UFP exposure-related disease.

7.2. *Introduction*

Environmental factors and their interaction with the human genome are suspected as one of the main causes of chronic disease. Specifically, air pollution from fossil fuel emissions and biomass combustion is a key environmental risk factor for human health; with global disease burden estimates suggesting 3.1 million deaths in 2010 were attributable to ambient particulate exposure alone (Lim et al. 2012). Exposure to fine particle matter (PM_{2.5}, particulate matter with diameter ≤ 2.5 μm) has been linked to all-cause mortality, cardiovascular disease, cardiopulmonary outcomes and lung cancer mortality in epidemiological studies (Pope et al. 2002, Pope et al. 2009, Hart 2016). Acute exposure to pollutants from traffic-related sources have shown changes in systemic inflammation, coagulation, DNA methylation, gene expression and respiratory function (Baccarelli et al. 2009, Zuurbier et al. 2011, Fuller et al. 2015, Chu et al. 2016); however, underlying toxicological mechanisms and how they contribute to disease outcomes are not well characterized.

Recent studies suggest that oxidative stress and inflammation play a central role in health effects associated with fine particle exposure. PM_{2.5} has been shown to have redox activity, either through transportation of reactive oxygen species (ROS) on particles or through generation of ROS *in vivo* (Donaldson et al. 2001, Bates et al. 2015). In human populations, exposure has been associated with circulating biomarkers of oxidative stress, including increased malondialdehyde, thiobarbituric acid reactive substances and 8-hydroxy-2'-deoxyguanosine (Sorensen et al. 2003, Chuang et al. 2007). Although inconsistent, associations

with systemic inflammation in both acute and chronic exposure studies have been reported (Brugge et al. 2013), including pro-inflammatory cytokines (Pope et al. 2016), C-Reactive protein (Lanki et al. 2015) and tumor necrosis factor alpha receptor II (Lanki et al. 2015, Lane et al. 2016, Pope et al. 2016). While changes in oxidative stress and inflammation are consistent with the link between air pollution exposure and disease risk factors, many of these biomarkers lack specificity and are consistent with an immediate rather than chronic response.

The human metabolome, which includes all low molecular weight (<2000 dalton) chemical species present in a biological matrix, is a functional measure of the interaction between the genome, diet, environment and the biochemical processes required for life (Wishart et al. 2013). Identifying the metabolic changes associated with environmental exposures has the potential to improve understanding of biological response, providing new insight into the mechanisms underlying exposure-related disease pathobiology. A variety of analytical platforms are available for measuring the human metabolome (Scalbert et al. 2009). Untargeted chemical profiling techniques based upon ultra-high resolution mass spectrometry now make possible profiling of up to 20,000 chemical features. This analytical methodology provides sufficient metabolic characterization for precision medicine and exposome research, and has been identified as a central platform linking external exposure to internal dose, biological response and disease pathobiology (Walker et al. 2016).

Limited data are available on the metabolomic effects of air pollution, and more specifically ultrafine particles (UFP, < 0.1 μm). In the study by Breitner et

al. (2016), a Bayesian space-time hierarchical space-time model with 10 x 10 km grid was used to assess short-term PM_{2.5} and ozone exposure at residential address for a cardiac catheterization cohort. Metabolomic measures included 61 metabolites related to amino acid metabolism, fatty acid metabolism and total ketones, with only 10 used to assess association with PM_{2.5} and ozone. The results suggested a delayed association with decreased arginine and glycine based upon previous day elevated PM_{2.5} levels. Additional metabolic characterization of response to traffic-related pollution include a multi-platform assessment of lavage fluids following acute exposure to biodiesel exhaust (Surowiec et al. 2016) and occupational exposure of office workers, traffic cops and rickshaw drivers in India (Pradhan et al. 2016). Both studies showed metabolic changes following exposure to emissions from vehicle traffic. In the highest exposure group, Pradhan et al. (2016) measured elevated oxidative stress biomarkers, including increased malonyl dialdehyde and total ROS, while glutathione peroxidase, catalase and superoxide dismutase activity were decreased. Only one study has examined the relationship between long-term PM_{2.5} exposure and metabolic changes. Using a subset of the TwinsUK cohort, Menni et al. (2015) estimated exposure to ambient PM_{2.5} based upon residential addresses. Metabolic associations included four amino acid-related metabolites, benzoate, glycerate and the antioxidant α -tocopherol. These metabolites were also linked to a decline in lung function within a larger cohort, with results showing potential protective effects of α -tocopherol. While these studies provided insight into metabolic alterations associated with air pollution exposure, limitations included minimal

metabolome coverage, lack of individualized exposure levels and, excluding Menni et al. (2015), exposures were for acute time periods only.

In this pilot study, we performed high-resolution metabolomics (HRM) of plasma collected from participants enrolled in the Community Assessment of Freeway Exposure and Health Study (CAFEH) to assess metabolic alterations associated with traffic-related pollutants. Exposure to near-highway UFP was estimated with high spatial/temporal resolution (20m; 1 hour) and adjusted for time and activity to estimate individualized, year-long exposure levels (Lane et al. 2015). Exposure associated differences were first evaluated for central metabolic intermediates, followed by a metabolome-wide association study (MWAS) to identify systemic changes. We then determine the correlation structure of exposure-associated metabolomic features with independently measured biomarkers of inflammation and coagulation, providing insight into the relationship between changes related to disease pathobiology and the metabolome.

7.3. Methods

7.3.1. Study population

A subset of 59 participants from the CAFEH study were selected for HRM profiling. CAFEH is a community-based participatory research study of the relationship between air pollution and health effects in individuals living next to major highways. The core study involves measuring traffic-related pollutants in communities in or near Boston, Massachusetts. It was designed to test the hypothesis that chronic exposure to UFP is associated with biomarkers and cardiovascular health. Detailed methodology on the study population and design can be found elsewhere (Fuller et al. 2013, Lane et al. 2016). Briefly, recruitment occurred in near highway (≤ 500 from major interstates) and urban background areas ($\geq 1,000$ m) that including Somerville, Malden and the Boston neighborhoods of South Boston, Chinatown and Dorchester. Participants were recruited based upon age (≥ 40) and a geographically-weighted, random selection process. All participants provided informed consent and the study was approved by the Tufts University School of Medicine Institutional Review Board. Enrolled participants completed an in-home survey that included questions about demographics, recent illnesses, major cardiovascular diseases, hypertension, statin-use, insulin, oral hypoglycemics, smoking status, and micro-environment time-activity. Peripheral blood was drawn by registered nurses and stored at -80°C until use. Biomarker measures included tumor necrosis factor receptor II (TNF-RII), high-sensitivity C-reactive protein (CRP), interleukin-6 (IL-6) and fibrinogen, and were completed by immunoassay. For this pilot metabolomics

study, participant selection criteria included non-smokers, individuals with documented low (<19,000 particles/cm³) or high (>19,000 particles/cm³) UFP exposure and was restricted to residents of Somerville, Dorchester and South Boston to avoid confounding due to race differences in biological response.

7.3.2. *Exposure assessment*

The methodology for predicting time activity adjusted exposure levels is described in Patton et al. (2015), Lane et al. (2016). UFP levels throughout the Boston area were determined using the mobile Tufts Air Pollution Monitoring Laboratory, which consisted of a converted recreational vehicle equipped with a condensation particle counter (TSI Model 3775) providing fast-response measure of UFP as particle number concentration (PNC; particle size 4-3000 nm).

Measurements were made repeatedly over the course of the year along fixed routes in each study area during various times of the day, days of the week and seasons. Estimated ambient PNC for the study areas was accomplished using multivariable regression modeling to develop predictive models of hourly PNC with fine spatial resolution (20m). PNC exposures were then adjusted in hourly intervals to correct for time spent in different microenvironments over a year-long period (Lane et al. 2013), providing participant estimate of annual UFP exposure (particles per cm³).

7.3.3. *High-resolution metabolomics*

HRM was completed using well-established methods (Soltow et al. 2013). Samples were prepared and analyzed in batches of 20; each batch included duplicate analysis of pooled human plasma (QStd-3) for quality control purposes and reference standardization. Prior to analysis, plasma aliquots were removed from storage at -80°C and thawed on ice. Each cryotube was then vortexed briefly to ensure homogeneity, and 65 µL was transferred to a clean microfuge tube, and immediately treated with 130 µL of ice-cold LC-MS grade acetonitrile containing stable isotope internal standards selected to represent a range of metabolite physiochemical properties. Treated plasma was then equilibrated for 30 min on ice and centrifuged ($16.1 \times g$ at 4°C) for 10 minutes to remove precipitated proteins. The resulting supernatant was added to a 200 µL autosampler vial and maintained at 4°C until analysis (<22 h).

Sample extracts were analyzed using liquid chromatography and Fourier transform high-resolution mass spectrometry (Dionex Ultimate 3000, Q-Exactive HF, Thermo Scientific). The chromatography system was operated in a dual pump configuration that enabled parallel analyte separation and column flushing. For each sample, 10 µL aliquots were analyzed in triplicate using hydrophilic interaction liquid chromatography (HILIC) with electrospray ionization (ESI) source operated in positive mode and reverse phase chromatography (RPC) with ESI operated in negative mode. This use of complementary chromatography phases and ionization polarity has been shown to improve detection of endogenous and exogenous chemicals (Liu et al. 2016). Analyte separation was accomplished by HILIC using a 2.1 mm x 100 mm x 2.6 µm Accucore HILIC

column (Thermo Scientific) and an eluent gradient (A= 2% formic acid, B= water, C= acetonitrile) consisting of an initial 1.5 min period of 10% A, 10% B, 80% C, followed by linear increase to 10%A, 80% B, 10% C and then held for an additional 4 min, resulting in a total runtime of 10 min per injection. Reverse phase analyte separation was by 2.1 mm x 100 mm x 2.6 μ m Accucore C₁₈ column (Thermo Scientific) using an eluent gradient consisting of an initial 2 min period of 0.5%A, 94.5%B, 5%C, followed by a linear increase to 0.5%A, 4.5%B, 95%C at 6 min and held for the remaining 4 min. For both methods, mobile phase flow rate was held at 0.35 mL/min for the first 1.5 min, increased to 0.5 mL/min and held for the final 4 min.

The high-resolution mass spectrometer was operated at 120,000 resolution and mass-to-charge ratio (m/z) range 85-1275. Probe temperature, capillary temperature, sweep gas and S-Lens RF levels were maintained at 200°C, 300°C, 1 arbitrary units (AU), and 45, respectively, for both polarities. Additional source tune settings were optimized for sensitivity using a standard mixture, positive tune settings for sheath gas, auxiliary gas, sweep gas and spray voltage setting were 45 AU, 25 AU and 3.5 kV, respectively; negative settings were 30 AU, 5 AU and -3.0KV. Maximum C-trap injection times of 10 milliseconds and automatic gain control target of 1×10^6 were used for both polarities. During untargeted data acquisition, no exclusion or inclusion masses were selected, and data was acquired in MS1 mode only. Raw data files were then extracted using apLCMS (Yu et al. 2013) with modifications by xMSanalyzer (Uppal et al. 2013). Uniquely detected ions consisted of m/z , retention time and ion abundance, referred to as

m/z features. Prior to data analysis, *m/z* features were batch corrected using ComBat (Johnson et al. 2007) and filtered to remove those with coefficient of variation (CV) $\geq 75\%$ and greater than 20% non-detected values in both groups.

7.3.4. Reference standardization

To compare absolute concentrations of central metabolic intermediates and clinical health markers in this cohort, 79 metabolites previously confirmed by comparison to authentic reference standards or database spectra were quantified by reference standardization (Go et al. 2015). Using this approach, metabolite levels are first determined in Q-Std3 by methods of addition or comparison to NIST standard reference material 1950 (Metabolites in Frozen Human Plasma) (Quehenberger et al. 2010, Simon-Manso et al. 2013, Colas et al. 2014), providing a reference standard for each analyte. Average metabolite response was then determined based upon the most reliable adduct formed, and plasma concentrations were calculated by single point calibration via response factors (calculated as the ratio between the known concentration of the compound being quantified and ion intensity in Q-Std3 or NIST 1950). Metabolite concentrations were evaluated relative to the human metabolome database (HMDB) (Wishart et al. 2013) and data from the National Health and Nutrition Examination Survey (CDC and NCHS). Association with UFP exposure was tested using the Student's T test with $p < 0.05$ and Cohen's *d*.

7.3.5. Metabolome-wide association study of UFP exposure

Following reference standardization, we applied a multivariate metabolome wide association study (MWAS) framework to identify m/z features in the HILIC positive ionization and RPC negative ionization data. Partial least squares discriminant analysis (PLS-DA), which is a supervised multivariate dimensionality reduction method, was used for selection of exposure-associated features based upon low- and high-exposure classification (Wold et al. 2001). The optimal number of latent components was determined using the `plsgenomics` package in R. `MixOmics` was used to build the PLS-DA model and select discriminatory features based on variable importance for projection (VIP) score ≥ 2.0 (Le Cao et al. 2009). Model accuracy was evaluated by R^2/Q^2 for PLS components 1 and 2, 10-fold cross-validation accuracy and permuted accuracy based on data shuffling (random data predictive accuracy). Association with exposure was evaluated using Cohen's d .

7.3.6. Metabolite annotation

The m/z features were annotated using a three step process with final identification level based upon Schymanski et al. (2014). First, detected accurate mass m/z and retention time were compared to a database of approximately 150 metabolites with adduct mass and retention time confirmed by comparison of ion dissociation patterns to authentic reference standards (Level 1). Level 2 matches were not considered since MS^2 data was not collected for the majority of detected signals. Features not identified by comparison to the reference list were then annotated using HMDB in conjunction with `xMSannotator` (Uppal et al. 2016),

which uses a scoring system based upon correlation modularity clustering combined with isotopic, adduct and mass defect grouping to improve annotation of high-resolution mass spectrometry data (Level 3). The remaining m/z features were then matched using accurate mass alone and common adducts (Level 4) for positive and negative ESI using HMDB and METLIN (Smith et al. 2005); m/z values providing no matches in either of the databases were considered non-identifiable chemical signals (Level 5) and listed as accurate mass m/z and retention time. The non-identifiable signals are often present in untargeted HRM studies due to a number of reasons, including uncharacterized environmental exposures and unknown metabolic intermediates (Uppal et al. 2016). All metabolite matching was completed using 10 parts-per-million (ppm) accuracy ($(\Delta m/z)/\text{theoretical } m/z \times 10^6$).

7.3.7. Metabolome correlation and pathway enrichment

To identify systemic effects associated with UFP exposure, we combined a network based metabolome-wide correlation analysis of PLS-DA discriminatory metabolites with metabolic pathway enrichment. Correlated m/z features in the raw data were evaluated using Pearson rank correlation coefficient and corresponding p with the PLS-DA selected features were calculated pairwise using the R package MetabNet (Uppal et al. 2015). The m/z features with Pearson $|r| \geq 0.6$ and Benjamini-Hochberg $\text{FDR} \leq 5\%$ (Benjamini and Hochberg 1995) were selected for visualization in Cytoscape (Su et al. 2014) and metabolic pathway enrichment using Mummichog (Li et al. 2013).

Table 7.1. Population characteristics and UFP exposure levels

Characteristic	Total	Low exposure (PNC <19000)	High exposure (PNC >19000)
Number of individuals	59	28	31
Mean PNC \pm SD	20,066 \pm 4,452	16,125 \pm 2,051	23,625 \pm 2,636
Age			
\leq 50	25	15	10
$>$ 50	34	13	21
Active smokers	0	0	0
Race			
White Non-Hispanic	40	20	20
Other	19	8	11
BMI			
Normal (< 25)	22	12	10
Overweight (25-30)	19	7	12
Obese (> 30)	18	9	9
C-Reactive Protein (mg/L) \pm SD	2.74 \pm 3.6	2.88 \pm 3.9	2.62 \pm 3.4
IL-6 (pg/mL) \pm SD	2.09 \pm 2.3	1.8 \pm 2.3	2.3 \pm 2.4
TNF-RII (pg/mL) \pm SD	2741 \pm 1588	2535 \pm 1583	2927 \pm 1594
Fibrinogen (mg/dL) \pm SD	462 \pm 109	444 \pm 101	479 \pm 116

7.3.8. Correlation network with inflammatory and coagulation biomarkers

Network integration analysis of the metabolome correlation network and the independently measured biomarkers was completed using PLS regression combined with the *network()* function available in the R package mixOmics. Correlations with Pearson $|r| \geq 0.3$ were then selected for visualization in Cytoscape and additional characterization of cluster metabolites. Since HILIC positive and C₁₈ negative data were combined for this analysis, annotated metabolites from Mummichog were used with metabolite set enrichment analysis to identify enriched pathways in each cluster (Xia and Wishart 2010).

7.4. Results

7.4.1. Study population

Demographics of the 59 CAFEH participants selected for HRM profiling are provided in **Table 7.1**. Age, race, sex distribution and BMI were comparable for the two groups. With the exception of CRP, IL-6, TNF-RII and fibrinogen were greater in the high exposure group compared to the low exposure group; however, none were significantly different at $p < 0.05$. As expected, UFP levels were significantly different for the two groups ($p < 0.0001$).

7.4.2. High-resolution metabolomics

HRM profiling using HILIC chromatography with positive ESI detected 8,554 unique m/z features with mean replicate CV of 23%, while 6,337 were detected with average CV of 19% using RPC in negative ESI mode. Due to different adducts, retention characteristics and ionization efficiency for the two ionization modes, each dataset was tested for association with UFP independently. Prior to untargeted comparison of low and high-UFP exposure, m/z features detected in less than 80% of individuals in both groups were removed, resulting in 5,367 and 3,781 remaining for HILIC and RPC, respectively.

Table 7.2. Measured amino acid, amino acid metabolites and clinical health marker levels

Metabolite ^{a,b}	All participants Mean ± SD	TAA UFP <19,000 Mean ± SD	TAA UFP >19,000 Mean ± SD	Reference range	Cohen's <i>d</i>	<i>p</i>
Amino Acids						
Alanine	206 ± 145	178 ± 116	231 ± 164	259-607	0.37	0.16
Arginine*	35 ± 14	31 ± 13	38 ± 14	60-224	0.57	0.032
Asparagine	33 ± 16	33 ± 15	32 ± 17	16-57	-0.08	0.76
Aspartic acid*	10 ± 7	12 ± 7	8 ± 5	16-26	-0.65	0.014
Citrulline	27 ± 14	26 ± 13	29 ± 14	27-38	0.17	0.51
Cystine*	103 ± 46	88 ± 35	117 ± 51	36-90	0.68	0.012
Glutamic acid	43 ± 18	44 ± 18	41 ± 19	24-151	-0.17	0.52
Glutamine*	765 ± 293	679 ± 334	844 ± 228	396-645	0.58	0.029
Histidine	52 ± 20	53 ± 20	52 ± 20	75-143	-0.08	0.76
Leucine/Isoleucine	260 ± 202	265 ± 226	255 ± 182	155-355	-0.05	0.86
Lysine	162 ± 59	156 ± 70	167 ± 49	178-456	0.18	0.50
Methionine	23 ± 10	21 ± 5	25 ± 13	22-46	0.36	0.18
Ornithine	13 ± 6	14 ± 7	13 ± 5	53-135	-0.13	0.61
Phenylalanine	60 ± 16	58 ± 18	62 ± 14	48-169	0.21	0.41
Proline	213 ± 188	200 ± 172	225 ± 203	168-239	0.13	0.62
Serine	55 ± 17	54 ± 14	57 ± 20	42-238	0.19	0.47
Taurine	141 ± 42	144 ± 51	139 ± 33	42-198	-0.13	0.62
Threonine	122 ± 36	119 ± 34	124 ± 38	102-260	0.15	0.58
Tryptophan	58 ± 18	58 ± 17	58 ± 19	37-60	0	1.00
Tyrosine	59 ± 16	58 ± 17	60 ± 15	54-144	0.13	0.62
Valine	194 ± 117	185 ± 89	202 ± 139	178-260	0.15	0.58
Amino Acid Metabolites						
5-Hydroxy-l-tryptophan (nM)	17 ± 5	17 ± 6	16 ± 5	16-20	-0.09	0.73
Hippurate	1.2 ± 1.8	1.5 ± 2	0.9 ± 1.3	0-5	-0.30	0.25
Homocysteine	3 ± 5	3 ± 3	4 ± 6	6-12	0.18	0.50
Kynurenine	2.7 ± 1	2.8 ± 1.2	2.6 ± 0.8	0.7-3	-0.11	0.68
Methionine sulfoxide**	1.8 ± 0.5	1.6 ± 0.4	1.9 ± 0.5	3-5	0.72	0.008
Methyl-histidine	4 ± 3	4 ± 3	3 ± 3	0-12	-0.22	0.40
Acetyl-D-Tryptophan (nM)	148 ± 161	155 ± 164	142 ± 161	NA	-0.08	0.76
Oxoproline	40 ± 21	42 ± 22	39 ± 21	13-161	-0.18	0.49
Phenylpyruvic acid	1.2 ± 0.3	1.2 ± 0.3	1.3 ± 0.3	0.5-0.6	0.38	0.15
Clinical Markers						
Lactic acid	903 ± 745	797 ± 589	999 ± 861	740-2400	0.27	0.30
Glucose (mM)	4.3 ± 1.3	4.304 ± 1.4	4.2 ± 1.3	3.9-6.1	-0.05	0.86
Cortisol (nM)	414 ± 265	418 ± 302	410 ± 232	28-660	-0.03	0.92
Cholesterol (mM)	3.9 ± 2.5	3.7 ± 2.5	4.1 ± 2.5	4.5-7	0.14	0.58
Creatinine	88 ± 38	94 ± 40	83 ± 37	56-109	-0.29	0.27
Urea (mM)	3.6 ± 2.3	3.9 ± 2.6	3.3 ± 1.91	4-9	-0.27	0.30
Creatine	19 ± 14	17 ± 8	21 ± 17	8.4-65	0.36	0.18
Uric acid	241 ± 85	248 ± 79	234 ± 90	238-506	-0.16	0.55
Histamine (nM)	22 ± 9	20 ± 9	24 ± 9	0.31-2.2	0.41	0.12

a. Concentrations are given in μM unless noted otherwise

b. * $p < 0.05$; ** $p < 0.01$

Table 7.3. Measured fatty acids, lipids, nucleotides, TCA cycle metabolites and co-factors

Metabolite ^a	All participants Mean ± SD	TAA UFP <19,000 Mean ± SD	TAA UFP >19,000 Mean ± SD	Reference range	Cohen's <i>d</i>	<i>p</i>
Fatty Acid and Lipid Metabolism						
Carnitine	38 ± 22	42 ± 23	35 ± 21	26-79	-0.33	0.21
Lauric acid	25 ± 17	23 ± 13	27 ± 20	2-36	0.24	0.36
Acetyl-carnitine	2 ± 1	3 ± 2	2 ± 0.9	3-8	-0.38	0.14
Myristoleic acid	18 ± 16	17 ± 16	18 ± 15	2-19	0.06	0.82
Myristic acid	65 ± 73	83 ± 91	49 ± 49	58-248	-0.47	0.07
Palmitoleic acid	256 ± 224	243 ± 185	267 ± 257	105-454	0.11	0.68
Palmitic acid (mM)	1.8 ± 1.1	2.0 ± 1.3	1.6 ± 0.8	1.9-3	-0.36	0.17
Margaric acid	1.2 ± 1.1	1.2 ± 1.05	1.2 ± 1.1	1-3	0.01	0.98
gamma/alpha-Linolenic acid	54 ± 38	45 ± 26	63 ± 45	24-86	0.48	0.08
Linoleic acid (mM)	2.9 ± 1.6	2.7 ± 1.4	3.1 ± 1.7	2.6-4.6	0.28	0.28
Oleic acid (mM)	1.8 ± 1.2	1.6 ± 1.0	1.8 ± 1.3	1.4-3.2	0.19	0.46
Stearic acid	591 ± 431	628 ± 384	559 ± 474	515-939	-0.16	0.54
Eicosapentaenoic acid (nM)	447 ± 345	438 ± 369	454 ± 328	400-1800	0.05	0.86
Arachidonic acid	1001 ± 639	1080 ± 817	929 ± 422	538-1070	-0.23	0.37
homo-gamma-Linolenic acid	140 ± 75	132 ± 71	146 ± 79	98-232	0.18	0.49
11,14-Eicosadienoic acid (nM)	305 ± 246	294 ± 258	315 ± 239	250-580	0.09	0.74
Eicosenoic acid	22 ± 58	23 ± 49	21 ± 66	9-22	-0.03	0.91
Docosahexaenoic acid	333 ± 312	383 ± 419	288 ± 162	72-227	-0.30	0.25
Docosapentaenoic acid	53 ± 28	52 ± 29	54 ± 28	26-59	0.08	0.77
Adrenic acid (nM)	440 ± 321	453 ± 353	429 ± 294	500-1500	-0.07	0.77
Respolvin E2 (nM)	0.15 ± 0.16	0.14 ± 0.19	0.15 ± 0.14	2-11	0.04	0.89
Leukotriene B4 (pM)	5 ± 10	4.7 ± 12	5.1 ± 8	0-600	0.04	0.89
Glycerophosphocholine	1.1 ± 1	1.2 ± 1	1.1 ± 0.9	NA	-0.09	0.72
Sphingosine	0.95 ± 0.8	1 ± 0.8	0.9 ± 0.8	0.049-0.51	-0.12	0.64
Sphinganine (nM)	2 ± 4	2 ± 5	2 ± 4	10-11	-0.10	0.69
Nucleotide metabolism						
5,6-Dihydrouracil (nM)	742 ± 362	775 ± 394	712 ± 334	20-1600	-0.17	0.51
Hypoxanthine	3 ± 4	4 ± 5	2 ± 2	1.3-54.5	-0.37	0.15
Xanthine (nM)	560 ± 1135	720 ± 1641	416 ± 154	200-800	-0.26	0.31
Uridine	5 ± 3	5 ± 3	4 ± 2	2.9-3.3	-0.23	0.38
TCA Cycle Metabolites						
Pyruvic acid	33 ± 14	34 ± 16	32 ± 12	22-258	-0.19	0.47
Succinic acid	8 ± 1	8 ± 1	8 ± 1	0-32	-0.34	0.20
Malic acid	1.4 ± 0.7	1.5 ± 0.6	1.3 ± 0.7	NA	-0.31	0.23
Citric acid	98 ± 52	109 ± 58	89 ± 44	55-121	-0.39	0.13
Vitamins and Co-factors						
Pyridoxal (nM)	0.12 ± 0.12	0.1 ± 0.1	0.13 ± 0.15	0.2-0.3	0.21	0.43
Pantothenic acid	19 ± 29	20 ± 24	18 ± 33	4.5-5.3	-0.06	0.83
alpha-tocopherol	16 ± 9	16 ± 9	15 ± 9	18-44	-0.02	0.94
Tetrahydrofolic acid (nM)	9 ± 4	9 ± 4	8 ± 3	0-7	-0.11	0.67
Choline	1.3 ± 0.5	1.3 ± 0.5	1.3 ± 0.6	6.5-12.5	-0.01	0.97
Niacin	6 ± 14	4 ± 11	8 ± 16	43-55	0.34	0.20
Methylnicotinic acid (nM)	15 ± 34	19 ± 49	11 ± 8	NA	-0.24	0.35

a. Concentrations are given in μM unless noted otherwise

7.4.3. Reference standardization

To evaluate the metabolic effects of year-long UFP exposure, we first used a reference standardization approach to measure levels of central metabolic intermediates and test for differences between the high and low exposure groups. The results for amino acids, related metabolites and clinical health markers are provided in **Table 7.2**. Amino acids, including arginine, aspartic acid, cystine and glutamine differed between the two groups at $p < 0.05$. Although not meeting the p threshold of 0.05, Cohen's $|d| > 0.3$ for alanine, methionine and phenylpyruvic acid also suggest a small effect size for the association of UFP with these metabolites. Although none of the clinical health markers showed statistically significant differences at $p < 0.05$, the magnitude of the effect suggests the potential for alteration of creatinine, urea, creatine and histamine levels in the high exposure group. Mean metabolite concentrations outside of the typical ranges reported by HMDB included the amino acids alanine, arginine, aspartic acid, cystine, glutamine, histidine, lysine and ornithine; the amino acid metabolite phenylpyruvic acid, and the clinical health markers cholesterol, histamine and urea.

Reference standardization results for metabolites related to fatty acid metabolism, lipids, the TCA cycle, nucleotides and co-factor metabolism are provided in **Table 7.3**. None of the measured metabolites were different between the low- and high-exposure groups at $p < 0.05$. Comparison of the mean concentration for fatty acid and lipid-related metabolites to reference ranges showed acetyl-carnitine, adrenic acid, resolvin E2, and sphinganine were lower

while docohexanoic acid (DHA) and sphingosine were higher. Results for metabolites related to nucleotide metabolism and the TCA cycle were consistent with previously reported values in HMDB. Except for methylnicotinic acid, which did not have any reported values in HMDB, all measured vitamins and co-factors were outside of reference ranges. While not meeting the significance threshold of $p < 0.05$, the observed differences suggests positive associations with linolenic acid and niacin in the high-exposure group, while carnitine, acetyl-carnitine, myristic acid, palmitic acid, DHA, hypoxanthine, succinic acid, malic acid and citric acid were lower in the high exposure group.

7.4.4. Metabolome-wide association study of UFP exposure

To identify additional metabolic changes associated with UFP exposure, we applied PLS-DA to select discriminatory m/z features within the complete HRM dataset. Results from HILIC with positive ESI identified 178 m/z features that discriminated between exposure groups. VIP scores ranged from the threshold of 2.0 to 3.5, and the resulting model exhibited good classification accuracy (**Figure 7.1A**), with R^2/Q^2 values of 0.67/0.67 and 0.77/0.34 for PLS components 1 and 2, respectively. Support vector machine (SVM), 10-fold cross-validation balanced accuracy rate was 98.75% with mean permuted accuracy of 45.6%. Metabolite annotation identified 103 accurate mass m/z 's matching 136 unique chemicals. This included seven confirmed metabolites (Level 1), with alanine, cystine and glutamine having more than one adduct and/or isotope present; 42 metabolites identified based upon adduct and isotopic pairing (Level 3), and 85 based upon

matching accurate mass alone (Level 4). The remaining 70 detected m/z features provided no matches (Level 5). Identified metabolites included amino acids, metabolites related to endothelial function, indicators of oxidative stress, co-factors, and environmental chemicals consistent with polycyclic aromatic hydrocarbons (PAH) and nicotine exposure (**Table 7.4**). The complete results are provided in **Supplementary File 4**.

Comparison of m/z features from RPC with negative ESI also showed metabolic associations with UFP exposure. PLS-DA identified 138 m/z features discriminating between low- and high-exposure, with VIP scores ranging from the threshold of 2.0 to a maximum of 3.6. The selected features provided good classification accuracy (**Figure 7.2A**), with R^2/Q^2 of 0.73/0.72 and 0.83/0.38 for PLS components 1 and 2, respectively. The 10-fold cross-validation balanced accuracy rate was 98.75% with mean permuted accuracy of 55.2%. Annotation resulted in 74 m/z matching 83 unique metabolites. This included four confirmed metabolites, five metabolites meeting Level 3 identification criteria, and 74 matches based upon Level 4. The remaining 64 m/z features provided no matches (Level 5). Similar to the HILIC data, identified metabolites included amino acids, endothelial function related metabolites, co-factors and environmental chemicals arising from exposure to PAHs and nicotine (**Table 7.5**). Associations with sterols and glycan-related metabolites were also detected. The complete PLS-DA and annotation results for RPC with negative ionization are provided in **Supplementary File 4**.

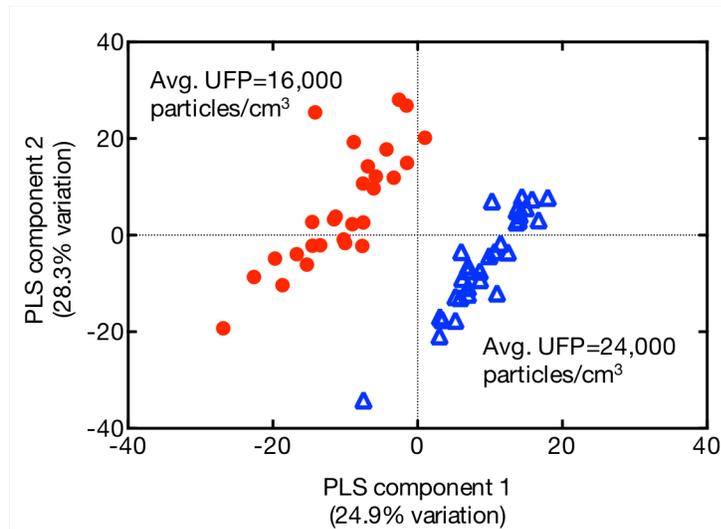
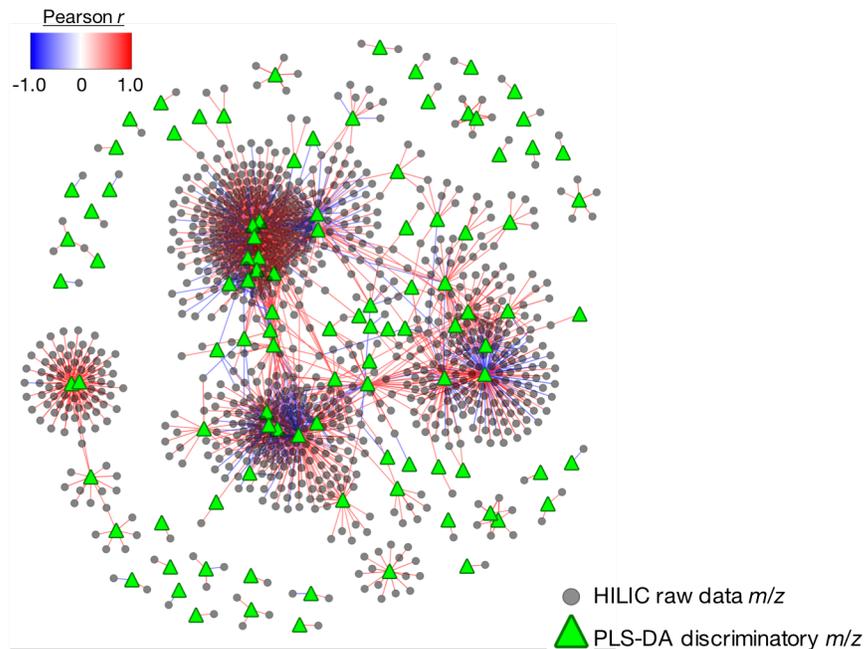
A**B**

Figure 7.1. Metabolome-wide association study of UFP exposure using HILIC HRM data. A) PLS-DA identified 178 *m/z* features discriminating CAFEH participants with low and high exposure, clearly separating the two groups. B) Metabolic network structure using PLS-DA selected features and raw HILIC data identified an additional 928 *m/z* features strongly correlated with the discriminatory metabolites. Correlation network data was then used to evaluate metabolic pathway enrichment.

Table 7.4. Relevant metabolites identified by cross-sectional PLS-DA of HILIC with positive ESI data

<i>m/z</i>	Retention time (s)	Adduct	Identity	Mass error (ppm)	Identification confidence ^a	VIP ^b	Cohen's <i>d</i>	<i>p</i>
134.0450	201	M+H	Aspartic acid	1.3	1	2.56	-0.73	0.007
279.2322	390	M+H	Alpha/gamma-Linolenic acid	1.1	1	2.22	-0.62	0.021
166.0534	222	M+H	Methionine sulfoxide	1.0	1	2.46	0.70	0.010
241.0314	258	M+H	Cystine	1.0	1	2.61	0.74	0.006
134.0190	180	M+2Na-H	Alanine	1.4	1	2.59	0.74	0.006
129.1024	250	M+H-H2O	Lysine	-3.4	1	2.71	0.78	0.004
169.0586	200	M+Na	Glutamine	1.4	1	3.35	1.01	2.98E-04
157.1225	491	M+H	4-Hydroxyphenol	1.0	3	2.17	0.60	0.024
225.1324	273	M+Na	Dimethylarginine	1.0	3	2.79	0.79	0.003
217.0821	193	M+ACN+Na	FAPy-adenine	5.9	4	2.58	-0.74	0.006
233.0788	144	M+ACN+H	gamma-Carboxyglutamic acid	8.6	4	2.40	-0.68	0.012
261.1448	173	M+ACN+H	Pantothenic acid	1.3	4	2.01	0.54	0.037
376.0578	200	M+K	S-(Hydroxymethyl)glutathione	0.6	4	2.12	0.59	0.027
335.1255	135	M+H-H2O	Cotinine glucuronide	3.6	4	2.25	0.64	0.019
347.0650	201	M+2Na-H	Benzo(a)pyrene diol epoxide	-1.6	4	2.33	0.65	0.015
337.1047	364	M+Na	7,8-Dihydropterotic acid	8.3	4	2.55	0.72	0.007
279.1116	167	M+ACN+Na	Acetyl vitamin K5	4.2	4	3.17	0.94	6.76E-04

a. Identification confidence based upon Schymanski et al. (2014)

b. VIP=Variable importance projection score

Table 7.5. Relevant metabolites identified by cross-sectional PLS-DA of C₁₈ with negative ESI data

<i>m/z</i>	Retention time (s)	Adduct	Identity	Mass error (ppm)	Identification confidence ^a	VIP ^b	Cohen's <i>d</i>	<i>p</i>
197.0684	27	M+Na-2H	Cotinine	6.2	1	2.43	-0.69	0.013
132.0305	31	M-H	Aspartic acid	1.7	1	2.48	-0.68	0.011
128.0355	28	M-H ₂ O-H	Glutamic acid	5.9	1	2.03	-0.55	0.039
118.0512	25	M-H	Threonine	1.9	1	2.06	0.55	0.036
151.0404	306	M-H	3,4-Dihydroxyphenylacetaldehyde	1.9	3	2.14	-0.57	0.029
157.0257	156	M-H	Dihydrorootic acid	1.4	4	2.15	-0.59	0.028
272.4254	57	M-3H	3-Sialyl Lewis	2.2	4	2.17	-0.59	0.027
145.0509	49	M-H ₂ O-H	Fucose	5.4	4	2.01	-0.55	0.041
107.0504	102	M-H	Cresol	1.2	4	2.01	0.53	0.041
289.1216	372	M-H	trans-dimethylbenz(a)anthracene -5,6-dihydrodiol	6.16	4	2.01	0.54	0.041
494.1544	478	M+Cl	5-Methyltetrahydrofolic acid	3.3	4	2.05	0.55	0.037
165.0924	363	M+Cl	Agmatine	7.0	4	2.43	0.68	0.013
157.0509	340	M+Na-2H	Tetrahydropteridine	8.5	4	2.67	0.73	0.006
445.2942	384	M-H	24-Oxo-1alpha2325-trihydroxyvitamin D3	3.9	4	2.82	0.78	0.003
475.1970	368	M-H	2-Methoxyestrone 3-glucuronide	-0.7	4	2.80	0.78	0.004
322.9821	133	M+Na-2H	Ellagic acid	3.6	4	2.83	0.79	3.38E-03

a. Identification confidence based upon Schymanski et al. (2014)

b. VIP=Variable importance projection score

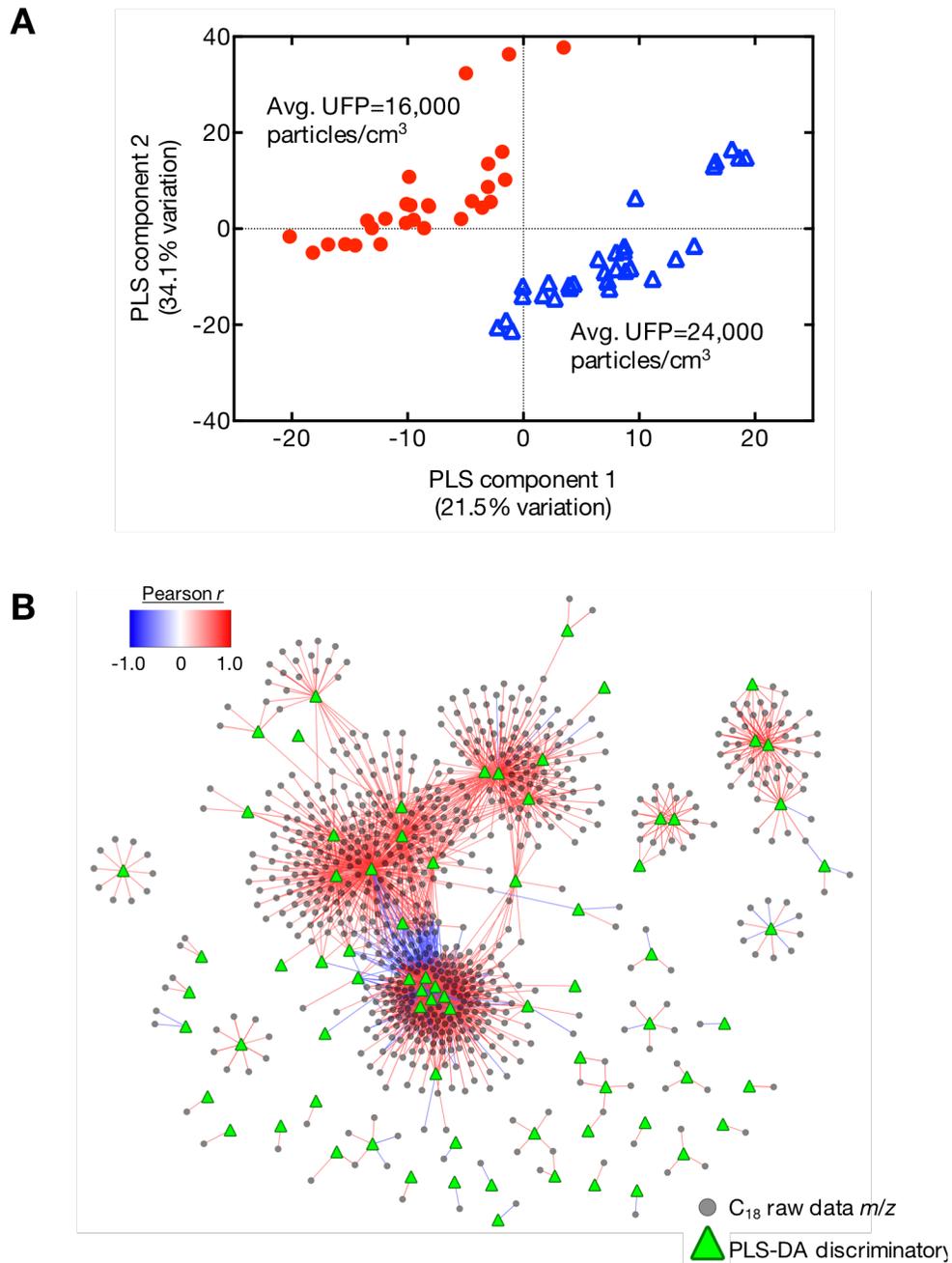


Figure 7.2. Metabolome-wide association study of UFP exposure using C₁₈ with negative ESI HRM data. A) PLS-DA identified 138 m/z features discriminating CAFEH participants with low and high exposure, providing clear separation between the two groups. B) The metabolic network structure was similar to HILIC, and included an additional 721 m/z features meeting the correlation and p thresholds.

