UCLA
UCLA Electronic Theses and Dissertations

Title

Metabolomic and Epigenomic Assessment of Air Pollution and Pesticides Exposure in California

Permalink

https://escholarship.org/uc/item/1h5654mk

Author

Yan, Qi

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Metabolomic and Epigenomic Assessment of Air Pollution and Pesticides Exposure in California

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Epidemiology

by

Qi Yan

2021

ABSTRACT OF THE DISSERTATION

Metabolomic and Epigenomic Assessment of Air Pollution and Pesticides Exposure in California

by

Qi Yan

Doctor of Philosophy in Epidemiology

University of California, Los Angeles, 2021

Professor Beate R. Ritz, Chair

Air pollutants and pesticides, two major and widespread environmental exposures, have been shown to increase the risk of various health outcomes such as birth outcomes, cardiovascular diseases, neurodevelopmental disorders, and neurodegenerative diseases. The cumulative toxic effects of chronic air pollution and pesticide exposures are mediated through various biological processes, such as oxidative stress and inflammation, epigenetic alterations, mitochondrial dysfunction, altered intercellular communication, altered microbiome communities, and impaired nervous system function. Mapping the exogeneous exposures to endogenous responses facilitates the ability to identify the nongenetic drivers of health and disease, and ultimately lead to more effective and efficient disease prevention. In this dissertation, we investigate metabolomic and epigenomic signals associated with traffic-related air pollution and chronic ambient pesticide exposures to better understand pathogenic mechanisms underlying chronic health effects.

We first investigated perturbations of the maternal serum metabolome in response to traffic-related air pollution. We retrieved stored mid-pregnancy serum samples from 160 mothers who lived in the Central Valley of California known for high air particulate levels. We estimated prenatal traffic-related air pollution exposure (carbon monoxide, nitric oxides, and particulate matter less than 2.5 microns) during the first trimester using the California Line Source Dispersion Model, version 4 (CALINE4) based on residential addresses recorded at birth. We used liquid chromatography-high resolution mass spectrometry (LC-HRMS) to obtain untargeted metabolic profiles. Multivariate and univariate analyses were conducted to select metabolomic features associated with air pollution exposure and pathway analyses identified biologic pathways related to air pollution exposure. In total, we identified 432 metabolomic features that discriminated between the high (n=98) and low air pollution exposed group (n=62). Pathway enrichment analysis for features associated with air pollution indicated that in maternal serum oxidative stress and inflammation related pathways were altered, including linoleate, leukotriene, and prostaglandin pathways.

We also investigated perturbations in the serum metabolome for organophosphates (OPs), organochlorines (OCs), and pyrethroids (PYRs), all pesticides that have been widely used in the agricultural regions of the Central Valley of California. We conducted high-resolution metabolomic profiling of serum samples from 176 older adults and estimated each participant's chronic ambient pesticide exposure (from 1974 to year of blood draw) to OPs, OCs, and PYRs with a geographic information system (GIS)-based model. We identified metabolites and metabolic pathways associated with one or multiple pesticide classes, including mitochondrial energy metabolism, fatty acid and lipid metabolism, and amino acid metabolism. Utilizing an

integrative network approach, we found that disturbances in the fatty acid beta-oxidation pathway are shared across all three pesticide classes.

Lastly, to investigate the systematic biological responses to chronic ambient OP exposure, we cross-sectionally integrated the methylome and metabolome measured in blood samples collected from older adults living in the Central Valley of California (n=176). Similarly, cumulative OP exposure over a ten-year period was estimated using a GIS-based model. The single-omics analyses showed both epigenomic and metabolomic signatures of OP as being enriched in the glycosphingolipid (GSL) biosynthesis pathway. Besides this common pathway, the metabolome and epigenome also exhibited distinct responses to OPs, with differently methylated CpGs being involved in intracellular membrane transport, cell adhesion, and carcinogenesis; and OP-related metabolites being involved in aromatic amino acids metabolism, neurotransmitter precursors, oxidative stress, and mitochondria function. Moreover, we illustrate possible interactions between these two molecular layers through metabolic processes and nutrient-sensing pathways when we integrated the epigenomic and metabolomic signals.

In summary, our studies linked macro-level population exposures to micro-level biologic responses affecting the metabolome and the epigenome, to provide insight into molecular mechanisms underlying the chronic toxicity of air pollution and pesticide exposure and their interaction with health outcomes, thereby providing much-needed information for the prediction, prevention, and treatment of a variety of diseases.

The dissertation of Qi Yan is approved.

Stefan Horvath

Alexandra M. Binder

Roch Nianogo

Beate R. Ritz, Committee Chair

University of California, Los Angeles

2021

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

12-oxo-LTB4	12-oxo Leukotriene B4
13-HODE	13-Hydroxyoctadecadienoic Acid
13-oxo-ODE	13-oxo-9,11-Octadecadienoic Acid
20-OH-LTE4	20-Hydroxy-Leukotriene E4
9-HPODE	13-Hydroperoxyoctadecadenoic Acid
AAA	Aromatic Amino Acids
AChE	Acetylcholinesterase
AD	Alzheimer's Disease
BBMRI-NL	Biobanking And Biomolecular Infrastructure of The Netherlands
Bicor	Biweight Midcorrelation
BIOS	Biobank-Based Integrative Omics Study
C18neg	C18 Column Coupled with Negative Ionization Mode
CA-PUR	California State-Mandated Pesticide Use Reports
CDPR	California Department of Pesticide Regulation
CMP	Cytidine-5'-Monophosphate
CO	Carbon Monoxide
COX	Cyclooxygenase
CPSP	California Prenatal Screening Program
CTD	The Comparative Toxicogenomics Database
CV	Coefficients of Variation
DDT	Dichlorodiphenyltrichloroethane

DHS	DNAse Hypersensitive Sites
eQTMs	Expression Quantitative Trait Methylation Loci
FDR	False Discovery Rate
GC-MS	Gas Chromatography-Mass Spectrometry
GDP	Guanosine 5'-Diphosphate
GIS	Geographic Information System
GM3	Monosialodihexosylganglioside
GO	Gene Ontology
GSL	Glycosphingolipid
HCB	Hexachlorobenzene
HILIC	Hydrophilic Interaction
HILICpos	HILIC Column Coupled with Positive Ionization Mode
HMDB	Human Metabolome Database
HPETE	Hydroperoxyeicosatetraenoic Acid
HRM	High-Resolution Metabolomics
KEGG	Kyoto Encyclopedia of Genes and Genomes
KORA	The Cooperative Health Research in The Region of Augsburg
LC-MS	Liquid Chromatography-Mass Spectrometry
LOX	Lipoxygenase
LPCs	Lysophosphatidylcholines
m/z	Mass-To-Charge Ratio

NAS	Normative Aging Study
NMR	Nuclear Magnetic Resonance
NO2	Nitrogen Dioxide
NOx	Nitrogen Oxides
O3	Ozone
OCs	Organochlorines
OPs	Organophosphates
p,p'-DDE	Dichlorodiphenyldichloroethylene
PAN	Pesticide Action Network
PBBs	Polybrominated Biphenyls
PCA	Principal Component Analysis
PD	Parkinson's Disease
PEG	The Parkinson's Environment and Gene Study
PFAS	Poly and Perfluoroalkyl Substances
PL	Phospholipid
PLA2	Phospholipase A2
PLS-DA	Partial Least Squares Discriminant Analysis
PM10	Particulate Matter Less Than 10 Micrometers
PM2.5	Particulate Matter Less Than 2.5 Micrometers
POPs	Persistent Organic Pollutants
PUFA	Polyunsaturated Fatty Acids
PYRs	Pyrethroids

QC	Quality Control
ROS	Reactive Oxygen Species
SVM	Support Vector Machine
VIP	Variable Importance in Projection

ACKNOWLEDGEMENTS

I know that I will look back on the years spent working towards this dissertation with great fondness, and there are a great number of people that must be thanked for making this trip and its outcome so thoroughly enjoyable.

First and foremost, I would like to thank my dissertation chair and Ph.D. advisor, Dr. Beate Ritz. I am very grateful for the opportunity to work in a highly synergistic and collaborative environment that Dr. Ritz has created. Dr. Ritz led me on the path of epidemiology almost six years ago and has provided me with tremendous support and encouragement ever since. She has always maintained faith in me even when my own faith waned. Dr. Ritz has served as an excellent role model for me and showed me what a scientist needs: integrity, courage, determination, humility, and a desire to explore.

I would like to thank the members of my dissertation committee. Dr. Steve Horvath, who radiates such an incredible enthusiasm for science that one cannot help but be motivated by it. In those moments of self-doubt, I often think of Dr. Horvath and his discovery of the "clock", which gives me the courage to keep going. Dr Alexandra Binder, for her incredibly timely and thorough guidance. Dr. Roch Nianogo, for all the inspiration in the field of Causal Inference. I would like to thank my other mentors - Dr. Zeyan Liew, who is not only my mentor, but also my true friend. I can always share my struggles with him and expect his generous support. Dr. Kimberly Paul, for her mentorship and her tremendous assistance. I would also like to express my sincere gratitude to all the collaborators, Dr. Dean Jones, Dr. Karan Uppal, and Dr. Douglas Walker, for their full support and trust in my work. I feel incredibly lucky to have had such amazing scientists and dedicated mentors to look up to and learn from.

Chapter 2 is a version of:

Yan, Q.; Liew, Z.; Uppal, K.; Cui, X.; Ling, C.; Heck, J. E.; von Ehrenstein, O. S.; Wu, J.; Walker, D. I.; Jones, D. P., Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. *Environ Int* **2019,** *130*, 104872.

Chapter 3 is a version of:

Yan, Q.; Paul, K. C.; Walker, D. I.; Furlong, M. A.; Del Rosario, I.; Yu, Y.; Zhang, K.; Cockburn, M. G.; Jones, D. P.; Ritz, B. R., High-Resolution Metabolomic Assessment of Pesticide Exposure in Central Valley, California. *Chemical Research in Toxicology* **2021**.

VITA

Education:

2015 B.Sc. Life Sciences, Peking University, Beijing, China.

Publications:

Yan, Q.; Paul, K. C.; Walker, D. I.; Furlong, M. A.; Del Rosario, I.; Yu, Y.; Zhang, K.; Cockburn, M. G.; Jones, D. P.; Ritz, B. R., High-Resolution Metabolomic Assessment of Pesticide Exposure in Central Valley, California. *Chemical Research in Toxicology* **2021**.

Kim, J. H.; **Yan, Q**.; Uppal, K.; Cui, X.; Ling, C.; Walker, D. I.; Heck, J. E.; von Ehrenstein, O. S.; Jones, D. P.; Ritz, B., Metabolomics analysis of maternal serum exposed to high air pollution during pregnancy and risk of autism spectrum disorder in offspring. *Environ Res* **2021**, 110823.

Yan, Q.; Paul, K. C.; Lu, A. T.; Kusters, C.; Binder, A. M.; Horvath, S.; Ritz, B., Epigenetic mutation load is weakly correlated with epigenetic age acceleration. *Aging* **2020**, *12*.

von Ehrenstein, O. S.; Cui, X.; **Yan, Q**.; Aralis, H.; Ritz, B., von Ehrenstein et. al respond to "Sibling comparison designs, are they worth the effort?". *American Journal of Epidemiology* **2020**.

von Ehrenstein, O. S.; Cui, X.; **Yan, Q**.; Aralis, H.; Ritz, B., Maternal Prenatal Smoking and Autism Spectrum Disorder in Offspring: a California Statewide Cohort and Sibling Study. *American Journal of Epidemiology* **2020**.

Ritz, B.; **Yan, Q**.; Uppal, K.; Liew, Z.; Cui, X.; Ling, C.; Inoue, K.; von Ehrenstein, O.; Walker, D. I.; Jones, D. P., Untargeted Metabolomics Screen of Mid-pregnancy Maternal Serum and Autism in Offspring. *Autism Res* **2020**.

Inoue, K.; **Yan, Q. (Co-first author)**; Arah, O. A.; Paul, K.; Walker, D. I.; Jones, D. P.; Ritz, B., Air Pollution and Adverse Pregnancy and Birth Outcomes: Mediation Analysis Using Metabolomic Profiles. *Current Environmental Health Reports* **2020,** *7* (3), 231-242.

Inoue, K.; Mayeda, E. R.; Paul, K. C.; Shih, I.-F.; **Yan, Q**.; Yu, Y.; Haan, M.; Ritz, B. R., The Association of Physical Activity with Cardiovascular Events and Mortality Mediated by Diabetes in Older Mexican Americans. *American Journal of Epidemiology* **2020**.

Furlong, M. A.; Paul, K. C.; **Yan, Q**.; Chuang, Y.-H.; Cockburn, M. G.; Bronstein, J. M.; Horvath, S.; Ritz, B., An epigenome-wide association study of ambient pyrethroid pesticide exposures in California's central valley. *International Journal of Hygiene and Environmental Health* **2020**, *229*, 113569.

Yan, Q.; Liew, Z.; Uppal, K.; Cui, X.; Ling, C.; Heck, J. E.; von Ehrenstein, O. S.; Wu, J.; Walker, D. I.; Jones, D. P., Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. *Environ Int* **2019**, *130*, 104872.

Tang, M.; Xu, C.; Chen, K.; **Yan, Q**.; Mao, W.; Liu, W.; Ritz, B., Hexachlorocyclohexane exposure alters the microbiome of colostrum in Chinese breastfeeding mothers. *Environmental Pollution* **2019,** *254*, 112900.

Ritz, B.; Liew, Z.; **Yan, Q**.; Cui, X.; Virk, J.; Ketzel, M.; Raaschou-Nielsen, O., Air pollution and Autism in Denmark. *Environ Epidemiol* **2018,** *2* (4).

1 Background and Introduction

1.1 Dissertation objectives

The aim of this dissertation is to investigate metabolomic and epigenomic signals associated with chronic ambient air pollution and pesticide exposures using population-based study design. The cumulative deleterious effects of these environmental exposures are mediated through a variety of biochemical mechanisms that can be comprehensively characterized with the rapid advances of omics technologies. In this dissertation, we hope to identify novel molecular fingerprints that reflect preclinical effects of chronic air pollution and pesticides among healthy population, which may offer opportunities for effective targets of disease prevention and control.

Briefly, the first project examines changes in specific metabolomic profiles associated with air pollution exposure in pregnant women lived in Central Valley, California, and investigates potential biological mechanisms related to adverse health effects from air pollution exposure during pregnancy. The second project examines perturbations in the serum metabolome in response to chronic ambient organophosphates (OPs), organochlorines (OCs), and pyrethroids (PYRs) pesticide exposures typical for older adults living in heavily agricultural regions of Central California, and obtains a holistic view of the mutual influence between metabolome and various pesticides. The third project examines epigenetic as well as metabolomic changes induced by chronic low-level OP exposure, and integrates the information from OP-related epigenomic and metabolomic signals to improve our understanding of molecular mechanisms involved in the human response to chronic OP exposures.

1.2 Air pollution and health background

Air pollution is any substance or mixture in the air that may have adverse effects on humans, animals, vegetation or materials (Kampa and Castanas 2008). In our study, we focus on traffic

related air pollutants including nitrogen oxides (NOx), carbon monoxide (CO), and particulate matter less than 2.5 microns (PM2.5). The main anthropogenic sources of NOx are mobile and stationary combustion sources, and the major source of CO is also road transport as it is a product of incomplete combustion (Kampa and Castanas 2008; Katsouyanni 2003). PM2.5, on the other hand, is a complex and varying mixture of particles suspended in the air with aerodynamic diameter smaller than 2.5 μm. The major sources of particulate pollution include factories, power plants, refuse incinerators, transportation related combustion (motor vehicles, airplanes, shipping, trains), construction activity, fires, and natural windblown dust.

Our study focuses on air pollution exposures in the California's Central Valley area. The California's Central Valley is known for having some of the nation's worst air quality, failing to meet federal health standards for both ozone and particulate pollution due to mountain ranges trapping air pollutants originating from heavy truck traffic, diesel-burning locomotives, tractors and irrigation pumps, and wood-burning. We previously have used CA's ambient air monitoring network extensively in our pregnancy outcome and childhood cancer studies (Heck et al. 2013; Ritz et al. 2007).

Exposure to traffic-related air pollution has been shown to be associated with a broad range of health effects such as genetic damage, inflammation, and oxidative stress response (Katsouyanni 2003; Loomis et al. 2013). Air pollution can potentially damage various organs and systems in the human body including the respiratory, cardiovascular, and nervous system (Kampa and Castanas 2008). Also, it is important to note that the pregnancy period is critical for human organ development and may be exceptionally vulnerable to air pollutants that can affect both maternal health and the development of fetus (Schell et al. 2006). Numerous epidemiological studies have linked maternal exposure to air pollution during pregnancy with pregnancy

complications and adverse birth outcomes such as preeclampsia, gestational diabetes, preterm birth and low birth weight (Dadvand et al. 2013; Eze et al. 2015; Pedersen et al. 2013; Pedersen et al. 2014; Ritz and Yu 1999; Ritz et al. 2000; Stieb et al. 2012). For example, studies have shown that exposure to high concentrations of carbon monoxide during the last trimester of pregnancy may increase the risk of being of low weight for term birth and that exposure to carbon monoxide and particulate matter <10 μm in aerodynamic diameter (PM10) either shortly after conception or before birth may trigger preterm birth (Ritz and Yu 1999; Ritz et al. 2000). Another study in Southern California linked more than 80,000 birth records with air pollution estimations during pregnancy and observed adverse effects of traffic-related nitrogen oxides (NOx) and PM2.5 exposures on preeclampsia (Wu et al. 2009).

In addition, several studies found prenatal traffic-related air pollution exposure to be associated with fetal development especially neurodevelopment disorders such as autism in offspring (Becerra et al. 2013; Glinianaia et al. 2004; Guxens et al. 2014; Suades-Gonzalez et al. 2015). For example, one case-control study in California found associations between autism and prenatal traffic-related air pollution exposure (Becerra et al. 2013). In another study conducted in LA, air toxics from traffic and industrial sources, including benzene, 1,3-butadiene, and lead, have been shown to increase autism risk (von Ehrenstein et al. 2014).

However, there is still not much data available to elucidate endogenous responses to adverse health effects of air pollution, even though it has been hypothesized that endocrine disruption, oxidative stress, inflammatory response, and DNA damage are major contributors (Hougaard et al. 2008; Kelly 2003; Risom et al. 2005). It is of great interests to investigate the mediated biological mechanisms of chronic air pollution exposure especially at pre-clinical stages, which may help prevent the development of related diseases.

1.3 Pesticide and health background

Pesticides include all classes of chemicals used to kill or repel insects, fungi, vegetation, and rodents (Sanborn et al. 2007). California leads the US in agricultural production, which is mainly concentrated in its Central Valley with more than a quarter of total US agricultural pesticides being used in this region. People living near agricultural fields have increased exposures to pesticides outdoors (Lu et al. 2000; Ward et al. 2006) as well as indoors in air (Foreman et al. 2000) or in dust (Curwin et al. 2007; Fenske et al. 2005). According to Pesticide Action Network (PAN) (Kegley et al. 2000), individual pesticides can be categorized by chemical classes such as organochlorines (OCs), organophosphates (OPs), and pyrethroids (PYRs).

OCs are a group of chlorinated compounds that are highly persistent in the environment (Jayaraj et al. 2016). The modes of action of OC pesticides include opening of sodium ion channels in neurons (Dong 2007), as well as binding with the GABA chloride ionophore complex producing a decreased uptake of chlorine ions in neurons (Zuluaga et al. 2016). Many OCs have been classified as endocrine disruptors, as they can reduce the efficiency of thyroid hormones. Chronic OC exposure can also affect neurobehavioral development and the reproductive system. Due to their environmental persistence, ability to bioaccumulate in humans, and adverse impacts on health, the U.S. started to ban the use of OCs for agricultural use starting in the 1970s. Use of some OCs extended into the late 1980s. Despite bans, some pesticides are still detectable in most of the US and its population (Orta et al. 2020).

OPs are one class of pesticides that work by damaging the acetylcholinesterase (AChE) which is a critical enzyme for controlling nervous system signaling in the body. Inhibition of AChE leads to accumulation of the neurotransmitter acetylcholine resulting in a variety of acute neurotoxic effects (Sultatos 1994). Chronic exposure has been shown to induce oxidative stress

leading to lipid peroxidation, DNA damage and protein oxidation (Zuluaga et al. 2016). Observational and experimental studies have shown detrimental health effects of OP exposures in adults ranging from effects due to acetylcholinesterase inhibition toxicity to long term and low-level effects such as neurodegeneration, such as seen in Parkinson's disease (PD) (Paul et al. 2018a).

Over the past decades, a less toxic insecticide, PYRs, has replaced OCs and OPs to become one of the most frequently used pesticides, and widely use in treated clothing, in mosquito control, and in agriculture. The main target of pyrethroids is the nervous system where it can disrupt sodium channels and also calcium channels in axons (Saillenfait et al. 2015). Epidemiological studies suggested that prenatal or postnatal exposure to pyrethroids is associated with low birth weight and higher risk of autism and developmental delay (Ding et al. 2015; Hanke et al. 2003; Shelton et al. 2014). Furthermore, pyrethroid can potentially contribute to neurological symptoms such as PD (Ritz et al. 2016).

Chronic exposure of OPs, PYRs, and OCs have been shown to have a wide range of secondary or off-target effects, such as oxidative stress, immunotoxicity, mitochondrial dysfunction, and regulation of neuronal apoptosis (Banerjee et al. 2001; Costa 2006; Karami-Mohajeri and Abdollahi 2011; Zuluaga et al. 2016). Numerous epidemiological studies, as well as in vitro and in vivo experimental evidence, suggests these three pesticide groups are associated with a wide range of adverse health effects including neurodegenerative disease, such as Parkinson's disease (PD) and Alzheimer's disease (AD) (Blair et al. 2015). As chronic pesticide exposures may have profound effects on various biological systems, characterizing metabolic response to these chemicals and how they influence human health systemically is important. Most previous studies measured pesticide exposure levels via pesticide biomarkers in blood or urine samples. Although for most past and currently used pesticides this is an appropriate way of

5

measuring short-term acute pesticide exposures, it is rarely appropriate for measuring long-term exposures due to the relatively short biological half-lives of most chemicals. For OPs and PYRs, the half-lives are within several days. Although OCs have longer physiologic half-lives of approximately 5-10 years (Jayaraj et al. 2016), it is still challenging to study these exposures as contributors to geriatric diseases when the most relevant exposures periods are decades in the past. To understand the detrimental biological effect of chronic long term and past pesticide exposure, it is important to use advanced geospatial exposure assessment models that can estimate such exposure over extended periods of time. Linking this kind of chronic exposure assessment to molecular profiles presents a great opportunity to estimate long-term health effects from chronic exposure in human populations.

1.4 Metabolomics background

The metabolome includes all low molecular weight (<2,000 Da) chemical species present in biological matrices, and it captures the interaction between the genome, diet, environment, and the biochemical processes required for life (Wishart et al. 2013). A recent study estimated that there are more than 1 million chemical compounds in the human metabolome (Uppal et al. 2016). The current untargeted high-resolution metabolomics platform is capable of simultaneously measuring more than 20,000 endogenous and exogenous chemical signals in biological samples, providing a unique opportunity to link environmental exposures to internal biological responses, thereby helping researchers to better understand the mechanisms of exposure-related diseases (Niedzwiecki et al. 2019).

Plenty of studies have demonstrated the utility of untargeted high-resolution metabolomics (HRM) for environmental epidemiological research (Jones 2016; Niedzwiecki et al. 2019; Stingone et al. 2017). The unbiased profiling of the metabolome helps the identification of new

markers of biological response and the generation of new hypotheses, while improving metabolomic coverage and lowering the costs for omics scale population-based research.

Several previous studies have used HRM as a tool to study the effect of air pollution on the human metabolome (Jin et al. 2021). For example, The Cooperative Health Research in the Region of Augsburg (KORA) study in Germany conducted targeted metabolomics for 138 metabolites in older adults' serum samples and found lysophosphatidylcholines to be associated with short-term air pollution exposures, especially nitrogen dioxide (NO2) (Ward-Caviness et al. 2016). A panel study of healthy college students characterized the metabolomic profile of plasma and saliva samples after traffic-related air pollution exposure and identified metabolites and pathways associated with oxidative stress, inflammation, and nucleic acid damage and repair (Liang et al. 2018). Two cross-sectional studies conducted in UK and Atlanta, respectively, found that oxidative stress and inflammation related metabolites were associated with both long term air pollution and (Menni et al. 2015; Walker et al. 2018). Another study estimated the association between long-term exposure to PM2.5, NO2, and Ozone with metabolites quantified in plasma samples among men from the Normative Aging Study (NAS). The perturbed metabolomic pathways found in the study are mainly related to inflammation and oxidative stress (Nassan et al. 2021).

Metabolomics has also been used to characterize biological disruptions caused by pesticides. Several studies used mouse models to investigate the effect of OPs exposure and found that exposure to OP mixtures (chlorpyrifos and carbaryl, or dichlorvos, dimethoate, acephate, phorate) induced perturbations in energy and lipid metabolism (Du et al. 2013; Wang et al. 2009; HP Wang et al. 2011). Another study focused on the prenatal and postnatal period and exposed female mice to single or a mixture of pesticides throughout gestation and lactation, and then used 1H NMR-based metabolomics to analyze the plasma of the offspring. The results showed that

7

exposure to each pesticide produced a specific metabolomic fingerprint in the adult offspring. Metabolites that discriminated between groups were glucose or lactate, choline, glycerophosphocholine and phosphocholine, potentially due to oxidative stress producing disturbances (Demur et al. 2013). Most previous studies using animal models relied on high-level pesticides exposure. Therefore, the biological mechanism underlying low-dose/daily-dose exposure to pesticides in the general population is unknown but of interest to evaluate the potential of low dose exposure to cause harm and chronic disease. One study based on the PELAGIE cohort (Brittany, France) assessed urine samples of 83 pregnant women in early pregnancy and associated their metabolomic profiles with the surface of land dedicated to agricultural cereal growing activities in their towns of residence as proxies for pesticides exposure. The results showed that exposure to complex pesticide mixtures from agricultural applications may have induced perturbations of glycine, threonine, lactate and glycerophosphocholine, again metabolites that are related to oxidative stress and energy metabolism (N. Bonvallot et al. 2013).

1.5 Epigenomics background

Another useful tool for investigating exposure-related biological processes is epigenomics. Gene expression can be modified through epigenetic changes without changing the DNA sequence itself. Epigenetics serves as a key mechanism controlling cell and tissue differentiation, and activation or inactivation of gene transcription through epigenetics can ultimately define the phenotype of a cell or tissue at specific developmental stages (Ho et al. 2012).

The environment is a major contributor to distinct epigenetic signatures. Furthermore, modifications of the epigenome can persist for a long time after the initial event which make epigenetics a great measurement of long-term exposure histories (Go and Jones 2016; Niedzwiecki et al. 2019). DNA methylation, as one type of epigenetic change, occurs at the CpG dinucleotides

in the cytosine C5 position. Current high-throughput assays based on parallel sequencing of DNA with bisulfite conversions such as the Illumina EPIC BeadChip can provide measures of up to 850,000 CpG sites. Epigenome wide association studies have been conducted to investigate the association between distinct CpGs methylation patterns and chemical exposures or diseases, providing insight into mechanisms underlying biological responses and disease (Hou et al. 2012; Niedzwiecki et al. 2019).

To date, A few population-based studies have been conducted to investigate the epigenomic fingerprints of pesticide exposure. For example, in the Agricultural Health Study, DNA methylation changes were associated with high pesticide exposure events in applicators and suggested that acute pesticide exposure may have the capacity to alter DNA methylation (Rusiecki et al. 2017). Our own lab investigated residents of central California and reported that OP-related epigenetic changes could be related to their intended mechanisms of action inhibiting acetylcholinesterase by suggesting differential methylation patterns for acetylcholine receptors in the chronically exposed populations (Paul et al. 2018a).

2 Maternal Serum Metabolome and Traffic-Related Air Pollution Exposure in Pregnancy

2.1 Introduction

Maternal exposures to air pollutants including nitrogen oxides (NO2, NOx), particulate matter (PM10, PM2.5), carbon monoxide (CO) and ozone (O3) and their mixtures represented by various sources have been associated with pregnancy complications and adverse birth outcomes such as preeclampsia, gestational diabetes, preterm birth and low birth weight (Dadvand et al. 2013; Eze et al. 2015; Pedersen et al. 2013; Pedersen et al. 2014; Stieb et al. 2012; Wu et al. 2009). It has also been shown to affect fetal development especially neurodevelopment in offspring (Becerra et al. 2013; Glinianaia et al. 2004; Guxens et al. 2014; Suades-Gonzalez et al. 2015). However, there is still not much data available to elucidate relevant biologic pathways and mechanisms underlying air pollution related adverse pregnancy outcomes, even though it has been hypothesized that endocrine disruption, oxidative stress, inflammatory response, and DNA damage are major contributors (Hougaard et al. 2008; Kelly 2003; Risom et al. 2005).

High resolution metabolomics (HRM) is a powerful analytical approach to profile thousands of chemicals in biological specimens and to identify metabolic disturbances. Recently, researchers have used HRM to study the effects of environmental exposures such as metals, tobacco smoking, and polycyclic aromatic hydrocarbons on human metabolism (Garcia-Sevillano et al. 2015; Gu et al. 2016; Jones et al. 2016; Wang et al. 2015).

There are currently no metabolomics studies that focus on air pollution exposure in pregnant women, however several previous studies have used HRM as a tool to study acute effects of air pollution on the human metabolome. The Cooperative Health Research in the Region of Augsburg (KORA) study in Germany conducted targeted metabolomics for 138 metabolites in older adults' serum samples and found lysophosphatidylcholines (LPCs) to be associated with

short-term air pollution exposures, especially nitrogen dioxide (NO_2) (Ward-Caviness et al. 2016). A randomized, crossover trial that enrolled 55 healthy college students showed that short-term exposure to particulate matter less than 2.5 microns (PM2.5) induced metabolic changes associated with stress hormone levels, insulin resistance, and markers of oxidative stress and inflammation (H Li et al. 2017). Two other experimental studies using a crossover design to assess short-term ambient air pollution exposure identified a variety of metabolomic features (van Veldhoven et al. 2018; Vlaanderen et al. 2017). One study utilized a multi-platform approach including gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), and nuclear magnetic resonance (NMR) to identify metabolites in bronchial wash and bronchoalveolar lavage fluid samples from fifteen healthy subjects exposed to biodiesel exhaust (Surowiec et al. 2016). Recently, a panel study of healthy college students characterized the metabolomic profile of plasma and saliva samples after traffic-related air pollution exposure and identified metabolites and pathways associated with oxidative stress, inflammation, and nucleic acid damage and repair (Liang et al. 2018). Most of the previous studies concentrated on short-term (hours or days) air pollution exposure, had relatively small sample sizes and assessed healthy adults.

Only two studies to date focused on long-term air pollution exposure. A cross-sectional study conducted within the TwinsUK cohort, found that oxidative stress and inflammation related metabolites such as α-tocopherol, benzoate, and glycine were associated with both long term air pollution (annual average PM2.5 concentration) and lung function (Menni et al. 2015). Another cross-sectional study with 59 healthy participants has linked year-long ultrafine particle exposure with metabolic variations related to antioxidant pathways and endothelial function (Walker et al. 2018).

We conducted an untargeted metabolomics study focusing on pregnant women since they and their offspring might be more susceptible to the adverse effects of air pollution. Specifically, in our study population, serum samples from 160 pregnant women were collected during the second trimesters of pregnancy, and average air pollution exposures (CO, nitric oxides (NOx), and PM2.5) during the first trimester of pregnancy were estimated based on residential addresses using an emissions based dispersion model– i.e. the California line source dispersion model, version 4 (CALINE4) (Benson 1984). We then utilized liquid chromatography with high-resolution mass spectrometry (LC-HRMS) to obtain maternal metabolomic profiles. The HRM platform we adopted has also been used in three previous studies that examined ambient air pollution exposure (Ladva et al. 2018; Liang et al. 2018; Walker et al. 2018), and been shown to be sensitive enough to capture metabolic perturbations in blood samples induced by air pollution. The aims of this study were to identify changes in specific metabolomic profiles associated with air pollution exposure in pregnant women, and to conduct comprehensive pathway analysis to investigate potential biological mechanisms related to adverse health effect from air pollution exposure during pregnancy.