7.4.5. *Metabolite correlation and pathway enrichment*

Following identification of PLS-DA selected m/z features, a metabolome-wide Pearson correlation analysis was completed by testing for additional associations within the raw data. Combined with pathway enrichment in Mummichog, this approach improves detection of metabolic changes associated with exposure while reducing false positives due to the use of a random sampling permutation approach to annotation and pathway mapping. The correlation networks for HILIC with positive ESI and RPC with negative ESI are provided in **Figure 7.1B** and **Figure 7.2B**, respectively. For HILIC, an additional 928 m/z features were correlated with the PLS-DA results. Clustering patterns suggest a large degree of centrality, with three large clusters exhibiting a high-degree of connectivity, with two additional separate clusters containing greater than ten m/z features. Pathway enrichment showed exposure-associated alterations in 34 different metabolic pathways that had 5 or more annotated metabolites within the correlation network (**Table 7.6**). These included pathways related to amino acids, oxidative stress, nucleotide metabolism, cofactors, xenobiotics and fatty acid metabolism.

A similar correlation network was obtained for the RPC data, which included an additional 721 m/z features. Clustering patterns included three large clusters exhibiting high connectivity and two separate clusters containing greater than 10 m/z features. Pathway enrichment results are shown in **Table 7.7**. Sixteen pathways were identified, including amino acid pathways, microbiome-related metabolites, fatty acid metabolism, cofactors, nucleotide metabolism and the TCA

cycle. Twelve pathways were enriched in both datasets, which included amino acid metabolism, microbiome pathways, co-factors, fatty acid and nucleotide metabolism.

Table 7.6. Metabolic pathways associated with exposure using HILIC with positive ESI data

Metabolic pathway ^a	Number of metabolites in correlation network	Total number of metabolites detected in pathway	Mummichog <i>p</i> -value
Alanine and Aspartate Metabolism	14	21	0.0008
Arginine and Proline Metabolism	17	36	0.0011
Ascorbate (Vitamin C) and Aldarate Metabolism	9	18	0.0020
Aspartate and asparagine metabolism	29	57	0.0008
Beta-Alanine metabolism	9	12	0.0008
Butanoate metabolism	12	21	0.0009
Carbon fixation	7	9	0.0009
Galactose metabolism	9	26	0.0237
Glutamate metabolism	9	11	0.0008
Glutathione Metabolism	5	10	0.0066
Glycerophospholipid metabolism	15	39	0.0050
Glycine, serine, alanine and threonine metabolism	25	50	0.0008
Glycolysis and Gluconeogenesis	16	32	0.0010
Glycosphingolipid metabolism	8	23	0.0260
Hexose phosphorylation	6	16	0.0227
Histidine metabolism	14	20	0.0008
Linoleate metabolism	7	19	0.0204
Lysine metabolism	14	22	0.0008
Methionine and cysteine metabolism	22	44	0.0008
Pentose phosphate pathway	11	33	0.0270
Porphyrin metabolism	7	21	0.0398
Propanoate metabolism	7	15	0.0046
Purine metabolism	18	48	0.0049
Pyrimidine metabolism	20	40	0.0009
Pyruvate Metabolism	8	15	0.0019
Selenoamino acid metabolism	7	15	0.0046
Sialic acid metabolism	7	20	0.0288
Tryptophan metabolism	26	62	0.0013
Tyrosine metabolism	31	83	0.0026
Urea cycle/amino group metabolism	18	44	0.0023
Valine, leucine and isoleucine degradation	14	28	0.0011
Vitamin B3 (nicotinate and nicotinamide) metabolism	10	17	0.0010
Vitamin B9 (folate) metabolism	6	12	0.0044
Xenobiotics metabolism	20	62	0.0256

a. Pathways in bold were also enriched in RPC with negative ESI data (Table 7.7).

Table 7.7. Metabolic pathways associated with exposure using C₁₈ with negative ESI data

Metabolic pathway ^a	Number of metabolites in correlation network	Total number of metabolites detected in pathway	Mummichog <i>p</i> -value
Alanine and Aspartate Metabolism	8	18	0.0083
Ascorbate (Vitamin C) and Aldarate Metabolism	9	19	0.0046
Beta-Alanine metabolism	7	12	0.0031
Butanoate metabolism	8	21	0.0238
De novo fatty acid biosynthesis	11	22	0.0026
Fatty acid activation	10	18	0.0022
Glutamate metabolism	7	11	0.0024
Histidine metabolism	9	20	0.0061
Leukotriene metabolism	13	34	0.0099
Linoleate metabolism	7	19	0.0382
Lysine metabolism	10	23	0.0061
Pentose and Glucuronate Interconversions	6	10	0.0038
Purine metabolism	19	39	0.0017
Pyrimidine metabolism	16	32	0.0018
TCA cycle	8	13	0.0022
Vitamin B3 (nicotinate and nicotinamide) metabolism	8	14	0.0027
Vitamin B9 (folate) metabolism	7	12	0.0031

a. Pathways in bold were also enriched in HILIC with positive ESI data (**Table 7.6**).

7.4.6. Metabolic correlation with inflammatory and endothelial biomarkers

To provide additional insight into how metabolic associations with UFP exposure are related to disease pathobiology, *m/z* features from the network correlation and pathway enrichment were tested for correlation with inflammatory (IL-6, TNF-RII, CRP) and coagulation (fibrinogen) biomarkers (**Figure 7.3**). Combining the results from HILIC with positive ESI and RPC with negative ESI, 197 *m/z* features were associated with the biomarkers, resulting in formation of four different clusters centered on each of the biomarkers. Comparison of the different clusters showed TNF-RII contained the largest number of *m/z* features (n=117), followed by CRP (n=29), IL-6 (n=27) and fibrinogen (n=23). Within this data driven network, fibrinogen appeared as a central hub linking the three inflammatory biomarkers. Pathway enrichment using annotated metabolites for the four clusters identified different metabolic associations with each biomarker. TNF-RII levels showed associations with pathways related to amino acid metabolism, co-factors and chemical detoxification; IL-6 pathways were related to purine and microbiome metabolites. CRP associated pathways included those related to protein synthesis and the degradation of purines, while arginine and proline metabolism were associated with the fibrinogen cluster.

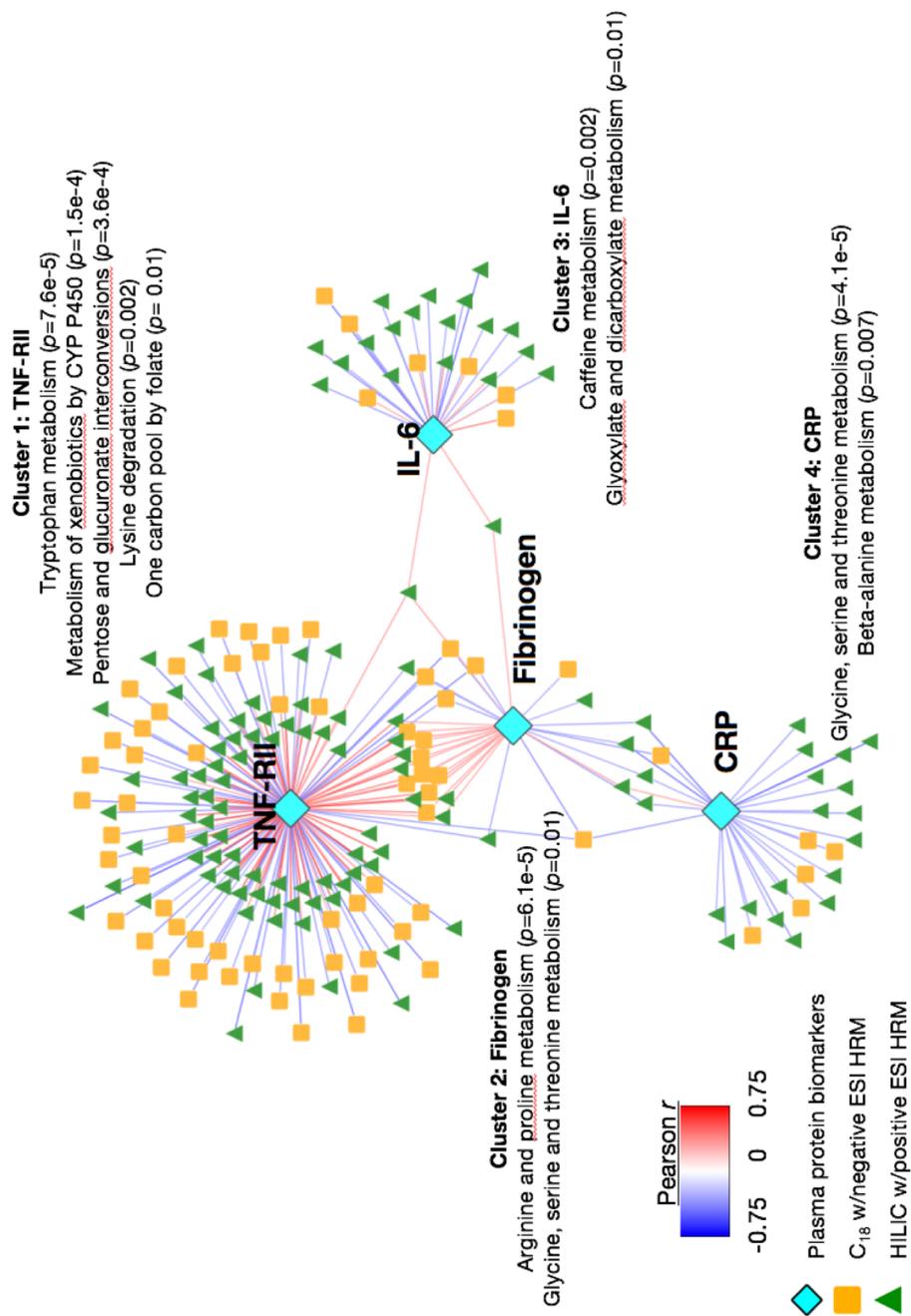


Figure 7.3. Correlation network structure of m/z features associated with UFP exposure from HILIC and C₁₈ HRM data with independently measured biomarkers of inflammation and coagulation. Associated pathways within each cluster was determined using metabolite set enrichment analysis. TNF-RII= Tumor necrosis factor-receptor II; IL-6= Interleukin 6; CRP= high-sensitivity C-reactive protein.

7.5. Discussion

Although long-term exposure to air pollution has been associated with cardiopulmonary diseases, the underlying mechanisms and mode of action are largely unknown. In this pilot study, we applied a hybrid targeted and untargeted HRM methodology to identify metabolic changes associated with averaged year-long exposure to near-highway UFP. Initial characterization of amino acids, clinical markers, lipid/fatty acid metabolites, co-factors and cellular respiration was completed to assess exposure-associated changes to key metabolic intermediates. This was followed by identification of global metabolomic differences between the two groups at both the metabolite and pathway level. Association of metabolic phenotype with independently measured biomarkers of coagulation and inflammation provides additional insight into biological alterations from exposure to traffic-related pollutants.

By providing absolute concentrations of metabolites, the use of reference standardization allows evaluation of physiological levels of central human metabolites and comparison to expected ranges (Go et al. 2015). In this study, concentrations of 79 metabolites were determined. Compared to HMDB, 26 were outside reported ranges, including 17 that were less than reported values. While this population was differentiated into low- and high-exposure groups, as a whole this cohort carries a higher burden of UFP exposure due to residential proximity to main highways within an urban area. Thus, exposure history could contribute to these differences; however, since demographic geographic information is not

provided for values reported in HMDB, it is not possible to evaluate if exposure alone could account for the observed differences.

Comparison of the metabolite levels between the low and high-exposure groups identified five metabolites significantly different at $p < 0.05$. Elevated methionine sulfoxide (MetSO) and cystine (CySS) are consistent with increased oxidative stress. In humans, the major extracellular thiol/disulfide redox couple is cysteine and its disulfide form, CySS (Go and Jones 2011). Increased oxidation of this couple and subsequent shift towards a more positive steady state redox potential has been shown to result in activation of pro-inflammatory cytokines (Iyer et al. 2009), regulate early events in atherosclerosis (Go and Jones 2005), increase mortality in patients with coronary artery disease (Patel et al. 2016) and is associated with endothelial function (Ashfaq et al. 2008). MetSO is the oxidation product of methionine with reactive oxygen species (ROS). Non-enzymatic oxidation of Met residues is a common post-translational protein modification showing oxidative damage and is a recognized marker of oxidative stress (Moskovitz et al. 1997, Mashima et al. 2003). Changes in arginine (Arg), aspartate (Asp) and glutamine (Gln) suggest alterations to critical processes for endothelial function. Gln is a precursor for *de novo* synthesis of Arg through intestinal uptake of Gln and release of citrulline, which is then converted to Arg at the kidney (Ligthart-Melis et al. 2008). Arg is a direct precursor of nitric oxide (NO), which is generated in the endothelium by endothelial NO synthase (Gornik and Creager 2004). Changes in NO production result in disruption to vascular homeostasis, and have been implicated in a wide range of pathological states

(Tousoulis et al. 2012). These findings are consistent with other metabolic profiling studies of air pollution. In the study by Breitner et al. (2016), short-term exposure to PM_{2.5} and ozone was associated with decreased Arg and elevated Asp in a cardiac cauterization cohort. Similarly, significant changes in Asp and Gln were detected in human lung lavage fluid following controlled exposure to biodiesel exhaust (Surowiec et al. 2016).

Cross-sectional comparison using all detected *m/z* features identified complementary metabolic changes associated with UFP. In addition to Gln, Asp, MetSO and CySS, PLS-DA identified other metabolic intermediates quantified by reference standardization, including linolenic acid, alanine (Ala), lysine (Lys), glutamate (Glu) and threonine (Thr). The observed changes for linolenic acid and Thr were consistent with associations observed in previous MWAS of benzo[a]pyrene (Walker et al. 2016). Decreases in linolenic acid and Glu with exposure provide additional evidence of oxidative stress related changes. Glu participates in a wide range of physiological processes, including cellular metabolism, neurotransmitter biosynthesis and protein catabolism. Glu is a precursor for glutathione, the second major thiol redox couple for antioxidant defense. Linolenic acid is an isomeric polyunsaturated fatty acid that includes an omega-3 (alpha-linolenic acid) and omega-6 (gamma-linolenic acid) species. Polyunsaturated fatty acids are particularly susceptible to lipid peroxidation, which occurs through oxidant attack at carbon-carbon bonds with subsequent hydrogen abstraction and oxygen insertion (Yin et al. 2011). Increased oxidative stress and formation of *in vivo* ROS has been well documented following

exposure to particulate matter (Lodovici and Bigagli 2011, Bates et al. 2015). Elevated 4-hydroxynonenal (4-HNE), a lipid peroxidation biomarker, supports increased formation of ROS associated with UFP. 4-HNE is a relatively stable, major aldehyde product formed during lipid peroxidation that exhibits cytotoxicity and reacts with a wide range of different biomolecules, including proteins, lipids and nucleic acids (Weber et al. 2013). Lipid peroxidation products have also been implicated as a pro-inflammatory signaling molecule and in the pathogenesis of multiple chronic diseases (Selley 1998, Akude et al. 2010).

Metabolites consistent with increased risk for endothelial dysfunction were also identified by HRM. Both dimethylarginine and agmatine have the potential to inhibit NO synthase and were increased in association with UFP exposure. Dimethylarginine exists in an asymmetric (ADMA) and symmetric (SDMA) form; due to the chromatography technique used in this study it was not possible to separate the two isomers and the measured intensity represents the sum of the two. Both ADMA and SDMA form through enzymatic methylation of Arg residues, which is catalyzed by *S-adenosylmethionine protein N-methyltransferase* and methylated by S-adenosylmethionine (SAMe), a critical substrate in sulfur amino acid and one-carbon metabolism. Interestingly, 5-methyltetrahydrofolic acid (5-MTHF), the most biologically active form of folic acid, was associated with exposure. Decreased levels of 5-MTHF have been implicated in dysregulation of homocysteine levels, which has been identified as a tentative biomarker of cardiovascular disease (Jacobsen 1998). SAMe is also a direct metabolite of methionine and precursor to homocysteine. Both forms of

methylarginine contribute to endothelial dysfunction. ADMA is a competitive inhibitor of NO synthase that can result in Arg deficiency and disruption to endothelial homeostasis (Sibal et al. 2010). ADMA has been linked to a number of different risk factors for cardiovascular disease, including hyperhomocysteinemia, hypercholesterolemia, obesity and inflammation (Stuhlinger et al. 2003, Eid et al. 2004). While not directly inhibiting NO, SDMA competes with Arg for endothelial transporters, and has been shown to reduce NO production in intact endothelial cells (Bode-Boger et al. 2006). SDMA has also been identified as a marker of renal function, survival rate following stroke and inflammation (Schepers et al. 2011, Luneburg et al. 2012). Agmatine, which is an endogenously produced decarboxylation product of Arg, inhibits inducible NO synthase and acts as a modulator of NO production (Raghavan and Dikshit 2004). Lys also competes with Arg transport during NO production. Finally, two clotting agents were also identified, which included gamma-carboxyglutamic acid and acetyl vitamin K5.

In addition to changes in endogenous metabolites, chemical biomarkers of exposure to air pollution showed a positive relationship with UFP exposure. These included two PAH metabolites, benzo(a)pyrene diol epoxide (BPDE) and trans-dimethylbenz(a)anthracene -5,6-dihydrodiol (tDMBA-DH). BPDE is a chemical carcinogen formed through oxidation of benzo[a]pyrene via a series of cytochrome P450 enzymes, resulting in a metabolically active species with the ability to bind to a wide range of biomolecules (Delaney and Essigmann 2008, Chung et al. 2010). Benzo[a]pyrene is an incomplete combustion product that is

found in cigarette smoke, grilled foods and automobile exhaust fumes. tDMBA-DH is a precursor to the metabolically active, dipole epoxide metabolite of dimethylbenz(a)anthracene (DMBA). Similar to BP, DMBA is a recognized carcinogen, with modes of action including formation of DNA adducts through binding to dAdo and dGuo DNA residues (Sobinoff et al. 2011). An *m/z* matching the methylphenol cresol was also detected. Cresols, which consist of three different isomers (para-, meta-, ortho-), are combustion products widely distributed throughout the environment. Sources include automobile exhaust, tobacco smoke and combustion of organic materials, with concentrations largely dependent on the distance from the source due degradation in the atmosphere (ATSDR 2008). Cresols are also biomarkers of exposure to toluene, a volatile component of automobile emissions. Although participant selection was biased towards non-smokers, the nicotine metabolites cotinine and cotinine-glucuronide were detected. To assess if any individuals were active smokers, cotinine levels were quantified using reference standardization. Excluding two individuals from the low UFP exposure group with cotinine levels of 4.9 and 1.8 ng/mL, for all other participants cotinine was either not detected or below 1 ng/mL. A plasma cotinine level greater than 10 ng/mL is typically used to identify active smokers. Thus, cotinine levels suggest only passive exposure to cigarette smoke.

The expression of metabolic phenotype arises from micro- and macro-scale interactions between different constituents, including small molecules, RNA, miRNA, DNA and proteins. Coordination of these systems is possible through different regulatory mechanisms and signaling pathways; dysregulation

detected as differentially expressed metabolites often represents biological changes transferred through connected hubs in a metabolic network (Albert 2005). Since disease typically manifests as disruption to homeostatic control, understanding systemic changes associated with exposure can improve insight into the mechanisms underlying exposure-related diseases. Thus, we used a network correlation approach to identify pathways associated with exposure and their relationship with independent measures of inflammation and coagulation. Pathway enrichment analyses showed systemic changes to a range biochemical processes, including pathways related to oxidative stress and endothelial function. These included sulfur amino acid metabolism, linoleate metabolism, glutamate and arginine/proline metabolism. In addition, seven metabolites from porphyrin metabolism were identified using HILIC. Porphyrins are critical co-factors for the synthesis of cytochrome proteins and hemoglobin; alterations in hemoglobin to levels outside normal physiological range has been associated with increased risk for cardiac events (Chonchol and Nielson 2008). Pathways consistent with metabolomic changes observed in obesity and insulin resistant individuals, including branched chain amino acids, gluconeogenesis/glycolysis, glutamate, glutamine, tyrosine and tryptophan metabolism (Newgard et al. 2009, Padberg et al. 2014). Additional metabolic alterations associated with UFP exposure important to human health include inflammatory processes (leukotriene metabolism), the microbiome (sialic acid metabolism, propionate metabolism and butanoate metabolism), mitochondrial bioenergetics (pentose phosphate pathway, fatty acid activation, TCA cycle, pyruvate metabolism, glycolysis and

gluconeogenesis), co-factor metabolism (vitamin B3, vitamin B9 and ascorbate/aldarate metabolism) and nucleotide metabolism (purine and pyrimidine metabolism). Thus, these results show exposure-associated changes were present for the biochemical processes that are central in maintaining homeostasis and health; additional characterization will be needed to evaluate if these alterations occur to an extent consistent with disease.

Proximity to traffic and exposure to UFP has previously been shown to be associated with biomarkers of systemic inflammation and coagulation, including IL-6, CRP, TNF-RII, von Willebrand Factor and fibrinogen (Brugge et al. 2013, Ruckerl et al. 2014); however, many of these findings are based upon short-term, acute exposures. All three inflammatory markers showed clustering with exposure associated *m/z* features and enrichment with metabolic pathways. Metabolic pathways associated with TNF-RII included essential amino acids, one-carbon metabolism and detoxification pathways. TNF-RII is one of two receptors that regulate the biological function of TNF α , a cytokine involved in systemic inflammation and acute reaction, among a number of other signaling processes. TNF α has been shown to suppress cytochrome P450 enzymes and down regulate genes modulated by the aryl-hydrocarbon receptor (AhR), which would have important implications for metabolism of exhaust related air pollutants, including PAHs. Individual species associated with TNF-RII included metabolites related to oxidative stress and sulfuramino acid metabolism, including methionine, 5-MTHF and glutathione disulfide, which is consistent with the link between oxidative stress and inflammation (Reuter et al. 2010). Unlike TNF-RII, CRP and IL-6

levels were only correlated with a limited number of *m/z* features. Enriched pathways associated with CRP were largely related to protein biosynthesis, while IL-6 showed correlation with purine-related metabolites and glyoxylate metabolism. The glyoxylate cycle is a carboxylic acid pathway primarily occurring in bacteria and fungi; association with IL-6 could be due to a recent infection or immune response. While IL-6 increases production of CRP during acute inflammation effects, there were no overlapping metabolites correlated with the IL-6 and CRP clusters. Based upon number of connections and position within the network, fibrinogen was a central hub within the correlation network connecting IL-6, TNF-RII and CRP (**Figure 7.3**). Consistent with the role of fibrinogen in endothelial function, arginine and proline metabolism were associated with the fibrinogen cluster. Fibrinogen is also a key regulator of inflammation, with fibrinogen detected in atherosclerotic plaques and contributing to secretion of IL-6 and TNF α . While the results from this correlation analyses are exploratory, characterizing the relationship between these biomarkers and metabolic features associated with UFP exposure suggests small molecules analysis has potential to identify additional mechanisms of air pollution-associated disease. Additional study in controlled settings is required to more thoroughly delineate these mechanisms.

We acknowledge several limitations of this study. First, participants in this study reside within 500 meters of major highways in an urban area. As a whole, this population carries a higher burden of exposure to traffic-related pollutants than residents in suburban and rural areas. Second, this HRM study was designed

as a pilot study to evaluate metabolic differences from UFP exposure. Participant selection was biased towards non-smokers and for cross-sectional comparison of low and high UFP exposures. Due to the small sample size, it was not possible to stratify based on confounders, including BMI and socioeconomic status. In addition, the results from this study are correlative in nature. We could not account for unknown confounders, nor identify the exact mechanism through which these metabolic associations occurred. Third, while confidence-based methods were used for annotation of metabolites, unidentifiable signals were not extensively characterized. Lack of reference standards and low abundance of many of the m/z features render characterization and identification of these features challenging. Despite these limitations, we identified UFP associated changes to the metabolic phenotype consistent with exposure to automotive emissions, oxidative stress and endothelial function, which are common risk factors for diseases related to air pollution exposure. The results demonstrate the use of high-resolution temporal-spatial models to classify annual exposure to traffic-related air pollution, and that HRM provides a viable platform for untargeted characterization of molecular mechanisms underlying environmental exposures. In addition, the results of this study represent the first HRM characterization of long-term exposure to air pollution using high-resolution models to estimate participant exposure burden. While acute exposures provide insight into mechanisms underlying the immediate biological response to environment, exposures are often ongoing. Improved understanding of how environmental pollutants influence metabolic phenotype can be used to develop

hypotheses that delineate underlying toxicological mechanisms and design interventions that reduce exposure-associated disease outcomes.

7.6. Conclusions

Exposure to traffic-related particulates has been associated with increased risk of adverse cardiovascular events and mortality. Underlying mechanisms appear to include systemic inflammation, endothelial dysfunction and oxidative stress; however, the precise biochemical pathways from exposure to effect are not known. In the present study, we applied high-resolution metabolomics to identify the metabolic effects of near-highway exposures. Time- and activity-adjusted UFP exposure was determined using high-resolution spatial-temporal models, and metabolic differences were evaluated for participants with low and high average, annual exposure. Using a hybrid targeted and untargeted metabolomic approach, we identified metabolic alterations associated with exposure that are consistent with inhibition of nitric oxide production and elevated oxidative stress. These alterations included a shift in the CyS/CySS couple towards a more positive steady-state redox potential, elevated dimethylarginine and elevated levels of 4-HNE, all of which are recognized risk factors for atherosclerosis, endothelial dysfunction and cardiovascular disease. In addition, oxidized metabolites of two PAH compounds causally linked to chemical carcinogenesis were elevated with UFP exposure, supporting the use of HRM as a central platform linking exposure to internal dose and biological response. Finally, metabolic pathway enrichment shows long-term UFP exposure leads to systemic changes in a wide range of

biochemical processes that have important implications for health. Taken together, the results from this study suggest that long-term exposure to air pollution leads to increased oxidative stress and changes to metabolic regulators of endothelial function. Additional study using controlled exposure models and validation cohorts will be needed to more fully elucidate how exposures contribute to these metabolic changes.

**Chapter 8. METABOLOME WIDE ASSOCIATION STUDY OF
POLYBROMINATED (PBB-153) AND POLYCHLORINATED (PCB-153)
BIPHENYLS**

Contributions to this work are as follows. D. Walker: Designed and executed all HRM data analysis, analyzed biomonitoring data, interpreted results, performed literature review, wrote manuscript, created figures and tables; M. Marder: Measured PBB-153 and PCB-153 levels in blood; Y. Yano: Assisted in measuring HRM and PBB-153 measures of plasma Y. Liang: Measured plasma lipids; D. Barr: Oversaw M. Marder and Y. Yano for PBB analysis; G. Miller: Initiated HRM project and provided resources; D. Jones: Supervised D. Walker, oversaw HRM measures and efforts; provided guidance on data analysis; read and edited chapter; M. Marcus: PI of original project, provided samples and population demographics, provided guidance on statistical design; K. Pennell: Supervised efforts of D. Walker, read and edited chapter.

8.1. Abstract

Background: Environmental exposure to persistent organic pollutants, including polychlorinated biphenyls and brominated flame-retardants, is ongoing.

Objectives: Identify metabolic phenotype of exposure to polybrominated biphenyl congener 153 (PBB-153) and polychlorinated biphenyl congener 153 (PCB-153) in two generations of participants enrolled in the Michigan PBB Registry.

Methods: Untargeted, high-resolution metabolomic profiling of plasma collected from 156 participants was completed using liquid chromatography with ultra-high resolution mass spectrometry. PBB-153 and PCB-153 levels were measured in the same individuals by tandem mass spectrometry and tested for dose-dependent changes using a metabolome-wide association study framework. Biological response was evaluated based upon annotated metabolites and metabolic pathway enrichment.

Results: While PCB-153 levels were consistent with similarly aged individuals, PBB-153 concentrations were elevated ($p < 0.0001$) in participants enrolled in the Michigan PBB Registry. Dose-dependent metabolic changes were detected for both PBB-153 and PCB-153 in the two generations. PBB-153 exposure was correlated with changes in pathways related to neurotransmitter biosynthesis, cellular respiration, essential fatty acids and polyamine metabolism, while PCB

associations showed additional changes to neurotransmitter pathways, fatty acid and lipid metabolism. Metabolic alterations for both pollutants were consistent with pathways associated with neurodegenerative diseases and previously identified metabolomic markers of Parkinson's disease. PBB-153 and PCB-153 alterations in fatty acid profiles showed similar patterns when compared to an independent cohort.

Conclusions: Exposure to PBB-153 and PCB-153 results in metabolic variations consistent with neurodegenerative disease and fatty acid metabolism.

8.2. *Introduction*

Due to their widespread use and environmental persistence, exposure to polychlorinated biphenyls (PCBs) and brominated flame-retardants is prevalent throughout the United States. Biomonitoring studies have shown detectable levels are still present in blood, breast milk, adipose and brain tissue, even though manufacture and use of many of these compounds has been phased out (Barr et al. 2005, Sjodin et al. 2008, Hopf et al. 2009). Exposure to the developing fetus can occur through maternal transfer *in utero*, and accumulation in lipid-rich breast milk can lead to continued exposure throughout infancy and critical developmental periods (Joseph et al. 2009). While the toxic effects of these compounds have been well documented in model systems, uncertainty exists in the role of environmental exposure and human health.

PCBs and BFRs are associated with a range of diseases. The International Agency for Research on Cancer (IARC) lists PCBs as carcinogenic to humans (Group 1), with evidence suggesting increased risk of malignant melanoma, non-Hodgkin lymphoma and cancer of the breast following exposure (IARC 2015). PCBs have also been identified as possible obesogens (La Merrill et al. 2013), exhibit endocrine disrupting behavior (Bonefeld-Jorgensen 2010), and alter thyroid function (Persky et al. 2001). In addition, exposure to PCBs has shown association with Parkinson's disease (PD) in both epidemiology studies and animal models (Caudle et al. 2006, Steenland et al. 2006). Less information is known about the BFRs. Bromine analogues of PCBs, polybrominated biphenyls (PBB) have been recognized as probably carcinogenic to humans (Group 2A) by

IARC (IARC 2015), showing association with digestive system cancer and elevated risk for lymphoma (Hoque et al. 1998). Polybrominated diphenyl ethers (PBDEs), which largely replaced PBBs in commercial flame retardant mixtures, have exhibited neurotoxicity (Birnbaum and Staskal 2004), endocrine disruption and immunotoxicity (Fowles et al. 1994, Makey et al. 2016).

Much of what is known about health effects of PBBs in humans has been learned from the Michigan PBB Registry. In 1973-1974, Michigan residents were exposed to PBB due to inadvertent contamination of livestock feed with the brominated flame retardant, Firemaster. As a result, contaminated milk, beef, and other farm products were sold to Michigan residents, resulting in substantial exposure to PBB (Fries 1985). To address public health concerns, the Michigan Department of Community Health enrolled approximately 4,000 individuals for long-term health monitoring. The study was eventually expanded to include offspring of those exposed during the original contamination and their children, with data collection still ongoing (Small et al. 2009). Findings in the original cohort have included elevated risk for breast cancer (Henderson et al. 1995, Terrell et al. 2016), thyroid dysfunction (Anderson et al. 1978) and decreased immune function (Bekesi et al. 1979). Endocrine effects of daughters born to mothers who consumed contaminated food have also been documented, including below median Apgar scores (Terrell et al. 2015), dose related increase in miscarriages (Small et al. 2011) and earlier age of menarche and appearance of pubic hair (Blanck et al. 2000). Male offspring with high exposure *in utero* have reported slower growth (Small et al. 2009) and increased genitourinary problems

(Small et al. 2009). In addition, altered sex ratio has been associated with the father's exposure (Terrell et al. 2009). Thus, this cohort provides the opportunity to evaluate the biological response of exposure to PBBs to delineate mechanisms underlying exposure-associated diseases.

Study of metabolic response to exposure shows environmental factors contribute greatly to the metabolic phenotype (Jeanneret et al. 2014, Salihovic et al. 2015, Wang et al. 2015, Walker et al. 2016, Walker et al. 2016, Walker et al. 2016, Carrizo et al. 2017). High-resolution metabolomics (HRM) profiling of biological specimens such as blood and urine provides a central platform linking exposure to internal dose and biological response (Walker et al. 2016, Walker et al. 2016). The use of untargeted high-resolution mass spectrometry with advanced computational methods shows that routine detection of upwards of 20,000 chemicals is possible, which includes endogenous metabolites, microbiome-related chemicals, lipids, compounds arising from diet, environmental chemicals and pharmaceuticals (Uppal et al. 2016). While originally developed as a platform for precision medicine, HRM allows the application of metabolome wide association (MWAS) framework to identify the metabolic variations associated with exposure, providing a means to assess mechanisms underlying chemical toxicity, the biological changes that lead to disease and metabolic phenotypes for retrospective classification of exposure. Therefore, applying these methods to the study of PCB and BFR exposure in humans can provide improved understanding in how these chemicals have the potential to influence human health.

In this study, we used HRM and an MWAS framework to identify dose dependent variations in the plasma metabolome associated with plasma biomarkers of PBB and PCB exposure. Participants were selected from the Michigan PBB Registry, and grouped based upon whether they were born before or after PBB entered the food chain. Metabolic alterations were identified by testing for correlation with plasma levels of the *ortho*-substituted PBB congener 153 (2,2',4,4',5,5'-hexabromobiphenyl) and PCB congener 153 (2,2',4,4',5,5'-hexachlorobiphenyl) in each age group. We then characterized biological response using metabolite annotation and pathway enrichment, providing new insight into how exposure to persistent organic pollutants alters the metabolome.

8.3. Methods

8.3.1. Study population

In 1976, the Michigan Department of Community Health enrolled approximately 4,000 farmers, chemical workers, and others with PBB exposure risk to participate in the Michigan PBB Registry. At enrollment, participants completed detailed questionnaires to capture demographic, health, and lifestyle information. In addition, a blood sample was collected from most participants. The cohort has been followed prospectively since enrollment, with periodic updates of health status and collection of additional blood samples from original members, children and grandchildren. All original study participants and related offspring are considered members of the registry.

For this study, a subset of 174 individuals from the Michigan PBB Registry was selected for measure of PBB-153, PCB-153 and HRM profiling. Plasma collection occurred over the course of 2011-2014, and included a multi-generational cohort consisting of 1) participants over the age of 16 when the livestock feed contamination occurred 2) participants aged 0-16 (F_0) and 3) conceived after but born to parents residing in the area (F_1). Blood samples were collected in liquid EDTA, centrifuged, aliquoted and stored at -80°C . Participants provided informed consent and all protocols were approved by the Emory University IRB and Michigan Department of Community Health.

8.3.2. Plasma PBB-153 and PCB-153 levels

Plasma levels of PBB-153 and PCB-153 were determined using validated methods based on isotope-dilution gas chromatography-tandem mass spectrometry (Marder et al. 2016). Briefly, 1 mL aliquots of plasma were spiked with 50 μ L of 10 ng/mL internal standard solution containing ^{13}C -PBB-153 and ^{13}C -PCB-153, treated with 2 mL of a 1:1 (vol/vol) formic acid:water solution and extracted twice with 5 mL of hexane. The extract was then eluted through an acidified silica column, brought to dryness and reconstituted in 50 μ L of isooctane.

Analysis was performed using an Agilent 7890A gas chromatograph (GC) interfaced to an Agilent 7000B triple-quad mass spectrometer with electron ionization source (Agilent Technologies, Santa Clara CA). A 2 μ L aliquot of sample extract was injected into a heated, splitless inlet maintained at 325°C connected to a ZB-5HT column (15m \times 0.250 ID \times 0.10 μ m film thickness, Phenomenex, Torrance, CA) with high-purity helium carrier gas at a constant flow of 2.25 mL/min. Analyte separation was accomplished using the following temperature program: 90°C (0.1 min), ramped to 340°C (20°C/min) and held for 5 min, resulting in a total runtime of 17.6 min. The mass spectrometer was operated in multiple-reaction-monitoring mode, and the following transitions at collision energy of 40 eV were used for quantification: PBB-153: 467.8 \rightarrow 307.9; PCB-153: 359.7 \rightarrow 289.9. Concentrations were determined by relative response of the integrated peak areas using the stable isotopic standards added prior to sample preparation. Calculated limits of detection (LOD) were 0.002 and 0.0016

ng/mL for PBB-153 and PCB-153, respectively. Non-detects were replaced with $0.5 \times \text{LOD}$ prior to data analyses.

8.3.3. High-resolution metabolomics

Untargeted metabolic profiling was completed following protocols developed in prior studies (Go et al. 2015, Walker et al. 2016). Briefly, plasma aliquots were removed from storage at -80°C , and thawed on ice, upon which 65 μL of plasma was added to 130 μL of acetonitrile containing a mixture of stable isotopic standards, vortexed, and allowed to equilibrate for 30 minutes. Following protein precipitation, triplicate 10 μL aliquots were analyzed by reverse-phase C_{18} liquid chromatography (Targa 100 mm x 2.1mm x 2.6 μm , Higgins Analytical Inc) with detection by high-resolution mass spectrometry (Q-Exactive, Thermo Scientific, San Jose, CA). Three replicates of each sample preparation were injected and analyte separation was accomplished using water, acetonitrile and solution A (2% [v/v] formic acid in water) mobile phases operating under the following gradient: initial 2 min period of 5% water, 15% acetonitrile, 80% solution A, followed by linear increase to 5% water, 95% acetonitrile, 0% solution A at 6 min held for an additional 4 min. Mobile phase flow rate was held at 0.35 mL/min for 6 min, and then increased to 0.5 mL/min. The high-resolution mass spectrometer was equipped with an electrospray ionization source operated in positive ion mode with spray voltage of 4.5 kV, probe, capillary temperature 275°C , sheath gas flow 45 (arbitrary units), auxiliary gas flow 5 (arbitrary units) and S-lens RF level of 69. Resolution was set at 70,000 (FWHM) and mass-to-charge (m/z) scan range

85-1275. Samples were analyzed in batches of 20, in addition to a quality control (QC) pooled reference sample included at the beginning and end for quantification and standardization.