2.2 Methods

2.2.1 Study population

We utilized the California birth records to randomly select mothers of children born between 2005 and 2010 from a previous case-control study of autism, all of whom were controls matched to children with autism in a 1:10 ratio by sex and birth year. All eligible mothers lived in the California Central Valley region according to the residential address recorded on the birth certificate (n = 1,466). In addition, mothers were eligible for the present study if their live born autism-free children did not have a birth defect recorded, were born at a gestational age between

21-46 weeks and had a recorded birth weight between 500g and 6800g (n = 1,433). The California central valley is known for its extremely high use of agricultural pesticides. Here, we restricted our analyses to mothers who according to our geographic information system (GIS) and state mandated pesticide use report based assessments had not been exposed to suspected neurotoxic pesticides (organophosphates, pyrethroids, glyphosates, fungicides, or neonicotinoids) (Rull and Ritz 2003). Considering the limited number of serum samples available for analysis via HRM and the generally high air pollution exposure levels in the study region, we further selected mothers based on the air pollution levels we modeled at residences to increase the statistical power of our study. Specifically, we used the population of 1,433 women originally randomly selected from birth certificates to estimate percentiles of air pollution exposure, and selected mothers from the extremes of the exposure distribution, considering women as "highly exposed" in the first trimester of pregnancy if three highly correlated CALINE4 modeled traffic related air pollutant levels (CO, NOx, and PM2.5) fell at or above the 75th percentile of the distribution, and as "low exposed" if their air pollutant levels fell below the 35th percentile. This was necessary because we did not have enough subjects available for a balanced selection of exposed and unexposed at the extremes of the air pollution distribution. Furthermore, mothers had to have utilized the California Prenatal Screening Program (CPSP) and their mid-pregnancy blood serum sample be stored and available (n = 160). About 74% of California women participate in the CPSP, which is a sequential serum screening program for birth defects during pregnancy that collects blood samples in mid-pregnancy (around the 16th week of gestation) (Cunningham and Tompkinison 1999). Maternal serum samples were collected by obstetrical care service providers and mailed to the state laboratory for analyses when subjects participated in the CPSP and specimens leftover after testing were stored at -20 °C at the California Biobank.

From birth records we obtained information on child sex, birth year, maternal age, race/ethnicity, education, and parity. Other potential confounders such as BMI or diet were not available in the birth records limiting this type of study.

2.2.2 Air pollution estimation

Residential addresses as recorded on birth certificates were geocoded using OpenSource geocoding software (Goldberg et al. 2008). We estimated each woman's average air pollutant exposure (i.e. CO, NOx, and PM2.5) during the first trimester, using a modified version of the CALINE4 that estimates pollution from sources within 1500 m of residential locations as described in detail elsewhere (Heck et al. 2013). In brief, input data for the prediction process included roadway geometry, traffic counts, emission factors, and meteorological parameters (wind speed, wind direction, temperature, stability class, and mixing heights). Year and season (winter and summer) and emission factors for CO, NOx and PM2.5 were obtained from the EMFAC2011 vehicle emissions model (California Air Resources Board 2013). CALINE4 predictions in this study do not incorporate background levels of pollutants, thus solely represent the contribution from local traffic emissions (Benson 1989; Broderick et al. 2005; Levitin et al. 2005; Marmur and Mamane 2003; Wu et al. 2016).

Subjects were classified based on their air pollution levels such that 98 pregnancies were considered "highly exposed" in the first trimester, while 62 were "low exposed" (Supplemental Table 2-S1).

2.2.3 High-resolution metabolomics

HRM profiling was completed according to established methods (Walker et al. 2018; Walker et al. 2019b). Serum samples were transported from California Biobank to Emory and stored at -80 °C. Batches of 40 serum samples were removed from storage and thawed on ice. Each

sample was then thoroughly vortexed, and 65 μL of serum was treated with 130 μL of LC-MS grade acetonitrile. The extract was equilibrated for 30 min on ice and centrifuged at 16,100 × g for 10 min to remove precipitated proteins. The resulting supernatant was transferred to an autosampler containing a low volume insert and maintained at 4ºC until analysis (<24 h). NIST 1950 (Simon-Manso et al. 2013) was analyzed at the beginning and end of the entire analytical run and for additional quality control (QC) two replicate pooled human plasma samples were analyzed at the beginning, middle, and end of each batch of 40 samples for normalization and batch effect evaluation.

Sample extracts were analyzed in triplicate using a dual column, dual polarity approach that includes hydrophilic interaction (HILIC) chromatography with positive ESI and C18 chromatography with negative ESI (Ultimate 3000, Q-Exactive HF, Thermo Fisher, *m/z* range 85-1275) (Walker et al. 2018). Following a 10 μL sample injection, HILIC separation was accomplished using a 2.1 cm x 5 cm x 2.5 μm HILIC column (Waters XBridge BEH Amide XP HILIC) and acetonitrile gradient (A= water, B= acetonitrile, C= 2% formic acid) consisting of an initial 1.5 minute period of 22.5% A, 75% B, 2.5% C, followed by linear increase to 77.5% A, 20% B, 2.5% C at 4 min and hold for 1 minute. Separation by C18 was with 2.1 cm x 5 cm x 3 μm column (Higgins endcapped C18) with C= 10mM ammonium and the following gradient: initial 0.5 minute period of 60% A, 35% B, 5% C, followed by linear increase to 0% A, 95% B, 5% C at 1.5 min and then held for an additional 3 minutes. Mobile phase flow rate was held at 0.4 mL/min for 1.5 minutes, and then increased to 0.5 mL/min. The mass spectrometer was operated using ESI mode at a resolution of 120,000 and mass-to-charge ratio (*m/z*) range 85-1275. 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. Raw data files were extracted and aligned using *apLCMS* (Yu et al. 2009) 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).

2.2.4 Statistical analyses

For three replicates of each feature, intensities were summarized using the median, except when more than 50% of the replicates were missing (in this instance the value was set to missing). Metabolomic data was then filtered to keep only features present in at least 80% of one comparison group and >50% of all samples together. After filtering, missing values were imputed by one-half of the lowest signal detected for that feature across all samples.

Feature intensities were log2 transformed before analyses. To control for potential confounding, we used residuals of intensities derived from linear regression against potential confounders including maternal age, maternal race/ethnicity, and maternal education.

We performed partial least squares discriminant analysis (PLS-DA) to identify features associated with air pollution exposure. PLS-DA is a supervised, multivariate analysis approach for dimensionality reduction that maximizes covariance between intensities of metabolomic features and air pollution exposure (Wold et al. 2001). We selected features with a Variable Importance in Projection (VIP) scores >= 2 (Le Cao et al. 2009). Fold change was calculated as the ratio of raw intensities between high and low-exposed groups. To evaluate the performance of selected features, we conducted 10-fold cross-validation tests utilizing the support vector machine (SVM) and

calculated the classification accuracy of the selected features. All feature selection approaches were implemented with the R package *mixOmics* v6.3.1.

2.2.5 Pathway analysis and annotation

Discriminative features selected by PLS-DA were annotated using xMSannotator (Uppal et al. 2017). Accurate mass *m/z* for adducts formed under positive/negative ESI mode was matched to the Human Metabolome Database (HMDB), Kyoto Encyclopedia of Genes and Genomes (KEGG), and LipidMaps with a mass error threshold of 10 ppm. xMSannotator also takes into consideration correlation of intensities and retention time, and assigns confidence scores based on a multilevel scoring algorithm which ensures the accuracy of annotation.

In order to identify enriched metabolic pathways comparing high and low exposed groups, we conducted pathway enrichment analysis utilizing *Mummichog* (v. 1.0.10) (Li et al. 2013). All discriminating features previously selected by PLS-DA with VIP >= 2 were included in this pathway enrichment analysis. *Mummichog* is a novel pathway and module enrichment analysis algorithm designed specifically for high-resolution liquid chromatography-mass spectrometry. Although the tentative annotation results in *mummichog* may include false positives, the enriched pathways inferred by the algorithm have been proven to be valid and to reflect real biological activity (S. Li et al. 2016; Shuzhao Li et al. 2016; S Li et al. 2017; Uppal et al. 2016). All metabolites annotated by *mummichog* were required to present in at least their primary adduct (M+H or M-H for positive and negative mode respectively) to reduce the false positive match rate. The p-value threshold we relied on was 0.05. Only enriched pathways with at least 3 overlapping metabolites were kept for further evaluation and interpretation.

Some discriminating features were further confirmed by matching the accurate mass *m/z* and retention time to authentic chemical standards analyzed using LC-MS/MS. The error tolerance was 5 ppm for *m/z* and 30 seconds for retention time.

2.3 Results

Demographics of the 160 subjects are provided in Table 2-1. Mothers who were younger, less educated, or of Hispanic origin were more likely to be highly exposed to air pollution in this study population.

In total, we detected 14,555 features (6,139 in HILIC column and 8,416 in C18 column), but after filtering for missing values, 8,995 features remained (4,038 in HILIC column and 4,957 in C18 column). We identified 181 and 251 unique metabolomic features that were associated with air pollution exposure during pregnancy from both HILIC and C18 column respectively (Figure 2-1 and Supplemental Figure 2-S1 to 2-S2) using PLS-DA and adjusting for maternal age, ethnicity and education while setting the VIP scores greater than 2. The balanced classification rate derived from 10-fold CV showed that both sets of discriminatory features effectively separate the classes (82.5% and 70.4% for HILIC and C18 respectively).

Annotation of PLS-DA discriminatory features included 187 *m/z* features matching to one or more unique metabolites with medium or high confidence scores. Because compounds were present in the databases used for annotation that share the same chemical formula but different structure, we confirmed the annotation results by matching the retention time and *m/z* to authenticated chemical standards verified by tandem mass spectrometry. In total, we confirmed 6 metabolites including serine, creatinine, L-histidine, myo-inositol, linoleic acid, and heptadecanoic acid (confidence level 1). Their chemical identities are shown in Table 2-2. Among them,

creatinine was positively associated, while four other metabolites were negatively associated with air pollution.

Using *mummichog*, we examined whether the features that were selected by PLS-DA were enriched within specific metabolic pathways. The result indicated that 24 metabolic pathways were differentially enriched with a P-value < 0.05 (Table 2-3). Tentative annotation results of metabolites in each pathway are provided in Supplemental Table 2-S2 to Supplemental Table 2-S28. Nine are lipid-related metabolic pathways including fatty acid activation, de novo fatty acid biosynthesis, and glycosphingolipid metabolism. Changes in these pathways may indicate associations between air pollution exposure and oxidative stress; features in the linoleate pathway were annotated as oxidative stress biomarkers or antioxidants including linoleic acid, 13-hydroxyoctadecadienoic acid (13-HODE), 13-oxo-9,11-octadecadienoic acid (13-oxo-ODE), gamma-linolenic acid, and azelaic acid. Linoleic acid was further confirmed by an authentic standard (confidence level 1). Disruption of eicosanoid metabolism such as the leukotriene metabolism and prostaglandin formation from arachidonate indicates an inflammatory response. Putative annotations in the leukotriene metabolism pathway included 12-oxo leukotriene B4 (12-oxo-LTB4) and 20-hydroxy-leukotriene E4 (20-OH-LTE4), all of which showed increased abundance in the high air pollution exposure group. Other pathways associated with oxidative stress included vitamin E metabolism, xenobiotic metabolism, urea cycle/amino group metabolism, and some amino acid metabolism pathways such as histidine metabolism and lysine metabolism. Perturbations were also observed in nucleotide metabolism and several carbohydrate metabolism related pathways.

2.4 Discussion

The untargeted HRM approach we employed provided us with an opportunity to explore the relationship between traffic related air pollution exposures and metabolomic signatures in maternal blood in mid-pregnancy. Measuring more than 10,000 metabolomic features in serum samples using HRM and comparing metabolomic profiles according to exposure, we illustrated that high air pollution exposures during pregnancy are linked to widespread perturbations in the maternal serum metabolome. Sorting these features according to biologic pathways allowed us to identify potential mechanisms by which air pollution may affect pregnant women and fetuses. Of particular importance is that the enriched pathways identified were related to oxidative stress and inflammatory reactions, which have been implicated in many pathological conditions, including adverse birth outcomes such as preterm birth, intrauterine growth restriction, low birth weight, and preeclampsia (Sultana et al. 2017) as well as neurodevelopmental disorders (Patterson 2009).

We observed alterations in fatty acid metabolism, phospholipid metabolism, linoleate metabolism, and eicosanoids including leukotriene and prostaglandin metabolism (see Table 2-3 and Figure 2-2). Traffic-related air pollutants may act directly as free radicals or generate free radicals and cause oxidative stress (Kelly 2003). One of the primary targets of reactive oxygen species derived from air pollutants is the cell membrane. Oxidative stress can induce the activation of phospholipase A2 (PLA2) which then hydrolyze phospholipid (PL) from the cell membrane to generate polyunsaturated free fatty acid and lyso-PL (Anthonymuthu et al. 2018; Sato et al. 2016). Major polyunsaturated fatty acids (PUFA) released through this step include linoleic acid and arachidonate acid. Due to the existence of double carbon bonds, these fatty acids can be subsequently oxidized by oxygenase (Sato et al. 2016). As one of the omega-6 polyunsaturated fatty acids, linoleic acid generates 13-hydroperoxyoctadecadenoic acid (9-HPODE) and 13-

HPODE through lipoxygenases. These products then convert to 9-hydroxyoctadecadienoic acid (9-HODE) and 13-HODE (Tam et al. 2013), and both are potential biomarkers for oxidative stress affecting lipids. In our study, we observed associations with the linoleate pathway, specifically, air pollution exposure increased metabolomic features we matched to 13-HODE; while linoleic acid and 13-keto-9Z,11E-octadecadienoic acid (13-oxo-ODE), a downstream product derived from 13-HODE were decreased in the highly exposed group. In particular, we were able to confirm the identification of linoleic acid and gamma-linolenic acid within the linoleate pathway using authentic standards. Gamma-linolenic acid is an anti-inflammatory metabolite that can inhibit the biosynthesis of leukotriene B4 (Horrobin 1992; Liang et al. 2018). Although below the discriminative feature selection threshold (VIP=1.31) in our data, we found gamma-linolenic acid to be negatively associated with air pollution exposure, which is consistent with previous studies (Liang et al. 2018; Walker et al. 2018).

In addition to linoleic acid, arachidonic acid is another omega-6 PUFA affected by air pollution induced oxidative stress. As a key inflammatory intermediate, arachidonic acid is being released from cell membranes and converted to eicosanoids through several pathways including: the lipoxygenase (LOX) pathway, where arachidonic acid is dioxygenated to produce hydroperoxyeicosatetraenoic acid (HPETE) and then converted to leukotrienes and other lipoxins; and the cyclooxygenase (COX) pathway that produces prostaglandin (Tam et al. 2013). Both leukotrienes and prostaglandins are major proinflammatory mediators.

In our study, two of the features found to be enriched in the leukotriene pathways were annotated as 12-oxo-leukotriene B4 and 10,11-dihydro-12-oxo-leukotriene B4. 12-oxo-LTB4 was found to be lower in the exposed group, while 10,11-dihydro-12-oxo LTB4 was increased. Both are downstream metabolites of LTB4, but possibly due to the rapid metabolism of 12-oxo-LTB4

to 10,11-dihydro-12-oxo-LTB4 by reductase (Powell et al. 1996; Wainwright and Powell 1991) only the end product in this pathway is found to be increased. LTB4 is a potent lipid chemoattractant that can induce an inflammatory response. It is synthesized through the LOX pathway and recruits leukocytes leading to inflammatory reactions (Ford-Hutchinson et al. 1980). Due to their pro-inflammatory effect, altered levels of LTB4 and its metabolites during pregnancy have been linked to various adverse health outcomes (e.g., asthma) and birth complications such as preterm birth and preeclampsia (Biagi et al. 1990; Busse 1998; Romero et al. 2007). Recently, a lipidomic study conducted in 197 mother-newborn pairs measured the association between in-utero PM2.5 exposure and oxylipin profiles in newborns. Consistent with our study, they also found significant differences in cord blood levels for metabolites derived from the LOX pathway (Martens et al. 2017).

Prostaglandins are a type of eicosanoid derived from arachidonic acids through COX. *Mummichog* annotated 8 features to be enriched in prostaglandin pathways including multiple prostaglandins and related metabolites. Previously, animal and human studies have linked excessive prostaglandins with NO2 inhalation (Yan et al. 2016), ozone (Peden 1999; Peden 2001), PM2.5, and sulfate (W Li et al. 2016). Increased levels of 8-epi-prostaglandin F2a are widely used as indicators of lipid peroxidation (Tacconelli et al. 2010), and the pro-inflammatory effects of prostaglandins are well documented (Funk 2001). Prostaglandins play important roles during pregnancy including vasodilatation and uterine contraction. Perturbations of the prostaglandin pathway are most likely responsible for abnormal placental and uterine blood flow. Several in-vivo and in-vitro studies have associated prostaglandins with preeclampsia (Kaaja et al. 1995; Ogburn et al. 1984).

In addition to lipid-related pathways, vitamin E metabolism is also associated with oxidative stress. Vitamin E, including tocopherols and tocotrienols, are lipid-soluble antioxidants and lipid peroxy radical scavengers (Chow 1991; DellaPenna and Mène-Saffrané 2011; Noctor et al. 2015). Vitamin E can protect membrane lipids from oxidation by physically or chemically quenching singlet oxygen or by donating a hydrogen atom to the PUFA peroxy radical thereby interrupting the peroxidative reaction (DellaPenna and Mène-Saffrané 2011). Two previous metabolomics studies found the activity of vitamin E metabolism to be decreasing with air pollution exposure (Liang et al. 2018; Menni et al. 2015). In our study, metabolites of tocotrienols showed alterations with a decrease in antioxidant and increase in oxidized metabolites among the exposed. Within the Vitamin E metabolism pathway, a feature tentatively annotated as 13'-carboxy-alpha-tocotrienol (mz:452.2952, M-H) was found to be higher while a feature tentatively annotated as alpha-tocotrienol (mz:423.3265, M-H) was lower in the exposed group.

Several amino acid metabolism pathways were also enriched including the methionine, cysteine, and histidine pathways which are sulfur-containing amino acids that are readily oxidized (Berlett and Stadtman 1997; Pisoschi and Pop 2015). Previous studies described oxidation of methionine to be associated with ambient PM exposure in mice (Lai et al. 2016; Lee et al. 2014). In our samples, histidine and its pathway were found to be reduced among the air pollution exposed. Notably, we were able to annotate histidine with confidence level 1 using the authentic standard. Histidine has been reported to have anti-inflammatory effects, and previous studies found that histidine was negatively associated with inflammation and oxidative stress (Liang et al. 2018; Niu et al. 2012; Uchida 2003; Watanabe et al. 2008).

Numerous in-vitro and in-vivo studies have shown that exposure to air pollution induces oxidative stress and inflammatory reactions (Daher et al. 2014; Dick et al. 2003; Ghio et al. 2012;

Guerra et al. 2013; Happo et al. 2013). Moreover, the pregnancy period is especially vulnerable to oxidative stress due to the necessary increased energy expenditure (Nagiah et al. 2015). Two studies measuring specific biomarkers for oxidative stress indicated that air pollution exposure during pregnancy elevated oxidative stress responses (Anderson et al. 2018; Nagiah et al. 2015). Consistent with these results, our untargeted systematic approach also identified pathways predominantly enriched for oxidative stress and inflammatory responses. Maternal oxidative stress may cause damage to all major cellular elements and might especially affect the placenta and its function, therefore contributing to adverse birth outcomes such as spontaneous abortion, preeclampsia, intrauterine growth restriction, low birth weight, and preterm delivery (Al-Gubory et al. 2010; Duhig et al. 2016; Lavigne et al. 2018; Peter Stein et al. 2008).

Our study has some limitations. Although model-based air pollution estimation has a high spatial resolution, they do not necessarily translate into personal exposure because 1) there are inherent uncertainties in an emissions-based dispersion model; 2) emission sources other than local traffic are not addressed by the model; 3) lack of information about work location and subjects' time-activity. Thus, we did not estimate 'personal' air pollution exposure but rather the contributions by local traffic sources to personal exposure. Local traffic noise might be a confounder, if it influences the same metabolic pathways as air pollution. However, we recently found that traffic noise estimates were only moderately correlated with CALINE4 estimates of traffic related air pollution (r=0.4) at the residences in Northern CA counties (unpublished data); most likely because air pollution and noise physically behave differently and CALINE4 estimates take factors such as meteorology into consideration.

Besides air pollution exposure, active or passive smoking and diet may explain differences in metabolites. We did not have information on the subjects' smoking behavior or dietary intake

during pregnancy, however in order for smoking or dietary differences to have confounded our analyses, they would have had to be related to air pollution exposures. By controlling for maternal age, maternal race/ethnicity, and maternal education, we hope to have at least partially addressed potential confounding. Additionally, while there is no active or second-hand smoking information on the birth records from these years, we were able to identify cotinine levels in the serum by matching the accurate mass m/z (177.1022) and retention time (32.8s) to authentic standards. Previous studies have shown that cotinine is a reliable marker of smoking [4-6]. Cotinine levels among women with high or low traffic-related air pollution exposure were generally extremely low, indicating that the percentage of active or passive smokers in our study is low - as expected in California. Also, the intensities of cotinine were not statistically significantly different (t-test p-value=0.34) for exposed versus unexposed women's samples. Nevertheless, residual uncontrolled confounding is possible. Also, women with less than a high school education are likely more recent immigrants from Mexico (Hoggatt et al. 2012). If these women are introducing a 'healthy migrant' bias this may bias our results.

To assess whether air pollution impacts on metabolomic profiles differed by ethnicity, we also conducted analyses in subgroups of Hispanic and Non-Hispanic women. We found that high level of traffic-related air pollution was associated with 9 metabolic pathways in both subgroups; importantly, we identified linoleate metabolism as well as methionine and cysteine metabolism pathways in both groups consistent with the induction of oxidative stress.

Another limitation is an inherent challenge in the untargeted metabolomics analysis. Without further metabolite identification using tandem MS, we could only tentatively annotate the extracted features using computational approaches. Adopting a pathway and network analysis approach, we were able to improve annotation results, but there may still be some false matches

that could have influenced the interpretation. It is recommended to improve the identification of metabolites using either tandem MS or internal standards in future studies.

2.5 Conclusions

In summary, we applied HRM to identify perturbations in the serum metabolome associated with traffic-related air pollution exposure during pregnancy. We observed metabolic pathways consistent with oxidative stress and inflammatory reactions, which may contribute to adverse health outcomes in offspring and we corroborated previous results of metabolome studies conducted in air pollution exposed non-pregnant adults.

2.6 Tables and figures

Table 2-1. Demographic characteristics of mothers and children

	High Exposure n=98		Low Exposure n=62	
	n	**%**	**n**	**%**
Maternal age at time of pregnancy (years)				
<= 18	18	18.4	5	8.1
19-25	42	42.9	17	27.4
26-30	24	24.5	20	32.3
> 30	14	14.3	20	32.3
Maternal race/ethnicity				
Non-Hispanic White	11	11.2	21	33.9
Hispanic	70	71.4	30	48.5
Others (African American/Black, Asian)	17	17.4	11	17.7
Maternal education				
Less than 12th grade	47	48.0	11	17.7
High school graduate or equivalent	24	24.5	22	35.5
Some college	20	20.4	12	19.4
College or more	7	7.1	17	27.4
Mother born in the US				
Yes	61	62.2	46	74.2
No	37	37.8	15	24.2
Missing	0	0.0	1	1.6
Preterm birth				
Yes	12	12.2	7	11.3
No	86	87.8	55	88.7
Child sex				
Male	79	80.8	46	74.2
Female	19	19.2	16	25.8

Table 2-2. Confirmed [a] chemical identity of the discriminatory metabolomic features associated with high traffic-related air pollution

m/z	RT (s)	Adduct Form	Metabolite	Fold Change (Log2)	Column
106.0499	94.2	M+H[1+]	Serine	-0.05	HILIC
114.0662	47.0	M+H[1+]	Creatinine	0.45	HILIC
156.0767	112.7	M+H[1+]	L-Histidine	-0.05	HILIC
215.0328	28.6	M+Cl[1-]	Myo-Inositol	1.24	C18
269.2485	271.5	M-H[1-]	Heptadecanoic acid	-0.25	C18
281.2474	33.6	M+H[1+]	Linoleic acid	-0.15	HILIC

[a] Chemical identification was conducted by matching peaks by accurate mass and retention time to authentic reference standards in an in-house library run under identical conditions using tandem mass spectrometry.

Table 2-3. Enriched metabolic pathways associated with traffic-related air pollution

Column	Pathway	Overlap size	Pathway size	P-value
HILIC	Urea cycle/amino group metabolism	7	40	0.001
	Glycosphingolipid metabolism	5	22	0.001
	Histidine metabolism	4	17	0.002
	Glycerophospholipid metabolism	4	30	0.008
	Linoleate metabolism	3	19	0.010
	Glycine, serine, alanine and threonine	4	36	0.016
	Pyrimidine metabolism	3	30	0.043
C18	Fatty acid activation	10	16	0.000
	De novo fatty acid biosynthesis	7	15	0.000
	Glycosphingolipid metabolism	7	24	0.000
	Keratan sulfate degradation	3	6	0.000
	Fatty Acid Metabolism	4	13	0.000
	TCA cycle	3	14	0.004
	Prostaglandin formation from arachidonate	8	54	0.006
	Lysine metabolism	4	24	0.007
	Glycerophospholipid metabolism	6	40	0.007
	Xenobiotics metabolism	8	59	0.011
	Glycolysis and Gluconeogenesis	4	27	0.012
	Methionine and cysteine metabolism	6	44	0.012
	Fructose and mannose metabolism	3	19	0.014
	Vitamin E metabolism	4	29	0.016
	Butanoate metabolism	3	20	0.016
	Linoleate metabolism	3	20	0.016
	Phosphatidylinositol phosphate metabolism	3	22	0.023
	Purine metabolism	5	42	0.027
	Leukotriene metabolism	5	43	0.030
	Sialic acid metabolism	3	28	0.049

[a] P-value calculated by *mummichog* are gamma-adjusted p-values based on permutation tests by resampling from the reference list (Li et al. 2013).

Figure 1a. HILIC Type 1 Manhattan plot (VIP vs. mass-to-charge)

Figure 1b. C18 Type 1 Manhattan plot (VIP vs. mass-to-charge)

Figure 1c. HILIC Volcano plot

Figure 1d. C18 Volcano plot
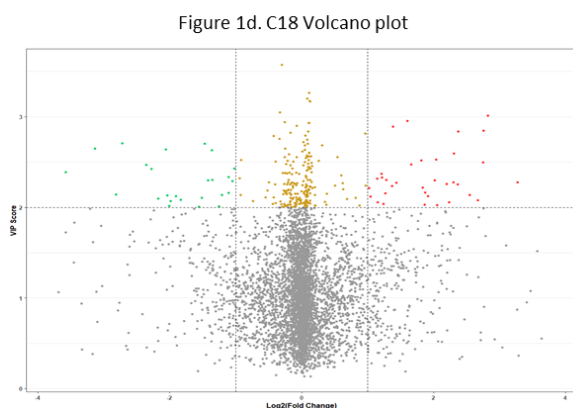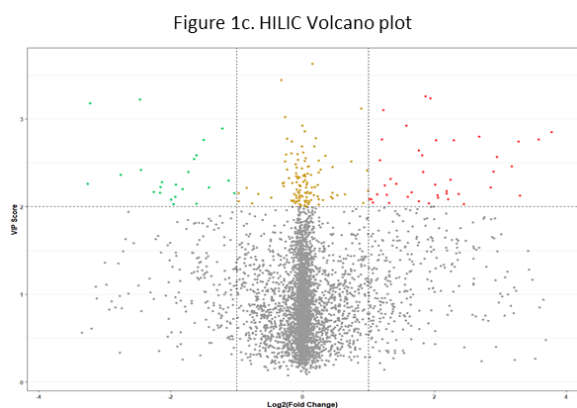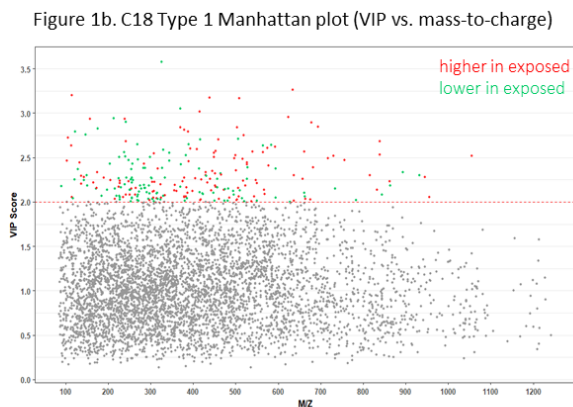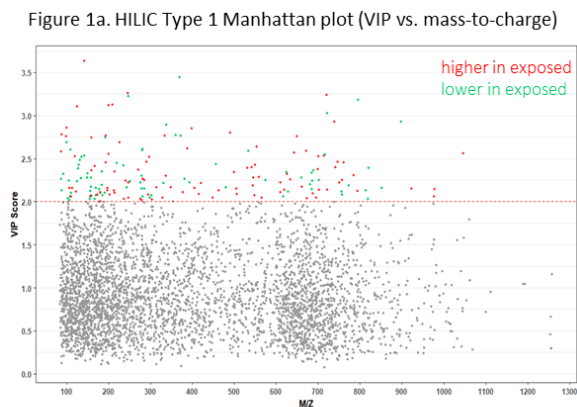
Figure 2-1. Identification of metabolomic features associated with air pollution exposure during pregnancy. A) Type 1 Manhattan plot for features in the HILIC column (positive ion mode), VIP score vs *m/z*. 181 *m/z* features were found above the VIP >= 2 threshold. Red dots represent the features that were higher in the high air pollution exposure group and the green dots represent the features that were lower in the high air pollution exposure group; B) Type 1 Manhattan plot for features in the C18 column (negative ion mode), VIP score vs mass-to-charge. 251 *m/z* features were found above the VIP >= 2 threshold; C) Volcano plot for features in the HILIC column; D) Volcano plot for features in the C18 column.
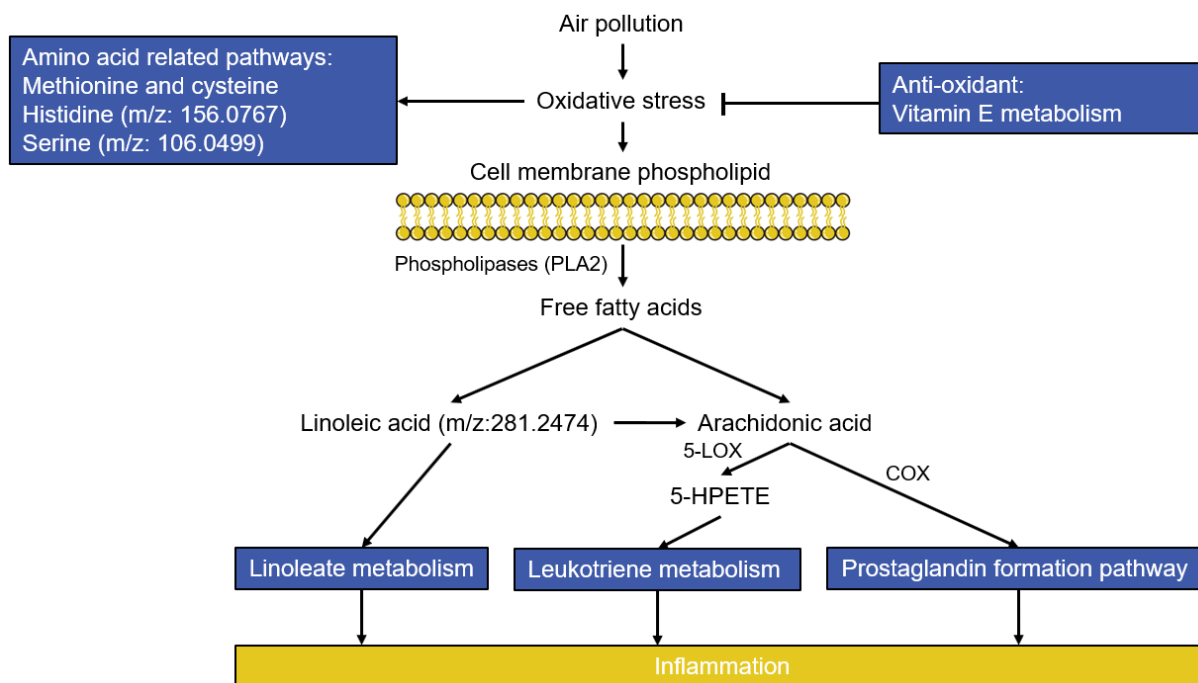
Figure 2-2. Enriched metabolic pathways in response to air pollution induced oxidative stress.

Pathways in blue boxes are significantly enriched pathways.

2.7 Supplemental materials

Supplemental Table 2-S1. Distribution of traffic-related air pollution among two exposure groups

Group	n	Air pollutant	Lower Quartile	Median	Upper Quartile	Mean
High exposed	99	CO (ppb)	37.23	45.27	60.91	49.90
		NOx (ppb)	3.74	4.91	7.78	6.39
		PM25 (mg/m^3)	0.25	0.33	0.48	0.42
Low exposed	62	CO (ppb)	2.30	3.48	5.19	3.66
		NOx (ppb)	0.20	0.33	0.46	0.34
		PM25 (mg/m^3)	0.01	0.02	0.03	0.02

Supplemental Table 2-S2. *Mummichog* **matches for fatty acid activation pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
143.1077	260.8	M-H[-]	Octanoic acid; Caprylic acid; Octylic acid
271.2277	192.5	M-H+O[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
287.2226	223.8	M+CH3COO[-]	Tetradecanoic acid; Tetradecanoate; Myristic acid
291.2096	256.5	M+Cl[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
301.2384	261.0	M+HCOO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
315.2541	259.7	M+CH3COO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
323.2215	240.7	M+HCOO[-]	(9Z,12Z,15Z)-Octadecatrienoic acid; alpha-Linolenic acid; 9,12,15-Octadecatrienoic acid; Linolenate; alpha-Linolenate
		M+HCOO[-]	(6Z,9Z,12Z)-Octadecatrienoic acid; 6,9,12-Octadecatrienoic acid; gamma-Linolenic acid; Gamolenic acid
326.2205	236.3	M+Na-2H[-]	(8Z,11Z,14Z)-Icosatrienoic acid; Dihomo-gamma-linolenic acid; (Z,Z,Z)-8,11,14-Eicosatrienoic acid; (Z,Z,Z)-8,11,14-Icosatrienoic acid; (Z,Z,Z)-8,11,14-Eicosatrienoate; (Z,Z,Z)-8,11,14-Icosatrienoate; 8,11,14-Eicosatrienoate; 8,11,14-Icosatrienoate
327.2541	260.9	M+HCOO[-]	(9Z)-Octadecenoic acid; (Z)-Octadec-9-enoic acid; Oleate; Oleic acid
329.2698	285.7	M+HCOO[-]	Octadecanoic acid; Stearate; Stearic acid
343.2854	287.4	M+CH3COO[-]	Octadecanoic acid; Stearate; Stearic acid
360.1641	88.4	M+Br[-]	vaccenic acid
		M+Br[-]	octadecenoate (n-C18:1)

Supplemental Table 2-S3. *Mummichog* matches for de novo fatty acid biosynthesis pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
271.2277	192.5	M-H+O[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
287.2226	223.8	M+CH3COO[-]	Tetradecanoic acid; Tetradecanoate; Myristic acid
291.2096	256.5	M+Cl[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
301.2384	261.0	M+HCOO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
315.2541	259.7	M+CH3COO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
323.2215	240.7	M+HCOO[-]	(9Z,12Z,15Z)-Octadecatrienoic acid; alpha-Linolenic acid; 9,12,15-Octadecatrienoic acid; Linolenate; alpha-Linolenate
		M+HCOO[-]	(6Z,9Z,12Z)-Octadecatrienoic acid; 6,9,12-Octadecatrienoic acid; gamma-Linolenic acid; Gamolenic acid
326.2205	236.3	M+Na-2H[-]	(8Z,11Z,14Z)-Icosatrienoic acid; Dihomo-gamma-linolenic acid; (Z,Z,Z)-8,11,14-Eicosatrienoic acid; (Z,Z,Z)-8,11,14-Icosatrienoic acid; (Z,Z,Z)-8,11,14-Eicosatrienoate; (Z,Z,Z)-8,11,14-Icosatrienoate; 8,11,14-Eicosatrienoate; 8,11,14-Icosatrienoate
327.2541	260.9	M+HCOO[-]	(9Z)-Octadecenoic acid; (Z)-Octadec-9-enoic acid; Oleate; Oleic acid
329.2698	285.7	M+HCOO[-]	Octadecanoic acid; Stearate; Stearic acid
343.2854	287.4	M+CH3COO[-]	Octadecanoic acid; Stearate; Stearic acid

Supplemental Table 2-S4. *Mummichog* **matches for fatty acid metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
271.2277	192.5	M-H+O[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
287.2226	223.8	M+CH3COO[-]	Tetradecanoic acid; Tetradecanoate; Myristic acid
291.2096	256.5	M+Cl[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
301.2384	261.0	M+HCOO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
315.2541	259.7	M+CH3COO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
329.2698	285.7	M+HCOO[-]	Octadecanoic acid; Stearate; Stearic acid
343.2854	287.4	M+CH3COO[-]	Octadecanoic acid; Stearate; Stearic acid
360.1641	88.4	M+Br[-]	octadecenoate (n-C18:1)

Supplemental Table 2-S5. *Mummichog* **matches for phosphatidylinositol phosphate metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
215.0328	28.6	M+Cl[-]	D-Galactose
		M+Cl[-]	myo-Inositol; D-myo-Inositol; 1D-myo-Inositol; L-myo-Inositol; 1L-myo-Inositol; meso-Inositol; Inositol; Dambose; Cyclohexitol; Meat sugar; Bios I
271.2277	192.5	M-H+O[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
291.2096	256.5	M+Cl[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
301.2384	261.0	M+HCOO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
315.2541	259.7	M+CH3COO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid

Supplemental Table 2-S6. *Mummichog* **matches for linoleate metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
137.0971	45.0	M-H2O-H[-]	NA
		M-H2O-H[-]	NA
275.2015	216.0	M-H2O-H[-]	13-OxoODE; 13-KODE; (9Z,11E)-13-Oxooctadeca-9,11-dienoic acid
		M-H2O-H[-]	NA
		M-H2O-H[-]	NA
323.2215	240.7	M+HCOO[-]	(6Z,9Z,12Z)-Octadecatrienoic acid; 6,9,12-Octadecatrienoic acid; gamma-Linolenic acid; Gamolenic acid

Supplemental Table 2-S7. *Mummichog* **matches for linoleate metabolism pathway using HILIC column with positive ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
106.0499	94.2	M+H+Na[2+]	Azelaic acid
281.2474	33.6	M+H[1+]	Linoleate; Linoleic acid; (9Z,12Z)-Octadecadienoic acid; 9-cis,12-cis-Octadecadienoate; 9-cis,12-cis-Octadecadienoic acid
281.2474	265.4	M+H[1+]	Linoleate; Linoleic acid; (9Z,12Z)-Octadecadienoic acid; 9-cis,12-cis-Octadecadienoate; 9-cis,12-cis-Octadecadienoic acid
		M+H[1+]	13(S)-HODE; (13S)-Hydroxyoctadecadienoic acid; (9Z, 11E)-(13S)-13-Hydroxyoctadeca-9,11-dienoic acid
297.2422	239.1	M+H[1+]	9(10)-EpOME; (9R,10S)-(12Z)-9,10-Epoxyoctadecenoic acid
		M+H[1+]	12(13)-EpOME; (12R,13S)-(9Z)-12,13-Epoxyoctadecenoic acid

Supplemental Table 2-S8. *Mummichog* matches for glycerophospholipid metabolism pathway using HILIC column with positive ion mode

m/z	RT (s)	Adduct Form	Tentative Match
106.0499	94.2	M+H[1+]	L-Serine; L-2-Amino-3-hydroxypropionic acid; L-3-Hydroxy-alanine; Serine
107.0532	93.6	M(C13)+H[1+]	L-Serine; L-2-Amino-3-hydroxypropionic acid; L-3-Hydroxy-alanine; Serine
184.0734	91.4	M[1+]	Choline phosphate; Phosphorylcholine; Phosphocholine; O-Phosphocholine
241.0086	71.8	M+HCOONa[1+]	sn-Glycerol 3-phosphate; Glycerophosphoric acid; D-Glycerol 1-phosphate; Glycerol-3-phosphate
281.2474	33.6	M+H[1+]	Linoleate; Linoleic acid; (9Z,12Z)-Octadecadienoic acid; 9-cis,12-cis-Octadecadienoate; 9-cis,12-cis-Octadecadienoic acid
281.2474	265.4	M+H[1+]	Linoleate; Linoleic acid; (9Z,12Z)-Octadecadienoic acid; 9-cis,12-cis-Octadecadienoate; 9-cis,12-cis-Octadecadienoic acid

Supplemental Table 2-S9. *Mummichog* matches for glycerophospholipid metabolism pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
90.0277	169.2	M(C13)-H[-]	Glycerone; Dihydroxyacetone; 1,3-Dihydroxyacetone; 1,3-Dihydroxy-2-propanone; 1,3-Dihydroxypropan-2-one
		M(C13)-H[-]	D-Glyceraldehyde
215.0328	28.6	M+Cl[-]	myo-Inositol; D-myo-Inositol; 1D-myo-Inositol; L-myo-Inositol; 1L-myo-Inositol; meso-Inositol; Inositol; Dambose; Cyclohexitol; Meat sugar; Bios I
		M+Cl[-]	Galactose
323.2215	240.7	M+HCOO[-]	(9Z,12Z,15Z)-Octadecatrienoic acid; alpha-Linolenic acid; 9,12,15-Octadecatrienoic acid; Linolenate; alpha-Linolenate
380.2556	287.8	M-H[-]	Sphinganine 1-phosphate; Dihydrosphingosine 1-phosphate
447.0477	146.7	M(S34)-H[-]	CDP-ethanolamine
		M(Cl37)-H[-]	CDP-ethanolamine
472.1108	116.7	M-H[-]	CMP-N-trimethyl-2-aminoethylphosphonate; CMP-2-trimethylaminoethylphosphonate

Supplemental Table 2-S10. *Mummichog* **matches for prostaglandin formation from arachidonate pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
370.1758	228.5	M+Cl37[-]	delta12-Prostaglandin J2
		M+Cl37[-]	NA
370.1931	87.7	M+Cl[-]	9-deoxy-delta12-PGD2
		M+HCOO[-]	NA
		M+HCOO[-]	NA
		M+HCOO[-]	Prostaglandin A2; PGA2; Medullin
379.2107	161.2	M+HCOO[-]	Prostaglandin C2; PGC2
		M+HCOO[-]	Prostaglandin B2; PGB2
		M+HCOO[-]	NA
		M+HCOO[-]	NA
		M+HCOO[-]	Prostaglandin J2; PGJ2
383.2249	203.2	M+K-2H[-]	Anandamide; Arachidonylethanolamide; N-(5Z,8Z,11Z,14Z-icosatetraenoyl)-ethanolamide
384.2449	200.3	M+Cl37[-]	Anandamide; Arachidonylethanolamide; N-(5Z,8Z,11Z,14Z-icosatetraenoyl)-ethanolamide
405.2280	202.7	M-H2O-H[-]	15-oxo-Prostaglandin E2 glyceryl ester
453.2853	249.0	M+CH3COO[-]	11-hydroxyeicosatetraenoate glyceryl ester
456.2240	169.2	M(C13)-H[-]	1-lyso-2-arachidonoyl-phosphatidate

Supplemental Table 2-S11. *Mummichog* **matches for glycine, serine, alanine and threonine metabolism pathway using HILIC column with positive ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
88.0393	97.6	M+H[1+]	2-Aminoacrylate; Dehydroalanine
106.0499	94.2	M+H[1+]	L-Serine; L-2-Amino-3-hydroxypropionic acid; L-3-Hydroxy-alanine; Serine
107.0532	93.6	M(C13)+H[1+]	L-Serine; L-2-Amino-3-hydroxypropionic acid; L-3-Hydroxy-alanine; Serine
134.0634	260.2	M(Cl37)+H[1+]	5-Aminolevulinate; 5-Amino-4-oxopentanoate; 5-Amino-4-oxovaleric acid
185.0076	285.0	M[1+]	O-Phospho-L-serine; L-O-Phosphoserine; 3-Phosphoserine; Dexfosfoserine

Supplemental Table 2-S12. *Mummichog* **matches for glycosphingolipid metabolism pathway using HILIC column with positive ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
106.0499	94.2	M+H[1+]	L-Serine; L-2-Amino-3-hydroxypropionic acid; L-3-Hydroxy-alanine; Serine
107.0532	93.6	M(C13)+H[1+]	L-Serine; L-2-Amino-3-hydroxypropionic acid; L-3-Hydroxy-alanine; Serine
184.0734	91.4	M[1+]	Choline phosphate; Phosphorylcholine; Phosphocholine; O-Phosphocholine
610.0854	26.1	M(Cl37)+H[1+]	UDP-N-acetyl-D-galactosamine
		M(Cl37)+H[1+]	UDP-N-acetyl-D-glucosamine; UDP-N-acetylglucosamine
		M(S34)+H[1+]	UDP-N-acetyl-D-galactosamine
		M(S34)+H[1+]	UDP-N-acetyl-D-glucosamine; UDP-N-acetylglucosamine
976.6181	30.2	M+HCOOK[1+]	Sulfatide; Galactosylceramidesulfate; Cerebroside 3-sulfate

Supplemental Table 2-S13. *Mummichog* **matches for glycosphingolipid metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
156.9810	136.5	M+CH3COO[-]	Sulfate; Sulfuric acid
215.0328	28.6	M+Cl[-]	D-Galactose
		M+Cl[-]	D-Glucose; Grape sugar; Dextrose
		M+Cl[-]	Galactose
		M+Cl[-]	beta-D-Galactose
271.2277	192.5	M-H+O[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
291.2096	256.5	M+Cl[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
301.2384	261.0	M+HCOO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
315.2541	259.7	M+CH3COO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
380.2556	287.8	M-H[-]	Sphinganine 1-phosphate; Dihydrosphingosine 1-phosphate

Supplemental Table 2-S14. *Mummichog* **matches for keratan sulfate degradation pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
156.9810	136.5	M+CH3COO[-]	Sulfate; Sulfuric acid
204.0860	27.6	M+ACN-H[-]	6-Deoxy-L-galactose; L-Fucose
215.0328	28.6	M+Cl[-]	D-Galactose

Supplemental Table 2-S15. *Mummichog* **matches for glycolysis and gluconeogenesis pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
90.0277	169.2	M(C13)-H[-]	(S)-Lactate; L-Lactate; L-Lactic acid
		M+Cl[-]	alpha-D-Glucose
215.0328	28.6	M+Cl[-]	D-Glucose; Grape sugar; Dextrose
		M+Cl[-]	beta-D-Glucose
227.0413	250.4	M+Na-2H[-]	Dihydrolipoamide; Dihydrothioctamide
248.0799	290.3	M-H[-]	S-Acetyldihydrolipoamide; 6-S-Acetyldihydrolipoamide

Supplemental Table 2-S16. *Mummichog* **matches for fructose and mannose metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
90.0277	169.2	M(C13)-H[-]	D-Glyceraldehyde
204.086	27.6	M+ACN-H[-]	6-Deoxy-L-galactose; L-Fucose
		M+Cl[-]	alpha-D-Glucose
215.0328	28.6	M+Cl[-]	D-Glucose; Grape sugar; Dextrose
		M+Cl[-]	D-Fructose; Levulose; Fruit sugar; D-arabino-Hexulose
		M+Cl[-]	D-Mannose; Mannose; Seminose; Carubinose

Supplemental Table 2-S17. *Mummichog* **matches for leukotriene metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
147.0493	21.4	M(C13)-H[-]	DL-Glutamate; DL-Glutaminic acid; 2-Aminoglutaric acid; Glutamate
370.1758	228.5	M+Cl37[-]	5-oxo-6E-12-epi-LTB4
		M+Cl37[-]	12-oxo-leukotriene B4
		M+Cl37[-]	5-oxo-6-trans-leukotriene B4
		M+Cl[-]	5,12-DiHETE
		M+Cl[-]	6E-12-epi-LTB4
		M+Cl[-]	10,11-dihydro-12-oxo-LTB4
370.1931	87.7	M+Cl[-]	6,7-dihydro-5-oxo-12-epi-LTB4
		M+Cl[-]	6-trans-leukotriene B4
		M+Cl[-]	6,7-dihydro-5-oxo-leukotriene B4
		M+Cl[-]	NA
456.2240	169.2	M(Cl37)-H[-]	20-Hydroxyleukotriene E4; 20-OH-LTE4; 20-OH-Leukotriene E4
		M(S34)-H[-]	20-Hydroxyleukotriene E4; 20-OH-LTE4; 20-OH-Leukotriene E4
464.1371	146.7	M+Br[-]	10,11-dihydro-20-trihydroxy-leukotriene B4

Supplemental Table 2-S18. *Mummichog* **matches for sialic acid metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
215.0328	28.6	M+Cl[-]	D-Galactose
		M+Cl[-]	myo-Inositol; D-myo-Inositol; 1D-myo-Inositol; L-myo-Inositol; 1L-myo-Inositol; meso-Inositol; Inositol; Dambose; Cyclohexitol; Meat sugar; Bios I
		M+Cl[-]	alpha-D-Glucose
		M+Cl[-]	Galactose
		M+Cl[-]	beta-D-Glucose
457.0549	174.2	M+Cl[-]	Lactose 6-phosphate
472.1108	116.7	M-H[-]	CMP-N-trimethyl-2-aminoethylphosphonate; CMP-2-trimethylaminoethylphosphonate

Supplemental Table 2-S19. *Mummichog* matches for TCA cycle pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
227.0413	250.4	M+Na-2H[-]	Dihydrolipoamide; Dihydrothioctamide
235.0085	26.9	M+HCOO[-]	Oxalosuccinate; Oxalosuccinic acid
379.2107	161.2	M+CH3COO[-]	Ubiquinol; QH2; CoQH2

Supplemental Table 2-S20. *Mummichog* matches for butanoate metabolism pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
143.0350	184.1	M+CH3COO[-]	3-Butynoate
147.0493	21.4	M(C13)-H[-]	L-Glutamate; L-Glutamic acid; L-Glutaminic acid; Glutamate
204.086	27.6	M+HCOO[-]	5-Acetamidopentanoate

Supplemental Table 2-S21. *Mummichog* matches for histidine metabolism pathway using HILIC column with positive ion mode

m/z	RT (s)	Adduct Form	Tentative Match
88.0393	97.6	M+2H[2+]	N-Formimino-L-glutamate; N-Formimidoyl-L-glutamate
111.0553	262.0	M+H[1+]	Imidazole-4-acetaldehyde; Imidazole acetaldehyde
128.0544	215.3	M(C13)+H[1+]	Imidazole-4-acetate; Imidazoleacetic acid; 4-Imidazoleacetate
156.0767	112.7	M+H[1+]	L-Histidine; (S)-alpha-Amino-1H-imidazole-4-propionic acid
157.0801	103.1	M(C13)+H[1+]	L-Histidine; (S)-alpha-Amino-1H-imidazole-4-propionic acid

Supplemental Table 2-S22. *Mummichog* matches for lysine metabolism pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
147.0493	21.4	M(C13)-H[-]	L-Glutamate; L-Glutamic acid; L-Glutaminic acid; Glutamate
197.0048	45.1	M+Cl37[-]	2-Oxoadipate; 2-Oxoadipic acid
204.0860	27.6	M+CH3COO[-]	L-2-Aminoadipate 6-semialdehyde; 2-Aminoadipate 6-semialdehyde
		M+CH3COO[-]	6-Amino-2-oxohexanoate; 2-Oxo-6-aminocaproate
227.0413	250.4	M+Na-2H[-]	Dihydrolipoamide; Dihydrothioctamide

Supplemental Table 2-S23. *Mummichog* matches for methionine and cysteine metabolism pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
113.9889	41.0	M+ACN-H[-]	hypothiocyanite
147.0493	21.4	M(C13)-H[-]	DL-Glutamate; DL-Glutaminic acid; 2-Aminoglutaric acid; Glutamate
		M(C13)-H[-]	L-Glutamate; L-Glutamic acid; L-Glutaminic acid; Glutamate
156.981	136.5	M+CH3COO[-]	Sulfate; Sulfuric acid
197.0048	45.1	M+Cl[-]	1,2-Dihydroxy-5-(methylthio)pent-1-en-3-one
248.0799	293	M-H2O-H[-]	Adenosine
287.062	291.9	M+Na-2H[-]	Adenosine

Supplemental Table 2-S24. *Mummichog* matches for purine metabolism pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
147.0493	21.4	M(C13)-H[-]	L-Glutamate; L-Glutamic acid; L-Glutaminic acid; Glutamate
227.0413	250.4	M+CH3COO[-]	Urate; Uric acid
248.0799	293.0	M-H2O-H[-]	Deoxyguanosine; 2'-Deoxyguanosine
		M-H2O-H[-]	Adenosine
249.9638	28.3	M+K-2H[-]	2-Deoxy-D-ribose 1-phosphate; 2-Deoxy-alpha-D-ribose 1-phosphate
287.062	291.9	M+Na-2H[-]	Deoxyguanosine; 2'-Deoxyguanosine
		M+Na-2H[-]	Adenosine

Supplemental Table 2-S25. *Mummichog* matches for pyrimidine metabolism pathway using HILIC column with positive ion mode

m/z	RT (s)	Adduct Form	Tentative Match
147.0493	21.4	M(C13)-H[-]	L-Glutamate; L-Glutamic acid; L-Glutaminic acid; Glutamate
227.0413	250.4	M+CH3COO[-]	Urate; Uric acid
248.0799	293.0	M-H2O-H[-]	Deoxyguanosine; 2'-Deoxyguanosine
		M-H2O-H[-]	Adenosine
249.9638	28.3	M+K-2H[-]	2-Deoxy-D-ribose 1-phosphate; 2-Deoxy-alpha-D-ribose 1-phosphate
287.062	291.9	M+Na-2H[-]	Deoxyguanosine; 2'-Deoxyguanosine
		M+Na-2H[-]	Adenosine

Supplemental Table 2-S26. *Mummichog* **matches for urea cycle/amino group metabolism pathway using HILIC column with positive ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
88.0757	224.1	M+H[1+]	4-Aminobutanal; 4-Aminobutyraldehyde; Butyraldehyde, 4-amino-
129.066	255.4	M+2H[2+]	Peptide 2-(3-carboxy-3-aminopropyl)-L-histidine; 2-(3-Carboxy-3-aminopropyl)-L-histidine; EF-2 2-(3-carboxy-3-aminopropyl)-L-histidine; Elongation factor 2 2-(3-carboxy-3-aminopropyl)-L-histidine
132.0808	47.2	M(S34)+H[1+]	N4-Acetylaminobutanal
134.0634	260.2	M(Cl37)+H[1+]	L-Glutamate 5-semialdehyde; L-Glutamate gamma-semialdehyde
190.9794	257.8	M+K[1+]	Thiopurine
271.04	73.1	M(C13)+H[1+]	N-Acetyl-L-glutamate 5-phosphate; N-Acetyl-L-glutamyl 5-phosphate
565.4028	34.9	M(Cl37)+H[1+]	3-Hexaprenyl-4,5-dihydroxybenzoate

Supplemental Table 2-S27. *Mummichog* **matches for vitamin E metabolism pathway using C18 column with negative ion mode**

m/z	RT (s)	Adduct Form	Tentative Match
370.1758	228.5	M+Cl37[-]	7'-carboxy-gama-chromanol
429.2859	174.4	M+ACN-H[-]	9'-carboxy-alpha-chromanol
452.2952	260.9	M-H[-]	13'-carboxy-alpha-tocotrienol
453.2853	249	M+ACN-H[-]	11'-carboxy-alpha-tocotrienol

Supplemental Table 2-S28. *Mummichog* matches for xenobiotics metabolism pathway using C18 column with negative ion mode

m/z	RT (s)	Adduct Form	Tentative Match
164.9273	31.9	M+Cl[-]	NA
		M(Cl37)-H[-]	Bromobenzene-2,3-oxide; Bromobenzene-2,3-epoxide
		M(Cl37)-H[-]	4-Bromophenol
		M(Cl37)-H[-]	Bromobenzene-3,4-oxide; Bromobenzene-3,4-epoxide
		M(S34)-H[-]	Bromobenzene-3,4-oxide; Bromobenzene-3,4-epoxide
172.9413	26.7	M(S34)-H[-]	Bromobenzene-2,3-oxide; Bromobenzene-2,3-epoxide
		M(S34)-H[-]	NA
		M+HCOO[-]	NA
		M(S34)-H[-]	4-Bromophenol
		M(Cl37)-H[-]	NA
271.2277	192.5	M-H+O[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
287.0620	291.9	M+Cl[-]	Benzpyrene; Benzo[a]pyrene
287.2226	223.8	M+CH3COO[-]	Tetradecanoic acid; Tetradecanoate; Myristic acid
291.2096	256.5	M+Cl[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
301.2384	261.0	M+HCOO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid
315.2541	259.7	M+CH3COO[-]	Hexadecanoic acid; Hexadecanoate; Hexadecylic acid; Palmitic acid; Palmitate; Cetylic acid

Supplemental Figure 2-S1. Type II Manhattan plots of associations between changes in maternal serum feature intensities and air pollution using the HILIC column with positive ion mode; VIP score vs retention time. Red dots represent the features that were up-regulated in the high exposed group and the green dots represent the features that were down-regulated in the high exposed group
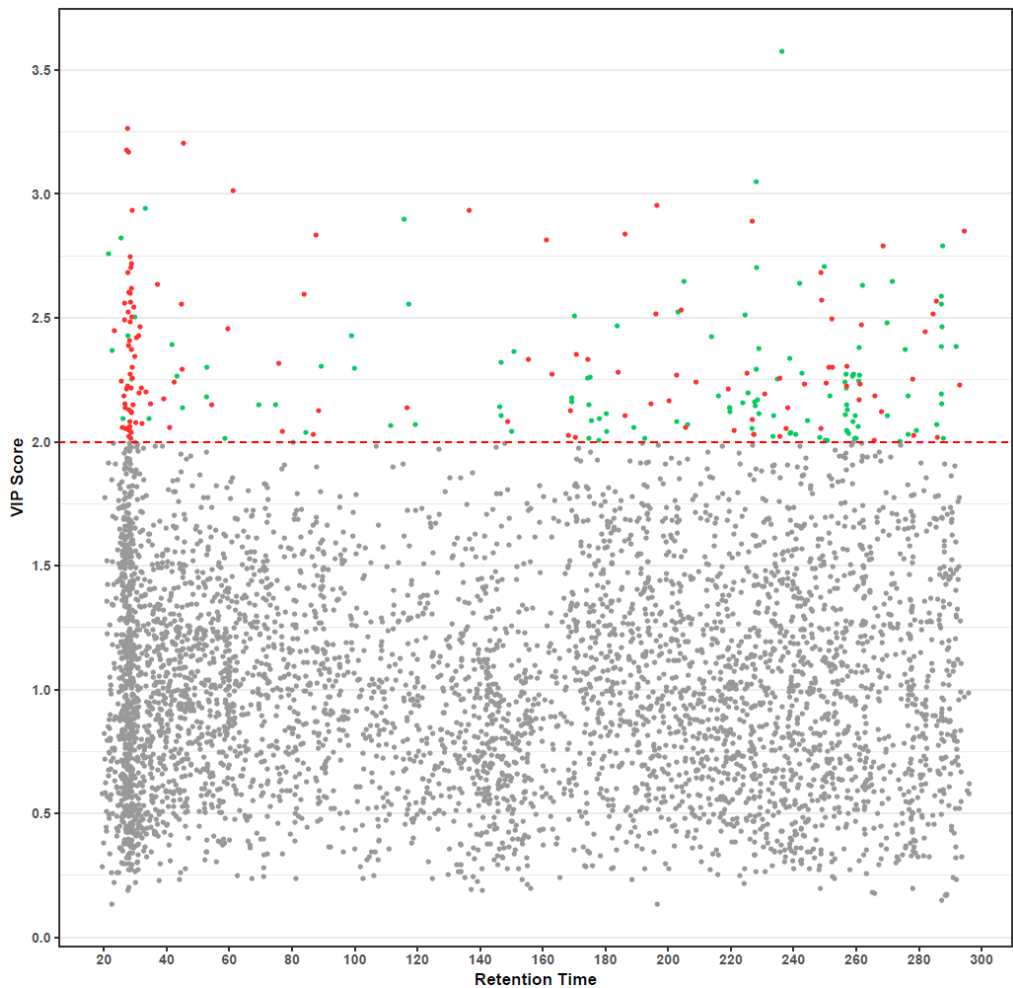
Supplemental Figure 2-S2. Type II Manhattan plots of associations between changes in maternal serum feature intensities and air pollution using the C18 column with negative ion mode; VIP score vs retention time. Red dots represent the features that were up-regulated in the high exposed group and the green dots represent the features that were down-regulated in the high exposed group.

3 High-resolution metabolomic assessment of pesticide exposure in Central Valley, California

3.1 Introduction

Agricultural pesticides are chemicals used for crop protection and combating animal pests and disease. California leads the U.S. in pesticide use, with a total of 205 million pounds of pesticides applied in 2017 (EPA 2017). In particular, the region of greatest pesticide use in California is the Central Valley area, including Fresno, Kern, and Tulare counties. Based on chemical properties, the main classes of insecticides used in California have historically included organophosphates (OPs), organochlorines (OCs), and pyrethroids (PYRs).

OPs are widely used insecticides, and they are designed to inhibit acetylcholinesterase enzyme activity, which leads to accumulation of the neurotransmitter acetylcholine and results in neurotoxic effects in the peripheral and central nervous system (Paul et al. 2018a; Terry Jr 2012; van der Plaat et al. 2018). OCs are a group of chlorinated compounds that are highly persistent organic pollutants (POPs) in the environment (Jayaraj et al. 2016). The modes of action of OC pesticides include opening of sodium ion channels in neurons (Dong 2007), as well as binding with the GABA chloride ionophore complex producing a decreased uptake of chlorine ions in neurons (Zuluaga et al. 2016). Due to the toxicity of OPs and the persistence of OCs their use in agriculture has been restricted and in recent decades other types of insecticides, including PYRs, have been used more widely (Jellali et al. 2018). The primary targets of PYRs are the membrane sodium channels but they also act on potassium, chloride, and calcium channels (Bradberry et al. 2005; Castellanos et al. 2018; Furlong et al. 2020). As a group they are known as axonic excitotoxins as they disturb sodium channels, leading to abnormal neural activity (Bradberry et al. 2005; Zuluaga et al. 2016). Furthermore, chronic exposure of OPs, PYRs, and OCs have been

48

shown to have a wide range of downstream secondary effects, including genotoxicity and DNA damage, oxidative stress, immunotoxicity, mitochondrial dysfunction, and regulation of neuronal apoptosis (Banerjee et al. 2001; Costa 2006; Hossain and Richardson 2011; Karami-Mohajeri and Abdollahi 2011; Patel et al. 2006; Zuluaga et al. 2016).

Numerous epidemiological studies, as well as in vitro and in vivo experimental evidence, suggests these three pesticide groups are associated with a wide range of adverse health effects including neurodegenerative disease, such as seen in Parkinson's disease (PD) and Alzheimer's disease (AD), diabetes, and cancers (Blair et al. 2015). OC and PYR exposures have also been associated with neurodevelopmental impairment (Ferreira et al. 2013; Furlong et al. 2017; Weisskopf et al. 2010).

As chronic pesticide exposures may have profound effects on various biological systems, characterizing metabolic response to these chemicals and how they influence human health systemically is important. High-resolution metabolomics (HRM) is a powerful analytical approach that can profile more than ten thousand endogenous and exogenous chemicals in biological specimens. Recently, HRM has been used to study the effects of environmental exposures such as metals, air pollution, tobacco smoking, and polycyclic aromatic hydrocarbons on human metabolism (Garcia-Sevillano et al. 2015; Gu et al. 2016; Jones 2016).