Upon injection of all study and quality control samples, mass spectral features with replicate coefficient of variation (CV) $\leq 100\%$ were extracted and aligned using apLCMS (Yu et al. 2013) with modifications by xMSanalyzer (Uppal et al. 2013) and batch effect correction by ComBat (Johnson et al. 2007). Detected chemical signals were defined by accurate mass-to-charge ratio (m/z), retention time and intensity, referred to as m/z features throughout. Prior to statistical analysis, replicate injections were averaged and m/z features not detected in $\geq 25\%$ of the participants were removed

8.3.4. Data analyses

Summary statistics were calculated using GraphPad Prism version 6.0 for MacOSX (GraphPad Software, La Jolla CA); metabolomic data processing and analyses were completed in R, version 3.1.2 (R Core Team 2014). To evaluate the relationship between PBB-153, PCB-153 and metabolic phenotype, we applied a network-based metabolome wide association study (MWAS) framework (Go et al. 2014, Uppal et al. 2015, Jones et al. 2016). Association between plasma levels of PBB-153, PCB-153 and detected m/z features was assessed using Spearman rank correlation coefficient, which was calculated separately for each generation with *MetabNet* (Uppal et al. 2015). Correlations exhibiting a Spearman $|r| \geq 0.3$ and $p < 0.01$ were selected for further characterization. We selected an unadjusted

significance threshold of 0.01 to balance Type I and Type II error. The use of unadjusted p values with effect size measures and pathway enrichment has been shown to protect against discarding metabolites with weak but important biological interactions (Go et al. 2014). For these m/z features, effect size variability was evaluated by estimating the 95% confidence intervals using a bootstrap approach (1,000 iterations) and R package RVAideMemoire. The resulting correlation networks were visualized in Cytoscape (Su et al. 2014).

8.3.5. Annotation and metabolic pathway enrichment

High-resolution mass spectrometry provides accurate mass measures of ion m/z , which is related to chemical monoisotopic mass, an intrinsic molecular property. The m/z features correlated with PBB-153 and PCB-153 exposure were first matched to a reference database of 120 metabolites previously confirmed with MS² and co-elution studies (Go et al. 2015). Additional m/z features not matching these metabolites were annotated based upon positive electrospray ionization adducts using the KEGG database (Kanehisa et al. 2012) and HMDB (Wishart et al. 2013). Identities were assigned using evidence scoring (Uppal et al. 2016) and ± 5 parts-per-million (ppm) mass tolerance ($\Delta m_{error}/\Delta m_{theoretical} \times 10^6$). Enriched metabolic pathways were selected using a Mummichog (Li et al. 2013) scoring threshold ≤ 0.05 and the requirement that two or more identified metabolites were present in that pathway. Due to the exploratory nature of this study and the low number of matches for each pathway, we included pathways meeting the scoring threshold and having two or more identified metabolites present. Since matching

Table 8.1. Population characteristics of participants selected for HRM profiling

Characteristic	Total	Age 0-16 during PBB event (F ₀)	Born following PBB event (F ₁)
Number of individuals	156	80	76
Age at blood draw, mean ± SD	39.7 ± 11.0	48.8 ± 4.8	30.2 ± 6.5
Age during contamination event	9.2 ± 4.8	9.2 ± 4.8	NA
Total lipids	629.9 ± 152.5	656.3 ± 153.1	602.2 ± 136.2
Sex, <i>n</i> (%)			
Male	68 (44%)	33 (41%)	35 (46%)
Female	88 (56%)	47 (59%)	41 (54%)
PBB-153			
Non-detects, <i>n</i> (%)	5 (7%)	0	5 (7%)
Plasma levels, mean ± SD (ng/mL)	0.8 ± 4.2	1.5 ± 5.9	0.1 ± 0.2
Plasma minimum - maximum (ng/mL)	ND - 50.4	0.02 - 50.4	ND - 1.5
Lipid normalized levels, 50 th percentile (ng/g)	17.26 ± 1161	53.2 ± 1612	5.3 ± 46.8
PCB-153			
Non-detects, <i>n</i> (%)	0%	0%	0%
Plasma levels, mean ± SD (ng/mL)	0.15 ± 0.19	0.22 ± 0.23	0.08 ± 0.08
Plasma minimum - maximum (ng/mL)	0.01 - 1.2	0.04 - 1.2	0.01 - 0.5
Lipid normalized levels, 50 th percentile (ng/g)	15.1 ± 31.0	20.5 ± 38.3	9.9 ± 14.4

is completed using accurate mass only for some of these pathways, it must be recognized this could result in incorrect annotations and false positive enrichment results.

8.3.6. Comparison to the National Health and Nutrition Examination Survey

To evaluate the exposure levels and metabolic associations observed in this study relative to a representative United States population, plasma fatty acid and lipid adjusted PBB-153 and PCB-153 concentrations were obtained from the 2003-2004 National Health and Nutrition Examination Survey (NHANES) and stratified into two age groups (20-39 and 40-59) for comparison (CDC and NCHS) (accessed January 21, 2017). For the 20-39 age group, 483 and 436 individuals were available with PBB-153 and PCB-153 measures, respectively; the 40-59 age group included 409 and 372. None of the individuals in either age group had measures of both pollutants available.

8.4. Results

8.4.1. Study population

Demographics for the selected participants from generations F₀ and F₁ are provided in **Table 8.1**. Though analyzed in the same instrument run, a subset of participants who were older than 16 when the contamination occurred were excluded from the study due to small sample size (n= 18) and higher risk of co-morbidities. Excluding age at blood draw, which had an 18.6 year difference in average age, total lipid measures and sex distribution were comparable across the two populations. For generation F₀, metabolomic profiling results were available for 80 participants; 76 individuals were available from F₁.

8.4.2. PBB-153 and PCB-153 plasma levels

Summary statistics for plasma levels of PBB-153 and PCB-153 are provided in **Table 8.1**; the distribution for F₀ and F₁ are shown in **Figure 8.1**. Skewness and kurtosis for all were greater than 2.5 and 7.2, respectively, for plasma levels of both analytes. Thus, non-parametric approaches were used in all statistical testing. PBB-153 and PCB-153 had a high rate of detection for both groups. PBB-153 was detected in all individuals from the F₀ group, while five individuals from F₁ had levels below the LOD. PCB-153 was detected in all individuals from both generations. For F₀, median (25th; 75th percentile) lipid adjusted concentrations of PBB-153 and PCB-153 were 53.2 (25.8; 92.1) and 20.5 (14.5; 38.2) ng/g, respectively. Levels detected in the F₁ group were lower, with median lipid-adjusted concentration of 5.3 (2.8; 9.2) and 9.9 (6.1; 16.1) ng/g for PBB-153 and

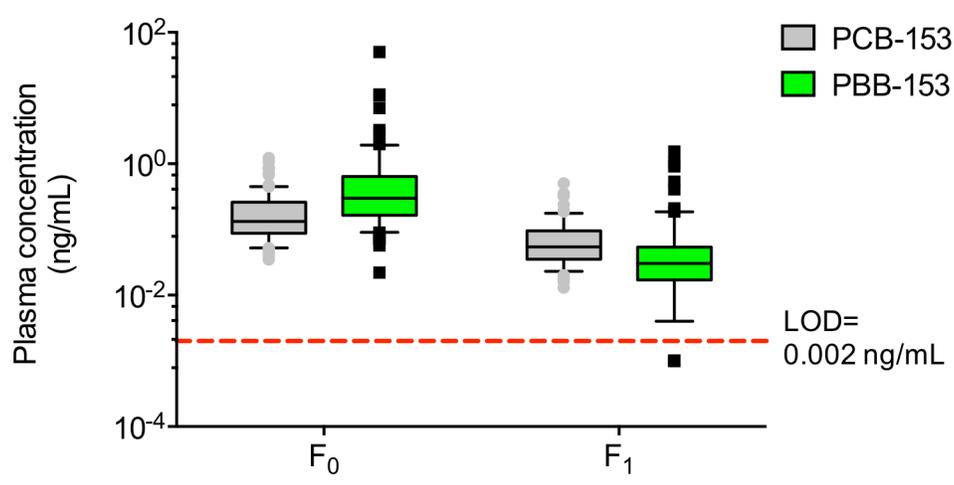


Figure 8.1. PCB-153 and PBB-153 congener levels (log scale) measured in plasma obtained from individuals in generation F₀ and F₁.

PCB-153, respectively. Spearman correlation between fresh-weight and lipid adjusted values showed high correlation for both PBB-153 (Spearman $r= 0.98$, $p< 0.0001$) and PCB-153 (Spearman $r= 0.96$, $p< 0.0001$); therefore, non-adjusted levels (ng/mL) were used in all further analyses, excluding comparison to NHANES data

To evaluate similarities in PBB-153 and PCB-153 exposure, we compared levels within and across the two generations. For both groups, correlation between the two congeners exhibited a positive Spearman correlation (F₀: $r= 0.50$, $p< 0.0001$; F₁: $r= 0.54$, $p< 0.0001$). Pearson correlation was also calculated for PBB-153 and PCB-153 to evaluate if the relationship between the two chemicals was linear. For both F₀ ($r= 0.07$, $p= 0.53$) and F₁ ($r= 0.21$, $p= 0.07$), Pearson correlation showed no association, suggesting a monotonic but non-linear relationship. Comparison across F₀ and F₁ indicated both species were elevated with age, with PBB-153 15-fold (Mann-Whitney $p< 0.0001$) and PCB-153 2.6-fold (Mann-Whitney $p< 0.0001$) higher in F₀.

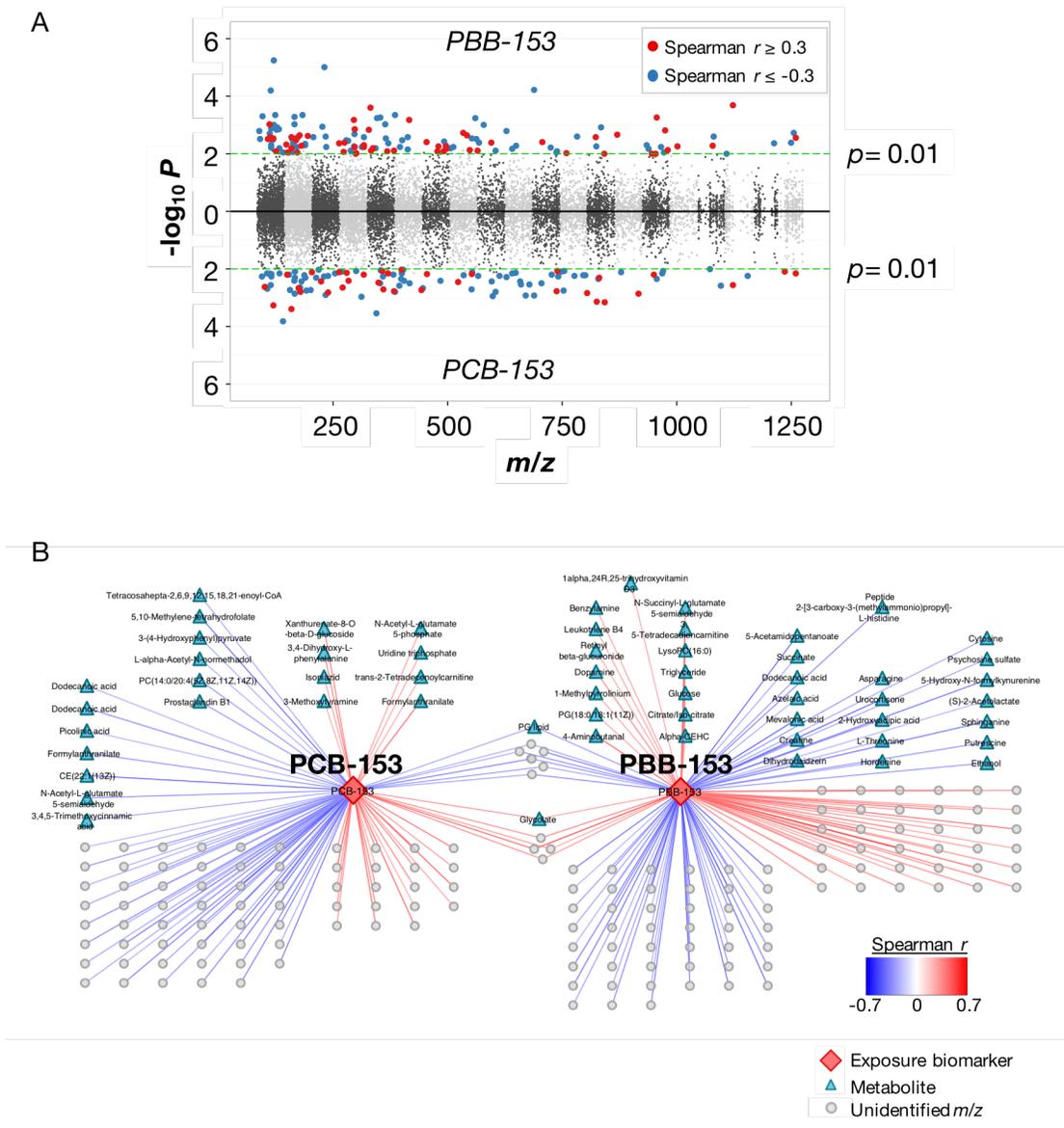


Figure 8.2. Results of PBB-153 and PCB-153 metabolome wide association study in F_0 . Metabolic changes associated with exposure were detected for both pollutants (A). Network correlation analysis showed only a limited number of metabolites were correlated with exposure to both PBB-153 and PCB-153 and different networks were obtained (B).

Table 8.2. Selected metabolites correlated with PBB-153 and PCB-153 in F₀

<i>m/z</i>	Time (s)	Identity	Spearman <i>r</i>	<i>P</i> -value	Spearman <i>r</i> 95% CI
<i>PBB-153</i>					
346.2665	584.6	Sphinganine	-0.54	0.0007	(-0.74, -0.28)
297.0437	110.2	5-Hydroxy-N-formylkynurenine	-0.53	0.0020	(-0.73, -0.24)
96.0562	50.7	Creatine	-0.42	0.0005	(-0.59, -0.22)
115.0505	591.7	Asparagine	-0.41	0.0013	(-0.64, -0.13)
89.1079	588.5	Putrescine	-0.41	0.0016	(-0.61, -0.16)
169.1083	589.6	Acetylcarnitine	-0.37	0.0032	(-0.59, -0.1)
112.0510	51.6	Cytosine	-0.34	0.0019	(-0.53, -0.11)
162.9983	348.9	Succinate	-0.33	0.0041	(-0.52, -0.11)
207.1745	476.8	25-Hydroxyvitamin D2	-0.33	0.0027	(-0.53, -0.12)
161.0925	315.9	Threonine	-0.32	0.0086	(-0.51, -0.1)
496.3404	524.3	1alpha,24R,25-trihydroxyvitamin D3	0.31	0.0053	(0.09, 0.50)
480.2956	586.6	Retinyl beta-glucuronide	0.31	0.0055	(0.07, 0.51)
169.1225	488.4	Leukotriene B4	0.32	0.0033	(0.09, 0.53)
275.0877	119.0	Citrate; Iso-Citrate	0.38	0.0039	(0.12, 0.58)
195.1129	312.9	Dopamine	0.44	0.0023	(0.16, 0.65)
<i>PCB-153</i>					
956.3670	451.0	5,10-Methylenetetrahydrofolate	-0.56	0.0053	(-0.81, -0.15)
156.0656	377.6	N-Acetyl-L-glutamate 5-semialdehyde	-0.43	0.0019	(-0.64, -0.21)
400.2467	80.1	Prostaglandin B1	-0.40	0.0070	(-0.66, -0.09)
239.0891	334.6	3,4,5-Trimethoxycinnamic acid	-0.34	0.0059	(-0.56, -0.08)
181.0502	113.4	3-(4-Hydroxyphenyl)pyruvate	-0.32	0.0038	(-0.5, -0.11)
106.0293	121.3	Picolinic acid	-0.30	0.0071	(-0.5, -0.08)
150.0901	52.1	3-Methoxytyramine	0.30	0.0066	(0.12, 0.5)
522.9324	64.2	Uridine triphosphate	0.33	0.0036	(0.09, 0.53)
271.0392	71.8	N-Acetyl-L-glutamate 5-phosphate	0.34	0.0023	(0.11, 0.52)
280.1291	521.8	3,4-Dihydroxy-L-phenylalanine	0.46	0.0076	(0.1, 0.72)

8.4.3. High-resolution metabolomic profiling

Following data extraction and alignment, 16,890 *m/z* features were detected with

mean triplicate CV of 38.1%. Participant samples were then grouped by

generation and *m/z* features filtered based on the $\geq 25\%$ missing value threshold.

For generation F₀, 8,376 *m/z* features remained after filtering while F₁ had 8,533

m/z features meeting the missing value threshold. Comparison of the two groups

showed 8,006 *m/z* features were detected in both groups, with 370 and 527 unique

to F₀ and F₁, respectively.

Table 8.3. Metabolic pathway enrichment results for *m/z* features correlated with PBB-153 and PCB-153 in F₀

Pathway	Number overlapping metabolites ^a	Total number metabolites ^b	Mummichog score ^c
<i>PBB-153 MWAS pathways</i>			
Butanoate metabolism	4	25	0.003
Heparan sulfate degradation	2	6	0.005
Chondroitin sulfate degradation	2	6	0.005
Glyoxylate and Dicarboxylate metabolism	2	9	0.011
Urea cycle/amino group metabolism	4	41	0.014
Glycosphingolipid metabolism	3	28	0.018
Aspartate and asparagine metabolism	5	62	0.021
TCA cycle	2	13	0.023
Hexose phosphorylation	2	15	0.031
Arginine and Proline Metabolism	3	35	0.039
<i>PCB-153 MWAS pathways</i>			
Tryptophan metabolism	4	61	0.005
Tyrosine metabolism	4	87	0.016
Urea cycle/amino group metabolism	2	41	0.044
Drug metabolism: Cytochrome P450	2	42	0.047

a. Number of metabolites meeting the correlation criteria

b. Total number of metabolites detected from corresponding pathway using HRM

c. See Li et al. (2013)

8.4.4. F₀ PBB-153 and PCB-153 MWAS

To assess metabolic perturbations due to PBB-153 and PCB-153 for F₀, we applied a Spearman correlation based MWAS to identify metabolites associated with the two chemicals based upon effect size measure (Spearman *r*) and $p \leq 0.01$. MWAS identified 139 and 113 *m/z* features associated with PBB-153 and PCB-153 (**Figure 8.2A**). Both positive and negative associations were present for both chemical species, with the majority negative associations (PBB-153: 81 (58%); PCB-153: 78 (69%)). Correlation network analysis showed different network structures for the two pollutants (**Figure 8.2B**), with 12 *m/z* features exhibiting the same direction of change with both PBB-153 and PCB-153.

Probable identities of the m/z features were determined using the criteria described above. Only a limited number of the m/z features matched known metabolites present in the KEGG and HMDB databases, with 65 unique metabolites matching 66 m/z features, 39 of which were associated with PBB-153 and 24 that were associated with PCB-153. Only glycolate (PBB-153 $r= 0.33$; PCB-153 $r= 0.39$), dodecanoic acid (PBB-153 $r= -0.40$; PCB-153 $r= -0.44$) and a phosphatidylglycerol lipid (PBB-153 $r= -0.38$; PCB-153 $r= -0.35$) were correlated with both chemical species. PBB-153 metabolite variations included co-factor metabolism, neurotransmitters, amino acids, pro-inflammatory signaling metabolites, nucleobases and mitochondrial bioenergetics (**Table 8.2**). Metabolites correlated with PCB-153 have similar biological functions and included metabolites related to cofactor metabolism, inflammation, amine metabolism and neurotransmitter precursors (**Table 8.2**); however, the specific species were different. The complete list of metabolite matches is provided in **Supplementary File 5**.

Metabolic pathway enrichment showed biological response was different for the two pollutants (**Table 8.3**). PBB-153 exposure was associated with changes to microbiome-related pathways, amino acid metabolism and catabolic pathways. Metabolic response to PCB-153 included pathways related to neurotransmitter precursors, nitrogen catabolism and drug metabolism. Only urea cycle/amino group metabolism was associated with exposure to both pollutants.

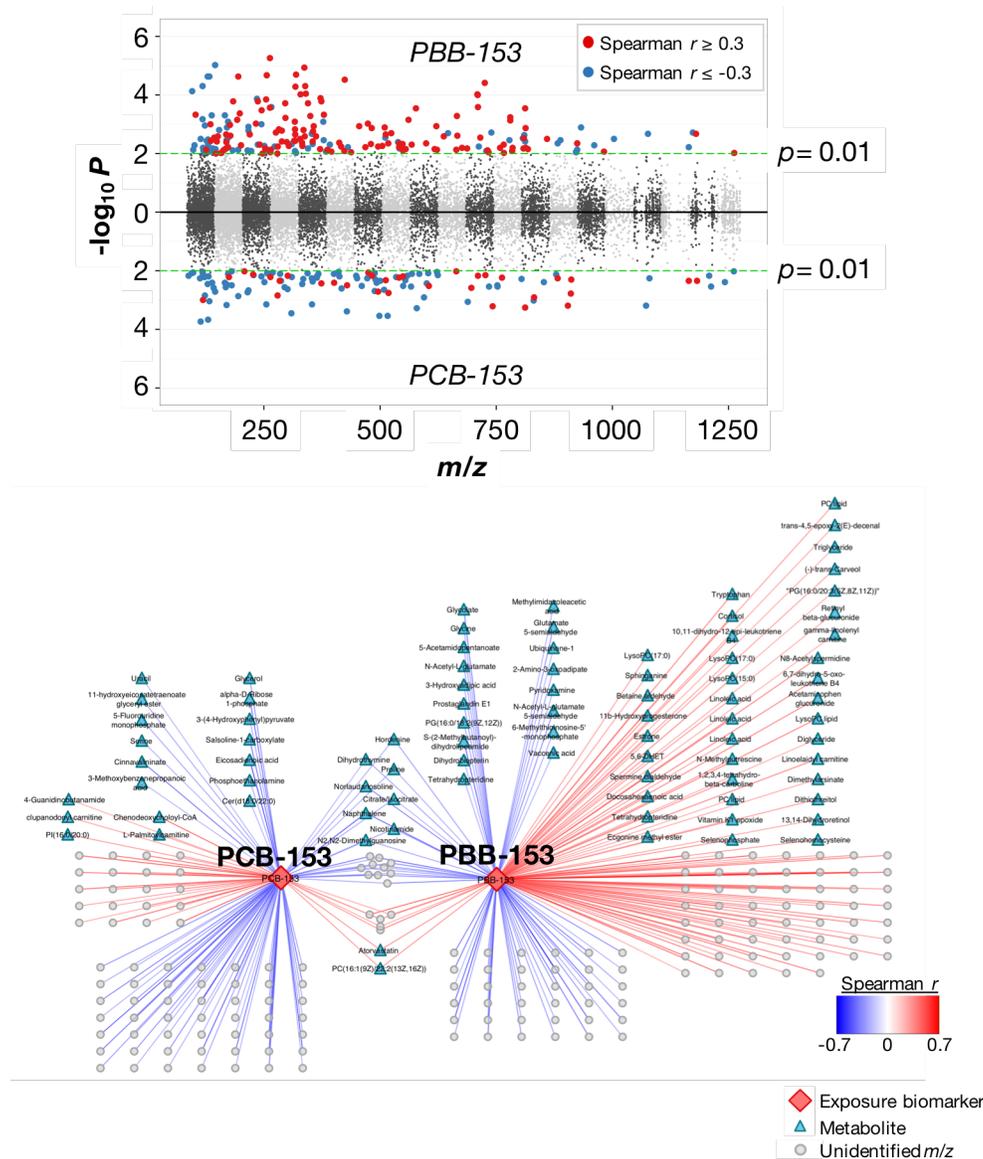


Figure 8.3. Results of PBB-153 and PCB-153 metabolome wide association study in F_1 . As was observed in the older generation, metabolic changes associated with exposure were detected for both pollutants (A). Network correlation analysis for PBB-153 and PCB-153 resulted in different correlation patterns for the two pollutants in F_1 , with only a small percentage correlated to both exposures (B).

Table 8.4. Selected metabolites correlated with PBB-153 and PCB-153 in F₁

<i>m/z</i>	Time (s)	Identity	Spearman <i>r</i>	<i>P</i> -value	Spearman <i>r</i> 95% CI
<i>PBB-153</i>					
118.0502	384.9	Glycolate	-0.51	4.9E-05	(-0.67, -0.29)
120.0026	65.3	Glycine	-0.35	0.0074	(-0.57, -0.08)
215.0144	134.8	Citrate; Iso-Citrate	-0.31	0.0076	(-0.52, -0.07)
257.1361	423.7	Dihydrobiopterin	-0.31	0.0072	(-0.51, -0.08)
133.0760	406.6	Pyridoxamine	-0.29	0.0086	(-0.54, -0.08)
147.0622	116.4	Betaine aldehyde	0.00	0.0096	(0.08, 0.59)
279.0889	517.3	Spermine dialdehyde	0.31	0.0097	(0.06, 0.52)
480.2956	586.6	Retinyl beta-glucuronide	0.32	0.0046	(0.12, 0.51)
284.2949	593.0	Sphinganine	0.36	0.0015	(0.15, 0.54)
524.2725	576.6	Tetrahydrobiopterin	0.36	0.0020	(0.15, 0.56)
246.9963	270.4	Selenohomocysteine	0.38	0.0098	(0.09, 0.63)
263.2369	566.4	Linoleic acid	0.41	0.0003	(0.2, 0.59)
766.4517	89.7	Cortisol	0.41	0.0049	(0.17, 0.6)
350.2302	540.2	11 β -Hydroxyprogesterone	0.42	0.0005	(0.2, 0.59)
103.1234	58.4	N-Methylputrescine	0.43	0.0005	(0.2, 0.61)
169.0753	522.5	Tryptophan	0.49	0.0066	(0.23, 0.68)
264.0901	521.0	N8-Acetylspermidine	0.50	0.0090	(0.09, 0.78)
<i>PCB-153</i>					
309.2806	567.9	Eicosadienoic acid	-0.69	0.0004	(-0.83, -0.44)
131.0107	130.7	Glycerol	-0.44	0.0067	(-0.73, -0.11)
142.0265	132.6	Phosphoethanolamine	-0.34	0.0027	(-0.53, -0.13)
367.2444	585.6	Urocortisol	-0.34	0.0079	(-0.58, -0.09)
88.0398	377.4	Serine	-0.32	0.0071	(-0.52, -0.08)
624.6248	93.1	Cer(d18:0/22:0)	-0.31	0.0074	(-0.52, -0.07)
95.0245	557.8	Uracil	-0.30	0.0085	(-0.52, -0.08)
401.3457	123.4	Palmitoylcarnitine	0.35	0.0040	(0.13, 0.53)
831.5731	435.2	PI(16:0/20:0)	0.39	0.0013	(0.16, 0.57)

8.4.5. F₁ PBB-153 and PCB-153 MWAS

MWAS of plasma PBB-153 and PCB-153 for F₁ detected a greater number of *m/z* features correlated with exposure. This included 307 *m/z* features total, with 208 and 124 associated with PBB-153 and PCB-153, respectively (**Figure 8.3A**).

Only 25 met the association threshold for both pollutants. Both positive and negative correlations were detected; however, unlike F₀, the majority of correlations with PBB-153 were positive (n=133; 64%) while the majority was negative for PCB-153 (n= 91, 73%). As a result, the correlation networks differed for the two chemicals (**Figure 8.3B**).

Table 8.5. Metabolic pathway enrichment results for *m/z* features correlated with PBB-153 and PCB-153 in F₁

Pathway	Number overlapping metabolites ^a	Total number metabolites ^b	Mummichog score ^c
<i>PBB-153 MWAS pathways</i>			
Linoleate metabolism	5	20	0.003
Prostaglandin formation from dihomo gamma-linoleic acid	2	6	0.017
Limonene and pinene degradation	2	6	0.017
Selenoamino acid metabolism	3	18	0.024
Fatty acid activation	3	18	0.024
Urea cycle/amino group metabolism	5	41	0.028
Alkaloid biosynthesis II	2	8	0.029
Glutathione Metabolism	2	9	0.037
Porphyrin metabolism	3	22	0.045
<i>PCB-153 MWAS pathways</i>			
Glycerophospholipid metabolism	3	42	0.005
Pyrimidine metabolism	3	48	0.007
Carnitine shuttle	2	21	0.009
Sialic acid metabolism	2	26	0.013
Glycosphingolipid metabolism	2	29	0.016
Tyrosine metabolism	3	89	0.040
Glycine, serine, alanine and threonine metabolism	2	49	0.048

a. Number of metabolites meeting the correlation criteria

b. Total number of metabolites detected from corresponding pathway using HRM

c. See Li et al. (2013)

Annotation of the *m/z* features associated with PBB-153 and PCB-153 exposure provided 94 matches (31%) total, which included 75 *m/z* matching 72 metabolites for PBB-153 and 29 metabolite matches for PCB-153. Of the 94 matches, 10 were associated with both pollutants, and included citrate/iso-citrate (PBB-153 $r = -0.31$; PCB-153 $r = -0.35$), proline (PBB-153 $r = -0.32$; PCB-153 $r = 0.33$), nicotinamide (PBB-153 $r = -0.36$; PCB-153 $r = 0.37$), naphthalene (PBB-153 $r = -0.57$; PCB-153 $r = -0.64$), dihydrothymine (PBB-153 $r = -0.37$; PCB-153 $r = 0.30$), hordenine (PBB-153 $r = -0.39$; PCB-153 $r = -0.33$), norlaudanosoline (PBB-153 $r = -0.49$; PCB-153 $r = -0.35$), N₂,N₂-dimethylguanosine (PBB-153 $r = -0.36$; PCB-153 $r = -0.32$), atorvastatin (PBB-153 $r = 0.42$; PCB-153 $r = 0.41$) and a

phosphocholine lipid (PBB-153 $r=0.41$; PCB-153 $r=0.39$). Metabolite matches for PBB-153 were consistent with stress and sex hormones, polyamines, one-carbon metabolism, amino acids, pro-inflammatory signaling lipids and essential dietary fatty acids (**Table 8.4**). PCB-153 metabolic variations included fatty acid metabolism, lipids, nucleobases and a metabolite of the stress hormone cortisol (**Table 8.4**). A complete list of annotated metabolites for both pollutants is provided in **Supplementary File 5**.

Similar to F_0 , F_1 metabolic pathway enrichment showed different biological response to the two pollutants (**Table 8.5**). PBB-153 exposure was associated with changes to fatty acid metabolism, pro-inflammatory signaling lipids, antioxidants and nitrogen catabolism pathways. Metabolic response to PCB-153 included lipid metabolism, neurotransmitter precursors and nucleotide metabolism. None of the enriched pathways were associated with exposure to both PBB-153 and PCB-153.

8.4.6. *Generational differences in exposure metabolic phenotype*

PBB-153 and PCB-153 blood levels were correlated with a range of biological variations in both generations. To assess generational differences in metabolic response to exposure, we compared the PBB-153 and PCB-153 metabolic phenotypes across the two generations. Since there were m/z features unique to each generation, we first evaluated if m/z features that were only detected in one of the age groups could explain the difference. For F_0 , 7 m/z features were unique to this generation, with four and three correlated with PBB-153 and PCB-153,

respectively. Excluding a match to 5,10-methylenetetrahydrofolate, none of the other m/z features matched entries in the KEGG and HMDB metabolite libraries. F_1 included 12 m/z features that were specific to this generation, which included eight associated with PBB-153 and four associated with PCB-153. Only two metabolites that were associated with PCB-153 matched known metabolites, which included α -D-ribose-1-phosphate and cinnavalinate.

Direct comparison of m/z features associated with PBB-153 across the two generations showed 11 were associated with this pollutant; however, only three matched known metabolites (**Table VI**). The late retention time (>300 s) of the unidentified m/z features are consistent with lipids and other lipophilic compounds. Although the absolute correlation coefficient values differed slightly across F_0 and F_1 , the direction of change was the same for all. Additional metabolites correlated with PBB-153 in both generations but detected as different adducts included glycolate, 5-acetamidopentanoate, citrate/iso-citrate, sphinganine and a phosphatidylglycerol lipid [PG(18:0/18:1(11Z))]. Only the urea cycle/amino group metabolism pathway was enriched in both generations.

PCB-153 correlations across F_0 and F_1 were similar to PBB-153. Only 5 of the m/z features were correlated with PCB-153 in both generations, which included four non-identifiable metabolites and 3-(4-hydroxyphenyl)pyruvate (**Table VI**). Excluding m/z 830.8173 (F_0 $r=0.35$; PCB-153 $r=-0.43$), the correlation pattern was the same for both generations. Tyrosine metabolism was the only metabolic pathway associated with exposure in F_0 and F_1 .

Table 8.6. Overlapping features correlated with PBB-153 or PCB-153 in both F₀ and F₁

<i>m/z</i>	Time (s)	Identity	F ₀ Spearman <i>r</i>	F ₁ Spearman <i>r</i>
<i>PBB-153</i>				
114.0738	390.7	No match	-0.43	-0.34
166.1227	491.2	Hordenine	-0.37	-0.39
127.0392	349.5	2-Hydroxyadipic acid	-0.36	-0.34
171.0347	438.7	No match	-0.31	-0.36
235.2058	294.8	No match	-0.31	-0.43
480.2956	586.6	Retinyl beta-glucuronide	0.31	0.32
317.6602	519.5	No match	0.31	0.35
317.6551	526.3	No match	0.31	0.33
416.3716	547.7	TG(16:0/16:0/18:2(9Z,12Z))[iso3]	0.38	0.31
263.2318	560.7	No match	0.39	0.55
548.3595	562.3	No match	0.45	0.53
<i>PCB-153</i>				
688.7503	95.5	No match	-0.35	-0.37
1072.7635	117.8	No match	-0.34	-0.53
114.0738	390.7	No match	-0.33	-0.42
181.0502	113.4	3-(4-Hydroxyphenyl)pyruvate	-0.32	-0.33
830.8173	57.4	No match	0.35	-0.43

8.5. Discussion

Metabolic phenotyping by HRM provides a unifying platform linking environmental exposure to internal dose and biological response in humans. The present study demonstrates the application of HRM to a multi-generational cohort exposed to high levels of PBB-153 due to food contamination and chronic exposure to the ubiquitous environmental pollutant, PCB-153. Plasma-based MWAS of the two pollutants showed dose-dependent variations in the metabolome for both PBB-153 and PCB-153. Although similar in structure, the metabolic phenotypes suggest response to the two pollutants differ at the metabolite and pathways level; however, the overall biological function of many of the alterations were consistent for both. Generational differences were also present; suggesting age, length and major route of exposure may contribute to metabolic changes from environmental exposures.

Biomonitoring of persistent organic pollutants, such as PBB-153 and PCB-153, provide a means to assess the current distribution within a population and prevalence of exposure. Although manufacture and use of PBB- and PCB-related commercial products is banned in the United States, almost all participants exhibited detectable levels of the two pollutants. For both F₀ and F₁, PBB-153 levels measured in this study are significantly higher than reported in NHANES (Mann-Whitney $p < 0.0001$). Relative to the 20-39 age group, the measured PBB-153 levels were 4.8-fold higher in the F₁ generation, while levels measured from F₀ were 41-fold higher relative to the 40-59 group. Similarly, the distribution of PBB-153 is consistently higher within this cohort when compared to NHANES,

which reported median (25th; 75th percentile) concentrations for age groups 20-39 and 40-59 of 1.6 (<LOD; 3.0) and 2.9 (1.8; 5.9) ng/g lipid, respectively (Sjodin et al. 2008). Unlike PBB-153, concentrations of PCB-153 were more consistent with the NHANES reported levels of 9.5 (6.3; 15.06) for ages 20-39 and 29.9 (18.6; 48.6) for ages 40-59. PCB-153 levels in the F₁ cohort showed no difference when compared to the NHANES younger age group (Mann-Whitney $p= 0.84$), while levels in F₀ were lower than measured in NHANES (Mann-Whitney $p= 0.006$).

For both pollutants, the NHANES data showed an age-dependent increase in concentration (Mann-Whitney $p < 0.0001$), which was consistent with the trends observed in this study. Association of age with exposure to persistent organic pollutants has been well documented, and is attributed to increased length of exposure (Sjodin et al. 2008, Hardell et al. 2010), changes in metabolism with age (Fangstrom et al. 2005) and exposure during peak emission periods (Quinn and Wania 2012). In this study, the magnitude of change across the two age groups suggests length of exposure and changes in metabolism cannot account for the age-related differences in PBB-153. Relative to the 20-39 age group, NHANES PBB-153 and PCB-153 levels increased 2.0- and 2.8-fold, respectively. In the results from this study, PCB-153 increased 2.5-fold in F₀ relative to F₁, while PBB-153 exhibited a 17.1-fold change with age. Compared to the NHANES data and PCB-153, the significantly greater increase in PBB-153 observed across the two generations is consistent with different routes and levels of exposure to PBB-153 for F₀ and F₁, with F₀ primarily exposure most likely through consumption of contaminated food and F₁ *in utero* (Small et al. 2011).

Taken together, these results show greater PBB-153 exposure in Michigan PBB registry participants than is experienced in the general population. Although the greatest levels were measured in F₀, elevated PBB-153 concentrations in F₁, which were born after the original PBB contamination, show the offspring of parents who consumed the contaminated food still experience elevated PBB exposure.

Metabolic responses to both PBB-153 and PCB-153 detected in the F₀ generation were consistent with biological changes underlying numerous chronic diseases. PBB-153 MWAS identified alterations in key metabolic intermediates in cellular respiration, which includes decreased creatine, acetyl-carnitine and succinate while citrate and glucose were increased with elevated PBB-153 levels. Additional metabolites arising from intermediates in the tricarboxylic acid (TCA) cycle, including asparagine and putrescine and suggest PBB-153 induces changes to mitochondrial energy metabolism. Mitochondrial dysfunction due to oxidative stress from environmental exposures is thought to underlie many age-related diseases, including neurodegenerative diseases, cardiovascular disease, diabetes, cancer and metabolic syndrome (Nicolson 2007, Swerdlow 2011, Lane et al. 2015). Mitochondrial energy production is the result of close coordination between the TCA cycle and electron transport chain, with disruption to either resulting in increased oxidative stress and dysfunction (Pieczenik and Neustadt 2007). Changes in the TCA cycle have been identified in Alzheimer's disease (AD) and environmental exposure models of Parkinson's disease (PD), which included increases in citrate and glucose levels that were consistent with the

results from this study (Shi and Gibson 2007, Lei et al. 2014). In addition to TCA alterations, both dopamine and putrescine were correlated with PBB-153 exposure. PD is characterized by progressive loss of dopamine neurons and subsequent reduction in dopamine input to the striatum (Hatcher et al. 2008), while previous metabolomic study of PD showed polyamine metabolism was associated with rapid disease progression (Roede et al. 2013). Metabolic pathway alterations were also consistent with changes to cellular respiration and the pathophysiology underlying neurodegenerative diseases. Gangliosides are found predominantly in nervous tissue and are suspected to contribute to PD pathogenesis (Grey et al. 2015). Although the populations from this study are too small to draw firm conclusions, these findings show exposure-related changes that are consistent with the pathophysiology underlying neurodegenerative diseases. In addition, mechanistic and epidemiology findings have identified exposure to persistent organic pollutants, including other brominated flame-retardants and PCBs, as possible etiological agents in PD and AD (Caudle et al. 2006, Hatcher et al. 2008, Caudle et al. 2012, Bradner et al. 2013, Richardson et al. 2014).

Although individual metabolites differed, metabolic changes associated with PCB-153 also showed alterations to neurotransmitter-related pathways, which included changes to both tryptophan and tyrosine metabolism. Metabolites associated with PCB-153 included 4-hydroxyphenylpyruvic acid (4-HPAA), 3,4-dihydroxy-L-phenylalanine (L-dopa) and 3-methoxytyramine (3-MTT). In humans, dopamine formation is accomplished through conversion of tyrosine (which can be synthesized from 4-HPAA by phenylalanine hydroxylase) to L-

dopa by tyrosine hydroxylase. L-dopa, which was positively associated with PCB-153 levels, is an immediate precursor to dopamine. Although dopamine was not associated with PCB-153, it was positively correlated with PBB-153 exposure, suggesting the possibility of interaction of both pollutants with this pathway.

Previous studies suggest dysregulation of TH can contribute to PD, AD and type 2 diabetes (Priyadarshini et al. 2012, Tabrez et al. 2012), with TH levels decreased in the substantia nigra and striatum regions of the brain. *In vitro* studies have also shown that exposure to ortho-substituted PCBs can reduce TH activity (Choksi et al. 1997); however, this effect was not observed in animal models (Caudle et al. 2006). 3-MTT is a methylated metabolite of dopamine and was elevated with exposure. This metabolite has been identified as a neuromodulator that can affect behavior and induce intracellular signaling by activation of trace amine-associated receptor 1 (TAAR-1) (Sotnikova et al. 2010). Thus, PCB-153 correlated metabolic variations in F₀ were consistent with pathways implicated in neurodegenerative diseases.