A large number of metabolomic studies have employed animal models or cell lines to study the toxicity of different pesticides (Du et al. 2014; Gu et al. 2019; Jellali et al. 2018; F Li et al. 2017; Xu et al. 2019; Yang et al. 2011; Yang et al. 2013). In recent years, metabolomics has also been applied to study the effect of OP, PYR, or OC pesticide exposure. For example, a cross-sectional study of 83 pregnant women characterizing the urine metabolome found that pesticide mixtures increased oxidative stress and disturbed energy metabolism (N. Bonvallot et al. 2013),

while a larger study of 750 pregnant women identified mitochondrial catabolic pathways as being associated with low-level exposure to OCs (Maitre et al. 2018). OP and OC metabolomic profiling of 102 Chinese pregnant women found OP and OC exposures may disrupt thyroid hormone metabolism and glyceraldehyde metabolism (Yang et al. 2020). A Swedish study investigated dichlorodiphenyldichloroethylene (p,p'-DDE) and hexachlorobenzene (HCB) exposure using serum metabolomic profiles from 1016 elderly participants and identified lipid metabolism as an essential metabolic response to OC pesticides (Salihovic et al. 2016). Other studies focusing on persistent organic pollutants in general indicated that POPs disturb amino acid, lipid and fatty acid, and carbohydrate metabolism (Carrizo et al. 2017; Valvi et al. 2020).

Most previous studies measured pesticide exposure levels via pesticide biomarkers in blood or in urine samples. While these biomarkers estimate exposure at the same time as the metabolome is being characterized, they may not be suitable for capturing the cumulative effects of low-level chronic past pesticide exposures or for chemicals with short biological half-lives. Also, POP exposures that occurred in the distant past may not be represented well in a recent blood draw, especially in older individuals. In addition, the majority of studies using biomarkers mostly focused on one or a few specific chemicals instead of integrating various types of pesticides. Here, we conducted an untargeted metabolomics study of exposure to three groups of pesticides (OPs, PYRs, and OCs) using metabolomes from serum samples from 176 older adults who were recruited as control subjects into a study of Parkinson's disease in the California Central Valley. Chronic pesticide exposure from ambient sources at homes and workplaces over decades was estimated with a geographic information system (GIS) based model for each type of pesticide. We utilized the same liquid chromatography high-resolution mass spectrometry (LC-HRMS) platform that has been used in several previous studies to examine the metabolomic response to environmental

toxicants including OCs (Li et al. 2019; Walker et al. 2019a). This platform has been shown to be sensitive enough to capture metabolic perturbations associated with low-level environmental exposures, a goal of our study. Furthermore, by conducting an integrative network analysis, we aim to provide a comprehensive view of the way in which these pesticides may influence the human metabolome and how this may result in potential adverse health effects.

3.2 Methods

3.2.1 Study population and samples collection

The Parkinson's Environment and Gene Study (PEG) is a community-based case-control study of Parkinson's disease (PD) etiology in agricultural regions of the California central valley that recruited control subjects from Kern, Tulare, or Fresno counties between 2001 to 2007 (Ritz et al. 2016). Controls were randomly sampled either from Medicare rolls or from property tax assessor records listing residential parcels. Participation was limited to one person per household and eligibility criteria included being 35+ years of age, having lived in one of the three counties for the past five years or more, and not having received a Parkinsonism diagnosis. More detailed information has been provided elsewhere (Kang et al. 2005).

Here, we used blood samples collected at enrollment from 176 adult controls, as well as demographic information, lifestyle and medical data collected in standardized interviews. Blood samples were collected by study staff during field visits in local clinics, centrifuged, transferred on ice, and stored in a -80°C freezer at UCLA. Serum samples were shipped frozen to Emory University on dry ice for metabolomics analyses, where they were stored at -80 ^{0}C until analyses.

3.2.2 Pesticide assessment

Pesticide exposure assessment was performed as previously described (Paul et al. 2016). Briefly, we estimated ambient pesticide exposure for each participant based on a GIS approach

51

(Cockburn et al. 2011). Participants' residential and occupational address histories were used to assess proximity to commercial agricultural pesticide applications, which are reported in the California state-mandated pesticide use reports (CA-PUR). The CA-PUR is a state-wide registry of all commercial pesticide applications since 1974, and our GIS approach linked these data to land use surveys providing locations of specific crops, and eventually to participants' geocoded residential and occupational address histories.

Based on the California Department of Pesticide Regulation (CDPR) and the pesticide action network (PAN) pesticide database, we identified individual chemicals that were classified as one of three types of pesticides: organophosphates (OP), pyrethroids (PYR), and organochlorines (OC). Specifically, 36 different chemicals were classified as OPs, 14 chemicals as PYRs, and 8 chemicals as OCs. We calculated the total pounds of each chemical applied within 500 m of the addresses annually from 1974 to the year of blood draw, and then averaged the pounds per acre for each pesticide across residential and occupational addresses over the entire period. Exposures at both residential and occupational addresses were included, and each participant could have been exposed at both locations, only one, or neither. To calculate the total number of individual OP, PYR, and OC pesticides a participant had been exposed to, we dichotomized exposure to each of the individual chemicals based on each chemical's median exposure level, and then summed the number of chemicals that each participant was exposed to above the median, counting chemical exposures at both residences and workplaces.

3.2.3 High-resolution metabolomics

High-resolution metabolomics (HRM) profiling was completed according to established methods (KH Liu et al. 2020). Briefly, serum samples were collected during the interview and stored at -80 °C. Prior to HRM, batches of 40 serum samples were removed from storage and

thawed on ice. Each sample was then thoroughly vortexed, and plasma proteins were precipitated by diluting 65 μL of serum with 130 μL of LC-MS grade acetonitrile. We previously compared extraction efficiency using different volume equivalents of acetonitrile and methanol (Johnson et al. 2010). Comparison of protein removal using the Lowry method and by SDS-PAGE with Coomassie blue staining and densitometry showed a 2:1 ratio of acetonitrile removed 98% of plasma proteins, and has been used for all subsequent HRM studies. The extract was centrifuged, and the resulting supernatant was transferred to an autosampler vial containing a low volume insert and maintained at 4ºC until analysis (<24 h). To evaluate system performance, we used two separate quality assessment methods. Our first QC sample was NIST 1950 (Simon-Manso et al. 2013) which was analyzed at the beginning and end of the entire analytical run. The second QC sample (Q-Std) included commercially purchased plasma pooled from an unknown number of males and females. Q-Std was analyzed at the beginning, middle, and end of each batch of 40 samples for normalization and batch effect evaluation.

Sample extracts were analyzed in triplicate using a dual-column, dual-polarity approach that includes hydrophilic interaction (HILIC) chromatography with positive ESI and C18 chromatography with negative ESI (Ultimate 3000, Q-Exactive HF, Thermo Fisher, m/z range 85-1275) (Walker et al. 2018). We chose HILICpos and C18neg because dual chromatography HRM with positive and negative electrospray ionization provides the optimal number of reproducible ions detected and matches to known chemicals in the KEGG Human Metabolite database (Liu et al. 2016). Following a 10μL sample injection, HILIC separation was accomplished using a 2.1 cm × 5 cm × 2.5 μm HILIC column (Waters XBridge BEH Amide XP HILIC) and acetonitrile gradient (A = water, B = acetonitrile, C = 2% formic acid) consisting of an initial 1.5 min period of 22.5% A, 75% B, 2.5% C, followed by linear increase to 77.5% A, 20% B, 2.5% C at 4 min and hold for

1 min. Separation by C18 was with 2.1 cm × 5 cm × three μm column (Higgins endcapped C18) with C = 10 mM ammonium and the following gradient: initial 0.5 min period of 60% A, 35% B, 5% C, followed by linear increase to 0% A, 95% B, 5% C at 1.5 min and then held for an additional 3 min. Mobile phase flow rate was held at 0.4 mL/min for 1.5 min and then increased to 0.5 mL/min. The mass spectrometer was operated using ESI mode at a resolution of 120,000 and mass-to-charge ratio (m/z) range 85-1275. High-resolution detection of m/z features was accomplished by a maximum injection time of 10 milliseconds and an AGC target of 1×10^6. Raw data files were extracted and aligned using *apLCMS* (Yu et al. 2009) with modifications by *xMSanalyzer* (Uppal et al. 2013). Uniquely detected ions consisted of m/z, retention time (rt), and ion abundance, referred to as m/z features. Prior to data analysis, m/z features were batch corrected using a novel algorithm based on wavelet (Deng et al. 2019). For further analyses, we only included metabolomic features detected in > 25% of all plasma samples, with median coefficients of variation (CV) among technical replicates < 30% and Pearson correlation > 0.7. Following quality assessment, intensities of three replicates for each feature were summarized using the median value. In addition, we conducted a log2 transformation and auto-scaling. Missing values were imputed using k-nearest neighbors (k=10) (Troyanskaya et al. 2001) imputed in the *impute* R package.

3.2.4 Metabolome-wide association analysis

We are presenting the distributions of three pesticide classes and Pearson correlations for levels of pesticide exposure. All analyses were performed using R version 4.0.1.

We conducted partial least squares (PLS) regression to identify metabolomic features associated with each pesticide group (i.e., OPs, PYRs, and OCs). PLS regression is a supervised, multivariate analysis approach for dimensionality reduction that maximizes covariance between intensities of metabolomic features and exposures (Wold et al. 2001). To adjust for potential

confounders, we regressed each feature's intensity on potential confounding variables (age, sex, and race/ethnicity) and formed residuals as the input matrix. Features with Variable Importance in Projection (VIP) scores > 2 were selected as significant. Ten-fold cross-validation was used to assess the performance of the selected features. We further applied linear regression to test for linearity between m/z features and pesticides.

To assess the potential confounding effect of socioeconomic status, we conducted a sensitivity analysis by additionally adjusting for education.

To gain a holistic view of the relationship between pesticides and the serum metabolome, we further conducted integrated network analysis using *xMWAS* (Uppal et al. 2018). *xMWAS* calculates the pairwise association scores (approximation of the correlation coefficient) between each metabolomic feature and each pesticide using partial least squares (Le Cao et al. 2008), and generates a multi-data integrative network. Only associations with both the |association score| > 0.2 and $p < 0.05$ (by Student's t-test) were included in the network. Finally, a multilevel community detection algorithm (Blondel et al. 2008) was used to identify metabolite and exposure clusters that are tightly connected with each other, but sparsely connected with the rest of the network.

3.2.5 Annotation and pathway analysis

HRM provides accurate mass (±5 parts-per-million; ppm) measures of ion m/z, which can be related to chemical monoisotopic mass. Significant features selected by PLS regression were first matched to a reference database of authenticated chemical standards previously characterized in our laboratory. Metabolites in the in-house database were confirmed using MS/MS and authentic standards; the detailed process has been published previously (Go et al. 2015; KH Liu et al. 2020). The error tolerance was 5 ppm and 30 seconds for m/z and retention time, respectively.

Additional m/z features not matching these metabolites were then annotated by *xMSannotator* (Uppal et al. 2017). Accurate mass m/z for adducts formed under positive/negative ESI mode was matched to the Human Metabolome Database (HMDB), KEGG, and LipidMaps with a mass error threshold of 10 ppm. xMSannotator also considers the correlation of intensities and retention time and assigns confidence scores based on a multilevel scoring algorithm (0-3, a higher score representing higher-confidence result), ensuring annotation accuracy. Only results with an annotation score >2 were kept. The metabolite identification confidence levels (Schrimpe-Rutledge et al. 2016) were reported for all annotation results.

We conducted pathway enrichment analysis utilizing *mummichog* version 2 (Nathalie Bonvallot et al. 2013) to identify perturbed metabolic pathways associated with pesticide exposure for each class of pesticides. All features previously selected by PLS regression with VIP ≥2, and features included in the integrated network analysis were included in this pathway enrichment analysis. *Mummichog* uses a permutation-based framework that accounts for the complexity of untargeted mass spectral data. Although annotation results in *mummichog* may include false positives, the enriched pathways inferred by the algorithm have been shown to be valid and to reflect real biological activity (Uppal et al. 2016). All metabolites annotated by *mummichog* were required to present in at least their primary adduct (M+H or M-H for positive and negative mode, respectively) to reduce the false positive match rate. A pathway was considered significant if gamma adjusted *p-values* were smaller than 0.05. Only pathways that contained at least three discriminative metabolites were interpreted.

3.3 Results

3.3.1 Study population

Demographics of the 176 adult controls unaffected by Parkinson's disease and the distribution of pesticide exposures are provided in Table 3-1. Lists of chemicals classified as OPs, PYRs, or OCs can be found in Supplemental Table 3-S1. Approximately half of the participants included in the study were male (51%), and a majority white (86%) with a mean age of 66 years. Participants who did not have metabolomics data available did not differ from participants we included in terms of age, sex, ethnicity, or pesticide exposure. OP exposure counts ranged from 0 to 41; PYR from 0 to 10; and OC from 0 to 8. Exposure levels of pesticides in the 3 classes were moderately correlated (Supplemental Figure 3-S1).

3.3.2 Metabolome-wide analysis

We detected 12,925 metabolomic features from the HILIC column coupled with positive ionization mode (HILICpos) and 7,209 metabolomic features from the C18 column coupled with negative ionization mode (C18neg). After quality control steps, a total of 16,510 metabolomic features (10,959 HILICpos and 5,551 C18neg) were included in the analyses.

After adjusting for potential confounding variables (age, sex, race/ethnicity), PLS regression selected 389 metabolomic features (254 HILICpos features and 135 C18neg features) across three components that were associated with OP exposure (Supplemental Figure 3-S2). Among these, 226 features (58%) were positively associated with OP exposure. Furthermore, 517 metabolomic features (331 HILICpos features and 186 C18neg features) were associated with PYR pesticide, and 233 features (45%) increased with increasing levels of PYR exposure (Supplemental Figure 3-S2). We also identified 485 features (311 HILICpos features and 174 C18neg features) related to OCs. Approximately half of the features (229 features, 47%) were positively associated with OCs (Supplemental Figure 3-S2). A large proportion of metabolomic

features were uniquely associated with just one of the pesticide groups, but 72 statistically significant features were shared across all three pesticide classes (Figure 3-1).

Based on *xMSannotator*, in total, 15.9% of the HILIC pos features and 16.2% C18 neg features matched compounds in the HMDB, KEGG, and LipidMaps databases. We matched metabolomic features related to at least one of the pesticide groups to known metabolites based on authenticated chemical standards verified by tandem mass spectrometry (identification confidence level 1). In total, we confirmed the identities of 12 metabolites (Table 3-2), including metabolites involved in the carnitine shuttle pathway and beta-alanine metabolism.

Using *mummichog*, we examined whether the features that were selected by PLS regression were enriched within specific metabolic pathways. The result indicated that 33 metabolic pathways were significantly enriched with a *p-value* smaller than 0.05 (Figure 3-2, Supplemental Table 3-S2). List of putatively annotated metabolites within each pathway can be found in Supplemental Table 3-S3. The majority of pathways were uniquely associated with exposure to one of the pesticide groups, whereas beta-alanine metabolism and glycine, serine, alanine and threonine metabolism were linked with all three classes i.e., OPs, PYRs, and OCs. Carnitine shuttle, glutamate metabolism, glycolysis and gluconeogenesis, butanoate metabolism, and pyruvate metabolism were related to both OPs and OCs. The enriched pathways are involved in a wide range of metabolic functions, including lipid metabolism, central carbon metabolism, amino acid metabolism, neurotransmitter precursors, cofactors, and nucleotide metabolism.

To assess the potential for additional confounding, we adjusted for education information (Supplementary Figure 3-S3) which can serve as a surrogate for socioeconomic status. Enriched pathways identified in the sensitivity analysis were very similar to the original results.

3.3.3 Integrated network analysis

To identify systemic metabolic alterations associated with different pesticides, we adopted a network-based approach to integrate metabolomic features with the three pesticide classes. In total, 383 features were included in the network, with an association score greater than 0.2 and a $p < 0.05$ (Figure 3-3). Unlike the previous MWAS approach in which most of the features identified were uniquely associated with one pesticide, this network was more densely connected, i.e., 187 out of 383 features within the network were correlated with all three pesticides. We identified two separate clusters using a community detection algorithm, including one cluster centered on OPs and a second cluster that included PYRs and OCs (Figure 3-3). Both clusters included pathways related to fatty acid beta-oxidation. In addition, the OP cluster consisted of several pathways involved in amino acid metabolism and neurotransmitter precursors (Supplemental Table 3-S4 to Supplemental Table 3-S6).

3.4 Discussion

Our community-based study provided us with a unique opportunity to link GIS derived pesticide exposure data with a high-resolution metabolomics approach thus allowing the external and internal exposome to inform on each other. Specifically, we first used a complex GIS approach to combine almost a half century of California pesticide application records spatially with home and work address histories to generate measures of long-term pesticide exposure from agricultural applications in older adults. Then we combined these exposure histories with data from high-resolution metabolomics to interrogate links between complex patterns of pesticide exposure with metabolic responses. Our approach provided us with the opportunity to measure more than 10,000 metabolomic features in serum samples simultaneously, and we found that chronic exposure to pesticides in three classes – OPs, PYRs, or OCs – seemed to perturb a wide range of metabolic pathways. Employing MWAS for these three pesticide groups as well as an integrative network

approach we identified metabolic pathways distinctly associated with only one group of pesticides, as well as metabolic response patterns shared across the pesticide classes that included fatty acid metabolism, central carbon metabolism, and amino acid metabolism. Specifically, we found that the beta-alanine metabolism, and glycine, serine, alanine and threonine metabolism, were related to all three pesticides. Also, several fatty acid metabolic pathways were associated with both OPs and OCs. Network analysis showed that the fatty acid beta-oxidation pathway was a common pathway shared by all three pesticide groups. More generally, the alterations we observed in these pathways suggested that chronic pesticide exposure may result in oxidative stress, inflammatory reactions, and mitochondrial dysfunction. These findings are informative as various pesticide-related health outcomes, such as neurodegenerative disorders, diabetes, and cancer are related to these pathways (Barnham et al. 2004; Jenner 2003; Lin and Beal 2006; McGeer and McGeer 2004; Parrón et al. 2011).

Previous epidemiological and experimental studies have shown that chronic exposure to OPs, PYRs, and OCs result in various secondary toxic effects, including mitochondria dysfunction and oxidative stress (Karami-Mohajeri and Abdollahi 2011; Parrón et al. 2011; Zuluaga et al. 2016). In our study, we identified alterations in oxidative stress-induced fatty acid oxidation in peroxisomes as being associated with OP and PYR exposures, and glycerophospholipid metabolism as being associated with OC pesticide exposures. Based on our network analysis, di-unsaturated fatty acid beta-oxidation was related to all three exposure types. This finding is consistent with previous findings that abnormal fatty acid metabolism and glycerophospholipid metabolism is correlated with persistent organic pollutant exposure (Carrizo et al. 2017; Valvi et al. 2020). A study that focused on polybrominated and polychlorinated biphenyl POPs also found alterations in fatty acid metabolism, including fatty acid activation and glycerophospholipid

metabolism (Walker et al. 2019a). Moreover, the same fatty acid metabolism pathways were associated with other environmental exposures such as trichloroethylene, benzo[a]pyrene, or air pollution (Walker et al. 2016a; Walker et al. 2016b; Yan et al. 2019). Oxidative stress and disturbed fatty acid metabolism are thought to underlie many adverse health outcomes such as heart disease, diabetes, and neurodegenerative diseases including Alzheimer's disease (AD) and Parkinson's disease (PD) (Fillmore et al. 2014; Foley 1992; Lane et al. 2015). The pathophysiology of PD involves the depletion of dopamine neurons, which is a neuronal population that is particularly susceptible to oxidative stress owing to its low antioxidant capacity (Gangemi et al. 2016).

Pesticide-induced oxidative stress may lead to mitochondrial dysfunction and we observed enrichment of several pathways associated with mitochondrial energy metabolism. For example, the carnitine shuttle, short-chain fatty acid (butanoate) metabolism, glycolysis and gluconeogenesis pathway, and pyruvate metabolism were perturbed with higher OPs and OCs exposure. Fatty acid oxidation in the peroxisome was associated with both OPs and PYRs. Although the carnitine shuttle pathway was not significantly enriched with the PYR class pesticides, several acylcarnitines, including lauroylcarnitine and octanoylcarnitine were significantly decreased when PYR exposure levels increased. The carnitine shuttle refers to the process of transporting long-chain fatty acetyl coenzyme A (acyl-CoA) into mitochondria as acylcarnitines for fatty acid oxidation (Reuter and Evans 2012). This process is critical for energy supply in every tissue. Studies have shown that plasma acylcarnitine levels can serve as an indicator of mitochondrial function (Jarrell et al. 2020; Peng et al. 2018). Mitochondrial dysfunction contributes to diabetes, cardiovascular disease, cancer, metabolic syndrome, and chronic neurodegenerative disorders such as Parkinson's disease and Huntington's disease (Nicolson 2007; Soane et al. 2007; Swerdlow 2011).

61

There is also increasing evidence for immunotoxicity of certain pesticides (Furlong et al. 2020; Gangemi et al. 2016). In our study, several inflammation-related pathways, including histidine metabolism, arachidonic acid metabolism, and prostaglandins metabolism, were associated with OPs and PYRs. Histidine has anti-inflammatory effects and is negatively associated with inflammation and oxidative stress (Liang et al. 2018). Oxidative stress can also induce the activation of phospholipase A2 (PLA2) and generate polyunsaturated free fatty acid (PUFA), including linoleic acid and arachidonate acid (Anthonymuthu et al. 2018). Arachidonic acid is an inflammatory intermediator and can be converted to prostaglandins, which has a major proinflammatory effect. In accordance with our results, previous studies also linked inflammation-related pathways to environmental toxicants (Banks and Lein 2012).

Perturbed glutamate metabolism was related to increased OPs and OC exposures, which was consistent with the results obtained in some animal models (Li et al. 2014; Torres-Altoro et al. 2011). Glutamate is a precursor for glutathione, the second major thiol redox couple for antioxidant defense. Moreover, glutamate serves as an excitatory neurotransmitter in the central nervous system. Previous studies have demonstrated a link between glutamate-mediated excitotoxicity and degeneration of dopamine neurons (Meredith et al. 2009; Sonsalla et al. 1998), and glutamate may play a role in the development of PD. Furthermore, we also observed an association between tyrosine metabolism and OP exposure and tyrosine is a precursor of dopamine i.e., tyrosine is converted to L-dopa by tyrosine hydroxylase (Walker et al. 2019a). Previous studies suggested that alterations of tyrosine hydroxylase can contribute to neurodegenerative diseases, including PD and AD because it affects the biosynthesis of dopamine (Priyadarshini et al. 2012; Tabrez et al. 2012). Studies also linked these pesticides to an increased risk of PD and other

neurodegenerative disorders. Our data suggest that dysregulation of these neurotransmitter precursors is tentatively associated with OPs or OCs.

By utilizing an integrative network analysis, we obtained a holistic view of interactions between the metabolome and various pesticides. Even though there were distinct mechanisms of action for each pesticide class, most metabolic connections were shared across pesticide groups. A similar overlap of metabolomic responses has been described before for dichlorodiphenyltrichloroethane (DDT), poly and perfluoroalkyl substances (PFAS), oxychlordane, hexachlorobenzenes, and several polychlorinated biphenyl compounds; specifically, the metabolome-exposure network converged on oxidative stress, fatty acid metabolism, and mitochondrial energy metabolism (Li et al. 2019). Our findings and those of previous studies, might suggest general metabolic mechanisms an organism employs in response to the stress induced by various environmental toxicants.

Two amino acid metabolism pathways, beta-alanine metabolism, and glycine, serine, alanine and threonine metabolism, were found to be associated with all three pesticide groups. Beta-alanine is a non-essential amino acid the human body makes by converting pyruvate (Li et al. 2014). Alanine has various biological functions in different tissues. For example, beta-alanine functions as a neurotransmitter or a neuromodulator in the brain, while in skeletal muscle, it is a major energy source and a component of carnosine (Shetewy et al. 2016; Tiedje et al. 2010). Using authentic standards, we were able to identify both alanine and carnosine within the beta-alanine metabolism pathway. Alanine increased with higher levels of OPs and OCs, and carnosine was significantly and positively associated with OPs and PYRs. Similarly, increased alanine levels have been found in the brain of goldfish after exposure to PYRs (Li et al. 2014). Rats exposed to OP pesticides expressed an abnormal level of alanine aminotransferase due to disturbances in the

oxidative-reductive hepatocyte system (Lukaszewicz-Hussain and Moniuszko-Jakoniuk 2005). A human study found elevated urine alanine levels in pesticide applicators compared to non-applicators (Ch et al. 2019). Carnosine can function as an antioxidant that scavenges ROS (Baraniuk et al. 2013; Boldyrev et al. 1987). Similarly, glycine and serine also act as a cytoprotective agent (Ch et al. 2019). Therefore, the increased activity in alanine, glycine, and serine metabolism may indicate the activation of antioxidative defense mechanism in response to pesticide-induced oxidative stress, as has been shown in previous metabolomics studies of pesticide mixtures (N. Bonvallot et al. 2013). Vitamin C is a water-soluble antioxidant (Padayatty et al. 2003) and disturbances of vitamin C metabolism may also indicate responses to oxidative damage caused by OP exposure. Finally, there are several other amino acid metabolism and cofactor pathways related to OP pesticide exposure that are involved in oxidative stress and inflammation. Methionine is an essential amino acid that promotes ROS production, in line with chronic OP exposure inducing oxidative stress.

Our study has some limitations. Our GIS model-based chronic pesticide exposure measure allowed us to investigate low level exposure that occurred over decades; however, it is challenging to estimate the absolute exposure levels for each chemical we investigated. In addition, our pesticide measure does not necessarily translate into total pesticide exposure levels as other sources of pesticide exposures may also contribute. Despite the inherent challenges in GIS model-based exposure estimation, our pesticide measurements have been associated with a number of prior health outcomes such as PD, autism spectrum disorders, and neural tube defects (Rull et al. 2006; Shelton et al. 2014; A Wang et al. 2011). Also, our previous epigenomic studies of OPs and PYRs identified critical genetic pathways consistent with the pathophysiology of each pesticide type of action, in support of the validity of our exposure assessment model. Diet may also explain

differences in metabolites, and we did not have information on the participants' dietary intake; however, in order for dietary differences to have confounded our analyses, they would have had to be related to agricultural pesticide exposure, which would be more likely for dietary pesticide sources rather than ambient exposures due to agricultural applications near homes and workplaces. By controlling for age, sex, and race/ethnicity, we hope to have at least partially addressed potential confounding. Another limitation is the annotation of metabolomic features detected using untargeted analysis. Adopting a pathway and network analysis approach, we were able to improve annotation results, but there may still be incorrect matches that could have influenced the interpretation of our results. It is recommended to improve metabolite identification using either tandem MS or internal standards in future studies.

3.5 Conclusion

In summary, we utilized a high-resolution metabolomic approach to identify perturbations in the serum metabolome of older adults in response to long-term ambient OP, PYR, and OC exposures measured via a complex GIS approach and pesticide application records. We identified disturbances in metabolic pathways related to oxidative stress, inflammation, lipid and fatty acid metabolism, mitochondrial energy metabolism, and neurotransmitter precursors. Furthermore, by adopting integrative network analysis, we illustrated that different pesticides might share similar biochemical fingerprints at the metabolomic response level. Collectively, our study points to potential common molecular mechanisms of chronic pesticide toxicity, and thus may also shed light on the physiology these exposures affect that eventually can lead to various adverse pesticide-related health outcomes. Most importantly, our approach for the first time informs the field of exposome research by moving from macro-level population exposures to micro-level biologic

responses thereby generating hypotheses about toxicological mechanisms that can be studied in vivo.

3.6 Tables and figures

Table 3-1. Distribution of demographics and pesticide levels

	PEG study community controls (n = 176)
Age	
Min	35
Max	92
Mean (SD)	66.13 (13.40)
Sex	
Male (%)	90 (51)
Female (%)	86 (49)
Race/Ethnicity	
White (%)	159 (90)
Hispanic (%)	17 (10)
Education	
Less than high school	20 (11)
High school	32 (18)
More than high school	124 (71)
Organophosphate count	
Min	0
Max	41
Mean (SD)	7.58 (8.46)
Pyrethroid count	
Min	0
Max	10
Mean (SD)	0.94 (1.70)
Organochlorine count	
Min	0
Max	8
Mean (SD)	1.23 (1.72)

Table 3-2. Confirmed chemical identity of metabolomic features associated with OP, PYR, or OC pesticide

Name	KEGGID	m/z	time	OP Coefficient	OP P-value	PYR Coefficient	PYR P-value	OC Coefficient	OC P-value
L-histidine	C00135	154.0623	50.63	0.014	0.031	0.091	0.005	0.060	0.067
Lauroylcarnitine c12	CA1406	344.279	48.7	-0.013	0.070	-0.077	0.027	-0.067	0.055
Alanine	C00041	90.055	97.36	0.014	0.044	0.059	0.072	0.065	0.050
Octanoylcarnitine c8	C02838	288.2166	52.25	-0.011	0.114	-0.068	0.037	-0.068	0.039
Sphinganine	C00836	302.3049	46.12	-0.018	0.024	-0.044	0.255	-0.122	0.001
Cytidine 2,3-cyclic monophosphate	C02354	306.0489	118.2	-0.017	0.041	-0.039	0.338	-0.080	0.051
Reichsteins substance S	C05488	347.2207	47.92	0.016	0.078	0.091	0.043	0.087	0.054
Pyruvate	C00186	87.0087	38.65	0.019	0.032	0.064	0.131	0.065	0.133
O-acetyl-L-carnitine	C02571	204.123	70.68	-0.012	0.162	-0.069	0.107	-0.088	0.041
L-methionine	C00073	148.0439	47.27	0.012	0.156	0.100	0.013	0.053	0.197
Carnosine	C00386	227.1136	110.36	0.012	0.121	0.095	0.012	0.034	0.374
Omega-hydroxydodecanoic acid	C08317	217.1798	20.12	0.022	0.011	0.057	0.167	0.028	0.498

Figure 3-1. Venn diagram of pesticides-related HILICpos features (**a**) and C18neg features (**b**).
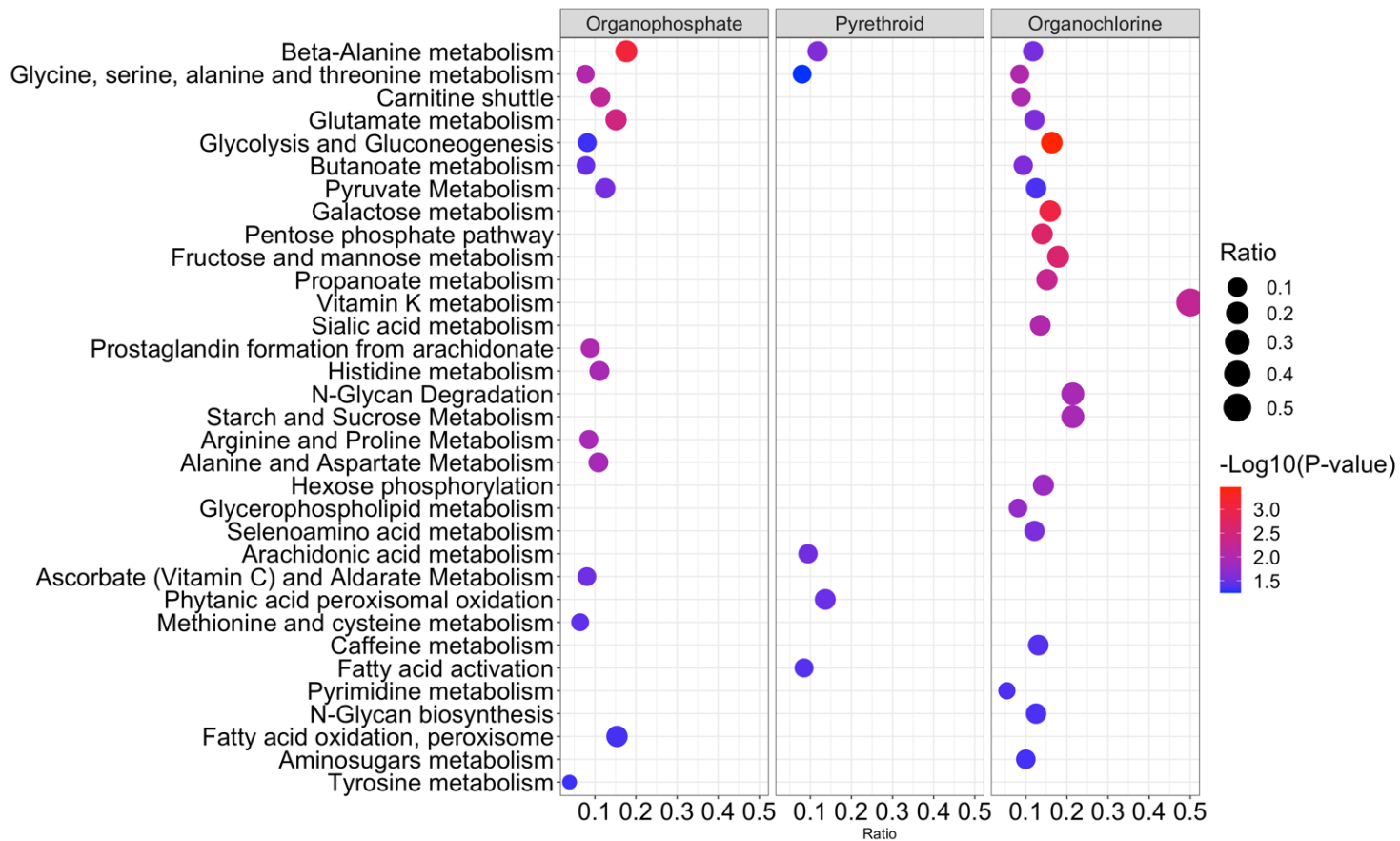
Figure 3-2. Metabolic pathways correlated with pesticides from a serum metabolome-wide association study. The vertical axis represents the pathways associated with pesticides. The circle radius is proportional to the number of correlated metabolite features within each pathway (ratio). The horizontal axis also represents the ratio. The color represents the negative log10 (*p-value*) of each pathway.
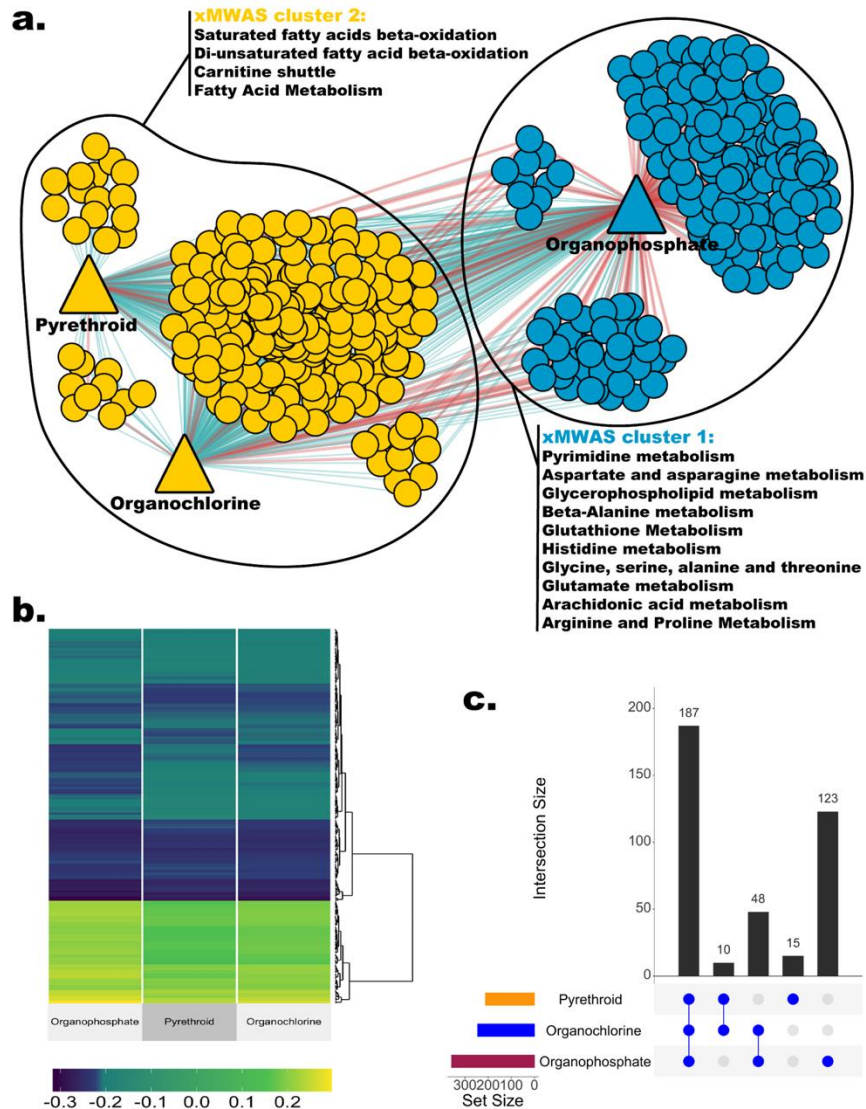
Figure 3-3. Integrative network analysis with xMWAS. (**a**) Metabolome-pesticide network revealed 2 metabolomic communities. Only metabolomic features with an association score > 0.2 and $p < 0.05$ were included in the network. Round nodes represent metabolomic features, and triangle nodes represent three pesticides. The edges' color represents the direction of association (red: positive, green: negative). Multilevel community detection algorithm identified clusters, represented by different node colors. (**b**) Heatmap of the association scores between metabolomic features and pesticides. Only features included in the network are shown here. (**c**) UpSet plot shows the number of features associated with each pesticide group and the intersection size.

71

3.7 Supplemental materials

Supplemental table 3-S1. List of chemicals within OP, PYR, or OC groups

Pesticide group	Name	Chemcode
OPs	Monocrotophos	52
	Bensulide	70
	Dicrotophos	72
	Trichlorfon	88
	Carbophenothion	110
	Ddvp	187
	S,S,S-Tributyl Phosphorotrithioate	190
	Dioxathion	192
	Diazinon	198
	Dimethoate	216
	Disulfoton	230
	Chlorpyrifos	253
	Ethion	268
	Merphos	293
	Azinphos-Methyl	314
	Phosmet	335
	Malathion	367
	Oxydemeton-Methyl	382
	Methyl Parathion	394
	Naled	418
	Parathion	459
	Phorate	478
	Phosalone	479
	Mevinphos	480
	Phosphamidon	482
	Sulfotep	558
	Demeton	566
	Tepp	577
	Ethephon	1626
	Leptophos	1676
	Acephate	1685
	Methidathion	1689
	Methamidophos	1697
	Dialifor	1799
	Fenamiphos	1857
	Profenofos	2042
PYRs	Fenvalerate	1963
	Permethrin	2008
	Phenothrin	2093
	Resmethrin	2119
	Flucythrinate	2168
	Cypermethrin	2171
	Tau-Fluvalinate	2195

	Cyfluthrin	2223
	Fenpropathrin	2234
	Lambda-Cyhalothrin	2297
	Bifenthrin	2300
	Esfenvalerate	2321
	Tralomethrin	2329
	(S)-Cypermethrin	3866
OCs	Chlordane	130
	Dieldrin	210
	Endosulfan	259
	Dicofol	346
	Lindane	359
	Methoxychlor	384
	Dienochlor	468
	Toxaphene	594

Supplemental Table 3-S2. *Mummichog* annotated metabolites within significantly enriched pathways

Pathway	Name	Adduct	M/Z	Time	VIP_OP	Coef_OP	VIP_PYR	Coef_PYR	VIP_OC	Coef_OC	Metabolite Annotation Confidence
Tyrosine metabolism	3-(4-Hydroxyphenyl)pyruvate	M+2H[2+]	91.0275	289.3	2.97	-0.020	4.06	-0.118	2.11	-0.065	3
	3-Methoxy-4-hydroxyphenylacetaldehyde	M[1+]	166.0645	261.2	2.04	-0.009	2.05	-0.039	1.72	-0.034	3
	dGDP	M(C13)+2H[2+]	215.0234	252.1	2.20	-0.017	1.94	-0.065	2.24	-0.081	3
	3'-monoiodo-L-thyronine	M(C13)+H[1+]	401.0086	61.7	2.61	-0.024	0.48	-0.017	2.26	-0.098	3
	3,5,3',5'-tetraiodo-L-thyronine-beta-D-glucuronoside	M+HCOONa[1+]	1020.704	59.7	2.57	0.008	1.80	0.022	2.42	0.034	3
	3-Methoxy-4-hydroxyphenylethyleneglycol	M+2H[2+]	93.0448	134.6	2.11	0.013	2.83	0.073	2.02	0.056	3
Nitrogen metabolism	L-Asparagine	M+H[1+]	133.0603	131.5	1.73	0.011	2.35	0.062	1.22	0.032	3
	L-Asparagine	M(C13)+H[1+]	134.064	112.7	1.58	0.010	2.14	0.054	1.14	0.028	3
	L-Asparagine	M+Na[1+]	155.0427	114.9	1.99	0.011	2.03	0.045	1.58	0.038	3
Vitamin B6 (pyridoxine) metabolism	Pyridoxal	M[1+]	167.0566	116.3	1.12	0.010	2.17	0.085	1.05	0.043	3
	Pyridoxamine	M[1+]	168.09	70.4	2.03	0.014	3.31	0.098	1.57	0.047	3
Pyrimidine metabolism	dTTP	M+2H[2+]	241.9996	178.9	3.10	-0.032	3.46	-0.163	2.29	-0.113	3
	ITP	M+2H[2+]	254.9993	97.3	1.02	-0.009	1.22	-0.046	2.02	-0.083	3
	CDP	M(C13)+H[1+]	405.025	296.9	1.12	-0.011	0.49	-0.022	2.39	-0.118	3
Carnitine shuttle	Tetradecanoylcarnitine	M+H_[+1]	373.3137	48	1.82	-0.017	1.50	-0.063	2.22	-0.099	3
	clupanodonyl carnitine	M+Na[1+]	496.3398	58.1	2.42	0.009	2.06	0.030	2.05	0.030	3
	tetracosapentaenoyl carnitine	M+Na[1+]	524.3714	57.1	2.16	0.009	1.66	0.024	2.07	0.035	3
	octadecenoyl carnitine	M+Na-2H[-]	445.3166	206	2.60	0.031	2.73	0.149	1.37	0.077	3
	stearidonyl carnitine	M+K-2H[-]	455.242	191.9	2.32	-0.016	3.05	-0.093	0.98	-0.030	3
	stearidonyl carnitine	M+K-2H[-]	455.2468	52	2.03	-0.021	1.23	-0.056	2.21	-0.107	3

	Linoelaidyl carnitine	M+K-2H[-]	459.2725	114.9	2.19	0.021	0.75	0.024	2.42	0.108	3
	cervonyl carnitine	M+K-2H[-]	507.2716	156.6	2.15	0.015	1.56	0.046	2.31	0.075	3
	pentadecanoyl Coenzyme A	M+Na[1+]	1010.28	77	2.15	0.018	2.66	0.099	2.79	0.112	3
Alanine and Aspartate Metabolism	Pyruvic acid	M-H	87.0087	38.6	2.02	0.019	1.54	0.064	1.67	0.065	1
Aminosugars metabolism	3-Oxopropanoate	M-H+O[-]	103.0037	128.1	1.92	-0.016	1.41	-0.053	2.66	-0.104	3
	N-Glycoloyl-neuraminate	M+Cl37[-]	362.0636	37	0.84	-0.007	1.69	-0.064	3.83	-0.150	3
Arachidonic acid metabolism	Glycerol	M+CH3COO[-]	151.0612	275	1.69	-0.014	2.27	-0.086	0.87	-0.031	3
	12 hydroxy arachidonic acid	M+CH3COO[-]	378.2392	241.7	1.82	-0.012	2.12	-0.068	1.81	-0.057	3
Arginine and Proline Metabolism	L-Methionine	M+Cl[-]	184.0194	36.6	2.12	0.015	3.68	0.111	1.68	0.051	3
	N-(L-Arginino)succinate	M+Cl37[-]	327.0905	36.5	2.31	-0.017	1.73	-0.053	1.38	-0.043	3
Ascorbate (Vitamin C) and Aldarate Metabolism	L-Gulonate	M-H2O-H[-]	177.0405	282.6	2.59	-0.016	3.68	-0.102	2.46	-0.071	3
	D-Glucarate	M-H+O[-]	225.0242	68.5	2.22	-0.012	1.15	-0.028	0.70	-0.017	3
	5,6-Dihydroxyindole-2-carboxylate	M+Cl37[-]	228.9944	28.4	2.69	0.024	0.73	0.030	1.57	0.066	3
Beta-Alanine metabolism	Beta-Alanine	M-H	88.0404	47	2.47	0.018	2.90	0.096	2.47	0.082	3
	Dihydroxyacetone	M+Cl[-]	125.0011	41.6	1.88	0.015	1.39	0.050	2.38	0.087	3
	L-Histidine	M-H[-]	154.0623	50.6	1.85	0.014	2.65	0.091	1.69	0.060	1
Butanoate metabolism	2-Methyl-3-oxopropanoate	M+CH3COO[-]	161.0456	278.2	0.68	-0.005	0.77	-0.025	2.50	-0.086	3
	4-Hydroxybutanoic acid	M+CH3COO[-]	163.0613	136.8	0.76	-0.007	0.53	-0.023	2.02	-0.089	3
Fatty acid activation	Octanoic acid	M+CH3COO[-]	203.129	91.6	2.82	-0.023	2.37	-0.087	1.93	-0.073	3
	pentadecanoate	M(C13)-H[-]	241.2147	219.1	1.85	-0.017	2.34	-0.104	0.98	-0.042	3
	Elaidic acid	M-H	281.2485	246	1.69	-0.012	2.00	-0.062	1.65	-0.050	1
	(9E)-Octadecenoic acid	M+HCOO[-]	327.2541	244.5	1.75	-0.012	2.09	-0.067	1.69	-0.052	3
	Phytanate	M-H+O[-]	327.2912	240.2	1.29	-0.011	2.65	-0.108	1.06	-0.044	3

	Compound	Adduct	m/z									
	(9E)-Octadecenoic acid	M+CH3COO[-]	341.2699	232.9	1.77	-0.012	2.08	-0.066	1.70	-0.052	3
Fructose and mannose metabolism	Galactose	M-2H[2-]	89.0244	41.1	1.95	0.018	1.50	0.062	2.36	0.098	3
Glycerophospholipid metabolism	3-beta-D-Galactosyl-sn-glycerol	M-H+O[-]	269.0878	45.7	2.33	0.017	1.42	0.046	2.04	0.066	3
	CMP-2-aminoethylphosphonate	M-H+O[-]	445.0537	76.8	1.40	-0.011	0.30	-0.007	2.17	-0.075	3
	CMP-2-aminoethylphosphonate	M+Na-2H[-]	450.037	48.6	1.49	-0.015	1.35	-0.062	2.00	-0.096	3
Glycine, serine, alanine and threonine metabolism	Phosphocreatine	M+Br[-]	289.9565	66.4	1.59	0.008	1.39	0.032	2.58	0.061	3
Glycolysis and Gluconeogenesis	Dihydroxyacetone	M+Cl37[-]	126.9982	42	2.01	0.016	1.53	0.057	2.45	0.092	3
	2,3-Bisphospho-D-glycerate	M-H+O[-]	280.948	77.7	2.52	-0.013	2.40	-0.057	2.56	-0.063	3
	S-acetyldihydrolipoyllysine	M+ACN-H[-]	289.1059	26.5	1.25	0.008	0.47	0.014	2.26	0.070	3
Methionine and cysteine metabolism	2-keto-4-methylthiobutyrate	M+Cl[-]	182.9888	54.7	2.11	0.016	0.96	0.034	1.89	0.069	3
	Adenosine 5'-phosphosulfate	M+K-2H[-]	462.9633	89.9	2.04	-0.010	1.86	-0.041	2.32	-0.052	3
N-Glycan biosynthesis	Isopentenyl diphosphate	M+Cl[-]	280.9741	183.1	1.92	-0.005	1.40	-0.016	2.07	-0.023	3
Pentose phosphate pathway	L-Gulonate	M+Na-2H[-]	216.0246	40.5	0.84	-0.008	1.24	-0.056	2.13	-0.100	3
Prostaglandin formation from arachidonate	Prostaglandin B1	M+HCOO[-]	381.2312	168.3	2.24	-0.022	1.20	-0.053	1.91	-0.087	3
	Prostaglandin E2 ethanolamide	M-H+O[-]	410.2575	200.6	2.28	-0.014	1.27	-0.036	2.06	-0.062	3
	(5Z)-(15S)-11alpha-Hydroxy-9,15-dioxoprostanoate	M+Br81[-]	433.1403	263.5	2.40	-0.013	3.21	-0.078	1.85	-0.045	3
	15-oxo-Prostaglandin E2 glyceryl ester	M+Cl37[-]	461.2079	175.2	3.47	0.025	2.90	0.094	2.87	0.096	3
Pyruvate Metabolism	Dihydroxyacetone	M-H[-]	90.0277	41.3	2.09	0.019	1.54	0.063	2.52	0.106	3
Selenoamino acid metabolism	Dithiothreitol	M-H+O[-]	168.9993	290.5	1.67	0.012	0.30	0.010	2.21	0.074	3
	Oxidized dithiothreitol	M+Br81[-]	232.9137	107.7	2.72	-0.013	3.46	-0.072	2.75	-0.059	3
Starch and Sucrose Metabolism	Cellobiose	M+Na-2H[-]	362.0806	48	0.84	-0.005	1.01	-0.038	2.01	-0.075	3

| Vitamin K metabolism | Phylloquinone | M+Br[-] | 529.2684 | 248.4 | 3.44 | -0.016 | 3.36 | -0.069 | 2.14 | -0.045 | 3 |
| | Vitamin K1 epoxide | M+Br[-] | 545.2683 | 186.1 | 1.91 | 0.018 | 1.25 | 0.050 | 2.70 | 0.118 | 3 |

Supplemental Table 3-S3. Enriched pathways associated with pesticide exposures

Pathways	Overlap size	Pathway size	Ratio	P-value	Platform	Pesticide
Tyrosine metabolism	6	156	0.038	0.04857	HILICpos	Organophosphate
Beta-Alanine metabolism	6	34	0.176	0.00076	C18neg	Organophosphate
Glutamate metabolism	5	33	0.152	0.00345	C18neg	Organophosphate
Carnitine shuttle	6	53	0.113	0.00588	C18neg	Organophosphate
Glycine, serine, alanine and threonine metabolism	9	117	0.077	0.00933	C18neg	Organophosphate
Prostaglandin formation from arachidonate	7	79	0.089	0.01017	C18neg	Organophosphate
Histidine metabolism	5	45	0.111	0.0116	C18neg	Organophosphate
Arginine and Proline Metabolism	7	82	0.085	0.01176	C18neg	Organophosphate
Alanine and Aspartate Metabolism	5	46	0.109	0.01227	C18neg	Organophosphate
Pyruvate Metabolism	3	24	0.125	0.02849	C18neg	Organophosphate
Ascorbate (Vitamin C) and Aldarate Metabolism	5	62	0.081	0.03151	C18neg	Organophosphate
Butanoate metabolism	5	64	0.078	0.03445	C18neg	Organophosphate
Methionine and cysteine metabolism	7	109	0.064	0.03756	C18neg	Organophosphate
Fatty acid oxidation, peroxisome	2	13	0.154	0.04554	C18neg	Organophosphate
Glycolysis and Gluconeogenesis	4	49	0.082	0.04823	C18neg	Organophosphate
Glycine, serine, alanine and threonine metabolism	6	75	0.080	0.05697	HILICpos	Pyrethroid
Beta-Alanine metabolism	4	34	0.118	0.02521	C18neg	Pyrethroid
Arachidonic acid metabolism	5	53	0.094	0.03059	C18neg	Pyrethroid
Phytanic acid peroxisomal oxidation	3	22	0.136	0.03353	C18neg	Pyrethroid
Fatty acid activation	5	59	0.085	0.0405	C18neg	Pyrethroid
Carnitine shuttle	4	45	0.089	0.0105	HILICpos	Organochlorine

Pyrimidine metabolism	4	74	0.054	0.04285	HILICpos	Organochlorine
Glycolysis and Gluconeogenesis	8	49	0.163	0.00034	C18neg	Organochlorine
Galactose metabolism	7	44	0.159	0.00092	C18neg	Organochlorine
Pentose phosphate pathway	7	50	0.140	0.00202	C18neg	Organochlorine
Fructose and mannose metabolism	5	28	0.179	0.00244	C18neg	Organochlorine
Propanoate metabolism	5	33	0.152	0.00521	C18neg	Organochlorine
Vitamin K metabolism	2	4	0.500	0.00563	C18neg	Organochlorine
Glycine, serine, alanine and threonine metabolism	10	117	0.085	0.0095	C18neg	Organochlorine
Sialic acid metabolism	5	37	0.135	0.0095	C18neg	Organochlorine
N-Glycan Degradation	3	14	0.214	0.01168	C18neg	Organochlorine
Starch and Sucrose Metabolism	3	14	0.214	0.01168	C18neg	Organochlorine
Hexose phosphorylation	4	28	0.143	0.01563	C18neg	Organochlorine
Glycerophospholipid metabolism	9	111	0.081	0.01773	C18neg	Organochlorine
Butanoate metabolism	6	64	0.094	0.02571	C18neg	Organochlorine
Glutamate metabolism	4	33	0.121	0.02647	C18neg	Organochlorine
Selenoamino acid metabolism	4	33	0.121	0.02647	C18neg	Organochlorine
Beta-Alanine metabolism	4	34	0.118	0.02857	C18neg	Organochlorine
Caffeine metabolism	3	23	0.130	0.04008	C18neg	Organochlorine
N-Glycan biosynthesis	3	24	0.125	0.04437	C18neg	Organochlorine
Pyruvate Metabolism	3	24	0.125	0.04437	C18neg	Organochlorine
Aminosugars metabolism	4	40	0.100	0.04638	C18neg	Organochlorine

Supplemental Table 3-S4. Enriched metabolic pathways associated with *xMWAS* cluster 1 (OP)

Pathways	Overlap size	Pathway size	P-value	Platform
Pyrimidine metabolism	3	74	0.01613	HILICpos
Aspartate and asparagine metabolism	3	98	0.0347	HILICpos
Glycerophospholipid metabolism	3	109	0.04563	HILICpos
Beta-Alanine metabolism	5	34	0.00017	C18neg
Glutathione Metabolism	3	12	0.00017	C18neg
Histidine metabolism	4	45	0.00109	C18neg
Glycine, serine, alanine and threonine metabolism	6	117	0.00109	C18neg
Glutamate metabolism	3	33	0.00412	C18neg
Arachidonic acid metabolism	3	53	0.0116	C18neg
Arginine and Proline Metabolism	3	82	0.03033	C18neg

Supplemental Table 3-S5. Enriched metabolic pathways associated with *xMWAS* cluster 2 (PYR, OC)

Pathways	Overlap size	Pathway size	P-value	Platform
Saturated fatty acids beta-oxidation	4	17	0.00025	HILICpos
Di-unsaturated fatty acid beta-oxidation	4	20	0.00034	HILICpos
Carnitine shuttle	3	45	0.01092	HILICpos
Fatty Acid Metabolism	3	50	0.01319	HILICpos
Pentose phosphate pathway	2	50	0.02017	C18neg
Ascorbate (Vitamin C) and Aldarate Metabolism	2	62	0.02823	C18neg
Butanoate metabolism	2	64	0.03017	C18neg

Supplemental Table 3-S6. Pathway enrichment analysis for features associated with all three pesticides in *xMWAS*

Pathways	Overlap size	Pathway size	P-value	Platform
Di-unsaturated fatty acid beta-oxidation	3	20	0.00143	HILICpos
Carnitine shuttle	3	45	0.00933	HILICpos
Selenoamino acid metabolism	2	33	0.0105	C18neg
Glutamate metabolism	2	33	0.0105	C18neg
Beta-Alanine metabolism	2	34	0.01084	C18neg
Histidine metabolism	2	45	0.01706	C18neg
Pentose phosphate pathway	2	50	0.02025	C18neg
Ascorbate (Vitamin C) and Aldarate Metabolism	2	62	0.02664	C18neg
Butanoate metabolism	2	64	0.02865	C18neg
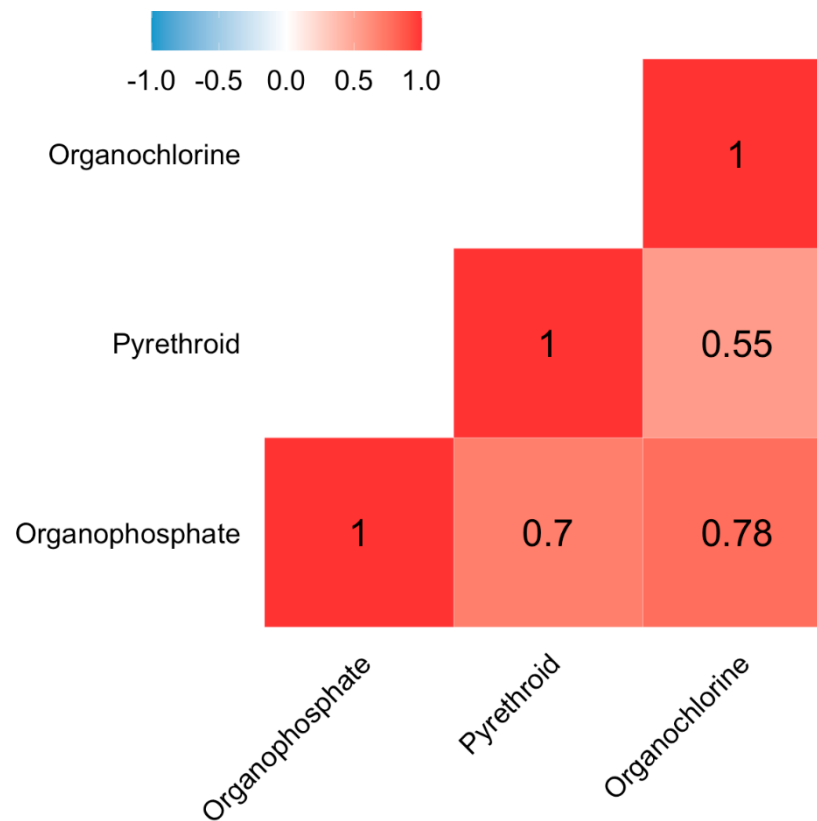
Supplemental Figure 3-S1. Heatmap of pairwise correlations of pesticide counts. The heatmap color-codes the pairwise *Pearson* correlations of organophosphates (OP), pyrethroids (PYR), and organochlorines (OC) counts. The shades of color (blue, white, and red) visualize correlation values from -1 to 1. Each square reports a *Pearson* correlation coefficient.
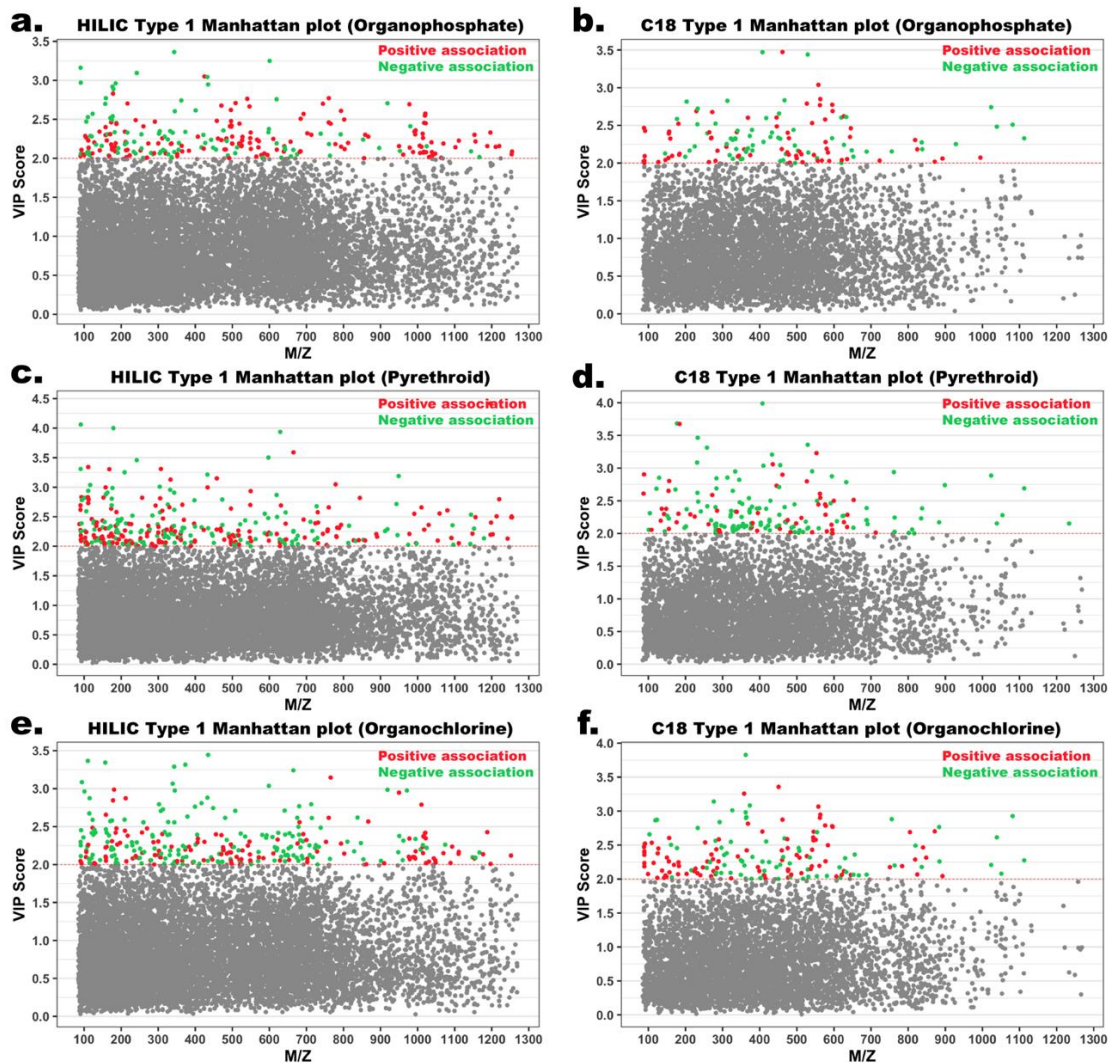
Supplemental Figure 3-S2. Identification of metabolomic features associated with pesticide exposure. (**a**) Type 1 Manhattan plot for OP-associated features in the HILIC column (positive ion mode), VIP score vs. m/z. Red dots represent features positively associated with OP exposure and green dots represent features negatively associated with OP exposure; **b**) Type 1 Manhattan plot for OP-associated features in the C18 column (negative ion mode), VIP score vs. mass-to-charge; **c**) Type 1 Manhattan plot for PYR-associated features in the HILIC column (positive ion mode); **b**) Type 1 Manhattan plot for PYR-associated features in the C18 column (negative ion mode); **e**) Type 1 Manhattan plot for OC-associated features in the HILIC column (positive ion mode); **f**) Type 1 Manhattan plot for OC-associated features in the C18 column (negative ion mode).
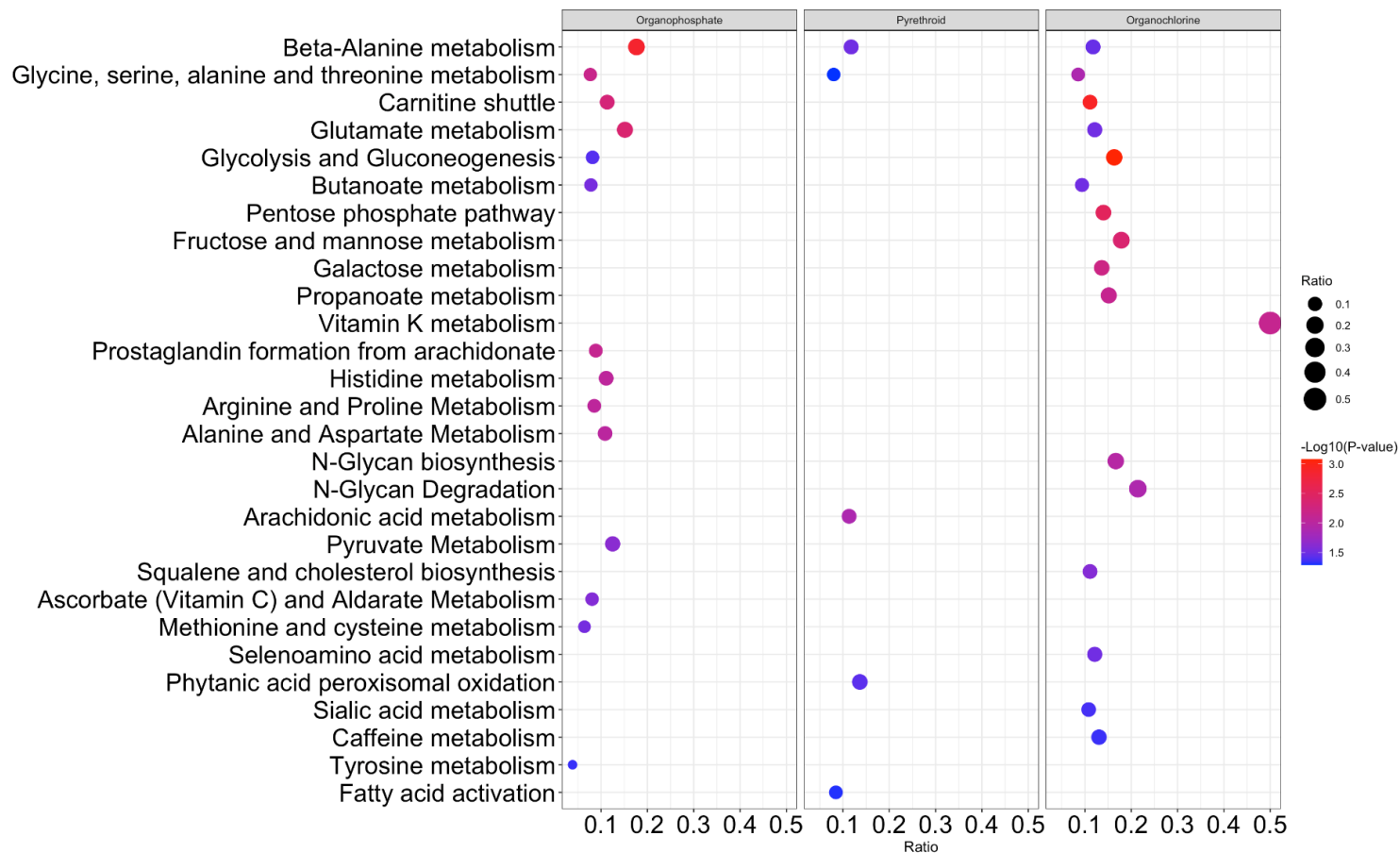
Supplemental Figure 3-S3. Enriched pathways identified from the sensitivity analysis by additionally adjusting for education. The vertical axis represents the pathways associated with pesticides. Circle radius is proportional to the number of correlated metabolite features within each pathway (ratio). The horizontal axis also represents the ratio. The color represents the negative log10 (*p-value*) of each pathway.