In F₁, PBB-153 metabolic variations also showed changes to neurotransmitter pathways, in addition to polyamine metabolism, mineralocorticoids, oxidative stress-related metabolites, co-factors and fatty acids. Dihydrobiopterin (BH₂), tetrahydrobiopterin (BH₄) and tryptophan were all associated with PBB-153 levels. BH₄ is an essential co-factor for biopterin-dependent aromatic amino acid hydroxylases, which include phenylalanine hydroxylase, tryptophan hydroxylase and TH. As discussed above, these enzymes provide critical functions in the biosynthesis of neurotransmitters, including

dopamine, serotonin, norepinephrine and epinephrine and dysregulation has been associated with both AD and PD. While BH₄ was positively associated with PBB-153 exposure, BH₂ was decreased. BH₂ is the oxidation product of BH₄ and produced during the conversion of phenylalanine, tryptophan and tyrosine to monoamine neurotransmitters. Polyamine metabolites associated with PBB-153 included spermine aldehyde, N-methylputrescine and N-acetylspermidine. As discussed above, polyamine metabolism was associated with rapid PD progression in a metabolomic study of PD (Roede et al. 2013), with N8-acetylspermidine significantly elevated in patients with rapid progression PD. Cerebral spinal fluid levels of polyamines have also been associated with PD (Paik et al. 2010), while increased PA excretion has been linked to neuronal cell death, traumatic brain injury and neuroinflammation (Poyhonen et al. 1990, Zahedi et al. 2010). Oxidative stress related metabolites included decreased pyridoxamine and elevated spermine dialdehyde. Pyridoxamine has been shown to inhibit oxygen radical, reduce lipid peroxidation and block delay progression of diabetic nephropathy by blocking oxidative pathways (Jain and Lim 2001, Dwyer et al. 2015). Spermine dialdehyde is the oxidized metabolite of spermine, which functions directly as a free radical scavenger and protects DNA from oxidative damage (Ha et al. 1998). Thus, these results suggest increased oxidative stress, possibly due to activation of the aryl hydrocarbon receptor (AHR). In addition, metabolic pathways altered in F₁ were consistent with those from F₀ and show similarities in pathways associated with neurodegenerative diseases.

Metabolic variations correlated with PCB-153 in F₁ also included tyrosine metabolism. 4-HPAA, which was decreased with PCB-153 exposure, is a precursor to tyrosine and formed through hydroxylation of phenylalanine by phenylalanine hydroxylase. 4-HPAA showed a similar relationship in F₀, providing additional evidence that exposure to PCB-153 interacts with catecholamine metabolism. Additional changes in metabolic pathways included fatty acid and lipid metabolism. These include glycerophospholipid metabolism, glycosphingolipid metabolism and carnitine shuttle. While these pathways were not associated with PCB-153 in F₀, other metabolomic studies of persistent organic pollutant exposures have shown similar results. In the study by Walker et al. (2016), an MWAS framework was used to test for benzo[a]pyrene (BaP) related changes in the metabolome. Using a pilot study of 30 individuals, untargeted HRM profiling detected changes in linoleate metabolism, carnitine shuttle, glycerophospholipid metabolism and prostaglandin formation from dihomogamma-linoleic acid, which was consistent with findings in model systems (Wang et al. 2014). Interestingly, sphinganine and linoleic acid, which were both negatively correlated with BaP, were also associated with PBB-153 exposure in the present study. Serum levels of the organochlorine pesticides p,p'-dichlorodiphenyldichloroethylene (p,p'-DDE) and hexachlorobenzene (HCB) have also shown association with variation in lipid metabolism, including fatty acids, glycerophospholipids, sphingolipids and glycerolipids. In the study by Salihovic et al. (2015), untargeted metabolomic profiles from an elderly cohort of 1,016 individuals in Sweden was compared to serum levels of p,p'-DDE and

HCB. Annotation of the untargeted data identified 204 metabolites, 16 of which were associated with exposure to the pollutants and included lipids related to cell signaling, energy regulation and membrane composition. Although changes in lipid metabolism were detected for both p,p'-DDE and HCB, the specific metabolites differed. This result is similar to the associations with PBB-153 and PCB-153, suggesting the detected metabolic changes in both studies are due to different molecular mechanisms. To test for changes in the metabolome with exposure to β -hexachlorohexane (β -HCH), HCB, p,p'-DDE and PCB congeners 28,138 and 153, Carrizo et al. (2017) used an untargeted metabolomic approach to compare low- and high-exposed individuals. Similar to the other studies, identified and unidentified metabolic features were consistent with lipid species, including sphingolipids and glycerophospholipids. In addition, changes to lipid, fatty acid and carnitine metabolism were also detected in response to occupational trichloroethylene exposure (Walker et al. 2016), suggesting perturbations to these pathways may be useful as surrogate measures of the exposome.

PBB-153 and PCB-153 associations with fatty acids in F₁ were consistent with previous studies showing exposure induced changes to fatty acid profiles (Hennig et al. 2002, Walker et al. 2016). Therefore, we evaluated if patterns of plasma fatty acid correlations with PBB-153 and PCB-153 in an independent cohort were similar to the results of this study. Plasma concentrations of PBB-153 and PCB-153 measured in the 2003-2004 NHANES dataset were tested for correlation with levels of 24 fatty acids that included saturated, monounsaturated and polyunsaturated fatty acids (**Table 8.7**) for the 20-39 and 40-59 age groups.

Only three polyunsaturated fatty acids, including docosahexaenoic, docosapentaenoic-3 and eicosapentaenoic, met the association criteria of Spearman $|r| \geq 0.3$ and $p < 0.05$. Additional omega-3 and omega-6 fatty acids showed significant associations with PBB-153 and PCB-153 levels at $p < 0.05$ but weaker correlations. These include α -linolenic, γ -linolenic and homo- γ -linolenic acid. Previous results have shown that co-planar PCBs can reduce synthesis of long-chain unsaturated fatty acids by inhibiting delta 5 and delta 6 desaturase activities in the liver, resulting in accumulation of omega-6 fatty acids (Matsusue et al. 1999). Linoleic acid has also been shown to potentiate co-planar PCB-related dysfunction to endothelial cells, which was hypothesized to be due to increased oxidative stress (Hennig et al. 1999). Arachidonic acid, which can be synthesized from homo- γ -linolenic acid and is a key intermediate in inflammation and cellular stress signaling, has also been shown to be negatively correlated with blood PCB levels during pregnancy (Grandjean and Weihe 2003). While this was not observed in the NHANES data, the results from Grandjean and Weihe (2003) provide additional evidence of PCB interaction with desaturase activities and the observed increase in linoleic acid in the present study. Thus, these results suggest exposure to PBB-153 and PCB-153 leads to disruption to essential fatty acid metabolism and has important implications for fatty liver, cardiovascular and metabolic diseases (Scorletti and Byrne 2013, Murphy et al. 2015).

Table 8.7. PBB-153 and PCB-153 correlation with circulating fatty acid levels in 2003-2004 NHANES

Fatty acid species	Lipid name	Classification	PBB-153, n=142		PCB-153, n=133		F ₀ PBB-153, n=147		F ₀ PCB-153, n=116	
			Ages 20-39	Ages 40-59	Ages 20-39	Ages 40-59	Ages 40-59	Ages 40-59		
			Spearman	Spearman	Spearman	Spearman	Spearman	Spearman	Spearman	Spearman
			r	p	r	p	r	p	r	p
cis-Vaccenic	C18:1 n-7	MUFA	0.05	0.4750	0.14	0.0737	0.16	0.0296	0.19	0.0205
Docosenoic	C22:1 n-9	MUFA	0.01	0.9406	0.02	0.8438	0.09	0.2525	0.10	0.2568
Eicosenoic	C20:1 n-9	MUFA	0.08	0.2788	0.17	0.0313	0.26	0.0005	0.18	0.0320
Myristoleic	C14:1 n-5	MUFA	0.17	0.0184	-0.02	0.8318	0.14	0.0691	0.06	0.5070
Nervonic	C24:1 n-9	MUFA	-0.11	0.1692	0.14	0.0757	0.02	0.7668	0.08	0.3755
Oleic	C18:1 n-9	MUFA	0.17	0.0235	0.19	0.0126	0.17	0.0185	0.18	0.0267
Palmitoleic	C16:1 n-7	MUFA	0.10	0.1564	0.06	0.4105	0.16	0.0344	0.16	0.0619
alpha-Linolenic	C18:3 n-3	PUFA	0.03	0.6982	0.24	0.0015	0.08	0.2784	0.18	0.0276
Arachidonic	C20:4 n-6	PUFA	0.13	0.0773	0.20	0.0099	0.04	0.6157	0.15	0.0689
Docosahexaenoic	C22:6 n-3	PUFA	-0.04	0.5534	0.31	3.9E-05	-0.12	0.1151	0.26	0.0019
Docosapentaenoic-3	C22:5 n-3	PUFA	0.15	0.0415	0.31	3.0E-05	0.03	0.6429	0.09	0.2767
Docosapentaenoic-6	C22:5 n-6	PUFA	0.04	0.6185	0.00	0.9912	0.16	0.0347	0.09	0.3091
Docosatetraenoic	C22:4 n-6	PUFA	0.19	0.0081	0.03	0.6875	0.21	0.0044	0.07	0.4032
Eicosadienoic	C20:2 n-6	PUFA	0.00	0.9814	0.11	0.1667	0.17	0.0186	0.18	0.0310
Eicosapentaenoic	C20:5 n-3	PUFA	0.07	0.3520	0.39	1.6E-07	-0.03	0.7135	0.24	0.0030
gamma-Linolenic	C18:3 n-6	PUFA	0.17	0.0207	0.16	0.0435	0.02	0.8357	0.05	0.5332
homo-gamma-Linolenic	C20:3 n-6	PUFA	0.02	0.8044	0.11	0.1374	0.01	0.8514	0.07	0.4328
Linoleic	C18:2 n-6	PUFA	0.04	0.5568	0.13	0.0917	0.05	0.5326	0.12	0.1518
Arachidic	C20:0	SFA	0.00	0.9571	0.12	0.1296	0.12	0.1077	0.05	0.5879
Docosanoic	C22:0	SFA	-0.01	0.8782	0.07	0.3694	0.01	0.8704	0.04	0.6145
Lignoceric	C24:0	SFA	-0.02	0.7924	0.17	0.0311	0.03	0.7092	0.12	0.1593
Myristic	C14:0	SFA	0.15	0.0447	0.10	0.1928	0.16	0.0371	0.12	0.1654
Palmitic	C16:0	SFA	0.13	0.0879	0.19	0.0120	0.16	0.0269	0.22	0.0089
Stearic	C18:0	SFA	0.12	0.0947	0.20	0.0100	0.11	0.1350	0.19	0.0217

We acknowledge limitations in this work. First, a limited sample population of 156 individuals from a specific geographical location with documented, high exposure to PBB was used, and an independent cohort was not available to replicate many of the metabolomic findings. Second, this study only focused on a single biomarker of PBB and PCB exposure. Both of the pollutants were commercially available as mixtures; previous studies have shown exposure to other PBB congeners is present in this area (Marder et al. 2016). Therefore, we could not account for the metabolic effects of co-exposures. Third, this study was largely focused on changes to endogenous metabolites listed in metabolomic databases. Due to the limited sample availability, we could not complete in-depth structural characterization of unidentified metabolites. Lack of reference standards and low abundance will make identification of these features challenging and they were not characterized for this study. Some of the enriched pathways only included 2-3 metabolites, and will require confirmation to verify identity is correct. Finally, the results of this study are correlative in nature. We could not account for unknown confounders, nor identify the exact mechanism through which these metabolic associations occurred. Many of the pathways altered in association with exposure occur in specific tissues, and it is unknown how distal changes are reflected in the blood metabolome. Despite these limitations, MWAS of both pollutants identified changes to pathways previously associated with PBB-153 and PCB-153 exposure. Metabolic changes consistent with diseases linked to these pollutants were also present, and the results show HRM can detect biological response to circulating levels of environmental

chemicals. In addition, comparison to an independent population showed similar patterns of exposure-related changes in fatty acids. Taken together, the results of this study show that HRM is a useful approach to link plasma exposure biomarkers to metabolic variations, providing insight into the mechanisms underlying exposure-related disease.

8.6. Conclusions

Although environmental exposures have been linked to a range of diseases, the underlying mechanisms and initiating events are not well understood. In this study, we applied HRM to identify biological response to PBB-153, a brominated flame retardant with high-exposure in the selected population due to accidental contamination during the 1970's and the ubiquitous environmental pollutant PCB-153. The results show HRM is able detect metabolic phenotypes of PBB-153 and PCB-153 exposure in both generations. Although metabolites differed among the groups, biochemical functions of many of the exposure-associated metabolic changes were similar and resulted in alterations to neurotransmitter synthesis, mitochondrial function and fatty acid pathways. In addition, the results show, on average, that participants in the Michigan PBB Registry who were born before or after the original contamination carry a significantly higher burden of PBB levels than experienced by the general United States population. While our study was limited to a high-exposure population, the metabolite variations detected in this study provide insight into how exposure to persistent organic pollutants influences metabolic phenotype. More generally, these results support HRM as a central

platform linking environmental and molecular epidemiology for understanding how exposure contributes to disease susceptibility.

**Chapter 9. DEPLOYMENT-ASSOCIATED EXPOSURE
SURVEILLANCE WITH HIGH-RESOLUTION
METABOLOMICS**

This chapter is taken from a manuscript that was published in the *Journal of Occupational and Environmental Medicine* (PMID: 27501099). Previous chapters in this thesis have focused on establishing HRM as a platform for linking exposure to internal dose and biological response. This manuscript builds upon these concepts by demonstrating how HRM can be incorporated into an environmental chemical surveillance and bioeffect monitoring system to identify environmental adverse events that occur in human populations. While this chapter is specific to the armed forces, the framework can easily be extended to other domains, including public health monitoring, precision medicine and occupational hazards. Co-authors who directly and contributed to this work and efforts are as follows. D. Walker: Designed and executed statistical analysis, performed literature review, wrote and combined sections, created figures and tables, approved final edits and submission; T. Mallon: Project PI, read and edited final document; P. Hopke: Oversaw targeted PAH measurements, read and edited final document; K. Uppal: Provided PAH metabolite annotation data; Y. Go: Read and edited final document; P. Rohrbeck: Project Co-I, provided orphan

samples for PAH and HRM analysis; K. Pennell: Supervised D. Walker, provided edits and guidance on manuscript, read and approved final draft; D. Jones: Project Co-I, oversaw D. Walker and HRM data collection, provided edits and guidance on manuscript, contributed to sections on unknowns, exposure memory and HRMS, read and approved final draft.

9.1. Abstract

Objective: Assess the suitability of high-resolution metabolomics (HRM) for measure of internal exposure and effect biomarkers from deployment related environmental hazards.

Methods: HRM provides extensive coverage of metabolism and data relevant to a broad spectrum of environmental exposures. This review briefly describes the analytic platform, workflow and recent applications of HRM as a prototype environmental exposure surveillance system.

Results: Building upon techniques available for contemporary occupational medicine and exposure sciences, HRM methods are able to integrate external exposures, internal body burden of environmental agents and relevant biological responses with health outcomes.

Conclusions: Systematic analysis of existing Department of Defense Serum Repository samples will provide a high-quality cross-sectional reference dataset for deployment-associated exposures while at the same time establishing a foundation for precision medicine.

9.2. Critical need for deployment-associated exposure assessment

In 2011, the United States Institute of Medicine (IOM) recommended that the Department of Defense (DoD) collect individual breathing zone samples and conduct long-term studies of troop health outcomes to address concerns about perceived health risks from exposure during deployment (Tollerud et al. 2011, Baird 2012, Lindler 2015). Realistically, there are inherent limits to exposure assessment in deployed settings. For example, the use of personal monitoring equipment limits mobility in active combat situations, logistics of sampler collection is challenging with large-scale troop movements, and assessment for biologically relevant dose requires additional molecular measurements. Furthermore, the post-exposure window of opportunity for measuring exposures or immediate consequences may range from hours to days for some agents. Therefore, valid and reliable measures are needed to characterize exposures that do not disrupt effective operation during deployment. Retrospective profiling of biological specimens collected pre- and post-deployment for biomarkers of exposure, effect and susceptibility provide a means of assessing the occurrence of chemical exposure related to poor health outcomes. Through the DoD Serum Repository (DoDSR), an extensive system exists for collection, cataloguing and storing of serum samples collected pre- and post-deployment from active duty armed forces personnel (Rubertone and Brundage 2002, Perdue et al. 2015). Incorporating chemical screening measures using serum samples collected under the current DoDSR framework could therefore be completed with minimum disruption to military operations.

To fully realize the potential benefits of environmental chemical surveillance and bioeffect monitoring in serum specimens collected from armed forces personnel, there is a need to identify biomarkers relevant to exposure during the deployment period. A number of methods are being evaluated to improve deployment-related exposure assessment using DoDSR samples (Lindler 2015, Mallon et al. 2016). For instance, biomarkers of combustion products, including dioxins, free and protein-bound polycyclic aromatic hydrocarbon (PAH) are being used to assess burn pit exposures (Xia et al. 2016); circulating micro-RNA (miRNA), which play an important role in gene expression (Gurtan and Sharp 2013), provide epigenetic measures of biological response to exposure (Vrijens et al. 2015); cytokines and cardiovascular health markers are being applied to assess pathophysiological changes during deployment, particularly those related to respiratory pathways. While providing a means to evaluate exposure and biological response to environmental hazards, the measurements discussed above are limited by the need to *a priori* select chemical targets. A more complete understanding of how environmental exposures contribute to disease susceptibility and progression is required to mitigate risk, develop effective treatment strategies and identify at risk populations. Currently, no unified method exists to characterize the cumulative contribution of environmental and chemical exposures in disease.

A variety of approaches using genomics, metabolomics, lipidomics, transcriptomics, and proteomics are being pursued to determine how external and internal exposures from the environment impact health (Bradburne et al. 2015,

DelRaso et al. 2015). The application of omic technologies in environmental health research has enabled a more comprehensive measure of the continuum from exposure to bioeffect occurring within human populations. Here, we focus on metabolomics as a general approach to provide a cost-effective means to measure body burden of known chemicals, detect unidentified exposures, and monitor a broad spectrum of metabolic perturbations that could predict potential adverse health outcomes. We first describe high-resolution metabolomics (HRM) using liquid chromatography (LC) coupled to high-resolution mass spectrometry (MS) as an analytical platform to simultaneously evaluate perturbations in metabolism and detect chemicals present at very low levels in biological samples (Park et al. 2012). This methodology has been developed for precision medicine (Johnson et al. 2010, Jones et al. 2012), and is being refined for use to sequence exposures as part of a human exposome project (Go et al. 2015). The exposome was introduced as a complement to the genome to account for environmental contributions to disease risk (Wild 2005, Rappaport and Smith 2010), and is defined as the cumulative measure of environmental influences and biological responses throughout the lifespan, including environmental, dietary, microbiome, behavioral, therapeutic and endogenous processes (Miller and Jones 2014). The second section of the review addresses the use of HRM for measurement of low-level chemical exposure biomarkers. The third section summarizes approaches to use HRM for deployment-associated exposure surveillance, where we address the challenges in using point measurements for reconstructing exposure history and the evolving concepts of exposure memory systems (Jones 2015). Within the

exposure memory framework, we discuss the need to pursue combined analyses of epigenetic changes and other biomarkers. Finally, we provide a brief perspective on opportunities and needs for development of HRM as an integral component of improved deployment exposure surveillance systems and the considerable societal benefit from having an in depth, cross-sectional reference database of high-quality metabolomics data in support of nation-wide precision medicine initiatives.

9.3. High-resolution metabolomics (HRM): advanced clinical chemistry

9.3.1. Mass spectrometry for metabolic profiling

Mass spectrometry (MS) involves ionization of chemicals in the gas phase with subsequent detection of mass-to-charge ratio (m/z) and ion intensity. Because molecular mass is an absolute property of the chemical, the method is powerful for measurement of endogenous metabolites and environmental chemicals. Multiple types of mass spectrometers are available and have been recently reviewed (Makarov et al. 2006, Marshall and Hendrickson 2008, Schymanski et al. 2015). Nutrients and intermediate metabolites are often present in the micromolar (μM) to millimolar (mM) concentration range, and many analytic methods are available for targeted analysis (Krumstiek et al. 2012, Menni et al. 2013, Rappaport et al. 2014, Scalbert et al. 2014). Using targeted approaches, analytes measured in biological samples are compared to a series of reference standards selected *a priori* and utilize optimized analytical workflows (including sample preparation and instrumental analysis) developed to maximize selectivity while providing sufficient sensitivity to detect metabolite levels observed over a prescribed concentration range (Sirimanne et al. 1996, Barr and Needham 2002, Sandau et al. 2003, Roberts et al. 2012). Such approaches are useful for routine evaluation of individual metabolic characteristics and changes in association with deployment; however, due to limited chemical coverage (50-100 analytes) and requirement that analytical targets are selected prior to analysis, there is the potential to not detect unidentified exposure markers and biosignatures occurring during deployment.

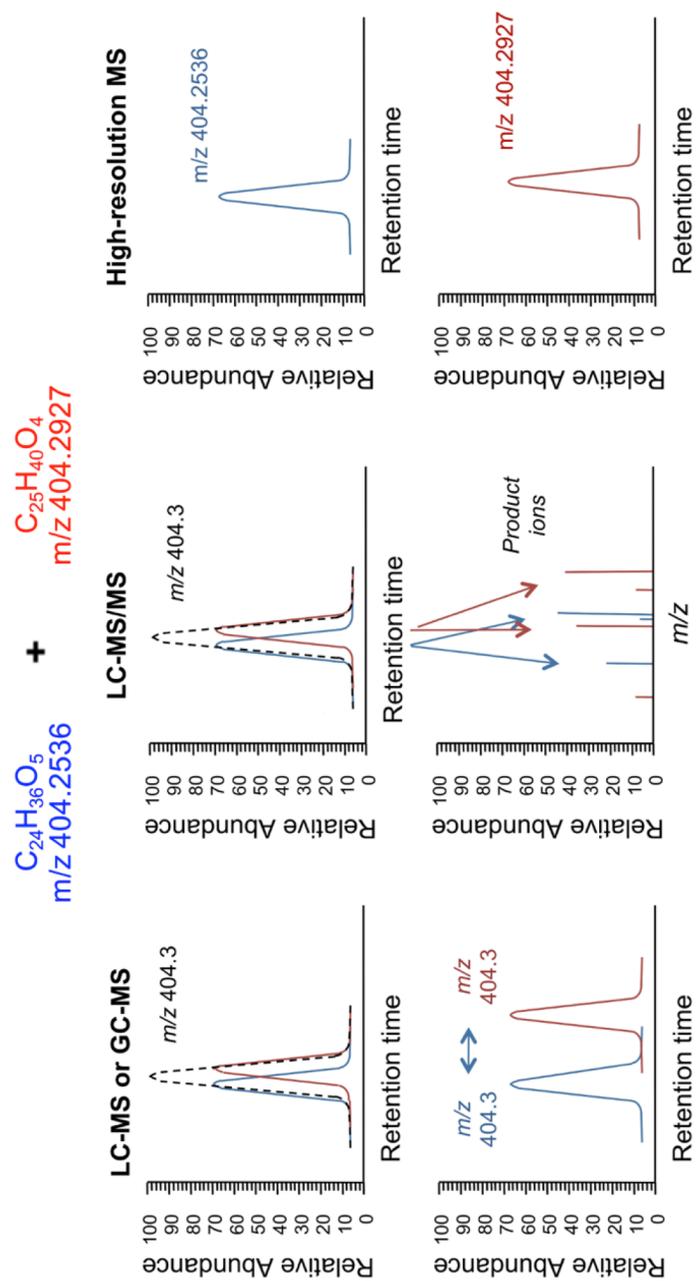


Figure 9.1. High-resolution MS supports untargeted measurement of metabolic chemicals by reducing requirements for chemical separation and sample preparation while providing improved capability to measure low abundance chemicals in biological samples. Adapted from Jones et al. (2012).

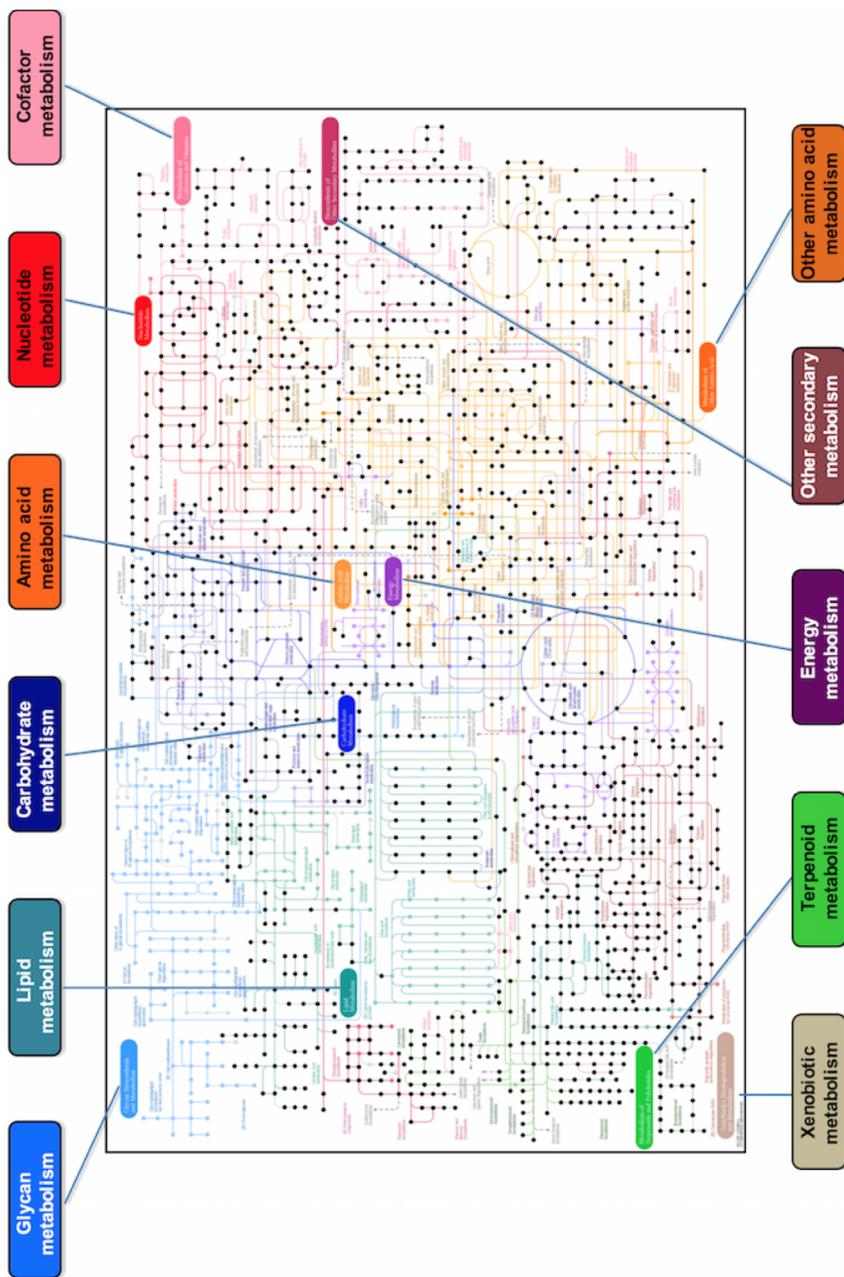


Figure 9.2. HRM profiling of plasma obtained from healthy individuals has indicated that metabolic intermediates from >80% of pathways present in the KEGG database can be detected (represented by black dots) using standardized sample preparation methods and dual-column chromatography. Adapted from Jones et al. 2012

Central metabolic pathways include approximately 2000 known metabolites, and improved coverage can be obtained using high performance liquid chromatography (HPLC) coupled to high-resolution MS for HRM analysis. The benefits of high-resolution MS instruments for screening are presented in **Figure 9.1**. Low resolution mass spectrometers, often utilized for targeted analysis, are capable of unit m/z accuracy (1 atomic mass unit (AMU)), which is not sufficient to distinguish compounds with very similar molecular mass, requiring chromatographic separation prior to detection (i.e., gas- or liquid chromatography (GC; LC). Tandem MS (MS/MS or MS²) involves combined use of mass spectrometry components to obtain m/z measurements on an ion and then subsequent measurement of m/z for product ions generated following ion dissociation, enabling quantification of specific chemicals based upon product ions even when the precursor ion is not separated from chemicals with very similar mass (Yan et al. 2008, Vogeser and Seger 2010). High-resolution accurate MS (e.g. Fourier Transform Ion Cyclotron Resonance (FTICR) or Orbitrap based mass spectrometers) resolve ions and measure m/z with 0.0002-0.005 AMU accuracy (where mass accuracy is commonly referred to as parts-per-million, defined as: relative m/z error $\times 10^6$). The high level of mass resolution achieved simplifies analyte separation requirements and provides improved capability to measure low abundance chemicals in complex biological samples. Hence, high-resolution MS provides identification and quantification of a broad spectrum of m/z features, facilitating the discovery of metabolic alterations since multiple metabolites in the same pathway can be measured simultaneously and tested for

enrichment (Xia and Wishart 2010, Li et al. 2013) while not requiring selection of specific analytical targets prior to analysis.

A broader view of metabolism includes more than a million biochemicals derived from the diet, microbiome and environment, as well as a large number of lipids, peptides, glycans, peptidolipids, glycolipids, and peptidoglycans (Jones et al. 2012, Wishart et al. 2013). These chemicals have the potential to act as specific markers of exposure, toxicity and/or disease. Systematic knowledge of these biomarkers is not available, but information is rapidly accumulating and may provide useful early disease or disease-risk indicators. Many of the m/z features detected by HRM platforms are currently unidentified and include uncharacterized complex carbohydrates, environmental chemicals and their metabolites, complex lipids and amino acid metabolites derived from covalently modified proteins (Bowen and Northen 2010, da Silva et al. 2015).

9.3.2. Quantification

High resolution MS offers an advantage for quantification of many metabolites since mass resolution is sufficient to separately quantify co-eluting ions when the accurate mass differs by >10 ppm (Johnson et al. 2010). Over 90% of metabolites in the Kyoto Encyclopedia of Genes and Genomes (KEGG) have accurate mass m/z values that differ by >10 ppm (Park et al. 2012). HRM profiling of blood plasma samples obtained from healthy individuals has indicated measure of metabolites from >80% of the pathways present in the KEGG database (**Figure 9.2**), enabling quantification with >85% accuracy based upon ion intensity in

terms of the integral of precursor ions (MS^1) (Roede et al. 2013, Go et al. 2015). Since HRM profiling is typically applied in an untargeted manner, analytical standards for absolute quantification are often not included within the analytical methodology; however, a number of strategies for quantification have been developed allowing determination of absolute concentrations. For example, different options are available using internal standardization with stable isotope dilution (Johnson et al. 2008), surrogate standardization (Greizerstein et al. 1997) or by external standardization with method of additions (Roede et al. 2013). Systematic comparison of these methods shows that reference standardization, a procedure using a quantitatively calibrated pooled reference material, such as National Institute of Standards and Technology SRM1950, can support quantification of thousands of chemicals in a single analysis (Go et al. 2015). Using reference standardization, quantification post-data acquisition is possible by referencing the pooled sample analyzed within each batch of samples. Known concentrations within the reference standard can be used to determine a chemical response factor and calculate analyte concentrations based on single-point calibration. Typically, intensity of the most reliable ion corresponding to a given chemical species is used for quantification, since combining multiple ions would result in error propagation. The benefit of reference standardization is that targeted quantification is only required in the reference sample, chemicals selected for quantification do not need to be selected *a priori* and population-wide estimates of plasma chemical concentrations can be determined without having to re-analyze samples using a targeted approach. Such results support development

of cumulative databases to evaluate time- and intensity-dependent changes in exposure and related metabolic perturbations.

9.3.3. Applications for precision medicine.

In addition to reproducibility and accurate quantification, other practical issues for routine use of HRM in precision medicine include cost and throughput. Assuming single instrument operation at 24 h/day for 250 days/year (which allows 25% time for holidays, vacations, servicing and repair) and a 3-year instrument lease, cost for triplicate analysis using dual chromatography (Soltow et al. 2013) is approximately \$125/sample and would enable analysis of 5500 samples per instrument-year. Due to improved mass resolution and increased scan speeds available with next-generation instruments, cost per sample can be reduced to \$50/sample and instrument-year throughput doubled by decreasing analytical run time to 5 minutes. Although detection of chemicals arising from exposure to environmental chemicals would be minimized, further cost reduction is also possible through focused analysis of high abundance metabolites (e.g., amino acids, lipids, vitamins/co-factors, fatty acids) that can be used for advanced clinical chemistry purposes. For example, routine analysis by HRM reliably detects approximately 1000 common, endogenous metabolites with coefficient of variation (CV) <10% (Jones et al. 2012, Uppal et al. 2013, Go et al. 2014, Go et al. 2015, Jones 2016). By limiting *m/z* detection to high abundance, endogenous metabolites, single replicate analysis on one column is sufficient, reducing analysis cost to approximately \$15 per sample. The resulting throughput for single

replicate analyses per instrument is 6 samples per hour, with analysis capacity of 150 samples/day (37,500 samples-per instrument year) possible. Thus, sufficient metabolic coverage for precision medicine purposes and detection of exposure related bioeffect could be obtained with sufficient data quality for lower costs using available instruments and appropriate automation.

The most important implication of these considerations is the potential feasibility of applying HRM to biological specimens collected for the DoDSR to obtain a comprehensive understanding of normal variations in metabolism and environmental exposures in young healthy adults from across the US. For instance, by selecting a random set of one million samples from the repository, one would be able to evaluate population differences due to geography, occupation, health habits, age, gender and disease, as well as trends over time. Using the cost structure discussed above, it would be possible to analyze one million samples (5 minute runtime with triplicate analyses) in five years using 20 instruments at a cost of \$10 million/year; focused analysis of high-abundance metabolites could be completed using five instruments at a cost of \$3 million/year. The data obtained from such an analyses would provide a resource for DoD researchers to evaluate possible health concerns and provide a large reference database of normal values for future precision medicine initiatives involving diet and metabolism.

9.4. HRM for environmental chemical analysis

9.4.1. Detection of low abundance chemicals

By increasing the number of replicate injections to three or more, the number of ions that can be reliably quantified in a biological sample is increased and noise reduction is improved. Improved sensitivity is important for measurement of environmental chemicals in biological samples, which are often present at four or five-orders-of-magnitude lower abundance than endogenous metabolites (**Figure 2.1**). A chemical present in only one sample out of 100 can be reliably measured if it is present in three replicate analyses of a sample (Go et al. 2015, Jones 2016, Walker et al. 2016), even though it is absent from the preceding 99 samples. With further relaxation of precision requirements, such as increasing CV thresholds from 50% to 100% and allowable number of samples with non-detected intensity values (which is justified by the usefulness of information on low abundance and low frequency exposures), >20,000 ions can be measured routinely on a single column; use of dual chromatography increases this to >30,000 ions. For high-abundance species, multiple adducts and isotopic forms will be detected based upon the ionization scheme used and molecule functional groups. The act of pairing different m/z features corresponding to the same chemical is referred to as deconvolution, and can be used as additional criteria for confirming identification. For example, the presence of an M+2 isotope corresponding to ^{37}Cl confirms the presence of at least one chlorine functional group, which can qualify whether the identification is correct based upon chemical formula. Numerous strategies exist for deconvoluting high-resolution mass spectra, with rule-based methods relying

on ion correlation, retention time matching, known adduct spacing, mass defect pairing and comparison of expected isotopic distributions to detected isotopes resulting in the lowest number of false groupings. Due to the low concentration of many chemicals present in biological samples (Rappaport et al. 2014), multiple features from highly abundant metabolic species represents a small portion of detected features and most low intensity ions appear to be exogenous chemicals detected as a single form. Thus, approximately >20,000 chemicals are uniquely detected using a dual column configuration. Although half of these ions do not match known chemicals in metabolomics databases, most are reproducible when analyzed on different LC-MS systems, indicating actual chemical signals and not instrumental noise or sample preparation artifacts. Therefore, HRM platforms, in addition to providing measure of endogenous metabolites, have multiple applications for profiling low abundance chemicals present in biological samples such as plasma (Frediani et al. 2014, Go et al. 2015), urine (Edmands et al. 2015, Khamis et al. 2015) or tissue biopsies (Dunn et al. 2012, Go et al. 2014).

9.4.2. Measurement of exposure and linking to body burden

Incorporating exposure information into population research has traditionally relied on monitoring approaches with varying levels of uncertainty (Lan et al. 2010, Chadeau-Hyam et al. 2011, O'Connell et al. 2014, Wambaugh et al. 2014, Lane et al. 2015). While these are useful measures for evaluating the occurrence of environmental exposure, they represent generalized estimates and cannot be used to assess internal exposure and biological relevance. Furthermore, their

implementation in active duty situations would negatively impact the efficacy of combat operations and daily functioning of a fighting force.

Measurement of exposure biomarkers within biological samples obtained from humans can reduce uncertainty in exposure assessment; however, interpretation within an environmental health context can be challenging. Models used for risk assessment (Patton et al. 2015) as well as toxicokinetic studies (Chiu et al. 2007, Sobus et al. 2010) provide frameworks for analysis, but more detailed studies are needed to interpret the meaning of spot measurements of environmental chemicals in single samples. Without knowledge of intensity, duration or route of exposure, inferences concerning abundance of chemicals in human samples are problematic. Similarly, for unidentified chemicals without time course data, biologic half-life, the apparent volume of distribution and other toxicokinetic parameters cannot be estimated. Thus, one cannot readily evaluate the likelihood that a measurement represents an acute or long-term exposure from a single point measurement. On the other hand, routine, periodic measurement can provide the ability to detect regional differences and time-dependent differences in specific chemicals. Ongoing measurements, even for a small fraction of individuals, can therefore approximate a real-time surveillance tool to detect new or unexpected exposures. Importantly, the power of such a tool increases with time as the surveillance database expands with different exposure scenarios.

As discussed above, there is a need to measure exposure biomarkers arising from a range of environmental chemicals. Targeted biomonitoring utilizes measurement in biological specimens to estimate body burden of specific

chemicals, providing information on internal dose and prevalence in a population. While biomonitoring has proven invaluable in population surveys, chemical coverage is often limited. The ability of HRM to advance measurement of exposure biomarkers in human samples is highlighted in **Figure 9.3A**. In this analysis, quantified PAH levels of anthracene measured using GC-MS in 30 non-identified serum samples (Walker et al. 2016, Xia et al. 2016) were tested for association with m/z ions detected by HRM matching known adduct masses of PAH metabolites from the KEGG database. Twenty-four metabolites detected using HILIC chromatography with positive electrospray ionization and 18 using reverse-phase chromatography with positive atmospheric chemical ionization (Liu et al. 2016) were found to be significantly associated with serum PAH levels, supporting the use of HRM for a more comprehensive measure of exposure biomarkers than available from targeted biomonitoring alone. Interestingly, anthracene levels were found to be predominantly associated with naphthalene metabolites (**Figure 9.3B**), suggesting possible co-exposures or synergistic metabolism. Successful applications of HRM in environmental chemical surveillance (Jamin et al. 2014, Roca et al. 2014, Schymanski et al. 2014, Bletsou et al. 2015, Go et al. 2015) further support the use of this approach in exposome and environmental health research.