4 Towards epigenomic and metabolomic profiles of chronic organophosphate exposure in residents of California's Central Valley

4.1 Introduction

Organophosphates (OPs) are non-persistent pesticides widely used for crop and agricultural food protection. The intended mechanism of action of OPs is damage to the peripheral and central nervous system of pests through the inhibition of acetylcholinesterase activity that leads to the neurotoxic accumulation of acetylcholine neurotransmitters (Terry Jr 2012; van der Plaat et al. 2018). While acute toxicity impacts mainly the cholinergic neurons of the nervous system, the toxicity of chronic and low-level exposure to OPs is likely mediated through a wide range of non-cholinergic mechanisms such as contributions to DNA damage, oxidative stress, immunotoxicity, mitochondrial dysfunction, and regulation of neuronal apoptosis (Androutsopoulos et al. 2013; Banerjee et al. 2001; Costa 2006; Farkhondeh et al. 2020; Zuluaga et al. 2016). Previous epidemiological studies have linked chronic low-level OP exposure to numerous adverse health effects, including neurodegenerative disease, such as seen in Parkinson's disease (PD) and Alzheimer's disease (AD), diabetes, and cancers (Blair et al. 2015; Clary and Ritz 2003; Cockburn et al. 2011; Narayan et al. 2013; Paul et al. 2018b).

Because of the rapid development of high-throughput analytical platforms such as high-resolution metabolomics (HRM) and DNA methylation arrays in the past decades, we now have an opportunity to thoroughly investigate molecular responses related to OP exposures. Metabolomics technologies can simultaneously detect more than 10,000 chemicals, including endogenous molecules and exogenous chemicals or toxins, making this approach particularly effective for assessing environmental exposures (Bloszies and Fiehn 2018). For example, a cross-sectional study characterized the urine metabolome of 83 pregnant women finding that pesticide

mixture exposures disturb energy metabolism (N. Bonvallot et al. 2013) and a study of 102 Chinese pregnant women suggested that OP and organochlorine (OC) exposures disrupt thyroid hormone metabolism and glyceraldehyde metabolism (Yang et al. 2020). Epigenomics provides another high-dimensional molecular approach that can improve our understanding of molecular mechanisms related to environmental pesticide exposures. For example, previously an in-vitro cancer cell line based study showed that three commonly used OP pesticides induce DNA methylation changes in 712 genes (Zhang et al. 2012). Furthermore, in the Agricultural Health Study, DNA methylation changes were associated with high pesticide exposure events in applicators suggesting that acute pesticide exposure may alter DNA methylation (Rusiecki et al. 2017). In our California study, we previously illustrated that OP-related epigenetic changes are consistent with their intended mechanisms of action inhibiting acetylcholinesterase (Paul et al. 2018a). To date, most studies using omics tools and investigating OP pesticides focused on a single molecular layer. Thus, it is yet unknown whether and how various molecular layers react to exposure or interact with each other generating a molecular cascade that mediates OP toxicity.

Here, we aim to elucidate epigenetic as well as metabolomic changes induced by chronic low-level OP exposure typical for older adults living in heavily agricultural regions of Central California. We will also attempt to integrate the information from OP-related epigenomic and metabolomic signals to improve our understanding of molecular mechanisms involved in the human response to chronic OP exposures. To accomplish this, we utilized serum samples from 176 control subjects recruited into a study of Parkinson's disease with DNA methylation and metabolomics data. In these residents we estimated chronic OP exposure from ambient sources at homes and workplaces using a geographic information system (GIS) based model and California pesticide use reporting data.

4.2 Methods

4.2.1 Study population and samples collection

The Parkinson's Environment and Gene Study (PEG) is a case-control study designed to investigate PD etiology in agricultural regions of the California central valley. The study recruited PD cases and population-based non-PD controls from Kern, Tulare, or Fresno counties in central California from 2001 to 2007. Controls were identified from residential parcel listings from property tax assessor records and screened for eligibility by mail or telephone; participation was limited to one person per household. Eligibility criteria for controls included being over 35 years of age at enrollment, having lived in one of the three counties for at least the past 5 years, and not having received a Parkinsonism diagnosis. More detailed recruitment information has been provided elsewhere (Kang et al. 2005).

For this study, we analyzed blood samples from 176 controls for whom we had both metabolomic and methylation data available. Demographic information was collected in standardized interviews.

4.2.2 OP exposure assessment

OP exposure assessment was performed as previously described (Furlong et al. 2020; Paul et al. 2018a). Briefly, we estimated ambient OP pesticide exposure for each participant from residential and occupational proximity to commercial agricultural pesticide applications using a geographic information systems (GIS) based model (Cockburn et al. 2011). This method links California state-mandated pesticide use reports (CA-PUR), a state-wide registry of all commercial pesticide applications, land use surveys providing locations of specific crops, and participants' geocoded residential and occupational addresses. All records are available back to 1974.

Participants in our study population were assigned exposure to 25 different chemicals classified as OPs, based on the California Department of Pesticide Regulation (CDPR) and the pesticide action network (PAN) pesticide database. We calculated the total pounds of each chemical applied within 500 m of the participant's residence and workplace based on their reported address histories for each year. For each individual OP pesticide, we averaged the pounds of pesticide applied per acre per year at each residential and occupational address over a ten-year period prior to enrollment and blood draw. As OPs have different toxicities that are not reflected in the amount (pounds) of pesticide applied, we then dichotomized this yearly-average pounds per acre at greater than the median levels, and opted for an overall OP count measure defined by summing the number of OP chemicals each participant was exposed to based on the dichotomized measure. Each participant could have been exposed at both residential and occupational addresses, only one, or neither Therefore, theoretically, the max OP count measurement is 50 if a participant were highly exposed to all OP chemicals at both addresses.

4.2.3 DNA methylation quantification and preprocess

For each participant, peripheral blood samples were collected, and we performed bisulfite conversion using the Zymo EZ DNA Methylation Kit (Zymo Research, Orange, CA, USA) as well as subsequent hybridization of the HumanMethylation450k Bead Chip (Illumina, San Diego, CA), and scanning (iScan, Illumina) according to the manufacturer's protocols by applying standard settings.

Raw methylation signal intensities were retrieved using the function *read.metharray.exp* of the *minfi* R package, followed by linear dye bias correction, noob background correction, and beta-mixture quantile normalization (BMIQ) for correcting probe design bias using the *minfi* and *ChAMP* R package (Aryee et al. 2014; Fortin et al. 2017; Fortin et al. 2014; Tian et al. 2017; Triche

et al. 2013). β-values were used in all analyses. One sample was identified as low quality due to low median methylated and unmethylated signal intensities across the entire array and removed from the study population. Detection p-values were derived using the function *detectionP* as the probability of the total signal (methylated + unmethylated) being detected above the background signal level, as estimated from negative-control probes. All in all, 845 probes with a detection p-value above 0.05 in at least 5% of samples were removed. Also, we removed 645 probes with a bead count <3 in at least 5% of samples; 11,334 probes on the X or Y chromosome; 7,306 probes containing a SNP at the CpG interrogation site and/or at the single nucleotide extension for 5% maf; and 27,332 cross-reactive probes. In total, 438,050 probes were included in downstream analyses.

White blood cell composition was imputed for each study participant using our online published DNA Methylation Age Calculator, https://dnamage.genetics.ucla.edu/. The following imputed blood cell counts were included in downstream analyses: CD4+ T, naïve CD8+ T, exhausted cytotoxic CD8+ T cells (defined as CD8 positive CD28 negative CD45R negative, CD8+CD28-CD45RA-), plasmablasts, and granulocytes. Naïve CD8+ T, exhausted cytotoxic CD8+ T cells, and plasmablasts were calculated based on the Horvath method (Horvath and Levine 2015). The remaining cell types were imputed using the Houseman method (Houseman et al. 2012). A potential batch effect (position on the array) was addressed using *combat* (Johnson et al. 2007).

4.2.4 High-resolution metabolomics

High-resolution metabolomics (HRM) profiling was completed according to established methods (Walker et al. 2018). Briefly, serum samples were collected during the interview and stored at -80 °C. Prior to HRM, batches of 40 serum samples were removed from storage and thawed on ice. Each sample was then thoroughly vortexed, and plasma proteins were precipitated

by diluting 65 µL of serum with 130 µL of LC-MS grade acetonitrile. The extract was centrifuged, and the resulting supernatant was transferred to an autosampler vial containing a low volume insert and maintained at 4ºC until analysis (<24 h). To evaluate system performance, we used two separate quality assessment methods. Our first QC sample was NIST 1950 (Simon-Manso et al. 2013) which was analyzed at the beginning and end of the entire analytical run. The second QC sample (Q-Std) included commercially purchased plasma pooled from an unknown number of males and females. Q-Std was analyzed at the beginning, middle, and end of each batch of 40 samples for normalization and batch effect evaluation.

Sample extracts were analyzed in triplicate using a dual-column, dual-polarity approach that includes hydrophilic interaction (HILIC) chromatography with positive ESI and C18 chromatography with negative ESI (Ultimate 3000, Q-Exactive HF, Thermo Fisher, m/z range 85-1275) (Liu et al. 2016). The mass spectrometer was operated using ESI mode at a resolution of 120,000 and mass-to-charge ratio (m/z) range 85-1275. High-resolution detection of m/z features was accomplished by a maximum injection time of 10 milliseconds and AGC target of 1×10^6. Raw data files were extracted and aligned using *apLCMS* (Yu et al. 2009) 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 a novel algorithm based on wavelet (Deng et al. 2019). For further analyses, we only included metabolomic features detected in > 25% of all plasma samples, with median coefficients of variation (CV) among technical replicates < 30% and Pearson correlation > 0.7. Following quality assessment, intensities of three replicates for each feature were summarized using the median value. In addition, we conducted log2 transformation and auto-scaling. Missing values

were imputed using k-nearest neighbors (k=10) (Troyanskaya et al. 2001) imputed in the *impute* R package.

4.2.5 Statistical analyses

<u>Epigenomic analyses</u>

The global mean methylation level was calculated by averaging genome-wide methylation levels across all 438,050 CpG sites. In addition, we calculated genomic region-specific mean methylation levels for CpGs within transcription start sites (TSS1500, TSS200), untranslated regions (5'UTR, 3'UTR), 1st Exon, and gene body. The association between OP exposure and global/genomic region-specific mean methylation levels were calculated using linear regression. We also conducted principal component analysis (PCA) for methylation data and the top 40 principal components were correlated with the covariates including age, sex, education, race/ethnicity, and cell compositions.

We identified CpG probes that are related to OP exposure by biweight midcorrelation (bicor) implemented in the *WGCNA* R package (Langfelder and Horvath 2008). Bicor is a median-based measurement of correlation that is robust to outliers (Langfelder and Horvath 2012). We adjusted for potential confounders including age, sex, race/ethnicity, education, and cell compositions (naïve CD8 cells, CD8+CD28-CD45RA- T cells, Plasma Blasts, CD4 T cells, and Granulocytes) by regressing out the effects of these factors and retaining the residuals of DNA methylation levels for analyses. We accounted for the multiple testing via the false discovery rate (FDR)-adjusted p-values.

CpGs were annotated using *IlluminaHumanMethylation450kanno.ilmn12.hg19* R package. The Comparative Toxicogenomics Database (CTD) is a publicly available database that identifies

interactions between chemicals and genes (Mattingly et al. 2006). We searched on CTD to identify any previously known interaction between OP chemicals and selected CpG genes.

In order to gain insight into the nature of the OP-related CpGs, we used the eFORGE 2.0 (Breeze et al. 2016), which tests for enrichment of cell-type-specific DNase hypersensitive sites (DHS). This analysis was performed for DHS profiled in specific tissues and cell types generated by Roadmap (Kundaje et al. 2015).

We conducted hypergeometric tests to test whether OP-related CpGs were randomly distributed across the genome or were more likely to be found in specific genomic regions (CpG islands, shores, shelves, or open sea area). We also conducted the gene ontology (GO) biological process, molecular function, and cellular component pathway enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, and disease ontology enrichment analysis using the *clusterprofiler* R package (Yu et al. 2012). GO terms were simplified for easier interpretation by calculating similarity of GO terms and removing those highly similar terms. In addition, we conducted genomic region enrichment analysis using a hypergeometric test to test for CpGs localization.

Finally, we examined whether the OP-related CpGs were expression quantitative trait methylation loci (eQTMs) according to the Biobank-based Integrative Omics Study (BIOS) database. BIOS database is part of the Biobanking and BioMolecular Infrastructure of the Netherlands (BBMRI-NL), which captured meQTLs, eQTLs, and eQTMs from a genome-wide database of 3841 Dutch blood samples.

Metabolomic analyses

To assess the relationship between the global metabolome and covariates, we first conducted principal component analysis (PCA) for metabolomic data, and the top 40 principal components were correlated with the age, sex, education, and race/ethnicity.

We identified metabolomic features (m/z features) that are related to OP exposure by using a combination of multivariate and univariate approaches (Saccenti et al. 2014). We first filtered out features likely not associated with OP exposure by sparse partial least square regression analysis (sPLS). sPLS is a sparse version of the supervised, multivariate PLS method with a Lasso penalization that combines variable selection and modeling in a one-step procedure (Le Cao et al. 2008). To adjust for potential confounders, we regressed the intensity of each feature on potential confounding variables (age, sex, education, and race/ethnicity) and formed residuals as the input matrix. The hyperparameters of the model were tuned by carrying out repeated cross-validations (50 x 5-fold CV). We then calculated bicor between OP exposure and metabolomic features selected by sPLS. We accounted for the multiple testing via the false discovery rate (FDR)-adjusted p-values.

HRM provides accurate mass (±5 parts-per-million; ppm) measures of ion m/z, which can be related to chemical monoisotopic mass. Discriminative features were first matched to a reference database of authenticated chemical standards previously derived with LC-MS/MS (Go et al. 2015). The error tolerance was 5 ppm and 30 seconds for m/z and retention time, respectively. Additional m/z features not matching these metabolites were then annotated by *xMSannotator* (Uppal et al. 2017). Accurate mass m/z for adducts formed under positive/negative ESI mode was matched to the Human Metabolome Database (HMDB), KEGG, and LipidMaps with a mass error threshold of 10 ppm. *xMSannotator* also considers the correlation of intensities and retention time and assigns confidence scores based on a multilevel scoring algorithm (0-3, a higher score

represents higher-confidence result), which ensures annotation accuracy. Only results with annotation scores greater than 2 were kept.

We conducted pathway enrichment analysis utilizing *mummichog* version 2 to identify perturbed metabolic pathways associated with OP exposure (Nathalie Bonvallot et al. 2013). Although the tentative annotation results in *mummichog* may include false positives, the enriched pathways inferred by the algorithm have been proven to be valid and to reflect real biological activity (S Li et al. 2017; Mallozzi et al. 2016; Uppal et al. 2016). All metabolites annotated by *mummichog* were required to present in at least their primary adduct (M + H or M-H for positive and negative mode, respectively) to reduce the false positive match rate. The p-value threshold we relied on was 0.05. To include all possible metabolomic features involved in OP-associated metabolic pathways, all features previously selected by sPLS were included in pathway enrichment analyses. A pathway was considered significant if gamma-adjusted p-values were smaller than 0.05. Only pathways that contained at least 3 discriminative metabolites were interpreted.

4.2.6 Linking the epigenomic and metabolomic features

To investigate the interplay between the metabolome and epigenome, we correlated OP-related CpGs and OP-related metabolomic features with the metabolome and epigenome, respectively. Specifically, we calculated pairwise bicor measures between CpGs and metabolomic features. We additionally adjusted for our OP count measure to remove potential confounding. Multiple testing corrections were conducted using the Benjamini-Hochberg procedure, and significantly correlated CpG-metabolomic feature pairs were selected for downstream functional annotations.

4.3 Results

4.3.1 Study Population

The general workflow of the project is shown in Figure 4-1. Demographics of the 176 healthy PEG1 controls are provided in Table 4-1. Approximately half of the participants included in the study were male (51%). Majority of the participants were elderly, with an average age of 66; and 90% of them were non-Hispanic whites. Participants who did not have both metabolomic and epigenomic data available were not different from participants who did with regards to age, sex, race/ethnicity, or exposure to OP pesticides.

4.3.2 Epigenomic analyses for OP exposure

<u>*Global methylation pattern*</u>

OP exposure status was not associated with global mean methylation levels (averaged genome-wide methylation levels across the 438,050 CpG sites) (p-value = 0.43). In addition, mean methylation levels of CpG probes within specific genomic regions (TSS1500, TSS200, 5'UTR, 1st Exon, gene body, 3'UTR) were not significantly associated with OP exposure.

As shown in Figure 4-2, correlations between principal components (PCs) from the methylation data with OP exposure and covariates suggested that estimated cell types were strongly correlated with the top PCs. OP exposure was significantly correlated with the first and several other PCs, including PC 19, 29, 34, and 39. Consistent with previous studies (Cardenas et al. 2019), our result suggested that top PCs reflect DNA methylation differences in cell types and heterogeneity associated with covariates such as age, sex, race/ethnicity, and education.

<u>*Epigenome-wide analyses*</u>

We conducted biweighted midcorrelation analyses to identify CpGs related to OP exposure. After adjustment for multiple testing, none of the CpGs were significant at FDR < 0.05 level. Using a loose threshold, we identified 10 CpGs at a raw p-value = 1E-6 level (Table 4-2, Supplemental Figure 4-S1) and an additional 398 CpGs with a raw p-value < 5E-4 (Figure 4-3a), mapping to 356

genes. Approximately half of these 408 CpGs associated with OP exposure with P-value < 5E-4 were hypermethylated (51.5% vs. 48.5%).

The most significant CpG site is cg26835374 (intergenic, chr1:142,618,826), for which methylation was positively correlated with OP exposure (r = 0.36, p = 7.65E-07). Other top hits include cg09763120 in *C11orf80* (r = 0.35, p = 1.87E-06), cg23037798 in *HNRNPR* (r = 0.35, p = 2.24E-06), cg17559110 in *AP2A2* (r = 0.34, p = 3.20E-06), and cg16471585 in *KSR2* (r = 0.34, p = 3.22E-06), which were all positively associated with the OP exposure (Table 4-2). Based on the Comparative Toxicogenomics Database, OP chemicals (including monocrotophos, dicrotophos, dichlorvos, diazinon, chlorpyrifos, malathion, and phosalone) can affect the expression level of 92 genes related to differentially methylated CpGs (Supplemental Table 4-S1).

Functional analysis

In order to gain insight into the nature of the 408 OP-related CpGs, we used the eFORGE v2 to test for the enrichment of tissue or cell-type specific DNase hypersensitive sites (DHS). None of the cell types was significantly enriched after multiple testing corrections. The strongest enrichment for the DHS was derived trophoblast cultured cells, fetal lung fibroblasts cell line, monocytes, and T cells, which we note is consistent with an OP pesticide-associated chronic inflammatory response (Supplemental Table 4-S2).

We conducted genomic region enrichment analysis using the hypergeometric test. The 408 differentially methylated CpGs in relation to OP exposure were significantly enriched for localization to CpG islands (44% of the 408 CpGs are in CpG islands, whereas 31% of all CpGs are in CpG islands, p-value = 1.0E-7), but not in shores, shelves, or open sea area.

GO analyses of the genes annotated to the 408 OP-related CpGs identified overrepresentation of cellular components involving the acetyltransferase complex (GO:0000123,

GO:0031248, and GO:1902493) (Supplemental Table 4-S3). KEGG pathway enrichment analysis did not identify any statistically significant pathways after adjustment for multiple testing (Figure 4-3b). The top enriched KEGG pathway is the glycosphingolipid (GSL) biosynthesis (hsa00604), followed by the bacterial invasion of epithelial cells (hsa05100). We also conducted disease ontology enrichment analysis and did not found any significantly enriched disease ontology. The majority of the top disease ontologies are related to cancer and skeletal muscle diseases (Figure 4-3b).

We also examined whether OP-related CpG sites were expression quantitative trait methylation loci (eQTMs), which means their methylation levels will affect the gene expression in blood. Based on the Biobank-based Integrative Omics Study (BIOS) database of the Biobanking and BioMolecular Infrastructure of the Netherlands (BBMRI-NL) (Bonder et al. 2017), we found that 3 of the 408 CpGs were identified as eQTMs, including cg17003212, cg26668919, and cg06639387 (Supplemental Table 4-S4).

Sensitivity analysis

We conducted a sensitivity analysis controlling for all covariates except cell compositions. In the sensitivity analysis, we identified 406 CpGs correlated with OP exposure with raw P-valve < 5E-4, 347 of which overlapped with the original result (Supplemental Table 4-S5). In addition, pathway enrichment analysis of these 406 CpGs showed the same set of enriched pathways as the original result.

4.3.3 Metabolomic analyses for OP exposure

Global metabolome pattern

Figure 4-4 displays the correlations between principal components (PCs) derived from the metabolomic data for OP exposure and covariates. The results suggest that covariates including

age, sex, and race/ethnicity are correlated with the top PCs. OP exposure was significantly correlated with several PCs – some of which were only correlated with OP exposures while others were correlated with OPs as well as all other covariates simultaneously - but not with the first PC. In line with our finding, previous studies (Li et al. 2019) also suggested that the influence of environmental exposure on the metabolome is not limited to the first component, and the small effect size might be masked by other variances.

Metabolome-wide analysis and pathway enrichment analysis

We detected 12925 metabolomic features from the HILIC column coupled with the positive ionization mode (HILICpos) and 7209 metabolomic features from the C18 column coupled with the negative ionization mode (C18neg). After quality control steps, a total of 16510 metabolomic features (10959 HILICpos and 5551 C18neg) were included in the analyses.

After adjusting for potential confounding variables (age, sex, and race/ethnicity), sPLS selected 1,367 metabolomic features (698 HILICpos features and 669 C18neg features) across 3 components that are associated with OP exposure status (Figure 4-5a, Supplemental Figure 4-S2). We then conducted linear regressions for each of the metabolomic features on OP exposure. After adjusting for multiple comparisons, none of the features had FDR smaller than 0.05. Using less stringent criteria, we identified 99 features (7 HILICpos features and 92 C18neg features) with FDR smaller than 0.2 (Supplemental Figure 4-S2 and 4-S3).

Using *mummichog*, we examined whether the features that sPLS selected were enriched within specific metabolic pathways. The result indicated that 14 metabolic pathways were differentially enriched with a p-value < 0.05 (Figure 4-5b, Supplemental Table 4-S6 and 4-S7). Top pathways included propanoate metabolism, fructose and mannose metabolism, cofactors such

as vitamin B5 and vitamin K, as well as lipid metabolism such as glycerophospholipid metabolism and glycosphingolipid biosynthesis.

In total, 15.9% of the HILIC pos features and 16.2% C18 neg features were tentatively annotated to one or more unique metabolites with medium or high confidence scores based on *xMSannotator*. We further matched the metabolites with previously confirmed chemical identities using MS2 spectra compared with authentic compounds analyzed under the identical experimental condition according to the Metabolomics Standards Initiative (MSI) level 1 criteria. In total, we confirmed 23 metabolites (Table 4-3). Among them, 14 were positively associated with OP exposure, while other metabolites were negatively associated with OP.

4.3.4 Multi-omics integration

Biweighted midcorrelations

We calculated pairwise Bicor correlations between 408 OP-related CpG sites and 99 OP-related metabolomic features. Approximately half of the pairs (50.5%) were positively correlated (Figure 4-6a). After multiple comparison correction, none of the CpGs-metabolite correlations remained significant at FDR = 0.05 level. One CpG (cg07385362) in UNKL was correlated with docosanamide (m/z = 374.3171, confidence level = 3) with FDR = 0.06 (bicor r = 0.35, see Figure 4-6b and Supplemental Table 4-S8). UNKL (Unk Like Zinc Finger) encodes a RING finger protein that may function in Rac signaling and is related to the innate immune system.

By correlating 408 OP-related CpGs with the whole metabolome, we additionally identified 1 CpG, cg08901151 in MAD1L1 as being significantly correlated with an unidentified metabolite (m/z = 261.1761) with FDR = 0.01 (bicor r = 0.43, see Figure 4-6b and Supplemental Table 4-S9). MAD1L1 (Mitotic Arrest Deficient 1 Like 1) is a component of the mitotic spindle-assembly checkpoint. By correlating 99 metabolomic features previously associated with OP

exposure with all CpGs not OP-related, we identified 2 significant pairs with FDR < 0.05 (Figure 4-6b). One CpG (cg09973663) in MYO1B was negatively correlated (bicor r = -0.43) with phenylacetic acid (m/z = 135.0452, confidence level = 1), and cg24365042 in ATF3 was positively correlated (bicor r = 0.43) with an unidentified metabolite (m/z = 241.2147). In addition, there were 515 pairs with FDR < 0.2 including 84 OP-related metabolomic features and 511 CpGs (Supplemental Table 4-S10). Pathway enrichment analyses showed that 515 CpG-metabolite pairs were involved in metabolic processes and nutrient-sensing, such as AMPK signaling pathway, glucagon signaling pathway, and type II diabetes mellitus (Figure 4-6c).

4.4 Discussion

Our study is mainly hypothesis-generating and serves as a proof of principle study for the usefulness of multi-omics approaches to address chronic, low level, and long-term environmental exposures and related health outcomes. Specifically, by linking GIS-modeled OP pesticide data with metabolomic and epigenomic data extracted from blood samples, we elucidated how different biological layers may respond to chronic low-level OP exposure. To obtain long-term OP exposure measurements for each participant, we combined California pesticide application records with home and work address histories. We then identified CpGs and metabolomic features and pathways significantly associated with OP exposure. The single-omics analyses showed both epigenomic and metabolomic signatures of OP as being enriched in the glycosphingolipid (GSL) biosynthesis pathway. Besides this common pathway, the metabolome and epigenome also exhibited distinct responses to OPs, with differently methylated CpGs being involved in intracellular membrane transport, cell adhesion, and carcinogenesis; and OP-related metabolites being involved in aromatic amino acids (AAA) metabolism, neurotransmitter precursors, oxidative stress, and mitochondria function. Moreover, we illustrated the possible interaction between these

101

two molecular layers through metabolic processes and nutrient-sensing pathways when we integrated the epigenomic and metabolomic signals. Our findings may facilitate a better understanding of the pathophysiology that eventually may lead to pesticide-related health outcomes (Barnham et al. 2004; Jenner 2003; Lin and Beal 2006; McGeer and McGeer 2004; Parrón et al. 2011).

4.4.1 The glycosphingolipid biosynthesis pathway identified in both epigenomic and metabolomic analyses

Both epigenomic and metabolomic analyses in our study suggest that chronic OP-exposure is involved in a wide range of molecular function and disease states. By comparing pathway enrichment results derived from single-omics analyses, we identified the GSL biosynthesis pathway as a shared mechanism identified with both our epigenomic and metabolomic approaches. OP-related CpGs within GSLs biosynthesis pathway were in *B4GALNT1*, *HEXA*, and *ST3GAL1* genes, respectively while metabolomic features involved in the pathway were annotated as cytidine-5'-monophosphate (CMP), galactose, and guanosine 5'-diphosphate (GDP). Previous studies have associated GSLs with environmental exposures such as air pollution and polybrominated biphenyls (PBBs) (Jeong et al. 2018; Walker et al. 2019a). GSLs play an important role in membrane organization and as signaling molecules. It has been shown that GSLs are involved in the regulation of cell properties such as adhesion, growth, apoptosis, and senescence (Garcia-Ruiz et al. 2015). Moreover, GSLs are potent regulators of the inflammatory process. Subtypes of GSLs are differentially expressed throughout the immune system and altered genes within the GSL pathway could be modulated through differentiation and activation of immune cells, suggesting a potential role for GSL in regulating immune functions. For example, monosialodihexosylganglioside (GM3) synthase *B4GALNT1* is differentially activated through

activation and maturation of human B cells (Zhang et al. 2019). In our study, we observed negative correlations between OP exposure and CpGs involved in GSL biosynthesis, which might suggest the alteration of the immune system in response to these pesticides. Interestingly, GSLs have also been implicated in a variety of neurodegenerative diseases (Kolter 2012). A previous study reported higher levels of GSLs in the plasma from PD patients compared to controls (Mielke et al. 2013), which is in line with the proposed role of neuroinflammation in the pathophysiology of PD (Horvath and Ritz 2015). Overall, our data suggest that OP exposure may alter the immune system in older adults and possibly this may contribute to neurodegenerative diseases such as PD. Future studies focusing on how GSLs pathways work together across different molecular layers in response to OP exposure are needed.