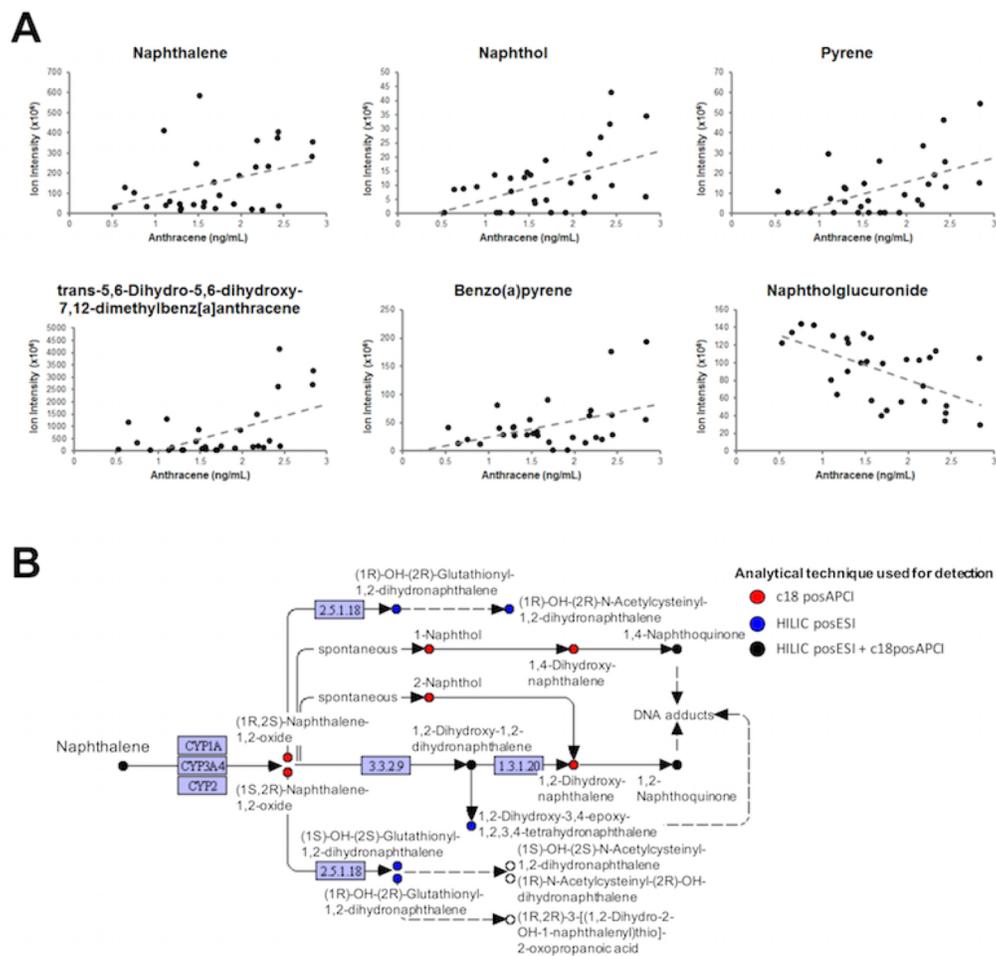


Figure 9.3. Cross-platform comparison of targeted polycyclic aromatic hydrocarbon (PAH) measurements using gas chromatography MS and suspect screening for PAH metabolites by HRM (A). Combining liquid chromatography and ionizations schemes maximized coverage of PAH metabolism (B).

9.4.3. Characterizing internal dose response.

Considerable data are available to show that environmental chemicals are present in all human blood and urine samples (Bonvallot et al. 2013, Jamin et al. 2014, Rappaport et al. 2014, Roca et al. 2014, Go et al. 2015, Wishart et al. 2015). As indicated above, detection of chemicals within the body does not provide information concerning the source of exposure, and relatively little information is available concerning the distribution chemicals within human tissues and rates of elimination. Furthermore, analytical methods are very sensitive so that detection alone does not provide information concerning risk from exposure and biological relevance.

In addition to measurement of chemical surveillance biomarkers, HRM profiling can be used to elucidate biological response to chemical exposures. Association of exposure biomarkers with HRM measured metabolic alterations provides insight into toxicant targets, biomarkers of effect and chemical biological fate, such as metabolism, distribution and excretion within human populations. Recent examples using human populations exposed to high levels of environmental chemicals highlight the ability of HRM in identifying biological response from exposures that occurred decades ago (Jeanneret et al. 2014). Thus, in addition to biomonitoring, measurement of metabolic variations by HRM can be used to identify biosignatures indicating a history of exposure. Additional studies applying metabolic profiling of human exposure to cadmium (Ellis et al. 2012, Gao et al. 2014), pesticides (Bonvallot et al. 2013), PAH exposure (Wang et al. 2015), welding fumes (Wei et al. 2013) and arsenic (Zhang et al. 2014) support

the use of retrospective chemical measurement for effect markers and evaluation of whether exposure has occurred at a biologically relevant dose.

Important to the development of HRM in routine screening for biomarkers of exposure and effect has been the creation of software tools enabling a systems biology approach to understanding exposure-associated metabolic perturbations. For example, MetabNet (Uppal et al. 2015), a software routine written in R, allows rapid analysis of metabolites and metabolic pathways correlating with individual chemicals in cross-sectional analyses. The analyses possible by MetabNet enable testing for correlations with environmental chemicals directly within the same metabolomics analyses, thereby avoiding the need for phenotypic biomarkers, which are often not available. For instance, HRM profiling was recently used for targeted measurement of the flame retardant triethylphosphate and untargeted metabolomics of plasma obtained from 150 healthy adults (Go et al. 2015). By testing all metabolites for correlations with triethylphosphate, one obtains a targeted metabolome wide association study (MWAS), shown in **Figure 9.4** as a Manhattan plot of $-\log p$ -value for the correlation as a function of retention time obtained using reverse phase LC. By evaluating feature association significance as a function of retention time, one can see that many of the ions correlated with the flame retardant exhibited retention times consistent with lipid species (retention time > 300 seconds), which was verified by database matches to phospholipids and sterols. Metabolites identified through MWAS can then be utilized for metabolic pathway enrichment analysis, providing a starting point to

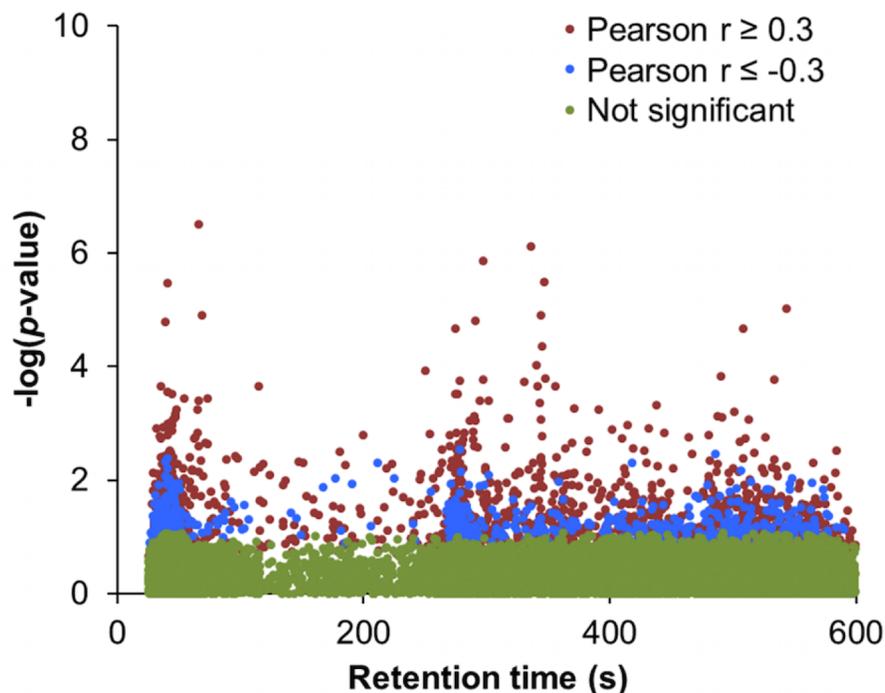


Figure 9.4. A targeted metabolome wide association study (MWAS) was completed by testing each metabolite for significant associations with triethylphosphate levels in plasma. Significant metabolites can be tested for pathway enrichment, providing a starting point in identifying perturbations linked to chemical exposure.

begin studies of metabolic perturbations linked to body measures of environmental chemicals.

9.4.4. *Exposure memory*

The need to include environment in understanding human disease led Christopher Wild to introduce the concept of the exposome in 2005 (Wild 2005), which he defined as “encompassing life-course environmental exposures (including lifestyle factors), from the pre-natal period onwards.” The exposome is envisioned as a complement to the genome, where a life course of exposure and interaction with the genome defines risk for disease development. Unlike the genome,

exposures are transient and change on both short and long-term time scales, making quantitative assessment challenging. A more tractable definition of the exposome was proposed by Miller and Jones (2014): “The cumulative measure of environmental influences and associated biological responses throughout the lifespan, including exposures from the environment, diet, behavior, and endogenous processes.” Exposures in this framework are not only limited to external chemical exposures, but also include processes internal to the body (host factors), wider socioeconomic influences and maladaptations to external influences, with the interaction of these factors linking environment to health and disease (Rappaport and Smith 2010, Wild 2012).

Accumulating research indicates that multiple exposure memory systems exist to allow adaptation to environmental challenges over a lifetime (Jones 2015). A consequence is that response can result in decreased flexibility and adaptability to subsequent challenges. Many of the body’s response to environmental insults are short-term and reversible; some, such as scarring, provide a long-term and sometimes permanent change in structure and function. For example, inhalation injury to the lungs can result in permanent change in structure and function, which may not be evident except in response to challenge. Computational methods are available to use the regularity/irregularity present in metabolomics data to measure health (Park et al. 2013) and, in principle; these methods could be used with HRM to study metabolic resilience following deployment. Such methods could be applied to screen service personnel for

exposure histories and identify individuals who would benefit from nutrition or medical intervention.

9.4.5. Addressing the dark matter of the exposome

A critical aspect of post-deployment surveillance lies in addressing unknown exposures. While contemporary MS methods are powerful, characterization of unidentified chemicals in human samples has seriously lagged behind the progress in understanding human genetics. At least half of the ions detected by HRM are uncharacterized; if one pushes data extraction to the limits of contemporary methods, the fraction of uncharacterized ions is approximately 80% (Johnson et al. 2010). This represents the dark matter of the exposome, which includes detected but uncharacterized analytes and a spectrum of unknown unknowns, for which little effort has been made to distinguish. Resources to address this daunting challenge are currently unavailable, as emphasis has traditionally been placed on analysis of a relatively small number of recognized hazardous chemicals. As a consequence, risks associated with uncharacterized chemicals are largely unknown.

HRM methods provide a useful approach to this challenging problem. Already, data are available within the Emory University Clinical Biomarkers Laboratory for more than 100,000 ions obtained from approximately 20,000 biological samples. All samples have been analyzed with rigorous standard operating procedures so that the accurate mass, retention time and ion intensity enable collation into a reference database. This type of database will provide a

resource for comparison to post-deployment measurement for the detection of new ions, which could represent unidentified or unknown exposures and capabilities are now available to establish a rigorous analytical framework to regularly monitor exposures. With appropriate commitment and resources, an analytical resource can be established to support routine exposure surveillance in armed forces personnel.

9.5. HRM in deployment exposure surveillance

9.5.1. Retrospective analysis using the DoDSR

For effective real-time analyses, databases representing normal exposures and ranges of unexpected exposures will be needed to provide reference values. Previous studies have established the integrity of samples stored within the DoDSR (Perdue et al. 2015, Walker et al. 2016) and therefore highlight the opportunity for further advancement of the repository as an environmental health resource. The recent development of reference standardization (Go et al. 2015) to obtain estimates of absolute concentrations of endogenous metabolites, health indicators and environmental chemicals, establishes an affordable approach. Additionally, the concepts are developed to use metabolic correlations for retrospective chemical identification (Uppal et al. 2015), where correlations of a metabolite are obtained in reference populations so that the correlation structure can be used to identify chemicals using data from previously analyzed samples. Due to the use of rigorous sample collection, preparation and analysis procedures, the data from these analyses can be stored in a cumulative database for future analysis and data mining. As new hazards and exposures are recognized, the data will be available for retrospective analysis of individual exposures, their trends and associated health outcomes.

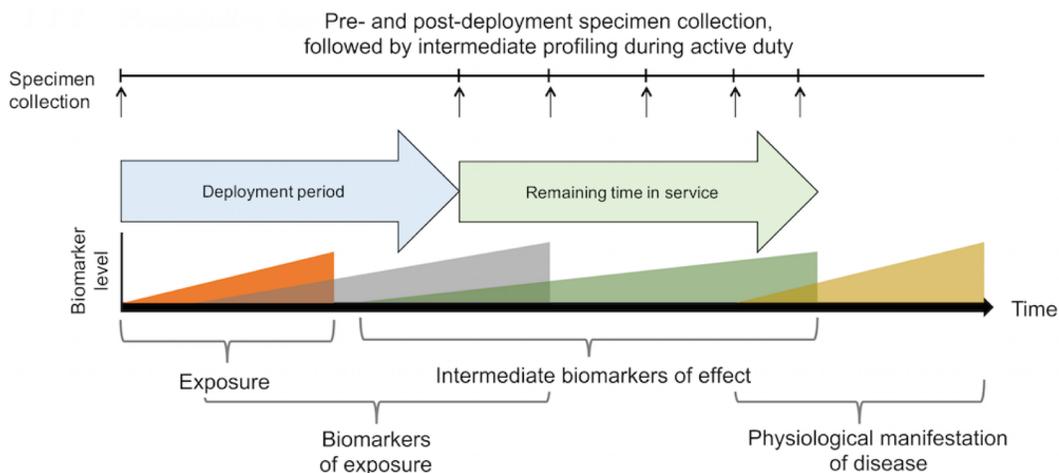


Figure 9.5. HRM profiling of specimens collected from active duty individuals could readily be integrated into DoDSR protocols. In this example, HRM can evaluate the presence of exposure and effect biomarkers, identifying individuals at risk for exposure related diseases.

HRM is sufficiently developed to allow implementation on a test basis for ongoing deployment surveillance. For instance, metabolic indicators of nutrition (vitamins, amino acids), liver function (bilirubin), renal function (creatinine, uric acid) and other health phenotypes can be readily measured using routine analysis. At the same time, a range of PAH metabolites, persistent halogenated chemicals, insecticides, herbicides, flame retardants and other chemicals are detected (Park et al. 2012, Roede et al. 2012, Osborn et al. 2013, Soltow et al. 2013, Go et al. 2014, Jamin et al. 2014, Go et al. 2015). Thus, implementation of routine HRM for exposure surveillance would provide a means to address uncertainties concerning exposure. An example framework for specimen collection and implementation of HRM profiling that is consistent with current DoDSR protocols is provided in **Figure 9.5**. In this framework, chemical profiling of samples obtained pre- and post-deployment and over the course of service will enable evaluation for

biomarkers of exposure, effect and poor health outcome. Routine detection of biomarkers will enable identification of individuals at risk for environment-associated disease, enabling intervention and preventive measures. Ultimately, such an analytical structure could facilitate improved management of risk associated with work in adverse environments.

9.5.2. Integration with targeted phenotypic and other omics platforms.

Many mechanistic studies show that the combination of HRM with other phenotypic platforms (e.g. genomics, proteomics, transcriptomics) provides a powerful approach to understand mechanisms of toxicity and disease (Wild 2012, Vineis et al. 2013, Wild et al. 2013, Gomez-Cabrero et al. 2014). Although deployment surveillance does not necessarily require mechanistic understanding, the integrative approaches may be suitable for improved risk assessment. For instance, transcriptome-metabolome wide association study (TMWAS) showed interaction hubs reflecting toxic reactions as well as early stress response and adaptive responses (Roede et al. 2014). Gene expression and metabolism in peripheral white blood cells may be sufficiently persistent to monitor adverse exposures for several weeks (van Leeuwen et al. 2008). Similarly, circulating miRNA in combination with metabolomics could provide a useful means to evaluate prior adverse exposures. The utility of integrative approaches was recently shown by the interaction of the lung microbiome with the metabolome measured in lung bronchoalveolar lavage fluid from healthy controls and HIV-1 infected individuals (Cribbs et al. 2016). In this study, increasing association

stringency for microbiome-metabolome interaction using all individuals simplified a large number of significant associations to the top two genera of bacteria and top 3 metabolites, with the associations centered on bacteria causing opportunistic infection (**Figure 9.6A-C**). Microbiome-metabolome associations were then evaluated based upon significance of association with HIV-1 status (**Figure 9.6D-F**). By decreasing the p -value threshold, clear visualization of the genera-metabolite hubs was obtained, which include the top 3 groups of bacteria specifically causing opportunistic infections in individuals with HIV-1. Thus, there are considerable opportunities to utilize HRM in combination with other powerful contemporary methods to enhance detection, understanding and management of post-deployment health risks.

9.6. HRM: summary and perspective

High-resolution metabolomics (HRM) provides an advanced clinical chemistry platform for precision medicine that could be of considerable utility for exposure surveillance of armed forces personnel. HRM not only provides extensive coverage of metabolism but also detects a broad spectrum of exposure biomarkers, including both known and currently unidentified chemicals. Building upon contemporary occupational medicine and exposure science, HRM can be integrated into environmental health research through connecting external exposures and health outcomes to internal body burden of environmental agents and respective biological responses. Analytic platforms, workflow and available applications establish the suitability of HRM for development into an

environmental exposure surveillance system. Systematic analysis of existing DoDSR samples using HRM would provide a high-quality cross-sectional reference dataset for deployment-associated exposures, while development of real-time analytical capabilities using HRM would provide a demonstration project for use in precision medicine. Furthermore, use of HRM is expected to improve the ability of healthcare practitioners to include exposure-related measurements in management and treatment of disease in active duty and retired armed forces personnel.

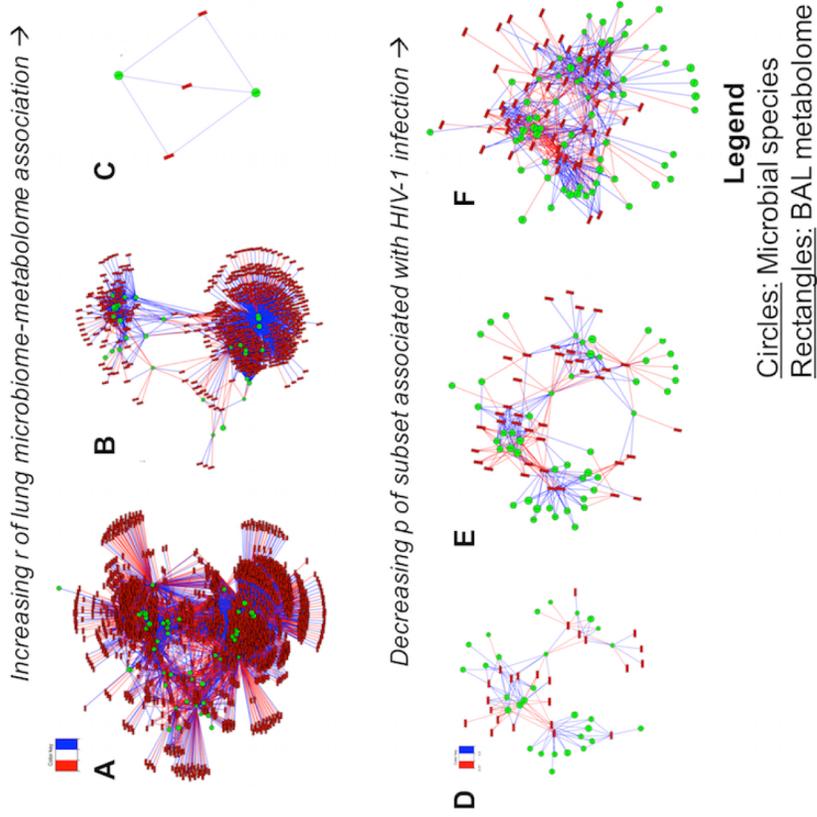


Figure 9.6. Microbiome-MWAS of bronchoalveolar lavage fluid in HIV-1: **A to C** represent increasing stringency to identify network associations with greatest r meeting significance criteria. **D to F** contains the network subset associated with HIV-1, in order of increasing leniency. Adapted from Cribbs et al. (2016).

Chapter 10. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

10.1. Conclusions

The environment is recognized as a critical determinant of human health. Even with the rapid growth in genomic technologies, current estimates suggest no more than 15% of diseases are due to heritability alone. Environment and gene-environment interaction are largely suspected to contribute to over 85% of disease burden; however, even the most advanced targeted platforms provide measure of no more than 200 exposure biomarkers at considerable expense, while routine genotyping can assess 10^5 - 10^6 genetic variants for \$200. To develop a balanced view of human health and disease risk there is a critical need to establish analytical methodologies that provide enhanced population screening for identifying exposures and associated biological response.

The major contribution of this thesis was to establish high-resolution metabolomics (HRM) as a central platform linking human chemical exposures to internal dose and biological response. This was accomplished through application of HRM to profile five independent cohorts with previously characterized exposure measures. Each study was selected to include different environmental exposures, timescales and resolution of exposure assessment. Application of

metabolome-wide association study (MWAS) was used to identify peripheral blood metabolic features associated with each exposure and characterized based upon metabolite source and function. The results demonstrate it is possible to detect metabolic variations arising from acute and chronic environmental exposures using HRM. These metabolic variations, which include biomarkers of environmental pollutants, biological response molecules and metabolic changes consistent with disease pathobiology, are the key functional inputs required to characterize environmental contributions to human disease and develop HRM-based population screening as a public health tool. Unlike targeted methodologies that would require multiple analyses and large sample volumes to measure both exposure and biological response markers, this thesis shows it is possible to obtain measures of both using a single instrumental analysis and small volume of material (65 μ L). The efficiency of HRM therefore greatly improves throughput, reduces cost and minimizes use of biological material, establishing HRM is a suitable analytical platform for broad scale environmental chemical surveillance and bioeffect monitoring in humans. Future recommendations for continued development of this framework include improved characterization of unknown chemical signals, development of reference databases containing HRM profiles from a range of health states, combined experimental and computational techniques for unambiguous signal annotation and creation of reference standard materials for large-scale quantitation. The goal of these recommendations is to provide a sufficiently robust analytical framework for use in national chemical surveillance and precision medicine programs, with continued development

expected to greatly enhance measure of the occurrence, distribution and magnitude of environmental exposures while advancing knowledge on mechanisms underlying environment-related diseases

In **Chapter 4**, the hypothesis was samples collected and stored at the Department of Defense Serum Repository (DoDSR) are of sufficient quality for HRM profiling. This study provided a critical test for using HRM to characterize endogenous metabolite levels and measure exposure associated metabolic variations. Results from this study showed select metabolic intermediates and clinical health indicators were largely within expected concentration ranges, providing evidence samples are of sufficient quality for quantitative measure of endogenous metabolites. Exposure-associated metabolic alterations were assessed by completing an MWAS using the HRM data and quantified levels of free serum benzo[a]pyrene, which was measured using targeted methodologies. Benzo[a]pyrene was selected as an exposure marker for testing due to its relevance to burn pit exposures, ubiquitous presence in the environment and well-established mechanisms of toxicity. Using a multivariate variable selection technique, metabolic features associated with benzo[a]pyrene exhibited dose-dependent hierarchical clustering patterns, demonstrating metabolic variations were consistent with an exposure-related metabolic phenotype. Characterization of these endogenous metabolites included pathways related to oxidative stress, lipid metabolism and fatty acids, which are consistent with the toxic mechanisms of benzo[a]pyrene and exposure response observed in both cell and animal models. These results demonstrate DoDSR samples are of sufficient quality for

HRM characterization and establishes the repository, which has collected over 50 million serum samples from 15 million armed forces personnel, as a tremendous resource for precision medicine research. Furthermore, this pilot study established the principle of using MWAS to assess metabolic changes of environmental exposures in HRM data. Results from this chapter were critical to the successful completion of this thesis.

Chapter 5 extended the MWAS framework to include acute occupational exposures to the chlorinated solvent, trichloroethylene (TCE). For this study, we hypothesized that external, respiratory exposure to TCE vapors during a full work shift would lead to dose-dependent variations in the metabolome. The objective for this study differed from **Chapter 4** in that we aimed to link a short-term, external measurement of exposure to metabolic changes using HRM. Testing for dose-dependent changes between control, low and high-exposure workers identified systemic metabolic response to trichloroethylene. Metabolic features associated with TCE exposure included a large number of unidentified chlorinated chemicals and alterations in endogenous metabolism consistent with known toxic targets, including renal, liver and immune systems. These variations were consistent with risk factors for disease associated with trichloroethylene exposure, even though the majority of exposure levels were below the U.S. Occupational Safety and Health Administration limit. Integration of significant metabolites with independently measured markers of disease and exposure biomarkers showed unidentified, chlorinated chemicals were correlated with immune function and kidney injury, while identified TCE metabolites did not. This suggests the

possibility of uncharacterized TCE intermediates mediating toxic effects, highlighting the strength of using untargeted approaches. Although limited to an occupational exposure cohort, the findings provide insight into the underlying toxic mechanisms and exposure related changes to metabolic phenotype for a common environmental contaminant. Taken together, this study represents a critical first demonstration of using HRM as a central platform to link external exposure to internal dose and biological response. By providing evidence of an acute response to short-term (8 hr.) exposure, it shows metabolic variations due to external stimuli can be detected using HRM.

In **Chapter 6**, we applied an integrated molecular approach to characterize the effects of daily and extended occupational exposure to diesel fuel exhaust. Due to a repeat sampling study design that included pre- and post-shift blood samples for two separate workdays, it was possible to test the hypothesis that acute (daily) and extended workplace exposures (estimated using week-averaged microenvironment pollutant levels) would lead to a dissimilar metabolic response. Three different measures of exhaust exposure were available, including organic carbon (OC), elemental carbon (EC) and PM_{2.5}, allowing us to evaluate the second hypothesis that different constituents of diesel exhaust would lead to dissimilar metabolic variations. MWAS of post-shift samples to test for daily metabolic variations related to exposure to the three pollutants showed limited changes, suggesting immediate response to diesel exhaust is minimal. PM_{2.5}, a common measure of exposure to traffic related pollutants, resulted in no significant associations using either the daily or week-averaged MWAS

framework. EC and OC results show a significant influence on the exposure metabolic phenotype, with 845 and 650 *m/z* features associated with the two pollutants. Comparison of these metabolic variations indicated a differential metabolic response, with all but one of the overlapping *m/z* features showing a positive association with EC and negative relationship with OC. While the exact cause of this difference could not be identified, it suggests exposures enriched in high OC particulates may include different co-exposures that do not have the detrimental effects of exhaust emissions with high EC. Characterization of the metabolites associated with EC showed alterations in pathways consistent with elevated lipid peroxidation products, biomarkers of oxidative stress, thrombotic signaling lipids, and metabolites associated with disruption to nitric oxide production. When compared to OC levels, many of the general metabolic pathways were similar, but included alterations to different metabolites or exhibited a different direction of change with exposure. These included alterations in antioxidant levels, oxidative stress biomarkers and intermediates in nitric oxide production. Metabolic associations with both pollutants were largely related to endothelial function and oxidative stress, which are consistent with increased risk of cardiovascular diseases.

Significant metabolic features for both pollutants also included exposure biomarkers of chemicals present in exhaust emissions. Identified environmental chemicals positively association with EC included volatile organic chemicals, their metabolites and the PAH naphthalene. These results suggest EC may be a good indicator of diesel exhaust, since many of these compounds are not particle

bound and are expected to largely be in the gaseous phase. The exposure biomarkers associated with OC largely support this hypothesis. Many of the detected VOCs showed a negative relationship with OC, indicating exposures enriched in OC may be removed from volatile emissions. However, dipole epoxide of benzo[a]pyrene and the nitronaphthalene, which are cytotoxic bioactive metabolites, were positively associated with OC levels. Thus, the results from this study were able to successfully detect biomarkers of exposure and biological changes consistent with diseases related to air pollution exposures. This builds upon **Chapter 5** by showing length of exposure can influence metabolic phenotype and HRM can be used to characterize exposures that extend beyond a single shift. To provide additional evidence of biological response to diesel exhaust, we used a data-driven network approach to integrate identified metabolic variations and peripheral blood gene expression. The gene expression results showed enrichment in pathways related to endothelial function, immune response, inflammation and oxidative stress, providing additional evidence of changes in processes implicated in cardiovascular disease. This represents the first integrated molecular response characterization of environmental exposures in humans and demonstrates how the combination of multiple omic datasets can be used to merge molecular and environmental epidemiology for a new paradigm of toxicological research in humans.

Chapter 7 built upon the results of **Chapter 6** by assessing long-term metabolic effects of exposure to traffic-related pollutants. In addition, this study represented a critical test of HRM to detect metabolic variations with high-

resolution measures of chronic exposure. Using a hybrid targeted and untargeted HRM approach, we tested the hypothesis that elevated ultrafine particulate (UFP) exposures for study participants would lead to changes in central metabolic intermediates and additional systemic effects. First, we used reference standardization to estimate concentrations of 79 metabolic intermediates that included amino acids, clinical health markers, fatty acids, lipids, nucleotides, TCA cycle metabolites and co-factors. Significantly different metabolites increased with high UFP included arginine, glutamine, cysteine and methionine sulfoxide, suggesting long-term exposure results to metabolites critical for endothelial function and oxidative stress. Expanding to detected *m/z* features using two different analytical modes identified an additional 316 features discriminating low from high-UFP. Annotation identified these as lipid peroxidation products, endogenous inhibitors of nitric oxide and environmental chemicals, including volatile organic chemicals and bioactive PAH metabolites. Many of these metabolites were shown to be associated with exposure to EC and/or OC in **Chapter 6**, supporting results from both studies and providing strong evidence that exposure to traffic-related pollutants leads to adverse changes in critical biochemical processes. Expanding the results using network correlation to identify additional metabolic variations within the raw HRM data identified 38 enriched metabolic pathways enriched UFP exposure. The diverse series of biological processes represented by these pathways supports systemic biological response from long term exposure to traffic related pollutants, which include alterations to inflammatory pathways, oxidative stress, endothelial function and

mitochondrial bioenergetics. A second analysis integrating network metabolites with biomarkers of endothelial function and inflammation identified fibrinogen as a central cluster linking the inflammatory markers IL-6, CRP and TNF-RII. Enriched pathways clustering around each biomarker were consistent with known signaling functions of the different biomarkers, including oxidative stress, xenobiotic metabolism and endothelial function. The use of these complementary measures, which is similar to the analysis completed in **Chapter 5**, provides new evidence of functional interactions between acute response markers and metabolites associated with long-term environmental exposures. Taken together, results from this study demonstrate metabolic variations associated with UFP are consistent with risk for diseases linked air pollution exposure. These alterations included a shift in the Cys/CySS couple towards a more positive steady-state redox potential, elevated dimethylarginine and elevated levels of 4-HNE, all of which are recognized risk factors for atherosclerosis, endothelial dysfunction and cardiovascular disease.

Chapter 9 extended the MWAS framework to include a multi-generational cohort selected from the Michigan PBB registry cohort with measured exposure biomarkers for two different pollutants: the brominated flame retardant PBB-153 and the polychlorinated biphenyl PCB-153. Because this cohort was uniquely exposed to PBB-153 during an accidental industrial contamination that occurred in the 1970's, it was possible to test how exposures that primarily occurred decades ago influence current metabolic phenotype and the contribution of different routes of exposure (dietary vs. *in utero*). By

comparison, PCB-153 exposures occur due to its ubiquitous presence in the environment and resulting bioaccumulation. MWAS of PCB-153 provided a means of evaluating how chronic exposure to a common persistent organic pollutant influenced metabolic variations and if biological response was similar for these two structurally related pollutants. This study differs from the other chapters present in this thesis in that it allowed us to test for biological response to exposures that have occurred over the course of decades, demonstrating that exposures occurring through different routes and developmental periods have implications for the metabolic phenotype. Targeted quantification of blood PCB-153 showed concentrations were comparable to national averages for both generations, suggesting exposure in this cohort is representative to most of the US population. Blood PBB-153 concentrations indicated registry participants carry a significantly higher burden of PBB-153 when compared to similar age groups. Thus, while the initial contamination occurred decades ago, exposure either as a child or through maternal transfer has led to a continued high internal exposure to this compound.

Using the quantified levels of PBB-153 and PCB-153 for MWAS, metabolic variations correlated with exposure were detected in both generations. For the older generation, correlations with metabolites related to mitochondrial dysfunction and neurotransmitter pathways were present, including metabolites from the TCA cycle, polyamine metabolism and catecholamines. Metabolic variations for the younger generation where exposure primarily occurred *in utero* were representative of similar biochemical processes. Co-factors for

neurotransmitter biosynthesis and recycling were associated with PBB-153, in addition to pathways related to polyamine metabolism, mineralocorticoids, oxidative stress-related metabolites, co-factors and fatty acids. N8-acetylspermidine, a metabolite previously associated with rapid progression Parkinsons Disease (PD) exhibited a positive association with PBB-153 levels, providing additional evidence of metabolic variations consistent with neurodegenerative diseases. Metabolites correlated with PCB-153 included neurotransmitter precursors and pathways related to lipid and fatty acid metabolism, which has been observed in previous metabolomic studies of PCB exposure. Many of the changes detected in both generations were consistent with the pathophysiological processes underlying neurodegenerative diseases. While this study could not establish a link between exposure and these diseases, previous studies have shown PCB-153 to be associated with increased risk for PD and the metabolic variations support a link between biological response and the metabolic changes underlying neurodegenerative processes. Since this study was observational, we used PBB-153, PCB-153 and free fatty acids data available from the National Health and Nutrition Examination Survey (NHANES) to see if the pollutant correlations with fatty acids could be reproduced. A limited number showed correlation with PCB-153 or PBB-153, suggesting exposure leads to disruption in essential fatty acid metabolism. This has important implications for processes underlying development of fatty liver, cardiovascular and metabolic diseases. Taken together, the results from this study demonstrate HRM can be used to link markers of persistent environmental pollutant exposure to metabolic

variations, even if the primary exposure occurred decades ago. These findings have important implications for using HRM in a precision medicine framework, since it establishes past events can lead to metabolic variations consistent with disease risk years to decades after it occurred.

In **Chapter 4**, the suitability of samples collected in the DoDSR was established for chemical profiling and identifying the exposure-associated metabolic phenotypes. The concept of using MWAS to link exposure and metabolic variations was extended in **Chapters 5-8** to varying levels of exposure timescales and measurement certainty to show that HRM can be used as a central platform for linking exposure to biological response and internal dose. This principle establishes HRM as a platform for environmental chemical surveillance and bioeffect monitoring by showing it is possible to measure metabolic variations indicating exposure has occurred and associated response, establishing biological relevance. Thus, capabilities are now available to establish a rigorous analytical framework to regularly monitor exposures. In **Chapter 9**, a framework for incorporating HRM technologies as a tool for monitoring the occurrence of adverse exposure events in deployed troops is presented. The high-throughput and low cost of HRM analyses will allow periodic screening of exposure biomarkers. This reduces the uncertainty available from a single point measure. Application for a small fraction of individuals can approximate a real-time surveillance tool to detect new or unexpected exposures, improving identification of environmental hazards and developing survey of chemical occurrence in human populations. Due to the availability of a large number of serum samples in the DoDSR, a

logical first application would be HRM analysis of the DoDSR. Using currently available technologies, it would be possible to perform a focused analysis of high-abundance metabolites in a one million samples in 5 years at a costs of \$3 million per year, providing comprehensive metabolic assessment of population differences due to geography, occupation, health habits, age, gender and disease, as well as trends over time. This is a very small percentage of the current DoD healthcare budget and could be easily offset by developing a virtual data repository rather than requiring physical storage of samples.

HRM can be incorporated in current DoDSR operations, providing a means of assessing the occurrence of adverse environmental events for deployed troops. Profiling of samples collected pre- and post-deployment can be tested for exposure biomarkers, biological response and metabolic changes consistent with increased risk. Thus, with appropriate commitment and resources, an analytical resource can be established to support routine exposure surveillance in armed forces personnel. The data would provide a valuable resource to evaluate possible health concerns and provide a large reference database of normal values for future precision medicine initiatives. Additional omic measures could greatly enhance the value of establishing an HRM repository for the DoDSR. Integrating data from the proteome, epigenome, microbiome and transcriptome with the environmental chemical and bioeffect measures available from HRM improves risk assessment of exposures occurring during deployment and in other occupational settings. This is demonstrated in by using an example of combining the metabolome of bronchial lavage fluid in patients with HIV-1 infection, which

identified network clusters of bacterium that contribute the most to infections in these immune compromised patients. Although conceptually developed for environmental chemical surveillance during deployment, the HRM framework could easily be extended to precision medicine initiatives, national surveys of chemical exposure and public health screening programs. This will support development of cumulative databases to evaluate time- and intensity-dependent changes in exposure and related metabolic perturbations in cross-sections of the US population. The subsequent data will provide key advances in hazard identification and understanding of the chemical burden within the US.

10.2. Recommendations for future research

Detailed future recommendations for continued development of HRM are provided below. In summary, these include delineation of metabolic mechanisms consistent with disease processes, improved characterization of unknown chemical signals, development of reference databases containing HRM profiles from a range of health states and exposure, combined experimental and computational techniques for unambiguous signal annotation and creation of reference standard materials for large-scale quantification. The goal of these recommendations is to provide a sufficiently robust analytical framework for use in national chemical surveillance and precision medicine programs, with continued development expected to greatly enhance measure of the occurrence, distribution and magnitude of environmental exposures while advancing knowledge on mechanisms underlying environment-related diseases

- Each study showed metabolic variations associated with exposure that was consistent with disease processes potentially linked to each of the pollutants. These included alterations to metabolites from renal, liver and immune functions associated with TCE exposure in **Chapter 5**, variations in biochemical processes indicating endothelial dysfunction and cardiovascular disease risk factors associated with traffic-related pollutants in **Chapters 6-7** and changes consistent with neurodegenerative diseases and mitochondrial dysfunction associated with persistent organic pollutants in **Chapter 8**. These findings have important implications for understanding the link to between exposure and initiation of diseases processed; however, these results are correlative in nature and do not establish causality. Future work should focus on applying HRM methodologies to characterize model systems that replicate the exposure paradigms described in this thesis. Replication of biological response will establish causation and allow more thorough investigation of the mechanisms identified in this thesis.
- In each study, many of the detected m/z features associated with environmental chemicals provided no plausible matches in metabolomic chemical databases. These signals are unidentifiable for a number of reasons, including the presence of uncharacterized metabolic intermediates, specificity of databases to endogenous metabolites and unknown environmental exposures. Application of HRM in the study of different disease processes has shown similar patterns, with up to 80% of

detected signals unidentified (Uppal et al. 2016). Thus, characterizing unidentified signals represents one of the key challenges in HRM profiling. While traditional analytical techniques such as ion dissociation patterns and NMR can be used to identify chemical formulas and structures, respectively, limited sample volume, time and low abundance make application challenging. Future work should focus on continued development of computational approaches that use data correlation structures to infer feature function and look at patterns of relationship with other commonly detected. This will be facilitated by establishment of cumulative databases that included aggregated HRM data from a diverse range of human samples.

- The ability to provide broad-scale coverage of the metabolome has only recently been established. As a result, the metabolic changes associated with different disease states and biological response to pollution is largely uncharacterized. Future areas of research should emphasize identifying specific metabolic phenotypes of different diseases and environmental stressors. Once these phenotypes have been established, the ability to identify diseases risk, metabolic mechanisms of disease and characterize exposures based on metabolic changes rather than exposure biomarkers will be greatly enhanced. This will also considerably enhance the use of HRM in precision medicine by allowing the comparison of the metabolic phenotype to changes consistent with disease and adverse health outcome risk.

- Although annotation of m/z features has been improved with the introduction of confidence-based approaches, ambiguity still exists. In many cases, multiple metabolites will match a single m/z , and absolute confirmation of identity cannot be assured without comparison of multiple independent feature parameters (m/z , retention time, dissociation pattern, etc.) to authentic reference standards. This is further complicated by isomeric and chiral molecules, which may not be separated using standard chromatography methods. Future efforts should continue to focus on development of computational approaches and reference databases that allow more efficient identification of detected metabolites. Improvement in identification can be accomplished through development of specific chromatographic and mass spectral landmarks that delineate regions within the HRM data containing known metabolites. Cumulative collection of identified metabolites within each region can then be developed independent of platforms.
- For the majority of the data presented in this thesis, m/z signal intensity was represented by area under the curve for each peak. This abundance is proportional to the concentration in the sample; however, it is unique to a given platform and influenced by physiochemical properties of the molecule. It is not possible to directly compare intensity to other systems and develop concentration ranges based on intensity alone. Reference standardization provides a framework for post-acquisition quantitation of HRM data using single point calibration. Currently available reference

materials lack comprehensive characterization of a large number of metabolites and environmental chemicals. Because these samples often consist of pools generated from a large number of individuals, environmental chemicals provide additional challenges and are often below detection limits. Development of reference samples that have quantitative data on a large number of endogenous metabolites and environmental chemicals (possibly added at detectable levels) will be critical for using HRM as a population screening platform. When absolute concentrations can be determined, it is possible to establish reference ranges and see if physiological concentrations are within healthy ranges or consistent with disease states. Providing this data for environmental chemicals will also greatly enhance risk assessment of environmental exposures.