4.4.2 Epigenomics and metabolomics suggested various distinct mechanisms in response to chronic OP exposure

We also observed various distinctive signals for OP exposures that were not shared across the two molecular layers. For example, based on epigenomics, we identified several CpGs annotated to genes related to the membrane trafficking, such as *AP2A2* (Adaptor Related Protein Complex 2 Subunit Alpha 2), *DNM3* (Dynamin 3), *Rab5A* (Member RAS Oncogene Family), *DYNC2H1* (Dynein Cytoplasmic 2 Heavy Chain 1), and *VPS45* (Vacuolar Protein Sorting 45 Homolog). It has been well-documented that dysfunction of intracellular membrane trafficking plays an important role in several neurodegenerative diseases such as AD, dementia, and especially PD (Abeliovich and Gitler 2016; Bandres-Ciga et al. 2019; Hunn et al. 2015). Many PD-related genes are directly involved in synaptic vesicle trafficking, and some well-known PD genes, such as *LRRK2*, have also been found to interact with genes such as *DNM3* and *AP2A2* (Bandres-Ciga et al. 2019; Beilina et al. 2014) involved in intracellular trafficking.

We also observed DNA methylation changes in genes related to carcinogenesis and tumor metastasis. Several top CpGs were located within genes involved in cell adhesion, such as *JUP* (Junction Plakoglobin), *LPP* (LIM Domain Containing Preferred Translocation Partner In Lipoma), and *PTK2* (Protein Tyrosine Kinase 2). Growing evidence suggests that chronic pesticide exposure may increase the risk of various cancers (Weichenthal et al. 2010). Furthermore, several in vitro and in vivo studies have shown that epigenetic mechanisms may play an important role in pesticide-related carcinogenesis (Rusiecki et al. 2017; Zhang et al. 2012).

In our metabolomics analysis, the most significantly enriched pathway was propanoate metabolism. Propanoate is a short-chain fatty acid derived from both endogenous and exogenous sources (Frye et al. 2016). Also, propanoate is a vital metabolite produced by the gut microbiome (Louis and Flint 2017). Propanoate metabolism has been reported to be associated with mitochondria function. Therefore, our result was in line with previous findings that chronic OP exposure may induce neurotoxicity mediates through mitochondria dysfunction (Farkhondeh et al. 2020; Leung and Meyer 2019), which may contribute to cardiovascular disease, cancer, metabolic syndrome, and chronic neurodegenerative disorders such as PD and Huntington's disease (Nicolson 2007; Soane et al. 2007; Swerdlow 2011).

We also identified enrichment for metabolomic pathways involved in aromatic amino acids (phenylalanine, tryptophan, and tyrosine) metabolism, and we were able to confirm several metabolites within these pathways, such as phenylacetic acid and tyrosine. Phenylalanine is a metabolic precursor of tyrosine. Furthermore, tyrosine is a precursor of dopamine i.e. tyrosine is converted to L-dopa by tyrosine hydroxylase (Y Liu et al. 2020). Previous studies suggested that alterations of tyrosine hydroxylase can contribute to PD risk because it affects the biosynthesis of dopamine (Priyadarshini et al. 2012; Tabrez et al. 2012). In addition, we also observed perturbed

glutamate metabolism related to OP exposure. Glutamate is an excitatory neurotransmitter in the central nervous system, and glutamate-mediated excitotoxicity has been linked to the etiology of PD (Meredith et al. 2009; Sonsalla et al. 1998). Collectively, our data suggest that OP-induced dysregulation of these neurotransmitter precursors could contributing to the etiology of PD.

Our metabolomics analysis also identified alterations in several oxidative stress-related pathways, including vitamin C metabolism and glycerophospholipid metabolism. Previous studies have associated fatty acid oxidation and glycerophospholipid metabolism with numerous environmental exposures that are known to cause oxidative stress, such as persistent organic pollutant exposures, trichloroethylene, benzo[a]pyrene, and air pollution (Carrizo et al. 2017; Walker et al. 2016a; Walker et al. 2016b; Yan et al. 2019). Oxidative stress is also thought to play an important role in adverse health outcomes such as diabetes and neurodegenerative diseases, including AD and PD (Fillmore et al. 2014; Foley 1992; Lane et al. 2015).

In summary, apart from the GSL biosynthesis pathway implicated by both the epigenome and metabolome as being altered by OP exposures, our approach also implicated distinct mechanisms within each omics layer in response to chronic OP exposure. The non-convergence of signals from the two molecular layers suggests that the epigenomic and metabolomic profiles reflect different aspects of the human system responding to chronic OP exposure. Interestingly, many of the pathways are indeed connected. For example, oxidative stress can influence membrane trafficking and eventually lead to neurodegenerative diseases (Lie and Nixon 2019). Thus, metabolomic and epigenomic profiles may inform on different aspects of a molecular cascade that reflects pathogenic mechanisms triggered by chronic OP exposure. It should be noted that other omics layers such as the transcriptome and proteome may be more proximal to the mechanisms invoked by the changes in the epigenome or metabolome. A cross-sectional study examining

biologic correlates of birthweight integrated the methylome with the transcriptome and metabolome, as well as a set of inflammatory proteins in cord blood samples and showed – as one would expect - stronger correlations between gene expression and methylation and also between gene expression and metabolites than between the more distal omics layers (Alfano et al. 2020).

4.4.3 The epigenome and metabolome were connected through pathways involving metabolic processes and nutrient sensing

To investigate how DNA methylation and metabolism are possibly mutually regulating each other in response to chronic OP exposure, we correlated OP-related signals identified within one omics layer with signals from the other layer while adjusting for OP exposure and other factors. We identified very few significantly associated CpG-metabolite pairs (at FDR < 0.05 level), possibly due to limited statistical power or the inherent properties of these two molecular layers. A previous study assessed 640 blood metabolites from 1814 participants of the Kooperative Gesundheitsforschung in der Region Augsburg (KORA) population for their association with 457,004 CpG sites. Not surprisingly, the effect of DNA methylation on the metabolome was not as strong as the effect of the genome, and many DNA methylation-metabolite associations were confounded by environmental factors (Petersen et al. 2014). In our integration analysis, we controlled for OP exposure and other potential confounders by calculating partial correlations and observed only weak associations between the two molecular layers which might indicate that the responses are mostly independent from each other. Also, it is worth noting that our DNA methylation data were extracted from whole blood, while metabolites found in the serum are largely subject to the metabolic function of other tissues, such as liver, kidney, or muscle (Petersen et al. 2014).

Due to the limited metabolite annotation ability and the risk of type I error, it is not appropriate to discuss correlations between specific molecules in detail as targeted analyses with larger sample size and experimental validation are required. Nevertheless, it is worthwhile noting that DNA methylation signals associated with OP-related metabolomic features were mainly involved in metabolic processes and nutrient-sensing. For example, AMPK (AMP-activated protein kinase) is one of the central regulators of cellular metabolism (Mihaylova and Shaw 2011). Also, AMPK is involved in the rate-limiting steps for fatty-acid and sterol synthesis (Sato et al. 1993). Our finding that AMPK signaling pathway is associated with OP-related metabolomic changes coincides with the molecular function of AMPK. Meanwhile, the glucagon signaling pathway and Type II diabetes mellitus pathway are involved in regulating energy and glucose homeostasis. Thus, the epigenetic pathways found associated with OP-related metabolomic signaling are well-known metabolic regulators and, correspondingly, can be regulated by metabolic pathways. Although not directly associated with OP exposure, these CpG sites and epigenetic pathways may regulate the metabolome's response to OPs or might be subject to the perturbations of the metabolism, and these pathways are also associated with diseases that have been related to OP-exposures previously. Although we cannot draw causal inferences from these epigenomic-metabolomic associations, our pilot study demonstrates the feasibility of using epigenomic-metabolomic integration in environmental epidemiological studies of long-term chronic low-level exposures.

4.4.4 Limitations

Our study has several limitations. First, our sample size was limited to 176 participants due to the availability of both metabolomics and epigenomics data. The relative sample size could limit the statistical power, especially when assessing the effect of low-level OP exposure on high-

dimensional multi-omics datasets containing many noises. Nonetheless, as a hypothesis-generating and proof of principle study, our result elucidated the feasibility of utilizing multiple omic datasets to study chronic environmental exposure. Second, although our GIS-based pesticide estimation allowed us to investigate chronic exposure over decades, assigning absolute exposure levels for each OP chemical to each participant is challenging. Also, ambient pesticide estimation does not reflect other sources of exposure, for example, diet. Still, it is unlikely for exposure from diet to be associated with ambient pesticides. Therefore, it will not bias the ambient pesticide-epigenome relationship. Another limitation is that our epigenomics and metabolomics were measured within the blood. Although the blood can be an ideal surrogate tissue because it is readily available, future studies using other tissues, such as cerebrospinal fluid, could provide additional information. Finally, the limited metabolite annotation ability could influence the interpretation of our results. Adopting pathway and network analyses could improve annotation results, but it is still necessary to improve metabolite identification using either tandem MS or internal standards in future studies.

4.5 Conclusion

In summary, we linked GIS-modeled chronic low-level OP pesticide data with blood metabolomic and epigenomic data in an older adult population. Our findings suggest that OP exposure may alter the glycosphingolipid biosynthesis pathway as implicated by epigenome and metabolome profiles we identified. In addition, DNA methylation changes were involved in membrane trafficking and cell adhesion related mechanisms, while OP-related metabolites were involved in aromatic amino acid metabolism, oxidative stress, and mitochondrial function. Integrating epigenomic and metabolomic signals, we found that the metabolome and epigenome may mutually regulate each other through signaling pathways involved in metabolic processes and

nutrient sensing. Thus, it seems feasible to use multi-omics integration in studies of chronic environmental exposures in humans to gain a better understanding of the pathophysiology involved in pesticide-related chronic health outcomes.

4.6 Tables and figures

Table 4-1. Distribution of demographics and OP pesticide level

	PEG1 (n = 176)
Age	
Min	35
Max	92
Mean (SD)	66.13 (13.40)
Sex	
Male (%)	90 (51)
Female (%)	86 (49)
Race/Ethnicity	
White (%)	159 (90)
Hispanic (%)	17 (10)
Education	
Less than 12 years (%)	20 (11)
12 years (%)	32 (18)
More than 12 years (%)	124 (71)
Organophosphate	
Min	0
Max	24
Mean (SD)	2.38 (4.00)

Table 4-2. Top 20 differentially methylated CpGs associated with OP exposure

ID	Bicor r	Unadjusted P-value	CHR	MAPINFO	UCSC gene	UCSC Group	Island
cg26835374	0.362511739	7.65E-07	1	142618826			Island
cg09763120	0.350369446	1.87E-06	11	66529822	C11orf80	Body	Island
cg23037798	0.347832816	2.24E-06	1	23650159	HNRNPR	Body	
cg17559110	0.342802282	3.20E-06	11	984447	AP2A2	Body	S_Shelf
cg16471585	0.342731919	3.22E-06	12	118199103	KSR2	Body	Island
cg06279535	0.339835049	3.94E-06	21	45135999			N_Shelf
cg13282252	0.338653294	4.28E-06	3	188581493	LPP	Body	
cg04724556	0.332028496	6.74E-06	15	44507997			
cg17339258	-0.329151283	8.18E-06	14	59655074	DAAM1	TSS1500	Island
cg02308245	-0.327067115	9.40E-06	16	85833143	COX4I1;EMC8	1stExon	Island
cg18642271	-0.324439596	1.12E-05	19	46180314	GIPR	Body	Island
cg08166232	-0.323879446	1.16E-05	12	89918718	WDR51B;GALNT4	Body;TSS200	N_Shore
cg19832597	-0.323842372	1.16E-05	7	30634132	GARS	TSS200	Island
cg16170490	0.3213646	1.37E-05	22	42302146	SREBF2	3'UTR	N_Shelf
cg16266672	0.318256415	1.67E-05	8	22461575	KIAA1967;C8orf58	TSS1500;3'UTR	N_Shore
cg12593608	-0.315103206	2.05E-05	15	72668042	C15orf34;HEXA	TSS1500;Body	Island
cg13639866	0.311990878	2.50E-05	14	101430211	SNORD114-8	TSS1500	
cg24350213	-0.311029792	2.65E-05	17	39943012	JUP	TSS200	Island
cg05008688	0.309917102	2.84E-05	5	121464372	ZNF474	TSS1500	
cg26016985	0.309374376	2.94E-05	8	141821152	PTK2	Body	

Chr: chromosome

Table 4-3. Confirmed[a] OP-related metabolites

m/z	RT (s)	Adduct Form	Metabolite	Bicor r[b]	Platform
110.0601	272.5	M+H	4-Aminophenol	-0.163	HILIC
142.0976	29.6	M+H	L-Histidinol	0.049	HILIC
153.0772	102.6	M+H	L-Arabitol	0.014	HILIC
181.0720	196.9	M+H	D-Galactose	-0.177	HILIC
227.1136	110.4	M+H	Carnosine	0.169	HILIC
306.0489	118.2	M+H	Cytidine 23-cyclic phosphate	-0.130	HILIC
347.2207	47.9	M+H	Cortexolone	0.157	HILIC
88.0404	47.0	M-H	Beta-Alanine	0.156	C18
88.9880	123.7	M-H	Oxalic acid	0.019	C18
116.0353	37.7	M-H	Acetylglycine	0.017	C18
129.0922	279.0	M-H	Heptanoic acid	-0.150	C18
135.0452	252.4	M-H	Phenylacetic acid	-0.192	C18
154.0623	50.6	M-H	L-Histidine	0.128	C18
158.0612	74.9	M-H	Indoleacetaldehyde	-0.002	C18
159.0650	89.8	M-H	Pimelic acid	0.000	C18
164.0717	47.2	M-H	L-Phenylalanine	0.016	C18
165.0558	177.7	M-H	3-(2-Hydroxyphenyl) propanoic acid	0.106	C18
177.0405	282.6	M-H	L-Gulonolactone	-0.127	C18
178.0510	41.8	M-H	Hippuric acid	0.157	C18
187.1340	38.1	M-H	9-Hydroxydecanoic acid	-0.048	C18
191.0198	42.7	M-H	Citric acid	0.001	C18
203.0825	46.6	M-H	L-Tryptophan	0.026	C18
583.2554	127.8	M-H	Bilirubin	-0.006	C18

[a] Chemical identification was conducted by matching peaks by accurate mass and retention time to authentic reference standards in an in-house library run under identical conditions using tandem mass spectrometry.
[b] Correlations derived from biweighted midcorrelations.
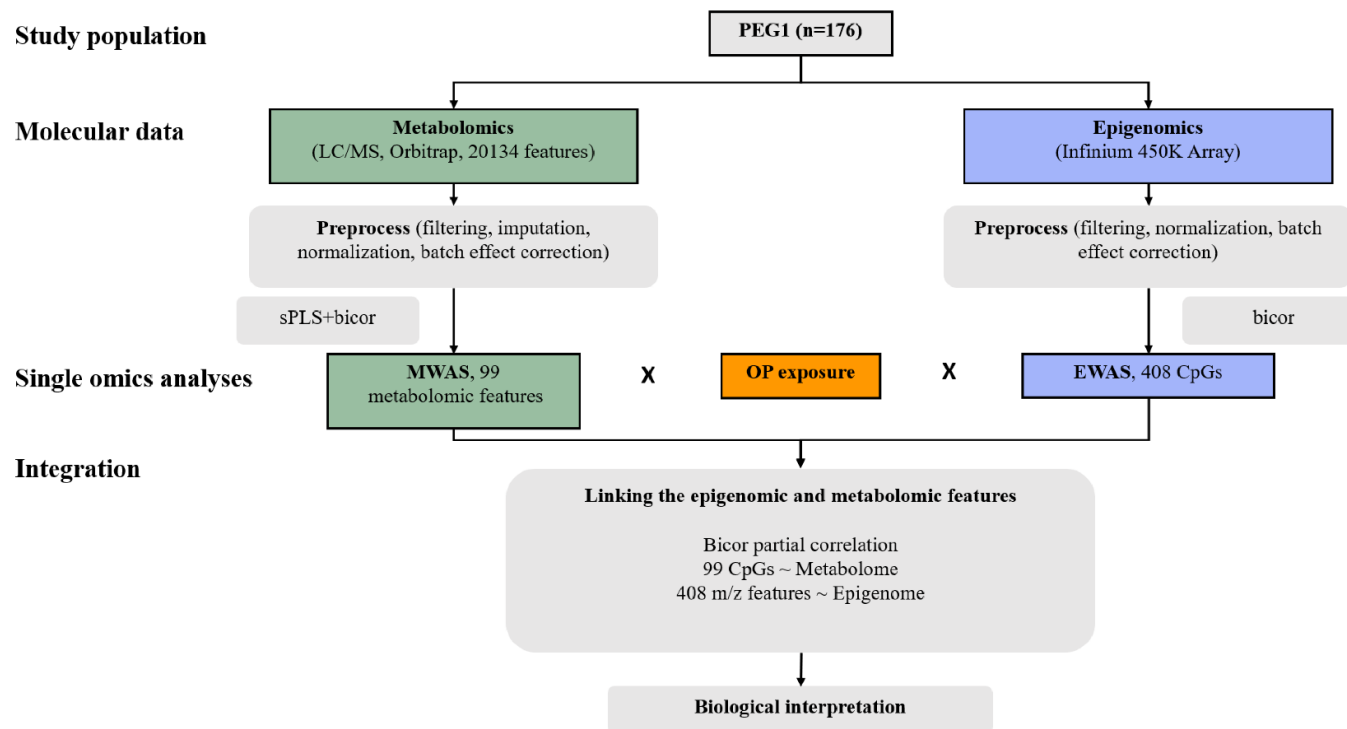
Figure 4-1. Analytical workflow. Our study population includes 176 control participants residing in the Central Valley of California, from 2001 to 2007. Ambient OP pesticide exposure for each participant was estimated in a GIS-based model. Metabolomic and epigenomic profiles were extracted from blood. We first identified CpGs and metabolomic features significantly associated with OP exposure, and subsequently integrated OP-related features from the two molecular layers by calculating pairwise bicor correlations.
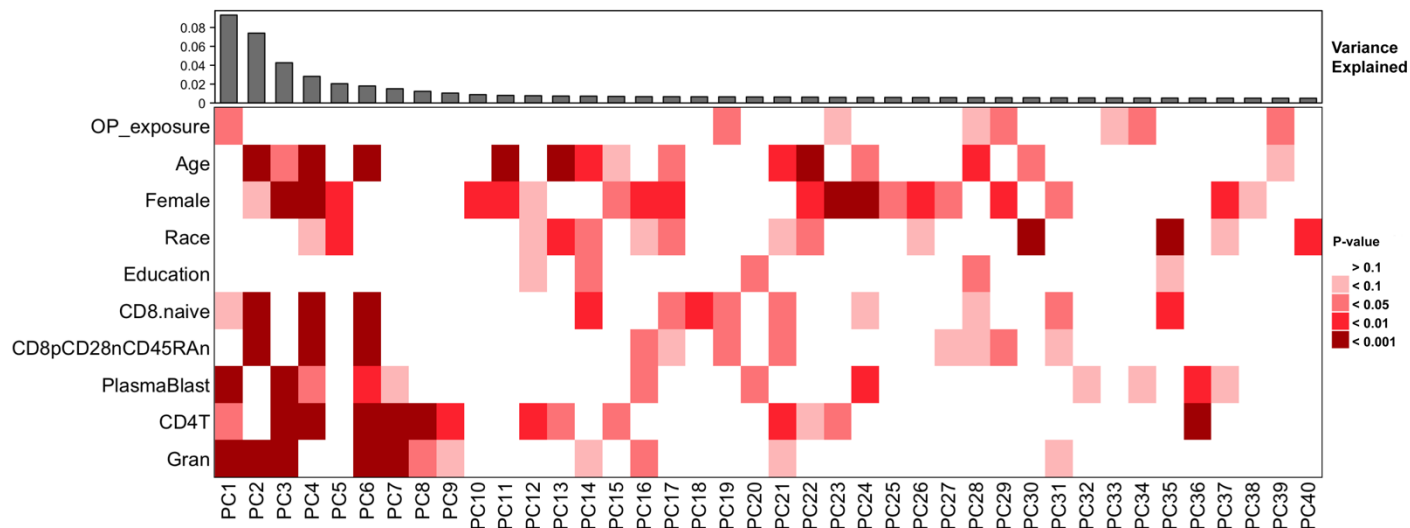
Figure 4-2. Associations with blood DNA methylation variability. Principal components (PCs) were correlated with OP exposure and covariates with bicor. Bicor p-values color-coded by smallest p-value (dark red; P < 0.001) to largest (blank; P > 0.10). The top 40 PCs explained approximately 50% of the variance in blood DNA methylation. The contribution of PCs to the variance explained is shown on the top.
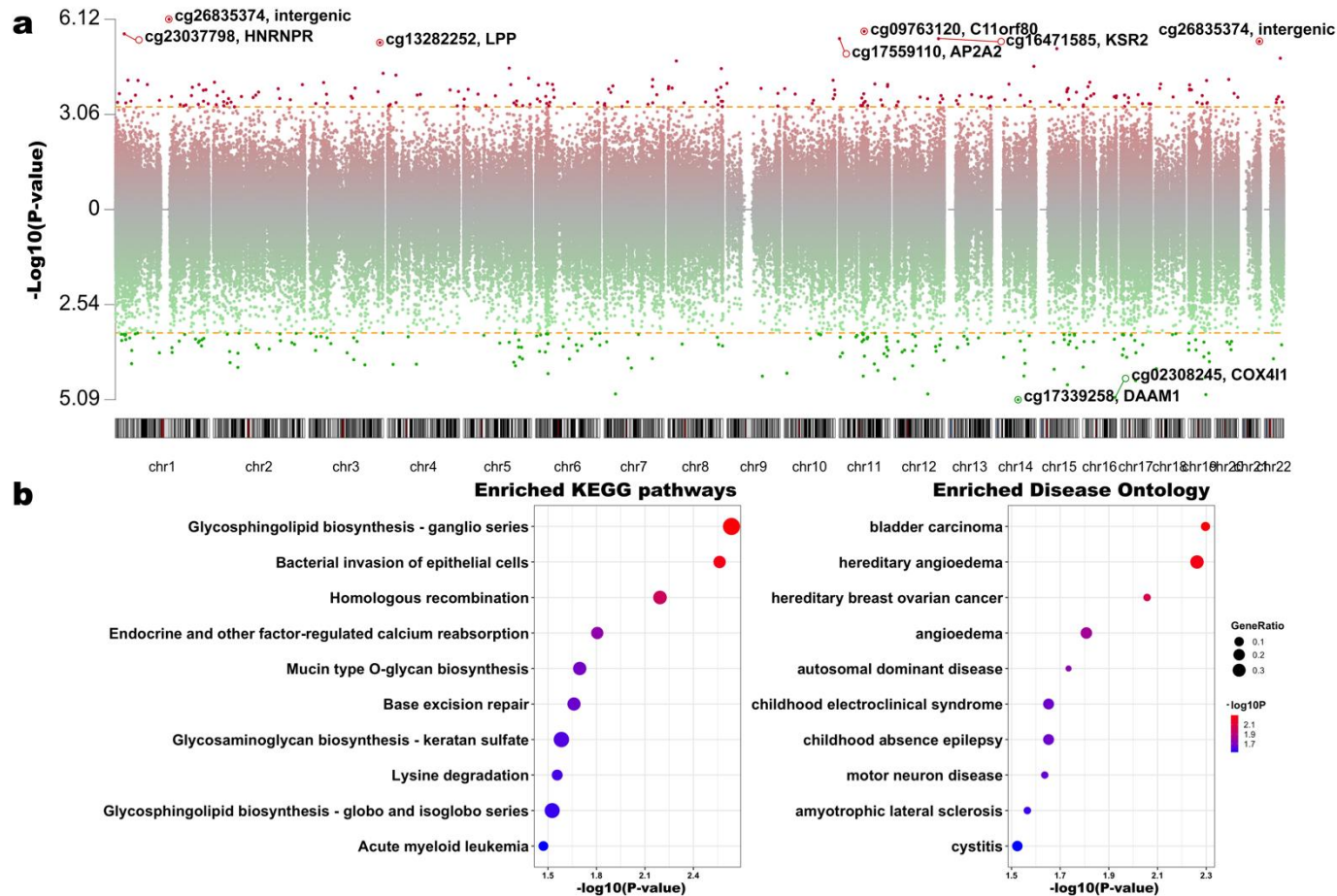
Figure 4-3. Identification of CpGs and genetic pathways associated with OP pesticides exposure. Manhattan plot (**a**) showing the significance of associations between methylation levels of CpGs and OP exposure during the 10 years prior to the blood draw. The y-axis shows the -log10(p-value) after adjustment for covariates and estimated cell proportions. Red dots represent CpG sites hypermethylated, and green dots represent CpG sites hypomethylated, respectively. And dot plots (**b**) show results of functional enrichment analyses. The left panel shows the top ten enriched KEGG pathways based on 408 OP-related CpGs. And the right panel shows the top ten enriched disease ontologies. The size of the dot represents the enrichment ratio.
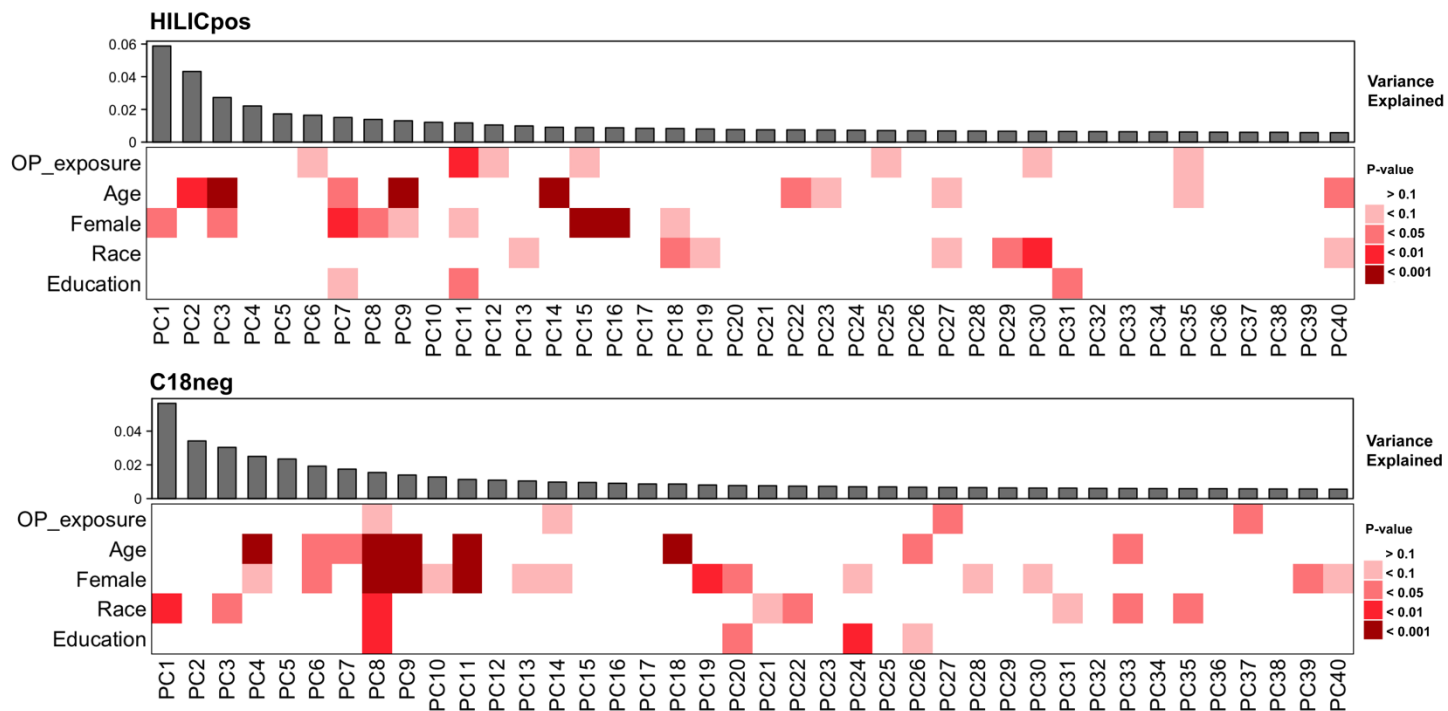
Figure 4-4. Associations with serum metabolome variability. Principal components (PCs) were correlated with OP exposure and covariates with bicor. Bicor p-values color-coded by smallest p-value (dark red; p < 0.001) to largest (blank; P > 0.10). The top 40 PCs explained approximately 50% of the variance of the blood DNA methylation data. The contribution of PCs is shown on the top.
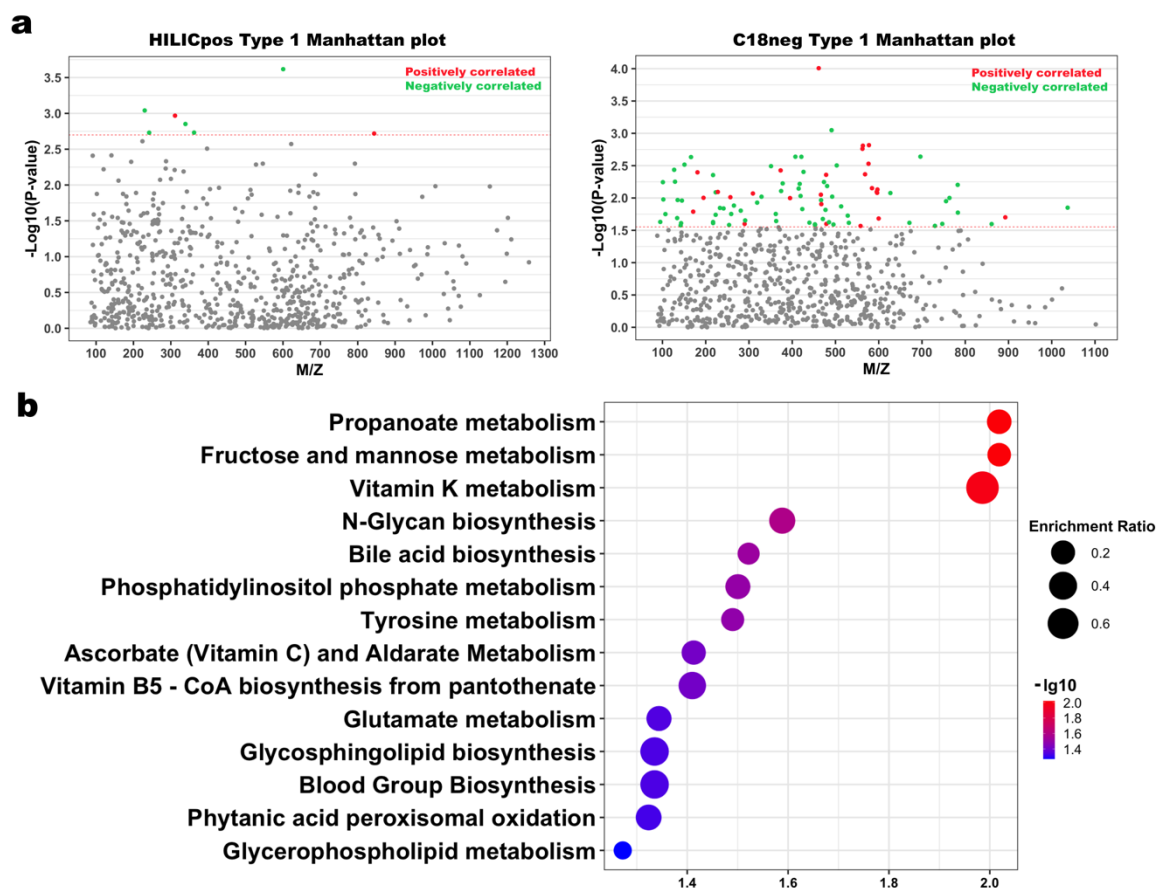
Figure 4-5. Identification of metabolomic features and pathways associated with OP pesticides exposure. (**a**) Left: Type 1 Manhattan plot for features in the HILIC column (positive ion mode), -log10(p-value) vs mass-to-charge (m/z). Only features selected by sparse PLS were included here. 7 m/z features were found with FDR < 0.2. Red dots represent the features that were positively correlated with OP, and the green dots represent the features that were negatively correlated with OP; Right: Type 1 Manhattan plot for features in the C18 column (negative ion mode), -log10(p-value) vs. mass-to-charge. 92 m/z features were found with FDR < 0.2. Dot plot (**b**) shows the results of mummichog pathway enrichment analyses. The size of the dot represents the enrichment ratio.
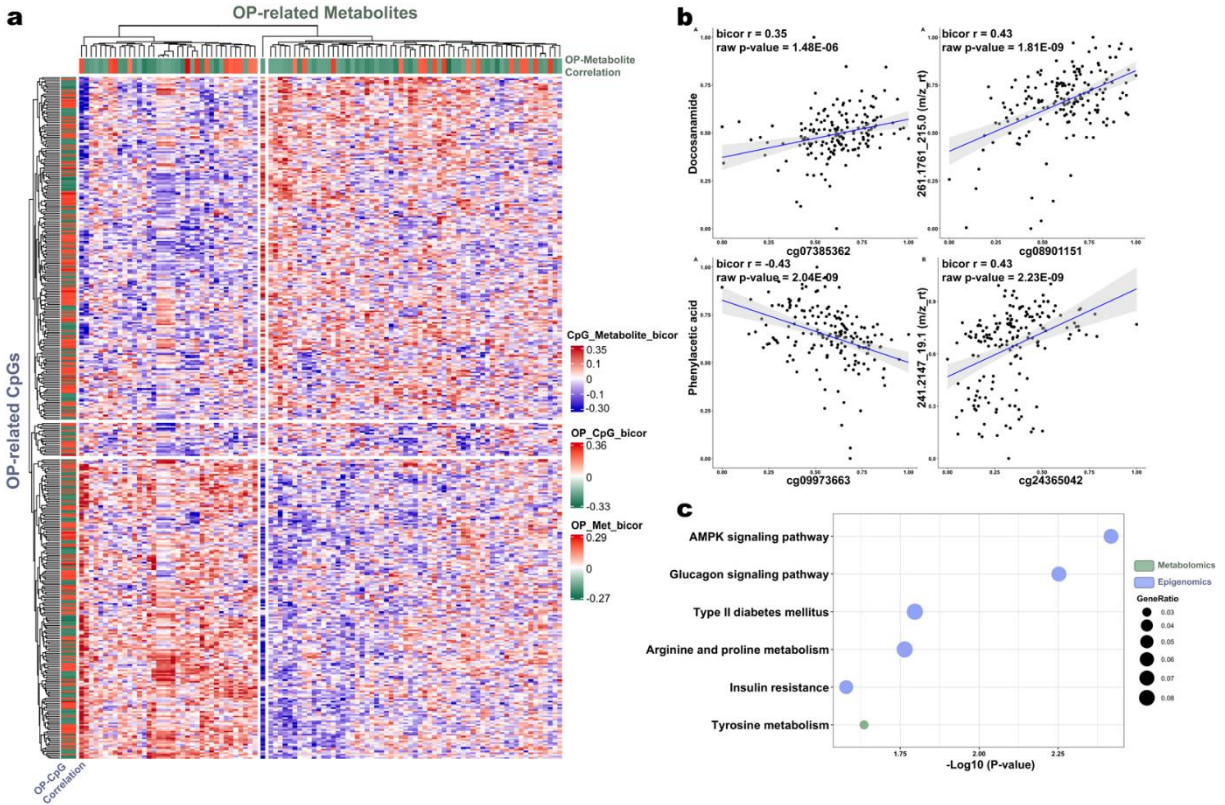
Figure 4-6. Metabolomics-driven and methylation-driven integration. (**a**) Heatmap showed the biweighted midcorrelations between OP exposure, 99 OP-related metabolomic features (columns), and 408 OP-related CpGs (rows). The color of the cell indicates the correlation between OP-related CpGs and OP-related metabolites; the row and column annotations indicate the correlation between OP and CpGs or metabolites, respectively. (**b**) scatter plots of CpG methylation levels and relative intensities of metabolomic features. Upper left: between OP-related CpGs and OP-related metabolomic features, we identified one pair, cg07385362 (*UNKL*) and docosanamide, with FDR = 0.06; upper left: between OP-related CpGs and the rest of the metabolome, we identified one pair, cg08901151 (*MAD1L1*) and an unidentified m/z feature (m/z = 261.1761), with FDR = 0.01; bottom: between OP-related metabolomic features and the rest of the methylome, we identified two pairs with FDR < 0.05, including cg09973663 (*MYO1B*) with phenylacetic acid, and cg24365042 (*ATF3*) with an unidentified metabolite (m/z = 241.2147). (**c**) Dot plot shows results of functional enrichment analyses. By correlating OP-related metabolomic features with the rest of the methylome, we identified 515 pairs with FDR < 0.2 involving 84 OP-related metabolomic features and 511 CpGs. We conducted KEGG pathway enrichment analysis and *mummichog* enrichment analysis to identify genetic and metabolomic pathways significantly enriched.