10.3. Contributions

The following publications include work taken from this thesis:

Papers published in peer-reviewed journals:

1. **Walker DI**, Uppal K, Zhang L, Vermeulen R, Smith M, Hu W, Purdue M, Tang X, Qiu C, Reiss B, Kim S, Li L, Pennell KD, Jones DP, Rothman N, Lan Q. High-resolution metabolomics of occupational exposure to trichloroethylene. *International Journal of Epidemiology*. 2016 45(5): 1517-1527. PMID: 27707868

2. **Walker DI**, Pennell KD, Uppal K, Xia X, Hopke PK, Utell MJ, Phipps RP, Sime PJ, Rohrbeck P, Mallon TM, Jones DP. (2016). Pilot metabolome-wide association study of benzo(a)pyrene in serum from military personnel. *Journal of Occupational and Environmental Medicine*. 58:S44-52. PMID: 27501104
3. **Walker DI**, Mallon TM, Hopke PK, Uppal K, Go YM, Rohrbeck P, Pennell KD, Jones DP. (2016). Deployment-associated exposure surveillance with high-resolution metabolomics. *Journal of Occupational and Environmental Medicine*. 58:S12-21. PMID: 27501099

Edited book chapters

1. **Walker DI**, Go YM, Liu K, Pennell KD, Jones DP. Population screening for biological and environmental properties of the human metabolic phenotype: Implications for personalized medicine. Eds. J Nicholson, A. Darzi, E Holmes, J Lindon. *In Metabolic Phenotyping in Personalized and Public Healthcare*. 2016 7:168-206
2. **Walker DI**, Pennell KD, Jones DP. (2017). The mitochondrial exposome. Eds. J. Dykens, Y. Will. *In Drug-induced Mitochondrial Dysfunction Volume 2*. Accepted

Manuscripts in preparation

1. **Walker DI**, Lane KJ, Uppal K, Liu K, Patton A, Durant JL, Jones DP, Brugge D, Pennell KD. Metabolomic assessment of yearlong exposure to near-highway ultrafine particles
2. **Walker DI**, Marder ME, Yano Y, Liang Y, Boyd-Barr D, Miller GW, Jones DP, Marcus M, Pennell KD. Metabolome wide association study of polybrominated (PBB-153) and polychlorinated (PCB-153) biphenyls
3. **Walker DI**, Hart JE, Uppal K, Patel C, Rudel R, Garshick E, Pennell KD, Laden F, Jones DP. Integrated molecular response of exposure to traffic-related pollutants in the US trucking industry

The following conference presentations include work taken from this thesis.

Oral presentations

1. **Walker DI**, High-resolution metabolomics: Linking molecular and environmental epidemiology for precision medicine research. American Occupational Health Conference, Denver CO. April 23-26, 2017
2. **Walker DI**, Zhang L, Vermeulen R, Smith M, Uppal K, Hu W, Pennell K, Rothmann N, Jones D, Lan Q. High-resolution metabolomics study of occupational exposure to trichloroethylene. Annual Meeting of the Society of Environmental Toxicology and Chemistry, Salt Lake City, UT. November 1-5, 2015
3. **Walker DI**, Yano Y, Uppal K, Boyd-Barr D, Terrell M, Marder M, Jones D, Marcus M. High resolution metabolomics to identify PBB 153 exposure associated metabolic alterations in-vivo. 11th International

Conference of the Metabolomics Society. San Francisco, CA. June 29-July 2, 2015

4. Jones DP, **Walker DI**, Uppal K, Tran T, Yu T, Li S, Pennell KD, Miller G. The million metabolome: Key to profiling the human chemical experience. Annual Meeting of the Society of Environmental Toxicology and Chemistry, Vancouver BC. November 9-13, 2014
5. **Walker DI**, Miller GW, Hu W, Gearing M, Levey AI, Pennell KD, Jones DP. A new paradigm for toxicity of environmental exposures: Targeted and un-targeted cross-platform correlation approach for human in-vivo toxicodynamics. Annual Meeting of the Society of Environmental Toxicology and Chemistry, Vancouver BC. November 9-13, 2014
6. **Walker DI**, Yano Y, Uppal K, Boyd-Barr D, Marder ME, Jones DP, Marcus M. "High performance metabolomics for exposure associated metabolic alterations in-vivo." Annual Meeting of the Southeast Society of Toxicology, Athens GA. October 23-24, 2014

Poster presentations

1. **Walker DI**, Lane KJ, Patton A, Collins C, Durant JL, Jones DP, Pennell KD, Brugge D. Metabolomic assessment of ultrafine particle exposure in the CAFEH cohort. International Society of Environmental Epidemiology Annual Meeting, Rome Italy, September 1-4, 2016.

2. **Walker, DI**, Pennell KD, Jones D. Non-targeted, Mass Spectrometry Based Chemical Profiling of Human Mitochondria. US EPA Non-Targeted Workshop, Durham NC, August 18-19, 2015
3. **Walker DI**, Utell M, Uppal K, Phipps RP, Hopke P, Xia X, Jones DP, Mallon TM. Integrated profiling framework for investigating PAH induced metabolic alterations. Annual Meeting of the Society of Toxicology, San Diego, CA. March 22-26, 2015
4. **Walker, DI**, Pennell KD, Jones DP. Detailing the Human Exposome: Mass Spectrometry Based Chemical Profiling of Human Mitochondria. Annual Meeting of the Society of Toxicology, Phoenix, AZ. March 24-27, 2014
5. **Walker DI**, Roede JR, Park Y, Factor SA, Miller GW, Jones DP, Pennell KD. Exposomics approach to understanding the metabolic profile of Parkinson's disease. American Chemical Society Fall Meeting. September 8-12, 2013

Appendix 1. LIST OF SUPPLEMENTARY FILES

1. **Supplementary File 1:** Complete list of *m/z* features associated with serum benzo[a]pyrene, Chapter 4.
2. **Supplementary File 2:** Regression and annotation results for trichloroethylene metabolome wide association study, Chapter 5.
3. **Supplementary File 3:** Gene and metabolite pathway enrichment results for molecular response network, Chapter 6.
4. **Supplementary File 4:** PLS-DA metabolome wide association study of ultra-fine particle exposure and annotation results for both HILIC and RPC results, Chapter 7.
5. **Supplementary File 5:** PBB-153 and PCB-153 MWAS and annotation results for both generations, Chapter 8.

REFERENCES

- Accardi, C. J., D. I. Walker, K. Uppal, A. A. Quyyumi, P. Rohrbeck, K. D. Pennell, C. T. Mallon, D. P. Jones (2016). "High-Resolution Metabolomics for Nutrition and Health Assessment of Armed Forces Personnel." J Occup Environ Med **58**(8 Suppl 1): S80-88.
- Akude, E., E. Zhrebetskaya, S. K. Roy Chowdhury, K. Girling, P. Fernyhough (2010). "4-Hydroxy-2-nonenal induces mitochondrial dysfunction and aberrant axonal outgrowth in adult sensory neurons that mimics features of diabetic neuropathy." Neurotox Res **17**(1): 28-38.
- Al-Daghri, N. M., M. S. Alokail, S. H. Abd-Alrahman, H. M. Draz (2014). "Polycyclic aromatic hydrocarbon distribution in serum of Saudi children using HPLC-FLD: marker elevations in children with asthma." Environ Sci Pollut Res Int **21**(20): 12085-12090.
- Alam, M. S., S. Zeraati-Rezaei, C. P. Stark, Z. Liang, H. Xu, R. M. Harrison (2016). "The characterisation of diesel exhaust particles - composition, size distribution and partitioning." Faraday Discuss **189**: 69-84.
- Albert, R. (2005). "Scale-free networks in cell biology." J Cell Sci **118**(Pt 21): 4947-4957.
- Amorim, L. C., E. M. Alvarez-Leite (1997). "Determination of o-cresol by gas chromatography and comparison with hippuric acid levels in urine samples of individuals exposed to toluene." J Toxicol Environ Health **50**(4): 401-407.
- Anderson, H. A., R. Lilis, I. J. Selikoff, K. D. Rosenman, J. A. Valciukas, S. Freedman (1978). "Unanticipated prevalence of symptoms among dairy farmers in Michigan and Wisconsin." Environ Health Perspect **23**: 217-226.
- Ang, J. E., V. Revell, A. Mann, S. Mantele, D. T. Otway, J. D. Johnston, A. E. Thumser, D. J. Skene, F. Raynaud (2012). "Identification of human plasma metabolites exhibiting time-of-day variation using an untargeted liquid chromatography-mass spectrometry metabolomic approach." Chronobiol Int **29**(7): 868-881.
- Aquilina, N. J., J. M. Delgado-Saborit, C. Meddings, S. Baker, R. M. Harrison, P. Jacob, 3rd, M. Wilson, L. Yu, M. Duan, N. L. Benowitz (2010). "Environmental and biological monitoring of exposures to PAHs and ETS in the general population." Environ Int **36**(7): 763-771.
- Aruoma, O. I., B. Halliwell, B. M. Hoey, J. Butler (1988). "The antioxidant action of taurine, hypotaurine and their metabolic precursors." Biochem J **256**(1): 251-255.
- Ashfaq, S., J. L. Abramson, D. P. Jones, S. D. Rhodes, W. S. Weintraub, W. C. Hooper, V. Vaccarino, R. W. Alexander, D. G. Harrison, A. A. Quyyumi (2008). "Endothelial function and aminothiols biomarkers of oxidative stress in healthy adults." Hypertension **52**(1): 80-85.
- Athersuch, T. J., H. C. Keun (2015). "Metabolic profiling in human exposome studies." Mutagenesis.

- ATSDR (2008). Toxicological Profile for Cresols. Atlanta, GA, Agency for Toxic Substances and Disease Registry, Division of Toxicology and Environmental Medicine/Applied Toxicology Branch.
- Avila, M. A., E. R. Garcia-Trevijano, S. C. Lu, F. J. Corrales, J. M. Mato (2004). "Methylthioadenosine." Int J Biochem Cell Biol **36**(11): 2125-2130.
- Baccarelli, A., R. O. Wright, V. Bollati, L. Tarantini, A. A. Litonjua, H. H. Suh, A. Zanobetti, D. Sparrow, P. S. Vokonas, J. Schwartz (2009). "Rapid DNA methylation changes after exposure to traffic particles." Am J Respir Crit Care Med **179**(7): 572-578.
- Baird, C. P. (2012). "Review of the Institute of Medicine report: long-term health consequences of exposure to burn pits in Iraq and Afghanistan." US Army Med Dep J: 43-47.
- Barnett, K., S. W. Mercer, M. Norbury, G. Watt, S. Wyke, B. Guthrie (2012). "Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study." Lancet **380**(9836): 37-43.
- Barr, D. B., L. L. Needham (2002). "Analytical methods for biological monitoring of exposure to pesticides: a review." Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences **778**(1-2): 5-29.
- Barr, D. B., R. Y. Wang, L. L. Needham (2005). "Biologic monitoring of exposure to environmental chemicals throughout the life stages: requirements and issues for consideration for the National Children's Study." Environ Health Perspect **113**(8): 1083-1091.
- Bartel, J., J. Krumsiek, K. Schramm, J. Adamski, C. Gieger, C. Herder, M. Carstensen, A. Peters, W. Rathmann, M. Roden, et al. (2015). "The Human Blood Metabolome-Transcriptome Interface." PLoS Genet **11**(6): e1005274.
- Bartoń, K. (2013). {MuMIn}: multi-model inference, {R} package version 1.9.13.
- Barton, S., S. L. Navarro, M. F. Buas, Y. Schwarz, H. Gu, D. Djukovic, D. Raftery, M. Kratz, M. L. Neuhauser, J. W. Lampe (2015). "Targeted plasma metabolome response to variations in dietary glycemic load in a randomized, controlled, crossover feeding trial in healthy adults." Food Funct **6**(9): 2949-2956.
- Bassig, B. A., L. Zhang, X. Tang, R. Vermeulen, M. Shen, M. T. Smith, C. Qiu, Y. Ge, Z. Ji, B. Reiss, et al. (2013). "Occupational exposure to trichloroethylene and serum concentrations of IL-6, IL-10, and TNF-alpha." Environ Mol Mutagen **54**(6): 450-454.
- Bates, D., M. Mächler, B. Bolker, S. Walker (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software **67**(1).
- Bates, J. T., R. J. Weber, J. Abrams, V. Verma, T. Fang, M. Klein, M. J. Strickland, S. E. Sarnat, H. H. Chang, J. A. Mulholland, et al. (2015). "Reactive Oxygen Species Generation Linked to Sources of Atmospheric Particulate Matter and Cardiorespiratory Effects." Environ Sci Technol **49**(22): 13605-13612.

- Beelen, R., G. Hoek, P. A. van den Brandt, R. A. Goldbohm, P. Fischer, L. J. Schouten, M. Jerrett, E. Hughes, B. Armstrong, B. Brunekreef (2008). "Long-term effects of traffic-related air pollution on mortality in a Dutch cohort (NLCS-AIR study)." Environ Health Perspect **116**(2): 196-202.
- Bekesi, J. G., H. A. Anderson, J. P. Roboz, J. Roboz, A. Fischbein, I. J. Selikoff, J. F. Holland (1979). "Immunologic dysfunction among PBB-exposed Michigan dairy farmers." Ann N Y Acad Sci **320**: 717-728.
- Benbrahim-Tallaa, L., R. A. Baan, Y. Grosse, B. Lauby-Secretan, F. El Ghissassi, V. Bouvard, N. Guha, D. Loomis, K. Straif, G. International Agency for Research on Cancer Monograph Working (2012). "Carcinogenicity of diesel-engine and gasoline-engine exhausts and some nitroarenes." Lancet Oncol **13**(7): 663-664.
- Benjamini, Y., Y. Hochberg (1995). "Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society Series B-Methodological **57**(1): 289-300.
- Birnbaum, L. S., D. F. Staskal (2004). "Brominated flame retardants: cause for concern?" Environ Health Perspect **112**(1): 9-17.
- Blanck, H. M., M. Marcus, P. E. Tolbert, C. Rubin, A. K. Henderson, V. S. Hertzberg, R. H. Zhang, L. Cameron (2000). "Age at menarche and tanner stage in girls exposed in utero and postnatally to polybrominated biphenyl." Epidemiology **11**(6): 641-647.
- Bletsou, A. A., J. Jeon, J. Hollender, E. Archontaki, N. S. Thomaidis (2015). "Targeted and non-targeted liquid chromatography-mass spectrometric workflows for identification of transformation products of emerging pollutants in the aquatic environment." TrAC Trends in Analytical Chemistry **66**: 32-44.
- Bode-Boger, S. M., F. Scalera, J. T. Kielstein, J. Martens-Lobenhoffer, G. Breithardt, M. Fobker, H. Reinecke (2006). "Symmetrical dimethylarginine: a new combined parameter for renal function and extent of coronary artery disease." J Am Soc Nephrol **17**(4): 1128-1134.
- Boggs, R. W. (1978). "Bioavailability of acetylated derivatives of methionine, threonine, and lysine." Adv Exp Med Biol **105**: 571-586.
- Bonefeld-Jorgensen, E. C. (2010). "Biomonitoring in Greenland: human biomarkers of exposure and effects - a short review." Rural Remote Health **10**(2): 1362.
- Bonvallot, N., M. Tremblay-Franco, C. Chevrier, C. Canlet, C. Warembourg, J. P. Cravedi, S. Cordier (2013). "Metabolomics tools for describing complex pesticide exposure in pregnant women in Brittany (France)." PLoS One **8**(5): e64433.
- Bowen, B. P., T. R. Northen (2010). "Dealing with the unknown: metabolomics and metabolite atlases." J Am Soc Mass Spectrom **21**(9): 1471-1476.
- Boyd, B., C. Basic, R. Bethem (2008). Trace quantitative analysis by mass spectrometry. Chichester, West Sussex, England ; Hoboken, N.J., John Wiley & Sons.

- Bradburne, C., D. Graham, H. M. Kingston, R. Brenner, M. Pamuku, L. Carruth (2015). "Overview of 'Omics Technologies for Military Occupational Health Surveillance and Medicine." Mil Med **180**(10 Suppl): 34-48.
- Bradner, J. M., T. A. Suragh, W. W. Wilson, C. R. Lazo, K. A. Stout, H. M. Kim, M. Z. Wang, D. I. Walker, K. D. Pennell, J. R. Richardson, et al. (2013). "Exposure to the polybrominated diphenyl ether mixture DE-71 damages the nigrostriatal dopamine system: role of dopamine handling in neurotoxicity." Exp Neurol **241**: 138-147.
- Branten, A. J., T. P. Mulder, W. H. Peters, K. J. Assmann, J. F. Wetzels (2000). "Urinary excretion of glutathione S transferases alpha and pi in patients with proteinuria: reflection of the site of tubular injury." Nephron **85**(2): 120-126.
- Brautbar, N., J. Williams, 2nd (2002). "Industrial solvents and liver toxicity: risk assessment, risk factors and mechanisms." Int J Hyg Environ Health **205**(6): 479-491.
- Breitner, S., A. Schneider, R. B. Devlin, C. K. Ward-Caviness, D. Diaz-Sanchez, L. M. Neas, W. E. Cascio, A. Peters, E. R. Hauser, S. H. Shah, et al. (2016). "Associations among plasma metabolite levels and short-term exposure to PM2.5 and ozone in a cardiac catheterization cohort." Environ Int **97**: 76-84.
- Breton, C. V., A. N. Marutani (2014). "Air Pollution and Epigenetics: Recent Findings." Current Environmental Health Reports **1**(1): 35-45.
- Brower, J. B., M. Doyle-Eisele, B. Moeller, S. Stirdivant, J. D. McDonald, M. J. Campen (2016). "Metabolomic changes in murine serum following inhalation exposure to gasoline and diesel engine emissions." Inhal Toxicol **28**(5): 241-250.
- Brown, A. A., F. B. Hu (2001). "Dietary modulation of endothelial function: implications for cardiovascular disease." Am J Clin Nutr **73**(4): 673-686.
- Brown, R. C., A. H. Lockwood, B. R. Sonawane (2005). "Neurodegenerative diseases: an overview of environmental risk factors." Environ Health Perspect **113**(9): 1250-1256.
- Brugge, D., K. Lane, L. T. Padro-Martinez, A. Stewart, K. Hoesterey, D. Weiss, D. D. Wang, J. I. Levy, A. P. Patton, W. Zamore, et al. (2013). "Highway proximity associated with cardiovascular disease risk: the influence of individual-level confounders and exposure misclassification." Environ Health **12**(1): 84.
- Bruning, T., H. M. Bolt (2000). "Renal toxicity and carcinogenicity of trichloroethylene: key results, mechanisms, and controversies." Crit Rev Toxicol **30**(3): 253-285.
- Burchiel, S. W., M. I. Luster (2001). "Signaling by environmental polycyclic aromatic hydrocarbons in human lymphocytes." Clin Immunol **98**(1): 2-10.
- Burgess, L. G., K. Uppal, D. I. Walker, R. M. Roberson, V. Tran, M. B. Parks, E. A. Wade, A. T. May, A. C. Umfress, K. L. Jarrell, et al. (2015). "Metabolome-Wide Association Study of Primary Open Angle Glaucoma." Invest Ophthalmol Vis Sci **56**(8): 5020-5028.

- Carrizo, D., O. P. Chevallier, J. V. Woodside, S. F. Brennan, M. M. Cantwell, G. Cuskelly, C. T. Elliott (2017). "Untargeted metabolomic analysis of human serum samples associated with exposure levels of Persistent organic pollutants indicate important perturbations in Sphingolipids and Glycerophospholipids levels." *Chemosphere* **168**: 731-738.
- Castro Cabezas, M., C. J. Halkes, S. Meijssen, A. J. van Oostrom, D. W. Erkelens (2001). "Diurnal triglyceride profiles: a novel approach to study triglyceride changes." *Atherosclerosis* **155**(1): 219-228.
- Caudle, W. M., T. S. Guillot, C. R. Lazo, G. W. Miller (2012). "Industrial toxicants and Parkinson's disease." *Neurotoxicology* **33**(2): 178-188.
- Caudle, W. M., J. R. Richardson, K. C. Delea, T. S. Guillot, M. Wang, K. D. Pennell, G. W. Miller (2006). "Polychlorinated biphenyl-induced reduction of dopamine transporter expression as a precursor to Parkinson's disease-associated dopamine toxicity." *Toxicol Sci* **92**(2): 490-499.
- Cavalieri, E. L., E. G. Rogan (1995). "Central role of radical cations in metabolic activation of polycyclic aromatic hydrocarbons." *Xenobiotica* **25**(7): 677-688.
- Cavuoto, P., M. F. Fenech (2012). "A review of methionine dependency and the role of methionine restriction in cancer growth control and life-span extension." *Cancer Treat Rev* **38**(6): 726-736.
- CDC, NCHS Laboratory Data: Brominated Flame Retardants (BFRs), Non-dioxin-like Polychlorinated Biphenyls, Fatty Acids - Plasma; Year 2003-2004. [National Health and Nutrition Examination Survey Data, Year 2003-2004. https://wwwn.cdc.gov/nchs/nhanes/Search/DataPage.aspx?Component=Laboratory&CycleBeginYear=2003](https://wwwn.cdc.gov/nchs/nhanes/Search/DataPage.aspx?Component=Laboratory&CycleBeginYear=2003); Date accessed: January 21, 2017, Centers for Disease Control and Prevention; National Center for Health Statistics.
- Centre, T. M. I. (2015). FoodDB Version 1.0.
- Chadeau-Hyam, M., T. J. Athersuch, H. C. Keun, M. De Iorio, T. M. Ebbels, M. Jenab, C. Sacerdote, S. J. Bruce, E. Holmes, P. Vineis (2011). "Meeting-in-the-middle using metabolic profiling - a strategy for the identification of intermediate biomarkers in cohort studies." *Biomarkers* **16**(1): 83-88.
- Chambers, M. C., B. Maclean, R. Burke, D. Amodei, D. L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, et al. (2012). "A cross-platform toolkit for mass spectrometry and proteomics." *Nat Biotechnol* **30**(10): 918-920.
- Chang, T., R. Wang, L. Wu (2005). "Methylglyoxal-induced nitric oxide and peroxynitrite production in vascular smooth muscle cells." *Free Radic Biol Med* **38**(2): 286-293.
- Chang, T., L. Wu (2006). "Methylglyoxal, oxidative stress, and hypertension." *Can J Physiol Pharmacol* **84**(12): 1229-1238.
- Charbotel, B., J. Fevotte, M. Hours, J. L. Martin, A. Bergeret (2006). "Case-control study on renal cell cancer and occupational exposure to trichloroethylene. Part II: Epidemiological aspects." *Ann Occup Hyg* **50**(8): 777-787.

- Cheng, Y., S. C. Austin, B. Rocca, B. H. Koller, T. M. Coffman, T. Grosser, J. A. Lawson, G. A. FitzGerald (2002). "Role of prostacyclin in the cardiovascular response to thromboxane A₂." Science **296**(5567): 539-541.
- Chiu, W. A., J. Jinot, C. S. Scott, S. L. Makris, G. S. Cooper, R. C. Dzubow, A. S. Bale, M. V. Evans, K. Z. Guyton, N. Keshava, et al. (2013). "Human health effects of trichloroethylene: key findings and scientific issues." Environ Health Perspect **121**(3): 303-311.
- Chiu, W. A., S. Micallef, A. C. Monster, F. Y. Bois (2007). "Toxicokinetics of inhaled trichloroethylene and tetrachloroethylene in humans at 1 ppm: empirical results and comparisons with previous studies." Toxicol Sci **95**(1): 23-36.
- Choksi, N. Y., P. R. Kodavanti, H. A. Tilson, R. G. Booth (1997). "Effects of polychlorinated biphenyls (PCBs) on brain tyrosine hydroxylase activity and dopamine synthesis in rats." Fundam Appl Toxicol **39**(1): 76-80.
- Chonchol, M., C. Nielson (2008). "Hemoglobin levels and coronary artery disease." Am Heart J **155**(3): 494-498.
- Chu, J. H., J. E. Hart, D. Chhabra, E. Garshick, B. A. Raby, F. Laden (2016). "Gene expression network analyses in response to air pollution exposures in the trucking industry." Environ Health **15**(1): 101.
- Chua, E. C., G. Shui, I. T. Lee, P. Lau, L. C. Tan, S. C. Yeo, B. D. Lam, S. Bulchand, S. A. Summers, K. Puvanendran, et al. (2013). "Extensive diversity in circadian regulation of plasma lipids and evidence for different circadian metabolic phenotypes in humans." Proc Natl Acad Sci U S A **110**(35): 14468-14473.
- Chuang, K. J., C. C. Chan, T. C. Su, C. T. Lee, C. S. Tang (2007). "The effect of urban air pollution on inflammation, oxidative stress, coagulation, and autonomic dysfunction in young adults." Am J Respir Crit Care Med **176**(4): 370-376.
- Chung, M. K., J. Riby, H. Li, A. T. Iavarone, E. R. Williams, Y. Zheng, S. M. Rappaport (2010). "A sandwich enzyme-linked immunosorbent assay for adducts of polycyclic aromatic hydrocarbons with human serum albumin." Anal Biochem **400**(1): 123-129.
- Colas, R. A., M. Shinohara, J. Dalli, N. Chiang, C. N. Serhan (2014). "Identification and signature profiles for pro-resolving and inflammatory lipid mediators in human tissue." Am J Physiol Cell Physiol **307**(1): C39-54.
- Cone, E. J., Y. H. Caplan, F. Moser, T. Robert, M. K. Shelby, D. L. Black (2009). "Normalization of urinary drug concentrations with specific gravity and creatinine." J Anal Toxicol **33**(1): 1-7.
- Cooper, G. S., S. L. Makris, P. J. Nietert, J. Jinot (2009). "Evidence of autoimmune-related effects of trichloroethylene exposure from studies in mice and humans." Environ Health Perspect **117**(5): 696-702.
- Cox, J. D., E. Cama, D. M. Colleluori, S. Pethe, J. L. Boucher, D. Mansuy, D. E. Ash, D. W. Christianson (2001). "Mechanistic and metabolic inferences

- from the binding of substrate analogues and products to arginase." Biochemistry **40**(9): 2689-2701.
- Cribbs, S. K., Y. Park, D. M. Guidot, G. S. Martin, L. A. Brown, J. Lennox, D. P. Jones (2014). "Metabolomics of bronchoalveolar lavage differentiate healthy HIV-1-infected subjects from controls." AIDS Res Hum Retroviruses **30**(6): 579-585.
- Cribbs, S. K., K. Uppal, S. Li, D. Jones, L. Q. Huang, L. Tipton, A. Fitch, R. Greenblatt, L. Kingsley, D. Guidot, et al. (2016). "Correlation of the Lung Microbiota with Metabolic Profiles in Bronchoalveolar Lavage Fluid in HIV Infection." Microbiome (in press).
- Cribbs, S. K., K. Uppal, S. Li, D. P. Jones, L. Huang, L. Tipton, A. Fitch, R. M. Greenblatt, L. Kingsley, D. M. Guidot, et al. (2016). "Correlation of the lung microbiota with metabolic profiles in bronchoalveolar lavage fluid in HIV infection." Microbiome **4**(1): 3.
- Crider, K. S., T. P. Yang, R. J. Berry, L. B. Bailey (2012). "Folate and DNA methylation: a review of molecular mechanisms and the evidence for folate's role." Adv Nutr **3**(1): 21-38.
- Cristalli, G., S. Costanzi, C. Lambertucci, G. Lupidi, S. Vittori, R. Volpini, E. Camaioni (2001). "Adenosine deaminase: functional implications and different classes of inhibitors." Med Res Rev **21**(2): 105-128.
- Curfs, D. M., L. Beckers, R. W. Godschalk, M. J. Gijbels, F. J. van Schooten (2003). "Modulation of plasma lipid levels affects benzo[a]pyrene-induced DNA damage in tissues of two hyperlipidemic mouse models." Environ Mol Mutagen **42**(4): 243-249.
- da Silva, R. R., P. C. Dorrestein, R. A. Quinn (2015). "Illuminating the dark matter in metabolomics." Proc Natl Acad Sci U S A.
- Dallmann, R., A. U. Viola, L. Tarokh, C. Cajochen, S. A. Brown (2012). "The human circadian metabolome." Proc Natl Acad Sci U S A **109**(7): 2625-2629.
- Davis, A. P., C. G. Murphy, R. Johnson, J. M. Lay, K. Lennon-Hopkins, C. Saraceni-Richards, D. Sciaky, B. L. King, M. C. Rosenstein, T. C. Wieggers, et al. (2013). "The Comparative Toxicogenomics Database: update 2013." Nucleic Acids Res **41**(Database issue): D1104-1114.
- Davis, M. E., T. J. Smith, F. Laden, J. E. Hart, L. M. Ryan, E. Garshick (2006). "Modeling particle exposure in U.S. trucking terminals." Environ Sci Technol **40**(13): 4226-4232.
- Deelen, J., M. Beekman, M. Capri, C. Franceschi, P. E. Slagboom (2013). "Identifying the genomic determinants of aging and longevity in human population studies: progress and challenges." Bioessays **35**(4): 386-396.
- Dejean, S., I. Gonzalez, K.-A. L. Cao, P. Monget, J. Coquery, F. Yao, B. Lique, F. Rohart (2014). mixOmics: Omics Data Integration Project. R package version 5.0-3. <http://cran.r-project.org/package=mixOmics>.
- Delaney, J. C., J. M. Essigmann (2008). "Biological properties of single chemical-DNA adducts: a twenty year perspective." Chem Res Toxicol **21**(1): 232-252.

- DelRaso, N. J., V. T. Chan, C. A. Mauzy, P. A. Shivanov (2015). "Air Force Research Laboratory Integrated Omics Research." Mil Med **180**(10 Suppl): 67-75.
- Diaz-Sanchez, D. (1997). "The role of diesel exhaust particles and their associated polyaromatic hydrocarbons in the induction of allergic airway disease." Allergy **52**(38 Suppl): 52-56; discussion 57-58.
- Dieterle, F., B. Riefke, G. Schlotterbeck, A. Ross, H. Senn, A. Amberg (2011). "NMR and MS methods for metabonomics." Methods Mol Biol **691**: 385-415.
- Ding, X., P. A. Murray (2005). "Cellular mechanisms of thromboxane A₂-mediated contraction in pulmonary veins." Am J Physiol Lung Cell Mol Physiol **289**(5): L825-833.
- Dionisio, K. L., A. M. Frame, M.-R. Goldsmith, J. F. Wambaugh, A. Liddell, T. Cathey, D. Smith, J. Vail, A. S. Ernstoff, P. Fantke, et al. (2015). "Exploring consumer exposure pathways and patterns of use for chemicals in the environment." Toxicology Reports **2**: 228-237.
- Donaldson, K., V. Stone, A. Seaton, W. MacNee (2001). "Ambient particle inhalation and the cardiovascular system: potential mechanisms." Environ Health Perspect **109 Suppl 4**: 523-527.
- Du, P., W. A. Kibbe, S. M. Lin (2008). "lumi: a pipeline for processing Illumina microarray." Bioinformatics **24**(13): 1547-1548.
- Dunn, W., M. Brown, S. Worton, K. Davies, R. Jones, D. Kell, A. P. Heazell (2012). "The metabolome of human placental tissue: investigation of first trimester tissue and changes related to preeclampsia in late pregnancy." Metabolomics **8**(4): 579-597.
- Dunn, W. B., W. Lin, D. Broadhurst, P. Begley, M. Brown, E. Zelena, A. A. Vaughan, A. Halsall, N. Harding, J. D. Knowles, et al. (2015). "Molecular phenotyping of a UK population: defining the human serum metabolome." Metabolomics **11**: 9-26.
- Dwyer, J. P., B. A. Greco, K. Umanath, D. Packham, J. W. Fox, R. Peterson, B. R. Broome, L. E. Greene, M. Sika, J. B. Lewis (2015). "Pyridoxamine dihydrochloride in diabetic nephropathy (PIONEER-CSG-17): lessons learned from a pilot study." Nephron **129**(1): 22-28.
- Edmands, W. M., P. Ferrari, J. A. Rothwell, S. Rinaldi, N. Slimani, D. K. Barupal, C. Biessy, M. Jenab, F. Clavel-Chapelon, G. Fagherazzi, et al. (2015). "Polyphenol metabolome in human urine and its association with intake of polyphenol-rich foods across European countries." Am J Clin Nutr **102**(4): 905-913.
- Eid, H. M., H. Arnesen, E. M. Hjerkin, T. Lyberg, I. Seljeflot (2004). "Relationship between obesity, smoking, and the endogenous nitric oxide synthase inhibitor, asymmetric dimethylarginine." Metabolism **53**(12): 1574-1579.
- Ellis, J. K., T. J. Athersuch, L. D. Thomas, F. Teichert, M. Perez-Trujillo, C. Svendsen, D. J. Spurgeon, R. Singh, L. Jarup, J. G. Bundy, et al. (2012). "Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population." BMC Med **10**: 61.

- Ernstgard, L., B. Sjogren, M. Warholm, G. Johanson (2003). "Sex differences in the toxicokinetics of inhaled solvent vapors in humans 2. 2-propanol." Toxicol Appl Pharmacol **193**(2): 158-167.
- Fang, Z. Z., K. W. Krausz, N. Tanaka, F. Li, A. Qu, J. R. Idle, F. J. Gonzalez (2013). "Metabolomics reveals trichloroacetate as a major contributor to trichloroethylene-induced metabolic alterations in mouse urine and serum." Arch Toxicol **87**(11): 1975-1987.
- Fangstrom, B., L. Hovander, A. Bignert, I. Athanassiadis, L. Linderholm, P. Grandjean, P. Weihe, A. Bergman (2005). "Concentrations of polybrominated diphenyl ethers, polychlorinated biphenyls, and polychlorobiphenyls in serum from pregnant Faroese women and their children 7 years later." Environ Sci Technol **39**(24): 9457-9463.
- Fitzpatrick, A. M., Y. Park, L. A. Brown, D. P. Jones (2014). "Children with severe asthma have unique oxidative stress-associated metabolomic profiles." J Allergy Clin Immunol **133**(1): 258-261 e251-258.
- Fobker, M. (2014). "Stability of glucose in plasma with different anticoagulants." Clin Chem Lab Med **52**(7): 1057-1060.
- Fontana, L., L. Partridge, V. D. Longo (2010). "Extending healthy life span--from yeast to humans." Science **328**(5976): 321-326.
- Fontana, M., L. Pecci, S. Dupre, D. Cavallini (2004). "Antioxidant properties of sulfinates: protective effect of hypotaurine on peroxynitrite-dependent damage." Neurochem Res **29**(1): 111-116.
- Fowles, J. R., A. Fairbrother, L. Baecher-Steppan, N. I. Kerkvliet (1994). "Immunologic and endocrine effects of the flame-retardant pentabromodiphenyl ether (DE-71) in C57BL/6J mice." Toxicology **86**(1-2): 49-61.
- Frediani, J. K., D. P. Jones, N. Tukvadze, K. Uppal, E. Sanikidze, M. Kipiani, V. T. Tran, G. Hebbbar, D. I. Walker, R. R. Kempker, et al. (2014). "Plasma metabolomics in human pulmonary tuberculosis disease: a pilot study." PLoS One **9**(10): e108854.
- Fries, G. F. (1985). "The PBB episode in Michigan: an overall appraisal." Crit Rev Toxicol **16**(2): 105-156.
- Fuller, C. H., A. P. Patton, K. Lane, M. B. Laws, A. Marden, E. Carrasco, J. Spengler, M. Mwamburi, W. Zamore, J. L. Durant, et al. (2013). "A community participatory study of cardiovascular health and exposure to near-highway air pollution: study design and methods." Rev Environ Health **28**(1): 21-35.
- Fuller, C. H., P. L. Williams, M. A. Mittleman, A. P. Patton, J. D. Spengler, D. Brugge (2015). "Response of biomarkers of inflammation and coagulation to short-term changes in central site, local, and predicted particle number concentrations." Ann Epidemiol **25**(7): 505-511.
- Gao, Y., Y. Lu, S. Huang, L. Gao, X. Liang, Y. Wu, J. Wang, Q. Huang, L. Tang, G. Wang, et al. (2014). "Identifying early urinary metabolic changes with long-term environmental exposure to cadmium by mass-spectrometry-based metabolomics." Environ Sci Technol **48**(11): 6409-6418.

- Garshick, E., F. Laden, J. E. Hart, B. Rosner, M. E. Davis, E. A. Eisen, T. J. Smith (2008). "Lung cancer and vehicle exhaust in trucking industry workers." Environ Health Perspect **116**(10): 1327-1332.
- Go, Y., D. I. Walker, Q. A. Soltow, K. Uppal, L. M. Wachtman, F. H. Strobel, K. Pennell, D. E. Promislow, D. P. Jones (2014). "Metabolome-wide association study of phenylalanine in plasma of common marmosets." Amino Acids.
- Go, Y. M., D. P. Jones (2005). "Intracellular proatherogenic events and cell adhesion modulated by extracellular thiol/disulfide redox state." Circulation **111**(22): 2973-2980.
- Go, Y. M., D. P. Jones (2011). "Cysteine/cystine redox signaling in cardiovascular disease." Free Radic Biol Med **50**(4): 495-509.
- Go, Y. M., D. P. Jones (2014). "Redox biology: interface of the exposome with the proteome, epigenome and genome." Redox Biol **2**: 358-360.
- Go, Y. M., Y. Liang, K. Uppal, Q. A. Soltow, D. E. Promislow, L. M. Wachtman, D. P. Jones (2015). "Metabolic Characterization of the Common Marmoset (*Callithrix jacchus*)." PLoS One **10**(11): e0142916.
- Go, Y. M., J. R. Roede, M. Orr, Y. Liang, D. P. Jones (2014). "Integrated redox proteomics and metabolomics of mitochondria to identify mechanisms of Cd toxicity." Toxicol Sci **139**(1): 59-73.
- Go, Y. M., K. Uppal, D. I. Walker, V. Tran, L. Dury, F. H. Strobel, H. Baubichon-Cortay, K. D. Pennell, J. R. Roede, D. P. Jones (2014). "Mitochondrial metabolomics using high-resolution Fourier-transform mass spectrometry." Methods Mol Biol **1198**: 43-73.
- Go, Y. M., D. I. Walker, Y. Liang, K. Uppal, Q. A. Soltow, V. Tran, F. Strobel, A. A. Quyyumi, T. R. Ziegler, K. D. Pennell, et al. (2015). "Reference Standardization for Mass Spectrometry and High-resolution Metabolomics Applications to Exposome Research." Toxicol Sci **148**(2): 531-543.
- Go, Y. M., D. I. Walker, Y. Liang, K. Uppal, Q. A. Soltow, V. Tran, F. Strobel, A. A. Quyyumi, T. R. Ziegler, K. D. Pennell, et al. (2015). "Reference Standardization for Mass Spectrometry and High-Resolution Metabolomics Applications to Exposome Research." Toxicol Sci.
- Go, Y. M., D. I. Walker, Q. A. Soltow, K. Uppal, L. M. Wachtman, F. H. Strobel, K. Pennell, D. E. Promislow, D. P. Jones (2014). "Metabolome-wide association study of phenylalanine in plasma of common marmosets." Amino Acids.
- Gochfeld, M. (2007). "Framework for gender differences in human and animal toxicology." Environ Res **104**(1): 4-21.
- Gomez-Cabrero, D., I. Abugessaisa, D. Maier, A. Teschendorff, M. Merckenschlager, A. Gisel, E. Ballestar, E. Bongcam-Rudloff, A. Conesa, J. Tegner (2014). "Data integration in the era of omics: current and future challenges." BMC Syst Biol **8 Suppl 2**: I1.
- Goodacre, R., S. Vaidyanathan, W. B. Dunn, G. G. Harrigan, D. B. Kell (2004). "Metabolomics by numbers: acquiring and understanding global metabolite data." Trends Biotechnol **22**(5): 245-252.

- Gooley, J. J., E. C. Chua (2014). "Diurnal regulation of lipid metabolism and applications of circadian lipidomics." J Genet Genomics **41**(5): 231-250.
- Gornik, H. L., M. A. Creager (2004). "Arginine and endothelial and vascular health." J Nutr **134**(10 Suppl): 2880S-2887S; discussion 2895S.
- Gouveia-Figueira, S., M. Karimpour, J. A. Bosson, A. Blomberg, J. Unosson, J. Pourazar, T. Sandstrom, A. F. Behndig, M. L. Nording (2017). "Mass spectrometry profiling of oxylipins, endocannabinoids, and N-acylethanolamines in human lung lavage fluids reveals responsiveness of prostaglandin E2 and associated lipid metabolites to biodiesel exhaust exposure." Anal Bioanal Chem **409**(11): 2967-2980.
- Grandjean, P., P. Weihe (2003). "Arachidonic acid status during pregnancy is associated with polychlorinated biphenyl exposure." Am J Clin Nutr **77**(3): 715-719.
- Greizerstein, H. B., P. Gigliotti, J. Vena, J. Freudenheim, P. J. Kostyniak (1997). "Standardization of a method for the routine analysis of polychlorinated biphenyl congeners and selected pesticides in human serum and milk." J Anal Toxicol **21**(7): 558-566.
- Grey, M., C. J. Dunning, R. Gaspar, C. Grey, P. Brundin, E. Sparr, S. Linse (2015). "Acceleration of alpha-Synuclein Aggregation by Exosomes." J Biol Chem **290**(5): 2969-2982.
- Guha, N., D. Loomis, Y. Grosse, B. Lauby-Secretan, F. El Ghissassi, V. Bouvard, L. Benbrahim-Tallaa, R. Baan, H. Mattock, K. Straif, et al. (2012). "Carcinogenicity of trichloroethylene, tetrachloroethylene, some other chlorinated solvents, and their metabolites." Lancet Oncol **13**(12): 1192-1193.
- Gurtan, A. M., P. A. Sharp (2013). "The role of miRNAs in regulating gene expression networks." J Mol Biol **425**(19): 3582-3600.
- Güven, A., A. Kayikci, K. Cam, P. Arbak, O. Balbay, M. Cam (2008). "Alterations in semen parameters of toll collectors working at motorways: does diesel exposure induce detrimental effects on semen?" Andrologia **40**(6): 346-351.
- Ha, H. C., N. S. Sirisoma, P. Kuppasamy, J. L. Zweier, P. M. Woster, R. A. Casero, Jr. (1998). "The natural polyamine spermine functions directly as a free radical scavenger." Proc Natl Acad Sci U S A **95**(19): 11140-11145.
- Hansen, J., M. Sallmen, A. I. Selden, A. Anttila, E. Pukkala, K. Andersson, I. L. Bryngelsson, O. Raaschou-Nielsen, J. H. Olsen, J. K. McLaughlin (2013). "Risk of cancer among workers exposed to trichloroethylene: analysis of three Nordic cohort studies." J Natl Cancer Inst **105**(12): 869-877.
- Hardell, E., M. Carlberg, M. Nordstrom, B. van Bavel (2010). "Time trends of persistent organic pollutants in Sweden during 1993-2007 and relation to age, gender, body mass index, breast-feeding and parity." Sci Total Environ **408**(20): 4412-4419.
- Hart, J. E. (2016). "Air pollution affects lung cancer survival." Thorax **71**(10): 875-876.

- Hatcher, J. M., K. C. Delea, J. R. Richardson, K. D. Pennell, G. W. Miller (2008). "Disruption of dopamine transport by DDT and its metabolites." Neurotoxicology **29**(4): 682-690.
- Hatcher, J. M., K. D. Pennell, G. W. Miller (2008). "Parkinson's disease and pesticides: a toxicological perspective." Trends Pharmacol Sci **29**(6): 322-329.
- Hatcher-Martin, J. M., M. Gearing, K. Steenland, A. I. Levey, G. W. Miller, K. D. Pennell (2012). "Association between polychlorinated biphenyls and Parkinson's disease neuropathology." Neurotoxicology **33**(5): 1298-1304.
- Haugen, A., G. Becher, C. Benestad, K. Vahakangas, G. E. Trivers, M. J. Newman, C. C. Harris (1986). "Determination of polycyclic aromatic hydrocarbons in the urine, benzo(a)pyrene diol epoxide-DNA adducts in lymphocyte DNA, and antibodies to the adducts in sera from coke oven workers exposed to measured amounts of polycyclic aromatic hydrocarbons in the work atmosphere." Cancer Res **46**(8): 4178-4183.
- Heller, R., A. Unbehauen, B. Schellenberg, B. Mayer, G. Werner-Felmayer, E. R. Werner (2001). "L-ascorbic acid potentiates endothelial nitric oxide synthesis via a chemical stabilization of tetrahydrobiopterin." J Biol Chem **276**(1): 40-47.
- Henderson, A. K., D. Rosen, G. L. Miller, L. W. Figgs, S. H. Zahm, S. M. Sieber, N. Rothman, H. E. Humphrey, T. Sinks (1995). "Breast cancer among women exposed to polybrominated biphenyls." Epidemiology **6**(5): 544-546.
- Hennig, B., B. D. Hammock, R. Slim, M. Toborek, V. Saraswathi, L. W. Robertson (2002). "PCB-induced oxidative stress in endothelial cells: modulation by nutrients." Int J Hyg Environ Health **205**(1-2): 95-102.
- Hennig, B., R. Slim, M. Toborek, L. W. Robertson (1999). "Linoleic acid amplifies polychlorinated biphenyl-mediated dysfunction of endothelial cells." J Biochem Mol Toxicol **13**(2): 83-91.
- Herve, M. (2017). RVAideMemoire: Diverse Basic Statistical and Graphical Functions.
- Hofmann, H., H. H. Schmidt (1995). "Thiol dependence of nitric oxide synthase." Biochemistry **34**(41): 13443-13452.
- Holmes, E., R. L. Loo, J. Stamler, M. Bictash, I. K. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, et al. (2008). "Human metabolic phenotype diversity and its association with diet and blood pressure." Nature **453**(7193): 396-400.
- Holmes, E., I. D. Wilson, J. K. Nicholson (2008). "Metabolic phenotyping in health and disease." Cell **134**(5): 714-717.
- Hopf, N. B., A. M. Ruder, P. Succop (2009). "Background levels of polychlorinated biphenyls in the U.S. population." Sci Total Environ **407**(24): 6109-6119.
- Hoque, A., A. J. Sigurdson, K. D. Burau, H. E. Humphrey, K. R. Hess, A. M. Sweeney (1998). "Cancer among a Michigan cohort exposed to polybrominated biphenyls in 1973." Epidemiology **9**(4): 373-378.

- Horai, H., M. Arita, S. Kanaya, Y. Nihei, T. Ikeda, K. Suwa, Y. Ojima, K. Tanaka, S. Tanaka, K. Aoshima, et al. (2010). "MassBank: a public repository for sharing mass spectral data for life sciences." J Mass Spectrom **45**(7): 703-714.
- IARC (2014). "Trichloroethylene." IARC Monographs on the evaluation of carcinogenic risks to humans; Vol. 106 International Agency for Research on Cancer.
- IARC (2015). Polychlorinated biphenyls and polybrominated biphenyls. IARC Monogr Eval Carcinog Risks Hum, International Agency for Research on Cancer. **107**.
- Iavicoli, I., A. Marinaccio, G. Carelli (2005). "Effects of occupational trichloroethylene exposure on cytokine levels in workers." J Occup Environ Med **47**(5): 453-457.
- Ibanez, C., C. Simo, V. Garcia-Canas, A. Cifuentes, M. Castro-Puyana (2013). "Metabolomics, peptidomics and proteomics applications of capillary electrophoresis-mass spectrometry in Foodomics: a review." Anal Chim Acta **802**: 1-13.
- International HapMap, C. (2003). "The International HapMap Project." Nature **426**(6968): 789-796.
- International HapMap, C., D. M. Altshuler, R. A. Gibbs, L. Peltonen, D. M. Altshuler, R. A. Gibbs, L. Peltonen, E. Dermitzakis, S. F. Schaffner, F. Yu, et al. (2010). "Integrating common and rare genetic variation in diverse human populations." Nature **467**(7311): 52-58.
- Ishizuka, T., M. Kawakami, T. Hidaka, Y. Matsuki, M. Takamizawa, K. Suzuki, A. Kurita, H. Nakamura (1998). "Stimulation with thromboxane A2 (TXA2) receptor agonist enhances ICAM-1, VCAM-1 or ELAM-1 expression by human vascular endothelial cells." Clin Exp Immunol **112**(3): 464-470.
- Iyer, S. S., C. J. Accardi, T. R. Ziegler, R. A. Blanco, J. D. Ritzenthaler, M. Rojas, J. Roman, D. P. Jones (2009). "Cysteine redox potential determines pro-inflammatory IL-1 β levels." PLoS One **4**(3): e5017.
- Jackson, T. S., A. Xu, J. A. Vita, J. F. Keaney, Jr. (1998). "Ascorbate prevents the interaction of superoxide and nitric oxide only at very high physiological concentrations." Circ Res **83**(9): 916-922.
- Jacobsen, D. W. (1998). "Homocysteine and vitamins in cardiovascular disease." Clin Chem **44**(8 Pt 2): 1833-1843.
- Jain, S. K., G. Lim (2001). "Pyridoxine and pyridoxamine inhibits superoxide radicals and prevents lipid peroxidation, protein glycosylation, and (Na⁺ + K⁺)-ATPase activity reduction in high glucose-treated human erythrocytes." Free Radic Biol Med **30**(3): 232-237.
- Jamin, E. L., N. Bonvallot, M. Tremblay-Franco, J. P. Cravedi, C. Chevrier, S. Cordier, L. Debrauwer (2014). "Untargeted profiling of pesticide metabolites by LC-HRMS: an exposomics tool for human exposure evaluation." Anal Bioanal Chem **406**(4): 1149-1161.
- Janssen, N. A., G. Hoek, M. Simic-Lawson, P. Fischer, L. van Bree, H. ten Brink, M. Keuken, R. W. Atkinson, H. R. Anderson, B. Brunekreef, et al. (2011).

- "Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5." Environ Health Perspect **119**(12): 1691-1699.
- Jeanneret, F., J. Boccard, F. Badoud, O. Sorg, D. Tonoli, D. Pelclova, S. Vlckova, D. N. Rutledge, C. F. Samer, D. Hochstrasser, et al. (2014). "Human urinary biomarkers of dioxin exposure: analysis by metabolomics and biologically driven data dimensionality reduction." Toxicol Lett **230**(2): 234-243.
- Jiang, R., M. J. Jones, F. Sava, M. S. Kobor, C. Carlsten (2014). "Short-term diesel exhaust inhalation in a controlled human crossover study is associated with changes in DNA methylation of circulating mononuclear cells in asthmatics." Part Fibre Toxicol **11**: 71.
- Jobst, K. J., L. Shen, E. J. Reiner, V. Y. Taguchi, P. A. Helm, R. McCrindle, S. Backus (2013). "The use of mass defect plots for the identification of (novel) halogenated contaminants in the environment." Anal Bioanal Chem **405**(10): 3289-3297.
- Johnson, J. M., F. H. Strobel, M. Reed, J. Pohl, D. P. Jones (2008). "A rapid LC-FTMS method for the analysis of cysteine, cystine and cysteine/cystine steady-state redox potential in human plasma." Clin Chim Acta **396**(1-2): 43-48.
- Johnson, J. M., T. Yu, F. H. Strobel, D. P. Jones (2010). "A practical approach to detect unique metabolic patterns for personalized medicine." Analyst **135**(11): 2864-2870.
- Johnson, W. E., C. Li, A. Rabinovic (2007). "Adjusting batch effects in microarray expression data using empirical Bayes methods." Biostatistics **8**(1): 118-127.
- Jones, D. P. (2006). "Extracellular redox state: refining the definition of oxidative stress in aging." Rejuvenation Res **9**(2): 169-181.
- Jones, D. P. (2015). "Redox theory of aging." Redox Biol **5**: 71-79.
- Jones, D. P. (2015). "Redox theory of aging." Redox biology **5**: 71-79.
- Jones, D. P. (2016). "Sequencing the exposome: A call to action." Toxicology Reports **3**: 29-45.
- Jones, D. P., Y. Park, T. R. Ziegler (2012). "Nutritional metabolomics: progress in addressing complexity in diet and health." Annu Rev Nutr **32**: 183-202.
- Jones, D. P., D. I. Walker, K. Uppal, P. Rohrbeck, C. T. Mallon, Y. M. Go (2016). "Metabolic Pathways and Networks Associated With Tobacco Use in Military Personnel." J Occup Environ Med **58**(8 Suppl 1): S111-116.
- Joseph, A. D., M. L. Terrell, C. M. Small, L. L. Cameron, M. Marcus (2009). "Assessing inter-generational transfer of a brominated flame retardant." J Environ Monit **11**(4): 802-807.
- Jove, M., I. Mate, A. Naudi, N. Mota-Martorell, M. Portero-Otin, M. De la Fuente, R. Pamplona (2015). "Human Aging Is a Metabolome-related Matter of Gender." J Gerontol A Biol Sci Med Sci.
- Jungnickel, H., S. Potratz, S. Baumann, P. Tarnow, M. von Bergen, A. Luch (2014). "Identification of lipidomic biomarkers for coexposure to subtoxic

- doses of benzo[a]pyrene and cadmium: the toxicological cascade biomarker approach." Environ Sci Technol **48**(17): 10423-10431.
- Jutel, M., K. Blaser, C. A. Akdis (2006). "The role of histamine in regulation of immune responses." Chem Immunol Allergy **91**: 174-187.
- Kanehisa, M., S. Goto, Y. Sato, M. Furumichi, M. Tanabe (2012). "KEGG for integration and interpretation of large-scale molecular data sets." Nucleic Acids Res **40**(Database issue): D109-114.
- Kanu, A. B., P. Dwivedi, M. Tam, L. Matz, H. H. Hill, Jr. (2008). "Ion mobility-mass spectrometry." J Mass Spectrom **43**(1): 1-22.
- Khamis, M. M., D. J. Adamko, A. El-Aneed (2015). "Mass spectrometric based approaches in urine metabolomics and biomarker discovery." Mass Spectrom Rev.
- Kilburn, K. H. (2002). "Is neurotoxicity associated with environmental trichloroethylene (TCE)?" Arch Environ Health **57**(2): 113-120.
- Kim, A. M., C. M. Tinggen, T. K. Woodruff (2010). "Sex bias in trials and treatment must end." Nature **465**(7299): 688-689.
- Kim, S., L. B. Collins, G. Boysen, J. A. Swenberg, A. Gold, L. M. Ball, B. U. Bradford, I. Rusyn (2009). "Liquid chromatography electrospray ionization tandem mass spectrometry analysis method for simultaneous detection of trichloroacetic acid, dichloroacetic acid, S-(1,2-dichlorovinyl)glutathione and S-(1,2-dichlorovinyl)-L-cysteine." Toxicology **262**(3): 230-238.
- Kizu, R., K. Okamura, A. Toriba, A. Mizokami, K. L. Burnstein, C. M. Klinge, K. Hayakawa (2003). "Antiandrogenic activities of diesel exhaust particle extracts in PC3/AR human prostate carcinoma cells." Toxicol Sci **76**(2): 299-309.
- Koizumi, H., A. Ohkawara (1992). "Adenosine deaminase activity in sera of patients with psoriasis, mycosis fungoides and adult T cell leukemia." Acta Derm Venereol **72**(6): 410-412.
- Kopff, M., I. Zakrzewska, J. Czernicki, J. Klem, M. Strzelczyk (1993). "Red blood cell adenosine deaminase activity in multiple sclerosis." Clin Chim Acta **214**(1): 97-101.
- Kovacic, P., R. Somanathan (2014). "Nitroaromatic compounds: Environmental toxicity, carcinogenicity, mutagenicity, therapy and mechanism." J Appl Toxicol **34**(8): 810-824.
- Krumsiek, J., K. Mittelstrass, K. Do, F. Stückler, J. Ried, J. Adamski, A. Peters, T. Illig, F. Kronenberg, N. Friedrich, et al. (2015). "Gender-specific pathway differences in the human serum metabolome." Metabolomics: 1-19.
- Krumsiek, J., K. Suhre, A. M. Evans, M. W. Mitchell, R. P. Mohny, M. V. Milburn, B. Wagele, W. Romisch-Margl, T. Illig, J. Adamski, et al. (2012). "Mining the unknown: a systems approach to metabolite identification combining genetic and metabolic information." PLoS Genet **8**(10): e1003005.
- Kwon, H., S. Oh, X. Jin, Y. J. An, S. Park (2015). "Cancer metabolomics in basic science perspective." Arch Pharm Res **38**(3): 372-380.

- La Merrill, M., C. Emond, M. J. Kim, J. P. Antignac, B. Le Bizec, K. Clement, L. S. Birnbaum, R. Barouki (2013). "Toxicological function of adipose tissue: focus on persistent organic pollutants." Environ Health Perspect **121**(2): 162-169.
- Lake, A. D., P. Novak, P. Shipkova, N. Aranibar, D. Robertson, M. D. Reily, Z. Lu, L. D. Lehman-McKeeman, N. J. Cherrington (2013). "Decreased hepatotoxic bile acid composition and altered synthesis in progressive human nonalcoholic fatty liver disease." Toxicol Appl Pharmacol **268**(2): 132-140.
- Lan, Q., L. Zhang, X. Tang, M. Shen, M. T. Smith, C. Qiu, Y. Ge, Z. Ji, J. Xiong, J. He, et al. (2010). "Occupational exposure to trichloroethylene is associated with a decline in lymphocyte subsets and soluble CD27 and CD30 markers." Carcinogenesis **31**(9): 1592-1596.
- Lane, K. J., M. Kangsen Scammell, J. I. Levy, C. H. Fuller, R. Parambi, W. Zamore, M. Mwamburi, D. Brugge (2013). "Positional error and time-activity patterns in near-highway proximity studies: an exposure misclassification analysis." Environ Health **12**(1): 75.
- Lane, K. J., J. I. Levy, M. K. Scammell, A. P. Patton, J. L. Durant, M. Mwamburi, W. Zamore, D. Brugge (2015). "Effect of time-activity adjustment on exposure assessment for traffic-related ultrafine particles." J Expo Sci Environ Epidemiol **25**(5): 506-516.
- Lane, K. J., J. I. Levy, M. K. Scammell, J. L. Peters, A. P. Patton, E. Reisner, L. Lowe, W. Zamore, J. L. Durant, D. Brugge (2016). "Association of modeled long-term personal exposure to ultrafine particles with inflammatory and coagulation biomarkers." Environ Int **92-93**: 173-182.
- Lane, R. K., T. Hilsabeck, S. L. Rea (2015). "The role of mitochondrial dysfunction in age-related diseases." Biochim Biophys Acta **1847**(11): 1387-1400.
- Lanki, T., R. Hampel, P. Tiittanen, S. Andrich, R. Beelen, B. Brunekreef, J. Dratva, U. De Faire, K. B. Fuks, B. Hoffmann, et al. (2015). "Air Pollution from Road Traffic and Systemic Inflammation in Adults: A Cross-Sectional Analysis in the European ESCAPE Project." Environ Health Perspect **123**(8): 785-791.
- Lash, L. H., J. W. Fisher, J. C. Lipscomb, J. C. Parker (2000). "Metabolism of trichloroethylene." Environ Health Perspect **108 Suppl 2**: 177-200.
- Lash, L. H., Y. Xu, A. A. Elfarra, R. J. Duescher, J. C. Parker (1995). "Glutathione-dependent metabolism of trichloroethylene in isolated liver and kidney cells of rats and its role in mitochondrial and cellular toxicity." Drug Metab Dispos **23**(8): 846-853.
- Le Cao, K. A., I. Gonzalez, S. Dejean (2009). "integrOmics: an R package to unravel relationships between two omics datasets." Bioinformatics **25**(21): 2855-2856.
- Le Cao, K. A., D. Rossouw, C. Robert-Granie, P. Besse (2008). "A sparse PLS for variable selection when integrating omics data." Stat Appl Genet Mol Biol **7**(1): Article 35.

- Lei, S., L. Zavala-Flores, A. Garcia-Garcia, R. Nandakumar, Y. Huang, N. Madayiputhiya, R. C. Stanton, E. D. Dodds, R. Powers, R. Franco (2014). "Alterations in energy/redox metabolism induced by mitochondrial and environmental toxins: a specific role for glucose-6-phosphate-dehydrogenase and the pentose phosphate pathway in paraquat toxicity." ACS Chem Biol **9**(9): 2032-2048.
- Leiva, A., B. Fuenzalida, F. Westermeier, F. Toledo, C. Salomon, J. Gutierrez, C. Sanhueza, F. Pardo, L. Sobrevia (2015). "Role for Tetrahydrobiopterin in the Fetoplacental Endothelial Dysfunction in Maternal Supraphysiological Hypercholesterolemia." Oxid Med Cell Longev **2015**: 5346327.
- Lertratanangkoon, K., J. M. Scimeca (1993). "Prevention of bromobenzene toxicity by N-acetylmethionine: correlation between toxicity and the impairment in O- and S-methylation of bromothiocatechols." Toxicol Appl Pharmacol **122**(2): 191-199.
- Levi, F., U. Schibler (2007). "Circadian rhythms: mechanisms and therapeutic implications." Annu Rev Pharmacol Toxicol **47**: 593-628.
- Li, S., Y. Park, S. Duraisingham, F. H. Strobel, N. Khan, Q. A. Soltow, D. P. Jones, B. Pulendran (2013). "Predicting network activity from high throughput metabolomics." PLoS Comput Biol **9**(7): e1003123.
- Li, S., A. Pozhitkov, R. A. Ryan, C. S. Manning, N. Brown-Peterson, M. Brouwer (2010). "Constructing a fish metabolic network model." Genome Biol **11**(11): R115.
- Li, T., J. Y. Chiang (2012). "Bile Acid signaling in liver metabolism and diseases." J Lipids **2012**: 754067.
- Libiseller, G., M. Dvorzak, U. Kleb, E. Gander, T. Eisenberg, F. Madeo, S. Neumann, G. Trausinger, F. Sinner, T. Pieber, et al. (2015). "IPO: a tool for automated optimization of XCMS parameters." BMC Bioinformatics **16**: 118.
- Ligthart-Melis, G. C., M. C. van de Poll, P. G. Boelens, C. H. Dejong, N. E. Deutz, P. A. van Leeuwen (2008). "Glutamine is an important precursor for de novo synthesis of arginine in humans." Am J Clin Nutr **87**(5): 1282-1289.
- Lim, S. S., T. Vos, A. D. Flaxman, G. Danaei, K. Shibuya, H. Adair-Rohani, M. Amann, H. R. Anderson, K. G. Andrews, M. Aryee, et al. (2012). "A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010." Lancet **380**(9859): 2224-2260.
- Lin, Z., J. R. Roede, C. He, D. P. Jones, N. M. Filipov (2014). "Short-term oral atrazine exposure alters the plasma metabolome of male C57BL/6 mice and disrupts alpha-linolenate, tryptophan, tyrosine and other major metabolic pathways." Toxicology **326**: 130-141.
- Lindler, L. E. (2015). "Enhancing the Department of Defense's Capability to Identify Environmental Exposures Into the 21st Century." Mil Med **180**(10 Suppl): 5-9.

- Liou, P. J., K. R. Smith (2013). "A discussion of exposure science in the 21st century: a vision and a strategy." Environ Health Perspect **121**(4): 405-409.
- Lipfert, F. W., J. D. Baty, J. P. Miller, R. E. Wyzga (2006). "PM2.5 constituents and related air quality variables as predictors of survival in a cohort of U.S. military veterans." Inhal Toxicol **18**(9): 645-657.
- Liquet, B., K. A. Le Cao, H. Hocini, R. Thiebaut (2012). "A novel approach for biomarker selection and the integration of repeated measures experiments from two assays." BMC Bioinformatics **13**: 325.
- Little, J. L., C. D. Cleven, S. D. Brown (2011). "Identification of "known unknowns" utilizing accurate mass data and chemical abstracts service databases." J Am Soc Mass Spectrom **22**(2): 348-359.
- Liu, K. H., D. I. Walker, K. Uppal, V. Tran, P. Rohrbeck, T. M. Mallon, D. Jones (2016). "High resolution metabolomics assessment of military personnel." Journal of Occupational and Environmental Medicine **Submitted; DoD Biomarkers Supplement**.
- Liu, K. H., D. I. Walker, K. Uppal, V. Tran, P. Rohrbeck, T. M. Mallon, D. P. Jones (2016). "High-Resolution Metabolomics Assessment of Military Personnel: Evaluating Analytical Strategies for Chemical Detection." J Occup Environ Med **58**(8 Suppl 1): S53-61.
- Liu, L. Y., M. A. Schaub, M. Sirota, A. J. Butte (2012). "Sex differences in disease risk from reported genome-wide association study findings." Hum Genet **131**(3): 353-364.
- Liu, M., D. Y. Choi, R. L. Hunter, J. D. Pandya, W. A. Cass, P. G. Sullivan, H. C. Kim, D. M. Gash, G. Bing (2010). "Trichloroethylene induces dopaminergic neurodegeneration in Fisher 344 rats." J Neurochem **112**(3): 773-783.
- Lodovici, M., E. Bigagli (2011). "Oxidative Stress and Air Pollution Exposure." Journal of Toxicology **2011**.
- Loos, M. (2015). nontarget: Detecting, Combining and Filtering Isotope, Adduct and Homologue Series Relations in High-Resolution Mass Spectrometry (HRMS) Data.
- Luiking, Y. C., M. Poeze, G. Ramsay, N. E. Deutz (2009). "Reduced citrulline production in sepsis is related to diminished de novo arginine and nitric oxide production." Am J Clin Nutr **89**(1): 142-152.
- Luneburg, N., R. A. von Holtzen, R. F. Topper, E. Schwedhelm, R. Maas, R. H. Boger (2012). "Symmetric dimethylarginine is a marker of detrimental outcome in the acute phase after ischaemic stroke: role of renal function." Clin Sci (Lond) **122**(3): 105-111.
- Makarov, A., E. Denisov, A. Kholomeev, W. Balschun, O. Lange, K. Strupat, S. Horning (2006). "Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer." Anal Chem **78**(7): 2113-2120.
- Makey, C. M., M. D. McClean, L. E. Braverman, E. N. Pearce, X. M. He, A. Sjodin, J. M. Weinberg, T. F. Webster (2016). "Polybrominated Diphenyl Ether Exposure and Thyroid Function Tests in North American Adults." Environ Health Perspect **124**(4): 420-425.

- Mallon, C. T., M. P. Rohrbeck, M. K. Haines, D. P. Jones, M. Utell, P. K. Hopke, R. P. Phipps, D. I. Walker, T. Thatcher, C. F. Woeller, et al. (2016). "Introduction to Department of Defense Research on Burn Pits, Biomarkers, and Health Outcomes Related to Deployment in Iraq and Afghanistan." *J Occup Environ Med* **58**(8 Suppl 1): S3-S11.
- Marder, M. E., P. Panuwet, R. E. Hunter, P. B. Ryan, M. Marcus, D. B. Barr (2016). "Quantification of Polybrominated and Polychlorinated Biphenyls in Human Matrices by Isotope-Dilution Gas Chromatography–Tandem Mass Spectrometry." *Journal of Analytical Toxicology* **40**(7): 511-518.
- Marshall, A. G., C. L. Hendrickson (2008). "High-resolution mass spectrometers." *Annu Rev Anal Chem (Palo Alto Calif)* **1**: 579-599.
- Mashima, R., T. Nakanishi-Ueda, Y. Yamamoto (2003). "Simultaneous determination of methionine sulfoxide and methionine in blood plasma using gas chromatography-mass spectrometry." *Anal Biochem* **313**(1): 28-33.
- Mathews, C. K., K. E. van Holde, D. R. Appling, S. J. Anthony-Cahill (2012). *Biochemistry*. Toronto, ON, Pearson.
- Matsusue, K., Y. Ishii, N. Ariyoshi, K. Oguri (1999). "A Highly Toxic Coplanar Polychlorinated Biphenyl Compound Suppresses $\Delta 5$ and $\Delta 6$ Desaturase Activities Which Play Key Roles in Arachidonic Acid Synthesis in Rat Liver†." *Chemical Research in Toxicology* **12**(12): 1158-1165.
- McGaw, E. A., K. W. Phinney, M. S. Lowenthal (2010). "Comparison of orthogonal liquid and gas chromatography-mass spectrometry platforms for the determination of amino acid concentrations in human plasma." *J Chromatogr A* **1217**(37): 5822-5831.
- McMichael, A. J., R. E. Woodruff, S. Hales (2006). "Climate change and human health: present and future risks." *Lancet* **367**(9513): 859-869.
- McNiven, E. M., J. B. German, C. M. Slupsky (2011). "Analytical metabolomics: nutritional opportunities for personalized health." *J Nutr Biochem* **22**(11): 995-1002.
- Menni, C., G. Kastenmuller, A. K. Petersen, J. T. Bell, M. Psatha, P. C. Tsai, C. Gieger, H. Schulz, I. Erte, S. John, et al. (2013). "Metabolomic markers reveal novel pathways of ageing and early development in human populations." *Int J Epidemiol* **42**(4): 1111-1119.
- Menni, C., S. J. Metrustry, R. P. Mohny, S. Beevers, B. Barratt, T. D. Spector, F. J. Kelly, A. M. Valdes (2015). "Circulating levels of antioxidant vitamins correlate with better lung function and reduced exposure to ambient pollution." *Am J Respir Crit Care Med* **191**(10): 1203-1207.
- Miller, G. W., D. P. Jones (2014). "The nature of nurture: refining the definition of the exposome." *Toxicol Sci* **137**(1): 1-2.
- Mittelstrass, K., J. S. Ried, Z. Yu, J. Krumsiek, C. Gieger, C. Prehn, W. Roemisch-Margl, A. Polonikov, A. Peters, F. J. Theis, et al. (2011). "Discovery of sexual dimorphisms in metabolic and genetic biomarkers." *PLoS Genet* **7**(8): e1002215.
- Moali, C., J. L. Boucher, M. A. Sari, D. J. Stuehr, D. Mansuy (1998). "Substrate specificity of NO synthases: detailed comparison of L-arginine, homo-L-

- arginine, their N omega-hydroxy derivatives, and N omega-hydroxynor-L-arginine." *Biochemistry* **37**(29): 10453-10460.
- Mohandas, R., R. J. Johnson (2008). "Uric acid levels increase risk for new-onset kidney disease." *J Am Soc Nephrol* **19**(12): 2251-2253.
- Mollace, V., C. Muscoli, E. Masini, S. Cuzzocrea, D. Salvemini (2005). "Modulation of prostaglandin biosynthesis by nitric oxide and nitric oxide donors." *Pharmacol Rev* **57**(2): 217-252.
- Moore, L. E., P. Boffetta, S. Karami, P. Brennan, P. S. Stewart, R. Hung, D. Zaridze, V. Matveev, V. Janout, H. Kollarova, et al. (2010). "Occupational trichloroethylene exposure and renal carcinoma risk: evidence of genetic susceptibility by reductive metabolism gene variants." *Cancer Res* **70**(16): 6527-6536.
- Moran, M. J., J. S. Zogorski, P. J. Squillace (2007). "Chlorinated solvents in groundwater of the United States." *Environ Sci Technol* **41**(1): 74-81.
- Moriwaki, Y., T. Yamamoto, K. Higashino (1999). "Enzymes involved in purine metabolism--a review of histochemical localization and functional implications." *Histol Histopathol* **14**(4): 1321-1340.
- Moskovitz, J., B. S. Berlett, J. M. Poston, E. R. Stadtman (1997). "The yeast peptide-methionine sulfoxide reductase functions as an antioxidant in vivo." *Proc Natl Acad Sci U S A* **94**(18): 9585-9589.
- Murphy, R. A., E. A. Yu, E. D. Ciappio, S. Mehta, M. I. McBurney (2015). "Suboptimal Plasma Long Chain n-3 Concentrations are Common among Adults in the United States, NHANES 2003-2004." *Nutrients* **7**(12): 10282-10289.
- Neal, M. S., J. Zhu, W. G. Foster (2008). "Quantification of benzo[a]pyrene and other PAHs in the serum and follicular fluid of smokers versus non-smokers." *Reprod Toxicol* **25**(1): 100-106.
- Neill, M. A., J. Aschner, F. Barr, M. L. Summar (2009). "Quantitative RT-PCR comparison of the urea and nitric oxide cycle gene transcripts in adult human tissues." *Mol Genet Metab* **97**(2): 121-127.
- Nel, A. E., D. Diaz-Sanchez, D. Ng, T. Hiura, A. Saxon (1998). "Enhancement of allergic inflammation by the interaction between diesel exhaust particles and the immune system." *J Allergy Clin Immunol* **102**(4 Pt 1): 539-554.
- Nemoto, N., S. Takayama (1983). "Effects of unsaturated fatty acids on metabolism of benzo[a]pyrene in an NADPH-fortified rat liver microsomal system." *Carcinogenesis* **4**(10): 1253-1257.
- Neophytou, A. M., J. E. Hart, J. M. Cavallari, T. J. Smith, D. W. Dockery, B. A. Coull, E. Garshick, F. Laden (2013). "Traffic-related exposures and biomarkers of systemic inflammation, endothelial activation and oxidative stress: a panel study in the US trucking industry." *Environ Health* **12**: 105.
- Neophytou, A. M., J. E. Hart, Y. Chang, J. J. Zhang, T. J. Smith, E. Garshick, F. Laden (2014). "Short-term traffic related exposures and biomarkers of nitro-PAH exposure and oxidative DNA damage." *Toxics* **2**(3): 377-390.
- Neujahr, D. C., K. Uppal, S. D. Force, F. Fernandez, C. Lawrence, A. Pickens, R. Bag, C. Lockard, A. D. Kirk, V. Tran, et al. (2014). "Bile acid aspiration

- associated with lung chemical profile linked to other biomarkers of injury after lung transplantation." Am J Transplant **14**(4): 841-848.
- Neveu, V., J. Perez-Jimenez, F. Vos, V. Crespy, L. du Chaffaut, L. Mennen, C. Knox, R. Eisner, J. Cruz, D. Wishart, et al. (2010). "Phenol-Explorer: an online comprehensive database on polyphenol contents in foods." Database (Oxford) **2010**: bap024.
- Newgard, C. B., J. An, J. R. Bain, M. J. Muehlbauer, R. D. Stevens, L. F. Lien, A. M. Haqq, S. H. Shah, M. Arlotto, C. A. Slentz, et al. (2009). "A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance." Cell Metab **9**(4): 311-326.
- Nicholson, J. K., E. Holmes, P. Elliott (2008). "The metabolome-wide association study: a new look at human disease risk factors." J Proteome Res **7**(9): 3637-3638.
- Nicolson, G. L. (2007). "Metabolic syndrome and mitochondrial function: molecular replacement and antioxidant supplements to prevent membrane peroxidation and restore mitochondrial function." J Cell Biochem **100**(6): 1352-1369.
- Niessen, W. M., P. Manini, R. Andreoli (2006). "Matrix effects in quantitative pesticide analysis using liquid chromatography-mass spectrometry." Mass Spectrom Rev **25**(6): 881-899.
- NIOSH (2003). Elemental Carbon (Diesel Particulate) 5040. NIOSH Manual of Analytical Methods, 4th Ed. M. Casinelli and O. C. PF. Cincinnati, OH, U.S Department of Health and Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH).
- Nobori, T., K. Takabayashi, P. Tran, L. Orvis, A. Batova, A. L. Yu, D. A. Carson (1996). "Genomic cloning of methylthioadenosine phosphorylase: a purine metabolic enzyme deficient in multiple different cancers." Proc Natl Acad Sci U S A **93**(12): 6203-6208.
- Noguchi, K., N. Hamadate, T. Matsuzaki, M. Sakanashi, J. Nakasone, T. Uchida, K. Arakaki, H. Kubota, S. Ishiuchi, H. Masuzaki, et al. (2011). "Increasing dihydrobiopterin causes dysfunction of endothelial nitric oxide synthase in rats in vivo." Am J Physiol Heart Circ Physiol **301**(3): H721-729.
- NRC (2000). Strategies to Protect the Health of Deployed U.S. Forces: Detecting, Characterizing, and Documenting Exposures. T. E. McKone, B. M. Huey, E. Downing and L. M. Duffy. Washington (DC), National Academies Press.
- NRC (2009). Science and Decisions: Advancing Risk Assessment. Science and Decisions: Advancing Risk Assessment. Washington (DC).
- NRC (2012). Exposure science in the 21st century : a vision and a strategy. Washington, D.C., National Academies Press.
- O'Brien, K. A., J. L. Griffin, A. J. Murray, L. M. Edwards (2015). "Mitochondrial responses to extreme environments: insights from metabolomics." Extrem Physiol Med **4**: 7.

- O'Connell, S. G., L. D. Kincl, K. A. Anderson (2014). "Silicone wristbands as personal passive samplers." *Environ Sci Technol* **48**(6): 3327-3335.
- O'Sullivan, A., M. J. Gibney, L. Brennan (2011). "Dietary intake patterns are reflected in metabolomic profiles: potential role in dietary assessment studies." *Am J Clin Nutr* **93**(2): 314-321.
- Oeder, S., T. Kanashova, O. Sippula, S. C. Sapcariu, T. Streibel, J. M. Arteaga-Salas, J. Passig, M. Dilger, H. R. Paur, C. Schlager, et al. (2015). "Particulate matter from both heavy fuel oil and diesel fuel shipping emissions show strong biological effects on human lung cells at realistic and comparable in vitro exposure conditions." *PLoS One* **10**(6): e0126536.
- Ohtoshi, T., H. Takizawa, H. Okazaki, S. Kawasaki, N. Takeuchi, K. Ohta, K. Ito (1998). "Diesel exhaust particles stimulate human airway epithelial cells to produce cytokines relevant to airway inflammation in vitro." *J Allergy Clin Immunol* **101**(6 Pt 1): 778-785.
- Osborn, M. P., Y. Park, M. B. Parks, L. G. Burgess, K. Uppal, K. Lee, D. P. Jones, M. A. Brantley, Jr. (2013). "Metabolome-wide association study of neovascular age-related macular degeneration." *PLoS One* **8**(8): e72737.
- Ostro, B., W. Y. Feng, R. Broadwin, S. Green, M. Lipsett (2007). "The effects of components of fine particulate air pollution on mortality in California: results from CALFINE." *Environ Health Perspect* **115**(1): 13-19.
- Padayatty, S. J., A. Katz, Y. Wang, P. Eck, O. Kwon, J. H. Lee, S. Chen, C. Corpe, A. Dutta, S. K. Dutta, et al. (2003). "Vitamin C as an antioxidant: evaluation of its role in disease prevention." *J Am Coll Nutr* **22**(1): 18-35.
- Padberg, I., E. Peter, S. Gonzalez-Maldonado, H. Witt, M. Mueller, T. Weis, B. Bethan, V. Liebenberg, J. Wiemer, H. A. Katus, et al. (2014). "A new metabolomic signature in type-2 diabetes mellitus and its pathophysiology." *PLoS One* **9**(1): e85082.
- Paik, M. J., Y. H. Ahn, P. H. Lee, H. Kang, C. B. Park, S. Choi, G. Lee (2010). "Polyamine patterns in the cerebrospinal fluid of patients with Parkinson's disease and multiple system atrophy." *Clin Chim Acta* **411**(19-20): 1532-1535.
- Park, E. J., J. Roh, M. S. Kang, S. N. Kim, Y. Kim, S. Choi (2011). "Biological responses to diesel exhaust particles (DEPs) depend on the physicochemical properties of the DEPs." *PLoS One* **6**(10): e26749.
- Park, Y., S. B. Kim, B. Wang, R. A. Blanco, N. A. Le, S. Wu, C. J. Accardi, R. W. Alexander, T. R. Ziegler, D. P. Jones (2009). "Individual variation in macronutrient regulation measured by proton magnetic resonance spectroscopy of human plasma." *Am J Physiol Regul Integr Comp Physiol* **297**(1): R202-209.
- Park, Y., K. Lee, T. R. Ziegler, G. S. Martin, G. Hebbar, B. Vidakovic, D. P. Jones (2013). "Multifractal analysis for nutritional assessment." *PLoS One* **8**(8): e69000.
- Park, Y. H., K. Lee, Q. A. Soltow, F. H. Strobel, K. L. Brigham, R. E. Parker, M. E. Wilson, R. L. Sutliff, K. G. Mansfield, L. M. Wachtman, et al. (2012). "High-performance metabolic profiling of plasma from seven mammalian

- species for simultaneous environmental chemical surveillance and bioeffect monitoring." Toxicology **295**(1-3): 47-55.
- Patel, C. J. (2016). "Analytical Complexity in Detection of Gene Variant-by-Environment Exposure Interactions in High-Throughput Genomic and Exposomic Research." Curr Environ Health Rep **3**(1): 64-72.
- Patel, C. J., J. Bhattacharya, A. J. Butte (2010). "An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus." PLoS One **5**(5): e10746.
- Patel, R. M., J. D. Roback, K. Uppal, T. Yu, D. P. Jones, C. D. Josephson (2014). "Metabolomics profile comparisons of irradiated and nonirradiated stored donor red blood cells." Transfusion.
- Patel, R. S., N. Ghasemzadeh, D. J. Eapen, S. Sher, S. Arshad, Y. A. Ko, E. Veledar, H. Samady, A. M. Zafari, L. Sperling, et al. (2016). "Novel Biomarker of Oxidative Stress Is Associated With Risk of Death in Patients With Coronary Artery Disease." Circulation **133**(4): 361-369.
- Patton, A. P., W. Zamore, E. N. Naumova, J. I. Levy, D. Brugge, J. L. Durant (2015). "Transferability and generalizability of regression models of ultrafine particles in urban neighborhoods in the Boston area." Environ Sci Technol **49**(10): 6051-6060.
- Perdue, C. L., A. A. Cost, M. V. Rubertone, L. E. Lindler, S. L. Ludwig (2015). "Description and utilization of the United States department of defense serum repository: a review of published studies, 1985-2012." PLoS One **10**(2): e0114857.
- Perdue, C. L., A. A. Eick-Cost, M. V. Rubertone (2015). "A Brief Description of the Operation of the DoD Serum Repository." Mil Med **180**(10 Suppl): 10-12.
- Peretz, A., E. C. Peck, T. K. Bammler, R. P. Beyer, J. H. Sullivan, C. A. Trenga, S. Srinouanprachnah, F. M. Farin, J. D. Kaufman (2007). "Diesel exhaust inhalation and assessment of peripheral blood mononuclear cell gene transcription effects: an exploratory study of healthy human volunteers." Inhal Toxicol **19**(14): 1107-1119.
- Persky, V., M. Turyk, H. A. Anderson, L. P. Hanrahan, C. Falk, D. N. Steenport, R. Chatterton, Jr., S. Freels, C. Great Lakes (2001). "The effects of PCB exposure and fish consumption on endogenous hormones." Environ Health Perspect **109**(12): 1275-1283.
- Phinney, K. W., G. Ballihaut, M. Bedner, B. S. Benford, J. E. Camara, S. J. Christopher, W. C. Davis, N. G. Dodder, G. Eppe, B. E. Lang, et al. (2013). "Development of a Standard Reference Material for metabolomics research." Anal Chem **85**(24): 11732-11738.
- Piecznik, S. R., J. Neustadt (2007). "Mitochondrial dysfunction and molecular pathways of disease." Exp Mol Pathol **83**(1): 84-92.
- Pleil, J. D., M. A. Stiegel, J. R. Sobus, S. Tabucchi, A. J. Ghio, M. C. Madden (2010). "Cumulative exposure assessment for trace-level polycyclic aromatic hydrocarbons (PAHs) using human blood and plasma analysis." J Chromatogr B Analyt Technol Biomed Life Sci **878**(21): 1753-1760.

- Pope, C. A., 3rd, A. Bhatnagar, J. P. McCracken, W. Abplanalp, D. J. Conklin, T. O'Toole (2016). "Exposure to Fine Particulate Air Pollution Is Associated With Endothelial Injury and Systemic Inflammation." Circ Res **119**(11): 1204-1214.
- Pope, C. A., 3rd, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, G. D. Thurston (2002). "Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution." JAMA **287**(9): 1132-1141.
- Pope, C. A., 3rd, M. Ezzati, D. W. Dockery (2009). "Fine-particulate air pollution and life expectancy in the United States." N Engl J Med **360**(4): 376-386.
- Poyhonen, M. J., U. M. Uusitalo, A. Kari, J. A. Takala, L. A. Alakuijala, T. O. Eloranta (1990). "Urinary excretion of polyamines: importance of circadian rhythm, age, sex, menstrual cycle, weight, and creatinine excretion." Am J Clin Nutr **52**(4): 746-751.
- Pradhan, S. N., A. Das, R. Meena, R. K. Nanda, P. Rajamani (2016). "Biofluid metabotyping of occupationally exposed subjects to air pollution demonstrates high oxidative stress and deregulated amino acid metabolism." Sci Rep **6**: 35972.
- Priyadarshini, M., M. A. Kamal, N. H. Greig, M. Reale, A. M. Abuzenadah, A. G. Chaudhary, G. A. Damanhour (2012). "Alzheimer's disease and type 2 diabetes: exploring the association to obesity and tyrosine hydroxylase." CNS Neurol Disord Drug Targets **11**(4): 482-489.
- Psychogios, N., D. D. Hau, J. Peng, A. C. Guo, R. Mandal, S. Bouatra, I. Sinelnikov, R. Krishnamurthy, R. Eisner, B. Gautam, et al. (2011). "The human serum metabolome." PLoS One **6**(2): e16957.
- Purdue, M. P., B. Bakke, P. Stewart, A. J. De Roos, M. Schenk, C. F. Lynch, L. Bernstein, L. M. Morton, J. R. Cerhan, R. K. Severson, et al. (2011). "A case-control study of occupational exposure to trichloroethylene and non-Hodgkin lymphoma." Environ Health Perspect **119**(2): 232-238.
- Qin, Y., M. Chen, W. Wu, B. Xu, R. Tang, X. Chen, G. Du, C. Lu, J. D. Meeker, Z. Zhou, et al. (2013). "Interactions between urinary 4-tert-octylphenol levels and metabolism enzyme gene variants on idiopathic male infertility." PLoS One **8**(3): e59398.
- Quehenberger, O., A. M. Armando, A. H. Brown, S. B. Milne, D. S. Myers, A. H. Merrill, S. Bandyopadhyay, K. N. Jones, S. Kelly, R. L. Shaner, et al. (2010). "Lipidomics reveals a remarkable diversity of lipids in human plasma." J Lipid Res **51**(11): 3299-3305.
- Quinn, C. L., F. Wania (2012). "Understanding differences in the body burden-age relationships of bioaccumulating contaminants based on population cross sections versus individuals." Environ Health Perspect **120**(4): 554-559.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. **Version 3.1.2, Pumpkin Helmet.**
- Raaschou-Nielsen, O., J. Hansen, J. K. McLaughlin, H. Kolstad, J. M. Christensen, R. E. Tarone, J. H. Olsen (2003). "Cancer risk among workers

- at Danish companies using trichloroethylene: a cohort study." Am J Epidemiol **158**(12): 1182-1192.
- Raghavan, S. A., M. Dikshit (2004). "Vascular regulation by the L-arginine metabolites, nitric oxide and agmatine." Pharmacol Res **49**(5): 397-414.
- Ramesh, A., A. Kumar, M. P. Aramandla, A. M. Nyanda (2014). "Polycyclic aromatic hydrocarbon residues in serum samples of autopsied individuals from Tennessee." Int J Environ Res Public Health **12**(1): 322-334.
- Rappaport, S. M., D. K. Barupal, D. Wishart, P. Vineis, A. Scalbert (2014). "The blood exposome and its role in discovering causes of disease." Environ Health Perspect **122**(8): 769-774.
- Rappaport, S. M., H. Li, H. Grigoryan, W. E. Funk, E. R. Williams (2012). "Adductomics: characterizing exposures to reactive electrophiles." Toxicol Lett **213**(1): 83-90.
- Rappaport, S. M., M. T. Smith (2010). "Epidemiology. Environment and disease risks." Science **330**(6003): 460-461.
- Reuter, S., S. C. Gupta, M. M. Chaturvedi, B. B. Aggarwal (2010). "Oxidative stress, inflammation, and cancer: how are they linked?" Free Radic Biol Med **49**(11): 1603-1616.
- Richardson, J. R., A. Roy, S. L. Shalat, R. T. von Stein, M. M. Hossain, B. Buckley, M. Gearing, A. I. Levey, D. C. German (2014). "Elevated serum pesticide levels and risk for Alzheimer disease." JAMA Neurol **71**(3): 284-290.
- Roback, J. D., C. D. Josephson, E. K. Waller, J. L. Newman, S. Karatela, K. Uppal, D. P. Jones, J. C. Zimring, L. J. Dumont (2014). "Metabolomics of ADSOL (AS-1) red blood cell storage." Transfus Med Rev **28**(2): 41-55.
- Roberts, L. D., A. L. Souza, R. E. Gerszten, C. B. Clish (2012). "Targeted metabolomics." Curr Protoc Mol Biol **Chapter 30**: Unit 30 32 31-24.
- Roca, M., N. Leon, A. Pastor, V. Yusa (2014). "Comprehensive analytical strategy for biomonitoring of pesticides in urine by liquid chromatography-orbitrap high resolution mass spectrometry." J Chromatogr A **1374**: 66-76.
- Roede, J. R., Y. Park, S. Li, F. H. Strobel, D. P. Jones (2012). "Detailed mitochondrial phenotyping by high resolution metabolomics." PLoS One **7**(3): e33020.
- Roede, J. R., K. Uppal, Y. Park, K. Lee, V. Tran, D. Walker, F. H. Strobel, S. L. Rhodes, B. Ritz, D. P. Jones (2013). "Serum metabolomics of slow vs. rapid motor progression Parkinson's disease: a pilot study." PLoS One **8**(10): e77629.
- Roede, J. R., K. Uppal, Y. Park, V. Tran, D. P. Jones (2014). "Transcriptome-metabolome wide association study (TMWAS) of maneb and paraquat neurotoxicity reveals network level interactions in toxicologic mechanism." Toxicology Reports **1**(0): 435-444.
- Roh, T., M. Y. Kwak, E. H. Kwak, D. H. Kim, E. Y. Han, J. Y. Bae, Y. Bang du, D. S. Lim, I. Y. Ahn, D. E. Jang, et al. (2012). "Chemopreventive mechanisms of methionine on inhibition of benzo(a)pyrene-DNA adducts formation in human hepatocellular carcinoma HepG2 cells." Toxicol Lett **208**(3): 232-238.

- Rubertone, M. V., J. F. Brundage (2002). "The Defense Medical Surveillance System and the Department of Defense serum repository: glimpses of the future of public health surveillance." Am J Public Health **92**(12): 1900-1904.
- Ruckerl, R., R. Hampel, S. Breitner, J. Cyrys, U. Kraus, J. Carter, L. Dailey, R. B. Devlin, D. Diaz-Sanchez, W. Koenig, et al. (2014). "Associations between ambient air pollution and blood markers of inflammation and coagulation/fibrinolysis in susceptible populations." Environ Int **70**: 32-49.
- Salihovic, S., A. Ganna, T. Fall, C. D. Broeckling, J. E. Prenni, B. van Bavel, P. M. Lind, E. Ingelsson, L. Lind (2015). "The metabolic fingerprint of p,p'-DDE and HCB exposure in humans." Environ Int **88**: 60-66.
- Sandau, C. D., A. Sjodin, M. D. Davis, J. R. Barr, V. L. Maggio, A. L. Waterman, K. E. Preston, J. L. Preau, Jr., D. B. Barr, L. L. Needham, et al. (2003). "Comprehensive solid-phase extraction method for persistent organic pollutants. Validation and application to the analysis of persistent chlorinated pesticides." Anal Chem **75**(1): 71-77.
- Sanmuganathan, P. S., P. Ghahramani, P. R. Jackson, E. J. Wallis, L. E. Ramsay (2001). "Aspirin for primary prevention of coronary heart disease: safety and absolute benefit related to coronary risk derived from meta-analysis of randomised trials." Heart **85**(3): 265-271.
- Sarnat, S. E., A. Winquist, J. J. Schauer, J. R. Turner, J. A. Sarnat (2015). "Fine particulate matter components and emergency department visits for cardiovascular and respiratory diseases in the St. Louis, Missouri-Illinois, metropolitan area." Environ Health Perspect **123**(5): 437-444.
- Scalbert, A., L. Brennan, O. Fiehn, T. Hankemeier, B. S. Kristal, B. van Ommen, E. Pujos-Guillot, E. Verheij, D. Wishart, S. Wopereis (2009). "Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research." Metabolomics **5**(4): 435-458.
- Scalbert, A., L. Brennan, C. Manach, C. Andres-Lacueva, L. O. Dragsted, J. Draper, S. M. Rappaport, J. J. van der Hoof, D. S. Wishart (2014). "The food metabolome: a window over dietary exposure." Am J Clin Nutr **99**(6): 1286-1308.
- Schauer, J. J. (2003). "Evaluation of elemental carbon as a marker for diesel particulate matter." J Expo Anal Environ Epidemiol **13**(6): 443-453.
- Schepers, E., D. V. Barreto, S. Liabeuf, G. Glorieux, S. Eloit, F. C. Barreto, Z. Massy, R. Vanholder, G. European Uremic Toxin Work (2011). "Symmetric dimethylarginine as a proinflammatory agent in chronic kidney disease." Clin J Am Soc Nephrol **6**(10): 2374-2383.
- Schulte, P. A., F. Perera, eds. (1993). Molecular Epidemiology, Principles and Practices. San Diego CA, Academic Press.
- Schwartz, D. A., J. H. Freedman, E. A. Linney (2004). "Environmental genomics: a key to understanding biology, pathophysiology and disease." Hum Mol Genet **13 Spec No 2**: R217-224.
- Schymanski, E. L., J. Jeon, R. Gulde, K. Fenner, M. Ruff, H. P. Singer, J. Hollender (2014). "Identifying small molecules via high resolution mass

- spectrometry: communicating confidence." Environ Sci Technol **48**(4): 2097-2098.
- Schymanski, E. L., H. P. Singer, P. Longree, M. Loos, M. Ruff, M. A. Stravs, C. Ripolles Vidal, J. Hollender (2014). "Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry." Environ Sci Technol **48**(3): 1811-1818.
- Schymanski, E. L., H. P. Singer, J. Slobodnik, I. M. Ipolyi, P. Oswald, M. Krauss, T. Schulze, P. Haglund, T. Letzel, S. Grosse, et al. (2015). "Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis." Anal Bioanal Chem **407**(21): 6237-6255.
- Scigelova, M., M. Hornshaw, A. Giannakopoulos, A. Makarov (2011). "Fourier transform mass spectrometry." Mol Cell Proteomics **10**(7): M111 009431.
- Scorletti, E., C. D. Byrne (2013). "Omega-3 fatty acids, hepatic lipid metabolism, and nonalcoholic fatty liver disease." Annu Rev Nutr **33**: 231-248.
- Scott, C. S., J. Jinot (2011). "Trichloroethylene and cancer: systematic and quantitative review of epidemiologic evidence for identifying hazards." Int J Environ Res Public Health **8**(11): 4238-4272.
- Segura, R. M., C. Pascual, I. Ocana, J. M. Martinez-Vazquez, E. Ribera, I. Ruiz, M. D. Pelegri (1989). "Adenosine deaminase in body fluids: a useful diagnostic tool in tuberculosis." Clin Biochem **22**(2): 141-148.
- Selley, M. L. (1998). "(E)-4-hydroxy-2-nonenal may be involved in the pathogenesis of Parkinson's disease." Free Radic Biol Med **25**(2): 169-174.
- Shi, Q., G. E. Gibson (2007). "Oxidative stress and transcriptional regulation in Alzheimer disease." Alzheimer Dis Assoc Disord **21**(4): 276-291.
- Shihabi, Z. K., K. S. Oles, C. P. McCormick, J. K. Penry (1992). "Serum and tissue carnitine assay based on dialysis." Clin Chem **38**(8 Pt 1): 1414-1417.
- Sibal, L., S. C. Agarwal, P. D. Home, R. H. Boger (2010). "The Role of Asymmetric Dimethylarginine (ADMA) in Endothelial Dysfunction and Cardiovascular Disease." Curr Cardiol Rev **6**(2): 82-90.
- Simile, M. M., S. Banni, E. Angioni, G. Carta, M. R. De Miglio, M. R. Muroli, D. F. Calvisi, A. Carru, R. M. Pascale, F. Feo (2001). "5'-Methylthioadenosine administration prevents lipid peroxidation and fibrogenesis induced in rat liver by carbon-tetrachloride intoxication." J Hepatol **34**(3): 386-394.
- Simon-Manso, Y., M. S. Lowenthal, L. E. Kilpatrick, M. L. Sampson, K. H. Telu, P. A. Rudnick, W. G. Mallard, D. W. Bearden, T. B. Schock, D. V. Tchekhovskoi, 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." Anal Chem **85**(24): 11725-11731.
- Sirimanne, S. R., J. R. Barr, D. G. Patterson, Jr., L. Ma (1996). "Quantification of polycyclic aromatic hydrocarbons and polychlorinated dibenzo-p-dioxins

- in human serum by combined micelle-mediated extraction (cloud-point extraction) and HPLC." Anal Chem **68**(9): 1556-1560.
- Sjodin, A., L. Y. Wong, R. S. Jones, A. Park, Y. Zhang, C. Hodge, E. Dipietro, C. McClure, W. Turner, L. L. Needham, et al. (2008). "Serum concentrations of polybrominated diphenyl ethers (PBDEs) and polybrominated biphenyl (PBB) in the United States population: 2003-2004." Environ Sci Technol **42**(4): 1377-1384.
- Skoglund, L. A., K. Ingebrigtsen, I. Nafstad, O. Aalen (1986). "Efficacy of paracetamol-esterified methionine versus cysteine or methionine on paracetamol-induced hepatic GSH depletion and plasma ALAT level in mice." Biochem Pharmacol **35**(18): 3071-3075.
- Small, C. M., J. J. DeCaro, M. L. Terrell, C. Dominguez, L. L. Cameron, J. Wirth, M. Marcus (2009). "Maternal exposure to a brominated flame retardant and genitourinary conditions in male offspring." Environ Health Perspect **117**(7): 1175-1179.
- Small, C. M., D. Murray, M. L. Terrell, M. Marcus (2011). "Reproductive outcomes among women exposed to a brominated flame retardant in utero." Arch Environ Occup Health **66**(4): 201-208.
- Small, C. M., M. L. Terrell, L. L. Cameron, J. Wirth, C. Monteilh, M. Marcus (2009). "In Utero Exposure to a Brominated Flame Retardant and Male Growth and Development." International Journal of Child and Adolescent Health **2**.
- Smith, B., C. A. Wong, E. J. Boyko, C. J. Phillips, G. D. Gackstetter, M. A. Ryan, T. C. Smith, T. Millennium Cohort Study (2012). "The effects of exposure to documented open-air burn pits on respiratory health among deployers of the Millennium Cohort Study." J Occup Environ Med **54**(6): 708-716.
- Smith, C. A., G. O'Maille, E. J. Want, C. Qin, S. A. Trauger, T. R. Brandon, D. E. Custodio, R. Abagyan, G. Siuzdak (2005). "METLIN: a metabolite mass spectral database." Ther Drug Monit **27**(6): 747-751.
- Smith, C. A., E. J. Want, G. O'Maille, R. Abagyan, G. Siuzdak (2006). "XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification." Anal Chem **78**(3): 779-787.
- Smith, T., M. S. Ghandour, P. L. Wood (2011). "Detection of N-acetyl methionine in human and murine brain and neuronal and glial derived cell lines." J Neurochem **118**(2): 187-194.
- Smyth, E. M. (2010). "Thromboxane and the thromboxane receptor in cardiovascular disease." Clin Lipidol **5**(2): 209-219.
- Smyth, J. F., D. G. Poplack, B. J. Holiman, B. G. Leventhal, G. Yarbro (1978). "Correlation of adenosine deaminase activity with cell surface markers in acute lymphoblastic leukemia." J Clin Invest **62**(3): 710-712.
- Snodgrass, J. J., W. R. Leonard, L. A. Tarskaia, V. P. Alekseev, V. G. Krivoshapkin (2005). "Basal metabolic rate in the Yakut (Sakha) of Siberia." Am J Hum Biol **17**(2): 155-172.
- Sobinoff, A. P., M. Mahony, B. Nixon, S. D. Roman, E. A. McLaughlin (2011). "Understanding the Villain: DMBA-induced preantral ovotoxicity

- involves selective follicular destruction and primordial follicle activation through PI3K/Akt and mTOR signaling." Toxicol Sci **123**(2): 563-575.
- Sobus, J. R., J. D. Pleil, M. D. McClean, R. F. Herrick, S. M. Rappaport (2010). "Biomarker variance component estimation for exposure surrogate selection and toxicokinetic inference." Toxicol Lett **199**(3): 247-253.
- Soltow, Q., F. H. Strobel, K. G. Mansfield, L. Wachtman, Y. Park, D. P. Jones (2013). "High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome." Metabolomics.
- Soltow, Q., F. H. Strobel, K. G. Mansfield, L. Wachtman, Y. Park, D. P. Jones (2013). "High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome." **Metabolomics**.
- Soltow, Q. A., F. H. Strobel, K. G. Mansfield, L. Wachtman, Y. Park, D. P. Jones (2013). "High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome." Metabolomics **9**(1 Suppl): S132-S143.
- Sonoda, H., H. Takase, Y. Dohi, G. Kimura (2011). "Uric acid levels predict future development of chronic kidney disease." Am J Nephrol **33**(4): 352-357.
- Sorensen, M., B. Daneshvar, M. Hansen, L. O. Dragsted, O. Hertel, L. Knudsen, S. Loft (2003). "Personal PM_{2.5} exposure and markers of oxidative stress in blood." Environ Health Perspect **111**(2): 161-166.
- Sotnikova, T. D., J. M. Beaulieu, S. Espinoza, B. Masri, X. Zhang, A. Salahpour, L. S. Barak, M. G. Caron, R. R. Gainetdinov (2010). "The dopamine metabolite 3-methoxytyramine is a neuromodulator." PLoS One **5**(10): e13452.
- Sperlingova, I., L. Dabrowska, V. Stransky, J. Kucera, M. Tichy (2007). "Human urine certified reference material CZ 6010: creatinine and toluene metabolites (hippuric acid and o-cresol) and a benzene metabolite (phenol)." Anal Bioanal Chem **387**(7): 2419-2424.
- Srivastava, S. K., X. Hu, H. Xia, S. Awasthi, S. Amin, S. V. Singh (1999). "Metabolic fate of glutathione conjugate of benzo[a]pyrene-(7R,8S)-diol (9S,10R)-epoxide in human liver." Arch Biochem Biophys **371**(2): 340-344.
- Steenland, K., M. J. Hein, R. T. Cassinelli, 2nd, M. M. Prince, N. B. Nilsen, E. A. Whelan, M. A. Waters, A. M. Ruder, T. M. Schnorr (2006). "Polychlorinated biphenyls and neurodegenerative disease mortality in an occupational cohort." Epidemiology **17**(1): 8-13.
- Stenner, R. D., J. L. Merdink, D. K. Stevens, D. L. Springer, R. J. Bull (1997). "Enterohepatic recirculation of trichloroethanol glucuronide as a significant source of trichloroacetic acid. Metabolites of trichloroethylene." Drug Metab Dispos **25**(5): 529-535.
- Strickland, P., D. Kang (1999). "Urinary 1-hydroxypyrene and other PAH metabolites as biomarkers of exposure to environmental PAH in air particulate matter." Toxicol Lett **108**(2-3): 191-199.

- Stuhlinger, M. C., R. K. Oka, E. E. Graf, I. Schmolzer, B. M. Upson, O. Kapoor, A. Szuba, M. R. Malinow, T. C. Wascher, O. Pachinger, et al. (2003). "Endothelial dysfunction induced by hyperhomocyst(e)inemia: role of asymmetric dimethylarginine." Circulation **108**(8): 933-938.
- Su, G., J. H. Morris, B. Demchak, G. D. Bader (2014). "Biological network exploration with cytoscape 3." Curr Protoc Bioinformatics **47**: 8 13 11-18 13 24.
- Sun, Y. V. (2014). "The Influences of Genetic and Environmental Factors on Methylome-wide Association Studies for Human Diseases." Curr Genet Med Rep **2**(4): 261-270.
- Surowiec, I., M. Karimpour, S. Gouveia-Figueira, J. Wu, J. Unosson, J. A. Bosson, A. Blomberg, J. Pourazar, T. Sandstrom, A. F. Behndig, et al. (2016). "Multi-platform metabolomics assays for human lung lavage fluids in an air pollution exposure study." Anal Bioanal Chem **408**(17): 4751-4764.
- Swerdlow, R. H. (2011). "Brain aging, Alzheimer's disease, and mitochondria." Biochim Biophys Acta **1812**(12): 1630-1639.
- Szema, A. M., M. C. Peters, K. M. Weissinger, C. A. Gagliano, J. J. Chen (2010). "New-onset asthma among soldiers serving in Iraq and Afghanistan." Allergy Asthma Proc **31**(5): 67-71.
- Szema, A. M., W. Salihi, K. Savary, J. J. Chen (2011). "Respiratory symptoms necessitating spirometry among soldiers with Iraq/Afghanistan war lung injury." J Occup Environ Med **53**(9): 961-965.
- Tabrez, S., N. R. Jabir, S. Shakil, N. H. Greig, Q. Alam, A. M. Abuzenadah, G. A. Damanhour, M. A. Kamal (2012). "A synopsis on the role of tyrosine hydroxylase in Parkinson's disease." CNS Neurol Disord Drug Targets **11**(4): 395-409.
- Tainio, M., J. T. Tuomisto, O. Hanninen, P. Aarnio, K. J. Koistinen, M. J. Jantunen, J. Pekkanen (2005). "Health effects caused by primary fine particulate matter (PM2.5) emitted from buses in the Helsinki metropolitan area, Finland." Risk Anal **25**(1): 151-160.
- Takeda, K., N. Tsukue, S. Yoshida (2004). "Endocrine-disrupting activity of chemicals in diesel exhaust and diesel exhaust particles." Environ Sci **11**(1): 33-45.
- Tanabe, M., M. Kanehisa (2012). "Using the KEGG database resource." Curr Protoc Bioinformatics **Chapter 1**: Unit 1 12.
- Terrell, M. L., A. K. Berzen, C. M. Small, L. L. Cameron, J. J. Wirth, M. Marcus (2009). "A cohort study of the association between secondary sex ratio and parental exposure to polybrominated biphenyl (PBB) and polychlorinated biphenyl (PCB)." Environ Health **8**: 35.
- Terrell, M. L., K. P. Hartnett, H. Lim, J. Wirth, M. Marcus (2015). "Maternal exposure to brominated flame retardants and infant Apgar scores." Chemosphere **118**: 178-186.
- Terrell, M. L., K. A. Rosenblatt, J. Wirth, L. L. Cameron, M. Marcus (2016). "Breast cancer among women in Michigan following exposure to brominated flame retardants." Occup Environ Med **73**(8): 564-567.

- Tice, R. R., C. P. Austin, R. J. Kavlock, J. R. Bucher (2013). "Improving the human hazard characterization of chemicals: a Tox21 update." Environ Health Perspect **121**(7): 756-765.
- Tissot van Patot, M. C., N. J. Serkova, M. Haschke, D. J. Kominsky, R. C. Roach, U. Christians, T. K. Henthorn, B. Honigman (2009). "Enhanced leukocyte HIF-1 α and HIF-1 DNA binding in humans after rapid ascent to 4300 m." Free Radic Biol Med **46**(11): 1551-1557.
- Tollerud, D., J. Balmes, A. Bhatnagar, E. Crouch, F. Dominici, E. Eisen, M. Fox, M. Frampton (2011). Long-Term Health Consequences of Exposure to Burn Pits in Iraq and Afghanistan. Washington DC.
- Tousoulis, D., A. M. Kampoli, C. Tentolouris, N. Papageorgiou, C. Stefanadis (2012). "The role of nitric oxide on endothelial function." Curr Vasc Pharmacol **10**(1): 4-18.
- Trushina, E., M. M. Mielke (2014). "Recent advances in the application of metabolomics to Alzheimer's Disease." Biochim Biophys Acta **1842**(8): 1232-1239.
- Tsai, J.-H., S.-Y. Chang, H.-L. Chiang (2012). "Volatile organic compounds from the exhaust of light-duty diesel vehicles." Atmospheric Environment **61**: 499-506.
- Uppal, K., Y.-M. Go, D. P. Jones (2017). "xMWAS: an R package for data-driven integration and differential network analysis." bioRxiv.
- Uppal, K., Q. A. Soltow, D. E. Promislow, L. M. Wachtman, A. A. Quyyumi, D. P. Jones (2015). "MetabNet: An R Package for Metabolic Association Analysis of High-Resolution Metabolomics Data." Front Bioeng Biotechnol **3**: 87.
- Uppal, K., Q. A. Soltow, F. H. Strobel, W. S. Pittard, K. M. Gernert, T. Yu, D. P. Jones (2013). "xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data." BMC Bioinformatics **14**: 15.
- Uppal, K., D. I. Walker, D. P. Jones (2016). "xMSannotator: an R package for network-based annotation of high-resolution metabolomics data." Anal Chem.
- Uppal, K., D. I. Walker, K. Liu, S. Li, Y. M. Go, D. P. Jones (2016). "Computational metabolomics: a framework for the million metabolome." Chem Res Toxicol.
- van den Berg, R. A., H. C. Hoefsloot, J. A. Westerhuis, A. K. Smilde, M. J. van der Werf (2006). "Centering, scaling, and transformations: improving the biological information content of metabolomics data." BMC Genomics **7**: 142.
- van Leeuwen, D. M., R. W. Gottschalk, G. Schoeters, N. A. van Larebeke, V. Nelen, W. F. Baeyens, J. C. Kleinjans, J. H. van Delft (2008). "Transcriptome analysis in peripheral blood of humans exposed to environmental carcinogens: a promising new biomarker in environmental health studies." Environ Health Perspect **116**(11): 1519-1525.

- Vasdev, S., V. Gill, P. Singal (2007). "Role of advanced glycation end products in hypertension and atherosclerosis: therapeutic implications." Cell Biochem Biophys **49**(1): 48-63.
- Vermeulen, R., L. Zhang, A. Spierenburg, X. Tang, J. V. Bonventre, B. Reiss, M. Shen, M. T. Smith, C. Qiu, Y. Ge, et al. (2012). "Elevated urinary levels of kidney injury molecule-1 among Chinese factory workers exposed to trichloroethylene." Carcinogenesis **33**(8): 1538-1541.
- Vinaixa, M., S. Samino, I. Saez, J. Duran, J. J. Guinovart, O. Yanes (2012). "A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data." Metabolites **2**(4): 775-795.
- Vincent, D. B., G. Jean-Loup, L. Renaud, L. Etienne (2008). "Fast unfolding of communities in large networks." Journal of Statistical Mechanics: Theory and Experiment **2008**(10): P10008.
- Vineis, P. (2015). "Exposomics: mathematics meets biology." Mutagenesis.
- Vineis, P., F. Perera (2007). "Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old." Cancer Epidemiol Biomarkers Prev **16**(10): 1954-1965.
- Vineis, P., P. Schulte, A. J. McMichael (2001). "Misconceptions about the use of genetic tests in populations." Lancet **357**(9257): 709-712.
- Vineis, P., K. van Veldhoven, M. Chadeau-Hyam, T. J. Athersuch (2013). "Advancing the application of omics-based biomarkers in environmental epidemiology." Environ Mol Mutagen **54**(7): 461-467.
- Vogesser, M., C. Seger (2010). "Pitfalls associated with the use of liquid chromatography-tandem mass spectrometry in the clinical laboratory." Clin Chem **56**(8): 1234-1244.
- Vrijens, K., V. Bollati, T. S. Nawrot (2015). "MicroRNAs as potential signatures of environmental exposure or effect: a systematic review." Environ Health Perspect **123**(5): 399-411.
- Vuckovic, D. (2012). "Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry." Anal Bioanal Chem **403**(6): 1523-1548.
- Walker, D. I., Y. Go, K. Liu, K. Pennell, D. Jones, Eds. (2016). Population Screening for Biological and Environmental Properties of the Human Metabolic Phenotype: Implications for Personalized Medicine. Metabolic Phenotyping in Personalized and Public Healthcare, Elsevier.
- Walker, D. I., K. D. Pennell, K. Uppal, X. Xia, P. K. Hopke, M. J. Utell, R. P. Phipps, P. J. Sime, P. Rohrbeck, C. T. Mallon, et al. (2016). "Pilot Metabolome-Wide Association Study of Benzo(a)pyrene in Serum From Military Personnel." J Occup Environ Med **58**(8 Suppl 1): S44-52.
- Walker, D. I., K. Uppal, L. Zhang, R. Vermeulen, M. Smith, W. Hu, M. P. Purdue, X. Tang, B. Reiss, S. Kim, et al. (2016). "High-resolution metabolomics of occupational exposure to trichloroethylene." Int J Epidemiol **45**(5): 1517-1527.
- Walsh, M. C., L. Brennan, E. Pujos-Guillot, J. L. Sebedio, A. Scalbert, A. Fagan, D. G. Higgins, M. J. Gibney (2007). "Influence of acute phytochemical

- intake on human urinary metabolomic profiles." *Am J Clin Nutr* **86**(6): 1687-1693.
- Walsh, M. C., G. A. McLoughlin, H. M. Roche, J. F. Ferguson, C. A. Drevon, W. H. Saris, J. A. Lovegrove, U. Riserus, J. Lopez-Miranda, C. Defoort, et al. (2014). "Impact of geographical region on urinary metabolomic and plasma fatty acid profiles in subjects with the metabolic syndrome across Europe: the LIPGENE study." *Br J Nutr* **111**(3): 424-431.
- Wambaugh, J. F., R. W. Setzer, D. M. Reif, S. Gangwal, J. Mitchell-Blackwood, J. A. Arnot, O. Joliet, A. Frame, J. Rabinowitz, T. B. Knudsen, et al. (2013). "High-throughput models for exposure-based chemical prioritization in the ExpoCast project." *Environ Sci Technol* **47**(15): 8479-8488.
- Wambaugh, J. F., A. Wang, K. L. Dionisio, A. Frame, P. Egeghy, R. Judson, R. W. Setzer (2014). "High throughput heuristics for prioritizing human exposure to environmental chemicals." *Environ Sci Technol* **48**(21): 12760-12767.
- Wang, X., J. Zhang, Q. Huang, A. Alamdar, M. Tian, L. Liu, H. Shen (2014). "Serum metabolomics analysis reveals impaired lipid metabolism in rats after oral exposure to benzo(a)pyrene." *Mol Biosyst*.
- Wang, Z., E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y. M. Chung, et al. (2011). "Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease." *Nature* **472**(7341): 57-63.
- Wang, Z., Y. Zheng, B. Zhao, Y. Zhang, Z. Liu, J. Xu, Y. Chen, Z. Yang, F. Wang, H. Wang, et al. (2015). "Human metabolic responses to chronic environmental polycyclic aromatic hydrocarbon exposure by a metabolomic approach." *J Proteome Res* **14**(6): 2583-2593.
- Ware, L. B., J. A. Magarik, N. Wickersham, G. Cunningham, T. W. Rice, B. W. Christman, A. P. Wheeler, G. R. Bernard, M. L. Summar (2013). "Low plasma citrulline levels are associated with acute respiratory distress syndrome in patients with severe sepsis." *Crit Care* **17**(1): R10.
- Weber, D., L. Milkovic, S. J. Bennett, H. R. Griffiths, N. Zarkovic, T. Grune (2013). "Measurement of HNE-protein adducts in human plasma and serum by ELISA-Comparison of two primary antibodies." *Redox Biol* **1**: 226-233.
- Wei, Y., Z. Wang, C. Y. Chang, T. Fan, L. Su, F. Chen, D. C. Christiani (2013). "Global metabolomic profiling reveals an association of metal fume exposure and plasma unsaturated fatty acids." *PLoS One* **8**(10): e77413.
- Whittle, B. J., S. Moncada, J. R. Vane (1978). "Comparison of the effects of prostacyclin (PGI₂), prostaglandin E₁ and D₂ on platelet aggregation in different species." *Prostaglandins* **16**(3): 373-388.
- Wild, C. P. (2005). "Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology." *Cancer Epidemiol Biomarkers Prev* **14**(8): 1847-1850.

- Wild, C. P. (2012). "The exposome: from concept to utility." Int J Epidemiol **41**(1): 24-32.
- Wild, C. P., A. Scalbert, Z. Herceg (2013). "Measuring the exposome: a powerful basis for evaluating environmental exposures and cancer risk." Environ Mol Mutagen **54**(7): 480-499.
- Wishart, D., D. Arndt, A. Pon, T. Sajed, A. C. Guo, Y. Djoumbou, C. Knox, M. Wilson, Y. Liang, J. Grant, et al. (2015). "T3DB: the toxic exposome database." Nucleic Acids Res **43**(Database issue): D928-934.
- Wishart, D. S., T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, et al. (2013). "HMDB 3.0--The Human Metabolome Database in 2013." Nucleic Acids Res **41**(Database issue): D801-807.
- Wishart, D. S., C. Knox, A. C. Guo, R. Eisner, N. Young, B. Gautam, D. D. Hau, N. Psychogios, E. Dong, S. Bouatra, et al. (2009). "HMDB: a knowledgebase for the human metabolome." Nucleic Acids Res **37**(Database issue): D603-610.
- Wold, S., M. Sjöström, L. Eriksson (2001). "PLS-regression: a basic tool of chemometrics." Chemometrics and Intelligent Laboratory Systems **58**(2): 109-130.
- Wurtz, P., A. S. Havulinna, P. Soininen, T. Tynkkynen, D. Prieto-Merino, T. Tillin, A. Ghorbani, A. Artati, Q. Wang, M. Tiainen, et al. (2015). "Metabolite profiling and cardiovascular event risk: a prospective study of 3 population-based cohorts." Circulation **131**(9): 774-785.
- Xia, J., D. S. Wishart (2010). "MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data." Nucleic Acids Res **38**(Web Server issue): W71-77.
- Xia, X., A. Carroll-Haddad, N. Brown, M. J. Utell, T. M. Mallon, P. K. Hopke (2016). "Polycyclic Aromatic Hydrocarbons and Polychlorinated Dibenzop-Dioxins/Dibenzofurans in Microliter Samples of Human Serum as Exposure Indicators." Journal of Occupational and Environmental Medicine Biomarkers Supplement.
- Yan, Z., N. Maher, R. Torres, C. Cotto, B. Hastings, M. Dasgupta, R. Hyman, N. Huebert, G. W. Caldwell (2008). "Isobaric metabolite interferences and the requirement for close examination of raw data in addition to stringent chromatographic separations in liquid chromatography/tandem mass spectrometric analysis of drugs in biological matrix." Rapid Commun Mass Spectrom **22**(13): 2021-2028.
- Yanosky, J. D., D. L. MacIntosh (2001). "A comparison of four gravimetric fine particle sampling methods." J Air Waste Manag Assoc **51**(6): 878-884.
- Yin, H., L. Xu, N. A. Porter (2011). "Free radical lipid peroxidation: mechanisms and analysis." Chem Rev **111**(10): 5944-5972.
- Yu, T., D. P. Jones (2014). "Improving peak detection in high-resolution LC/MS metabolomics data using preexisting knowledge and machine learning approach." Bioinformatics **30**(20): 2941-2948.

- Yu, T., Y. Park, J. M. Johnson, D. P. Jones (2009). "apLCMS--adaptive processing of high-resolution LC/MS data." Bioinformatics **25**(15): 1930-1936.
- Yu, T., Y. Park, S. Li, D. P. Jones (2013). "Hybrid feature detection and information accumulation using high-resolution LC-MS metabolomics data." J Proteome Res **12**(3): 1419-1427.
- Zahedi, K., F. Huttinger, R. Morrison, T. Murray-Stewart, R. A. Casero, K. I. Strauss (2010). "Polyamine catabolism is enhanced after traumatic brain injury." J Neurotrauma **27**(3): 515-525.
- Zaheer, F., J. T. Slevin (2011). "Trichloroethylene and Parkinson disease." Neurol Clin **29**(3): 657-665.
- Zhang, J., H. Shen, W. Xu, Y. Xia, D. B. Barr, X. Mu, X. Wang, L. Liu, Q. Huang, M. Tian (2014). "Urinary metabolomics revealed arsenic internal dose-related metabolic alterations: a proof-of-concept study in a Chinese male cohort." Environ Sci Technol **48**(20): 12265-12274.
- Zhang, L., B. A. Bassig, J. L. Mora, R. Vermeulen, Y. Ge, J. D. Curry, W. Hu, M. Shen, C. Qiu, Z. Ji, et al. (2013). "Alterations in serum immunoglobulin levels in workers occupationally exposed to trichloroethylene." Carcinogenesis **34**(4): 799-802.
- Zhang, L., F. L. Ye, T. Chen, Y. Mei, S. Z. Song (2011). "Trans, trans-muconic acid as a biomarker of occupational exposure to high-level benzene in China." J Occup Environ Med **53**(10): 1194-1198.
- Zhao, J., Y. Zhu, N. Hyun, D. Zeng, K. Uppal, V. T. Tran, T. Yu, D. Jones, J. He, E. T. Lee, et al. (2015). "Novel metabolic markers for the risk of diabetes development in American Indians." Diabetes Care **38**(2): 220-227.
- Zhao, J., Y. Zhu, K. Uppal, V. T. Tran, T. Yu, J. Lin, T. Matsuguchi, E. Blackburn, D. Jones, E. T. Lee, et al. (2014). "Metabolic profiles of biological aging in American Indians: the Strong Heart Family Study." Aging (Albany NY) **6**(3): 176-186.
- Zhong, J., O. Karlsson, G. Wang, J. Li, Y. Guo, X. Lin, M. Zemplyeni, M. Sanchez-Guerra, L. Trevisi, B. Urch, et al. (2017). "B vitamins attenuate the epigenetic effects of ambient fine particles in a pilot human intervention trial." Proc Natl Acad Sci U S A **114**(13): 3503-3508.
- Zhu, M., L. Ma, D. Zhang, K. Ray, W. Zhao, W. G. Humphreys, G. Skiles, M. Sanders, H. Zhang (2006). "Detection and characterization of metabolites in biological matrices using mass defect filtering of liquid chromatography/high resolution mass spectrometry data." Drug Metab Dispos **34**(10): 1722-1733.
- Zuurbier, M., G. Hoek, M. Oldenwening, K. Meliefste, P. van den Hazel, B. Brunekreef (2011). "Respiratory effects of commuters' exposure to air pollution in traffic." Epidemiology **22**(2): 219-227.