4.7 Supplemental materials

Supplemental Table 4-S1. Interaction between OP-related CpG genes and OP chemicals based on CTD database

Chemical Name	Gene Symbol	Interaction	PubMedID
Monocrotophos	APEX1	Monocrotophos results in increased expression of and affects the localization of APEX1 protein	28887667
Monocrotophos	APEX1	Monocrotophos results in increased expression of APEX1 mRNA	28887667
Monocrotophos	SREBF2	Monocrotophos results in increased expression of SREBF2 mRNA	32267039
Dicrotophos	AGPS	dicrotophos results in decreased expression of AGPS mRNA	28302478
Dicrotophos	AHDC1	dicrotophos results in increased expression of AHDC1 mRNA	28302478
Dicrotophos	AKAP9	dicrotophos results in decreased expression of AKAP9 mRNA	28302478
Dicrotophos	ANKRD11	dicrotophos results in increased expression of ANKRD11 mRNA	28302478
Dicrotophos	AP2A2	dicrotophos results in increased expression of AP2A2 mRNA	28302478
Dicrotophos	AP3D1	dicrotophos results in increased expression of AP3D1 mRNA	28302478
Dicrotophos	ASAP1	dicrotophos results in increased expression of ASAP1 mRNA	28302478
Dicrotophos	B4GALNT1	dicrotophos results in increased expression of B4GALNT1 mRNA	28302478
Dicrotophos	BLM	dicrotophos results in decreased expression of BLM mRNA	28302478
Dicrotophos	C1GALT1	dicrotophos results in decreased expression of C1GALT1 mRNA	28302478
Dicrotophos	C1QTNF6	dicrotophos results in increased expression of C1QTNF6 mRNA	28302478
Dicrotophos	CASZ1	dicrotophos results in increased expression of CASZ1 mRNA	28302478
Dicrotophos	CCDC126	dicrotophos results in decreased expression of CCDC126 mRNA	28302478
Dicrotophos	CHD6	dicrotophos results in increased expression of CHD6 mRNA	28302478
Dicrotophos	CNN2	dicrotophos results in increased expression of CNN2 mRNA	28302478
Dicrotophos	CNOT8	dicrotophos results in decreased expression of CNOT8 mRNA	28302478
Dicrotophos	COL9A3	dicrotophos results in decreased expression of COL9A3 mRNA	28302478
Dicrotophos	CREBBP	dicrotophos results in increased expression of CREBBP mRNA	28302478
Dicrotophos	FAM118B	dicrotophos results in decreased expression of FAM118B mRNA	28302478
Dicrotophos	FRAS1	dicrotophos results in increased expression of FRAS1 mRNA	28302478
Dicrotophos	GTF2H5	dicrotophos results in decreased expression of GTF2H5 mRNA	28302478
Dicrotophos	GTF3C6	dicrotophos results in decreased expression of GTF3C6 mRNA	28302478
Dicrotophos	HADHB	dicrotophos results in decreased expression of HADHB mRNA	28302478
Dicrotophos	HERC4	dicrotophos results in decreased expression of HERC4 mRNA	28302478
Dicrotophos	IFT140	dicrotophos results in increased expression of IFT140 mRNA	28302478
Dicrotophos	INCENP	dicrotophos results in increased expression of INCENP mRNA	28302478
Dicrotophos	JAKMIP2	dicrotophos results in decreased expression of JAKMIP2 mRNA	28302478
Dicrotophos	KDM6B	dicrotophos results in increased expression of KDM6B mRNA	28302478
Dicrotophos	KIF20B	dicrotophos results in decreased expression of KIF20B mRNA	28302478
Dicrotophos	LOX	dicrotophos results in increased expression of LOX mRNA	28302478
Dicrotophos	LPP	dicrotophos results in increased expression of LPP mRNA	28302478
Dicrotophos	NBPF1	dicrotophos results in increased expression of NBPF1 mRNA	28302478
Dicrotophos	NCLN	dicrotophos results in increased expression of NCLN mRNA	28302478
Dicrotophos	OMA1	dicrotophos results in decreased expression of OMA1 mRNA	28302478
Dicrotophos	OSMR	dicrotophos results in increased expression of OSMR mRNA	28302478
Dicrotophos	PAK1IP1	dicrotophos results in decreased expression of PAK1IP1 mRNA	28302478
Dicrotophos	PDHB	dicrotophos results in decreased expression of PDHB mRNA	28302478
Dicrotophos	PKD1	dicrotophos results in increased expression of PKD1 mRNA	28302478
Dicrotophos	PLEKHA6	dicrotophos results in increased expression of PLEKHA6 mRNA	28302478
Dicrotophos	POLD4	dicrotophos results in decreased expression of POLD4 mRNA	28302478
Dicrotophos	PPARD	dicrotophos results in increased expression of PPARD mRNA	28302478
Dicrotophos	PSMB1	dicrotophos results in decreased expression of PSMB1 mRNA	28302478
Dicrotophos	RAB1A	dicrotophos results in decreased expression of RAB1A mRNA	28302478
Dicrotophos	RAB8B	dicrotophos results in decreased expression of RAB8B mRNA	28302478
Dicrotophos	RARS2	dicrotophos results in decreased expression of RARS2 mRNA	28302478
Dicrotophos	RELA	dicrotophos results in increased expression of RELA mRNA	28302478
Dicrotophos	RIOK2	dicrotophos results in decreased expression of RIOK2 mRNA	28302478
Dicrotophos	RPL26	dicrotophos results in decreased expression of RPL26 mRNA	28302478
Dicrotophos	SERINC1	dicrotophos results in decreased expression of SERINC1 mRNA	28302478
Dicrotophos	SETD1B	dicrotophos results in increased expression of SETD1B mRNA	PubMedID
Dicrotophos	SLC9A1	dicrotophos results in increased expression of SLC9A1 mRNA	28302478
Dicrotophos	SMTN	dicrotophos results in increased expression of SMTN mRNA	28302478
Dicrotophos	SMYD3	dicrotophos results in decreased expression of SMYD3 mRNA	28302478
Dicrotophos	SNX13	dicrotophos results in decreased expression of SNX13 mRNA	28302478
Dicrotophos	SPCS2	dicrotophos results in decreased expression of SPCS2 mRNA	28302478
Dicrotophos	SPRED2	dicrotophos results in increased expression of SPRED2 mRNA	28302478
Dicrotophos	SREBF2	dicrotophos results in decreased expression of SREBF2 mRNA	28302478
Dicrotophos	SSH2	dicrotophos results in increased expression of SSH2 mRNA	28302478
Dicrotophos	TAF1D	dicrotophos results in decreased expression of TAF1D mRNA	28302478
Dicrotophos	TAPBP	dicrotophos results in increased expression of TAPBP mRNA	28302478
Dicrotophos	TERT	dicrotophos results in increased expression of TERT mRNA	28302478
Dicrotophos	TP53INP2	dicrotophos results in increased expression of TP53INP2 mRNA	28302478
Dicrotophos	TUBD1	dicrotophos results in decreased expression of TUBD1 mRNA	28302478
Dicrotophos	UNKL	dicrotophos results in increased expression of UNKL mRNA	28302478
Dicrotophos	VPS45	dicrotophos results in decreased expression of VPS45 mRNA	28302478
Dicrotophos	WASF2	dicrotophos results in increased expression of WASF2 mRNA	28302478
Dicrotophos	WDR92	dicrotophos results in decreased expression of WDR92 mRNA	28302478
Dicrotophos	YEATS2	dicrotophos results in increased expression of YEATS2 mRNA	28302478

Dicrotophos	ZFP36L2	dicrotophos results in increased expression of ZFP36L2 mRNA	28302478
Dicrotophos	ZNF680	dicrotophos results in decreased expression of ZNF680 mRNA	28302478
Dicrotophos	ZNF70	dicrotophos results in increased expression of ZNF70 mRNA	28302478
Dichlorvos	GZMA	Dichlorvos results in decreased expression of GZMA mRNA	16002202
Dichlorvos	GZMA	Dichlorvos results in decreased expression of GZMA protein	16002202
Diazinon	BMP6	Diazinon results in increased expression of BMP6 mRNA	29108742
Diazinon	CDKN1C	Diazinon affects the expression of CDKN1C mRNA	22546817
Diazinon	CX3CR1	Diazinon results in decreased expression of CX3CR1 mRNA	24356939
Diazinon	ETS1	Diazinon affects the expression of ETS1 mRNA	22546817
Diazinon	GNA11	Diazinon results in decreased expression of GNA11 mRNA	17452286
Diazinon	KIF20B	Diazinon affects the expression of KIF20B mRNA	22546817
Diazinon	LATS2	Diazinon affects the expression of LATS2 mRNA	22546817
Diazinon	PKD1	Diazinon affects the expression of PKD1 mRNA	22546817
Diazinon	PKIB	Diazinon results in increased expression of PKIB mRNA	17452286
Chlorpyrifos	AKAP9	[Copper co-treated with Chlorpyrifos] results in decreased expression of AKAP9 mRNA	21356183
Chlorpyrifos	APEX1	Chlorpyrifos results in increased expression of and affects the localization of APEX1 protein	28887667
Chlorpyrifos	APEX1	Chlorpyrifos results in increased expression of APEX1 mRNA	28887667
Chlorpyrifos	ATF3	Chlorpyrifos results in increased expression of ATF3 mRNA	21356183
Chlorpyrifos	CLIP3	Chlorpyrifos results in increased expression of CLIP3 protein	16790487
Chlorpyrifos	CPEB2	Chlorpyrifos results in decreased expression of CPEB2 mRNA	30621213
Chlorpyrifos	CPEB2	Chlorpyrifos results in increased expression of CPEB2 mRNA	30621213
Chlorpyrifos	DLGAP2	Chlorpyrifos results in decreased expression of DLGAP2 mRNA	18668222
Chlorpyrifos	FNTB	Chlorpyrifos results in increased expression of FNTB mRNA	18668222
Chlorpyrifos	GTF3C6	Chlorpyrifos affects the expression of GTF3C6 mRNA	27905518
Chlorpyrifos	GTF3C6	[Chlorpyrifos co-treated with Ethylenethiourea] results in decreased expression of GTF3C6 mRNA	27905518
Chlorpyrifos	HTRA2	Chlorpyrifos results in decreased expression of HTRA2 mRNA	20350560
Chlorpyrifos	NDUFS3	3-methyladenine inhibits the reaction [Chlorpyrifos results in decreased expression of NDUFS3 protein]	26070385
Chlorpyrifos	NDUFS3	Chlorpyrifos results in decreased expression of NDUFS3 protein	26070385
Chlorpyrifos	NDUFS3	PRKN protein inhibits the reaction [Chlorpyrifos results in decreased expression of NDUFS3 protein]	26070385
Chlorpyrifos	PHLDA2	Chlorpyrifos results in increased expression of PHLDA2 mRNA	21356183
Chlorpyrifos	PKIB	Chlorpyrifos results in increased expression of PKIB mRNA	17452286
Chlorpyrifos	PTK2	Chlorpyrifos results in increased expression of PTK2 mRNA	20691718
Chlorpyrifos	RELA	Chlorpyrifos results in decreased expression of RELA mRNA	30621213
Chlorpyrifos	SSH2	Chlorpyrifos results in increased expression of SSH2 mRNA	21356183
Chlorpyrifos	ST8SIA3	Chlorpyrifos results in decreased expression of ST8SIA3 mRNA	18668222
Chlorpyrifos	TTF1	[Copper co-treated with Chlorpyrifos] results in decreased expression of TTF1 mRNA	21356183
Chlorpyrifos	ZFP36L2	Chlorpyrifos results in decreased expression of ZFP36L2 mRNA	27905518
Chlorpyrifos	ZFP36L2	Chlorpyrifos results in increased expression of ZFP36L2 mRNA	30621213
Chlorpyrifos	ZFP36L2	Chlorpyrifos results in decreased expression of ZFP36L2 mRNA	27905518
Malathion	EPS8L1	Malathion results in increased expression of EPS8L1 mRNA	32069766
Malathion	RELA	Malathion results in increased expression of RELA mRNA	28987951
Malathion	RELA	Taurine inhibits the reaction [Malathion results in increased expression of RELA mRNA]	28987951
Phosalone	RELA	Ellagic Acid inhibits the reaction [phosalone results in increased expression of RELA protein]	27965107
Phosalone	RELA	phosalone results in increased expression of RELA protein	27965107

Supplemental Table 4-S2. OP-related CpGs enrichment analysis of tissue or cell-type specific DNase hypersensitive sites in Roadmap

Cell	Tissue	Datatype	Accession	Unadjusted P-value
E005 H1 BMP4 Derived Trophoblast Cultured Cells	ES Cell	DHS	E005	0.00379
E017 IMR90 fetal lung fibroblasts Cell Line	Lung	DHS	E017	0.00731
E029 Primary monocytes from peripheral blood	Blood	DHS	E029	0.00868
E028 Breast variant Human Mammary Epithelial Cells	Breast	DHS	E028	0.00983
E094 Gastric	Gastric	DHS	E094	0.0103
E046 Primary Natural Killer cells from peripheral	Blood	DHS	E046	0.0184
E033 Primary T cells from cord blood	Blood	DHS	E033	0.0272
E034 Primary T cells from peripheral blood	Blood	DHS	E034	0.0297
E056 Foreskin Fibroblast Primary Cells skin02	Skin	DHS	E056	0.0336
E100 Psoas Muscle	Psoas Muscle	DHS	E100	0.0459

Supplemental Table 4-S3. Top enriched GO terms based on 408 OP-related CpGs

KEGG ID	Pathway Name	Background Ratio	P Value	Q Value	Count	Type
GO:0000123	histone acetyltransferase complex	77/17403	5.83E-04	1.30E-01	7	CC
GO:0031248	protein acetyltransferase complex	86/17403	1.13E-03	1.30E-01	7	CC
GO:1902493	acetyltransferase complex	86/17403	1.13E-03	1.30E-01	7	CC
GO:0030914	STAGA complex	12/17403	1.25E-03	1.30E-01	3	CC
GO:0014704	intercalated disc	49/17403	2.14E-03	1.77E-01	5	CC
GO:0005795	Golgi stack	127/17403	2.66E-03	1.84E-01	8	CC
GO:0042575	DNA polymerase complex	20/17403	5.81E-03	3.11E-01	3	CC
GO:0001725	stress fiber	66/17403	7.77E-03	3.11E-01	5	CC
GO:0097517	contractile actin filament bundle	66/17403	7.77E-03	3.11E-01	5	CC
GO:0030117	membrane coat	94/17403	8.35E-03	3.11E-01	6	CC
GO:0048475	coated membrane	94/17403	8.35E-03	3.11E-01	6	CC
GO:0044291	cell-cell contact zone	69/17403	9.34E-03	3.11E-01	5	CC
GO:0070461	SAGA-type complex	24/17403	9.77E-03	3.11E-01	3	CC
GO:0032432	actin filament bundle	72/17403	1.11E-02	3.28E-01	5	CC
GO:0032580	Golgi cisterna membrane	74/17403	1.24E-02	3.43E-01	5	CC
GO:0031984	organelle subcompartment	346/17403	1.36E-02	3.50E-01	13	CC
GO:0016342	catenin complex	28/17403	1.50E-02	3.52E-01	3	CC
GO:0042641	actomyosin	78/17403	1.53E-02	3.52E-01	5	CC
GO:0098791	Golgi apparatus subcompartment	326/17403	2.00E-02	3.72E-01	12	CC
GO:0015629	actin cytoskeleton	484/17403	2.01E-02	3.72E-01	16	CC
GO:0060249	anatomical structure homeostasis	412/16658	5.50E-04	5.30E-01	18	BP

GO:0071214	cellular response to abiotic stimulus	314/16658	6.39E-04	5.30E-01	15	BP
GO:0104004	cellular response to environmental stimulus	314/16658	6.39E-04	5.30E-01	15	BP
GO:0061013	regulation of mRNA catabolic process	192/16658	7.78E-04	5.30E-01	11	BP
GO:0051489	regulation of filopodium assembly	41/16658	8.32E-04	5.30E-01	5	BP
GO:0035721	intraciliary retrograde transport	12/16658	1.15E-03	5.51E-01	3	BP
GO:0061614	pri-miRNA transcription by RNA polymerase II	46/16658	1.41E-03	5.51E-01	5	BP
GO:0036376	sodium ion export across plasma membrane	14/16658	1.85E-03	5.51E-01	3	BP
GO:0006363	termination of RNA polymerase I transcription	30/16658	2.00E-03	5.51E-01	4	BP
GO:0051291	protein heterooligomerization	125/16658	2.00E-03	5.51E-01	8	BP
GO:0031667	response to nutrient levels	470/16658	2.41E-03	5.51E-01	18	BP
GO:0042339	keratan sulfate metabolic process	32/16658	2.55E-03	5.51E-01	4	BP
GO:0055117	regulation of cardiac muscle contraction	77/16658	2.73E-03	5.51E-01	6	BP
GO:0006883	cellular sodium ion homeostasis	16/16658	2.77E-03	5.51E-01	3	BP
GO:0055119	relaxation of cardiac muscle	16/16658	2.77E-03	5.51E-01	3	BP
GO:0120032	regulation of plasma membrane bounded cell projection assembly	164/16658	3.07E-03	5.51E-01	9	BP
GO:0046847	filopodium assembly	55/16658	3.14E-03	5.51E-01	5	BP
GO:0035994	response to muscle stretch	17/16658	3.32E-03	5.51E-01	3	BP
GO:0060491	regulation of cell projection assembly	166/16658	3.33E-03	5.51E-01	9	BP
GO:0034660	ncRNA metabolic process	448/16658	3.46E-03	5.51E-01	17	BP
GO:0003730	mRNA 3'-UTR binding	82/16050	1.63E-04	8.49E-02	8	MF
GO:0035925	mRNA 3'-UTR AU-rich region binding	26/16050	1.37E-03	1.90E-01	4	MF
GO:0001098	basal transcription machinery binding	68/16050	1.83E-03	1.90E-01	6	MF
GO:0001099	basal RNA polymerase II transcription machinery binding	68/16050	1.83E-03	1.90E-01	6	MF
GO:0017091	AU-rich element binding	28/16050	1.83E-03	1.90E-01	4	MF

GO:0003924	GTPase activity	299/16050	4.86E-03	3.19E-01	13	MF
GO:0008373	sialyltransferase activity	20/16050	6.10E-03	3.19E-01	3	MF
GO:0045182	translation regulator activity	120/16050	8.00E-03	3.19E-01	7	MF
GO:0016757	transferase activity, transferring glycosyl groups	265/16050	1.27E-02	3.19E-01	11	MF
GO:0003713	transcription coactivator activity	305/16050	1.41E-02	3.19E-01	12	MF
GO:0005525	GTP binding	342/16050	1.41E-02	3.19E-01	13	MF
GO:0008236	serine-type peptidase activity	166/16050	1.41E-02	3.19E-01	8	MF
GO:0017166	vinculin binding	10/16050	1.46E-02	3.19E-01	2	MF
GO:0000993	RNA polymerase II complex binding	50/16050	1.48E-02	3.19E-01	4	MF
GO:0032550	purine ribonucleoside binding	346/16050	1.54E-02	3.19E-01	13	MF
GO:0019888	protein phosphatase regulator activity	77/16050	1.56E-02	3.19E-01	5	MF
GO:0051059	NF-kappaB binding	28/16050	1.57E-02	3.19E-01	3	MF
GO:0016825	hydrolase activity, acting on acid phosphorus-nitrogen bonds	170/16050	1.61E-02	3.19E-01	8	MF
GO:0017171	serine hydrolase activity	170/16050	1.61E-02	3.19E-01	8	MF
GO:0001883	purine nucleoside binding	349/16050	1.64E-02	3.19E-01	13	MF

Supplemental Table 4-S4. Identified eQTMs from 408 OP-related CpGs

PValue	CpG	CpG_Chr	CpG_ChrPos	CpG_gene	Gene_Chr	Gene_CenterChrPos	CisTrans	SNPType	AlleleAssessed	HGNCName
1.62E-09	cg17003212	10	97054967	TNXB	10	97050781	cis	C/T	C	PDLIM1
1.87E-06	cg26668919	6	35311486	TNXB	6	35310335	cis	C/T	C	PPARD
2.83E-04	cg06639387	1	228594973	TNXB	1	228823162	cis	C/T	C	FTH1P2

Supplemental Table 4-S5. Sensitivity analysis: top 20 differentially methylated CpGs associated with OP exposure, unadjusted for cell compositions

ID	Bicor r	Unadjusted P-value	CHR	MAPINFO	UCSC gene	UCSC Group
cg16471585	0.34634389	2.49E-06	12	118199103	KSR2	Body
cg09763120	0.34374536	3.00E-06	11	66529822	C11orf80	Body
cg23037798	0.34193852	3.40E-06	1	23650159	HNRNPR	Body
cg26835374	0.34029754	3.82E-06	1	142618826		
cg13282252	0.33754001	4.62E-06	3	188581493	LPP	Body
cg06279535	0.33642282	4.99E-06	21	45135999		
cg17339258	-0.3324684	6.54E-06	14	59655074	DAAM1	TSS1500
cg04724556	0.33051965	7.46E-06	15	44507997		
cg19832597	-0.3259911	1.01E-05	7	30634132	GARS	TSS200
cg05008688	0.32324406	1.21E-05	5	121464372	ZNF474	TSS1500
cg16170490	0.32132282	1.37E-05	22	42302146	SREBF2	3'UTR
cg17559110	0.32095833	1.40E-05	11	984447	AP2A2	Body
cg12593608	-0.3186597	1.63E-05	15	72668042	HEXA	TSS1500
cg13639866	0.31626334	1.90E-05	14	101430211	AL132709.8	TSS1500
cg08166232	-0.3146495	2.11E-05	12	89918718	GALNT4	Body
cg24350213	-0.3136159	2.25E-05	17	39943012	JUP	TSS200
cg16656979	-0.3133645	2.29E-05	15	37403242		
cg25470741	-0.3132502	2.30E-05	19	44617434	ZNF225	TSS200
cg26016985	0.31186559	2.51E-05	8	141821152	PTK2	Body
cg13699934	-0.3115614	2.56E-05	9	93956018		

Supplemental Table 4-S6. Enriched metabolic pathways associated with OP exposure

Pathways	Overlap size	Pathway size	p-value	platform
Propanoate metabolism	4	18	0.00958	HILICpos
Fructose and mannose metabolism	5	27	0.00958	HILICpos
Bile acid biosynthesis	5	36	0.03008	HILICpos
Glycerophospholipid metabolism	10	109	0.05344	HILICpos
Vitamin K metabolism	3	4	0.01034	C18neg
N-Glycan biosynthesis	7	24	0.0258	C18neg
Phosphatidylinositol phosphate metabolism	10	42	0.03159	C18neg
Tyrosine metabolism	37	221	0.03235	C18neg
Ascorbate (Vitamin C) and Aldarate Metabolism	13	62	0.03865	C18neg
Vitamin B5 - CoA biosynthesis from pantothenate	4	11	0.0389	C18neg
Glutamate metabolism	8	33	0.04529	C18neg
Glycosphingolipid biosynthesis	3	7	0.04621	C18neg
Blood Group Biosynthesis	3	7	0.04621	C18neg
Phytanic acid peroxisomal oxidation	6	22	0.04748	C18neg

pathway	Name	Chemical ID	Formula	ion	mz	time	coef	pvalue	Platform	Metabolite Annotation Confidence
Ascorbate (Vitamin C) and Aldarate Metabolism	Oxalic acid	C00209	C2H2O4	M-H[-]	88.988	123.7	0.019	8.06E-01	C18neg	1
	L-Gulonolactone	C01040	C6H10O6	M-H[-]	177.0405	282.6	-0.127	9.28E-02	C18neg	1
	Citric acid	C00158	C6H8O7	M-H[-]	191.0198	42.7	0.001	9.89E-01	C18neg	1
	Phenylacetaldehyde	C00601	C8H8O	M+Br[-]	198.9756	45.3	0.069	3.61E-01	C18neg	3
	cis-Aconitate	C00417	C6H6O6	M+Cl37[-]	210.9837	54.8	-0.131	8.21E-02	C18neg	3
	L-Gulonate	C00800	C6H12O7	M+HCOO[-]	241.0563	40.5	0.007	9.26E-01	C18neg	3
Bile acid biosynthesis	Calcidiol	C01561	C27H44O2	M+H+Na[2+]	212.1645	22.6	0.109	1.49E-01	HILICpos	3
	7alpha-Hydroxy-3-oxo-4-cholestenoate	C17337	C27H42O4	M+2H[2+]	216.1594	25.5	0.074	3.28E-01	HILICpos	3
	Cholesterol ester	C02530	C28H45O2R	M+2H[2+]	324.2893	30.3	0.144	5.60E-02	HILICpos	3
	3alpha,7alpha-Dihydroxy-5beta-cholestan-26-al	C05445	C27H46O3	M+HCOONa[1+]	487.3371	47.1	-0.089	2.39E-01	HILICpos	3
	3alpha,7alpha,24-Trihydroxy-5beta-cholestanoyl-CoA	C05448	C48H80N7O20P3S	M+2H[2+]	600.7209	79	-0.148	5.05E-02	HILICpos	3
Blood Group Biosynthesis	CMP	C00055	C9H14N3O8P	M-H+O[-]	338.0362	49.5	-0.019	8.01E-01	C18neg	3
	CMP	C00055	C9H14N3O8P	M+K-2H[-]	358.9951	67.8	0.121	1.09E-01	C18neg	3
	GDP	C00035	C10H15N5O11P2	M+Na-2H[-]	462.9889	38.1	-0.134	7.53E-02	C18neg	3
Fructose and mannose metabolism	L-Fucose 1-phosphate	C02985	C6H13O8P	M+2H[2+]	123.0247	250.3	-0.159	3.53E-02	HILICpos	3
	D-Glyceraldehyde	C00577	C3H6O3	M+HCOONa[1+]	159.0262	108.9	0.062	4.15E-01	HILICpos	3
	Galactose	C01582	C6H12O6	M+H[1+]	181.072	196.9	-0.177	1.84E-02	HILICpos	1
	GDP-D-mannose	C00096	C16H25N5O16P2	M(C13)+H[1+]	607.0916	43.8	0.080	2.89E-01	HILICpos	3
Glutamate metabolism	Bicarbonate	C00288	HCO3	M+Cl37[-]	97.9594	28.9	0.046	5.45E-01	C18neg	3
	2-Methyl-3-oxopropanoate	C00349	C4H6O3	M(C13)-H[-]	102.0277	24.3	-0.192	1.05E-02	C18neg	3
	Succinate	C00042	C4H6O4	M+Cl[-]	152.9959	36.5	0.099	1.92E-01	C18neg	3
	L-Glutamate	C00025	C5H9NO4	M+Cl37[-]	184.0194	36.6	0.021	7.86E-01	C18neg	3
	gamma-L-Glutamyl-L-cysteine	C00669	C8H14N2O5S	M+CH3COO[-]	309.0781	80.3	0.198	8.52E-03	C18neg	3
Glycosphingolipid biosynthesis - (neo)lactoseries/ganglioseries	CMP	C00055	C9H14N3O8P	M-H+O[-]	338.0362	49.5	-0.019	8.01E-01	C18neg	3
	CMP	C00055	C9H14N3O8P	M+K-2H[-]	358.9951	67.8	-0.121	1.09E-01	C18neg	3
	Galactose	C01582	C6H12O6	M+H[1+]	181.072	196.9	-0.177	1.84E-02	HILICpos	

Pathway	Metabolite	KEGG ID	Formula	Adduct	m/z	RT		p-value	Platform	
	GDP	C00035	C10H15N5O11P2	M+Na-2H[-]	462.9889	38.1	0.134	7.53E-02	C18neg	3
		C00129	C5H12O7P2	M-H[-]	244.9974	267.8	0.154	4.12E-02	C18neg	3
	Isopentenyl diphosphate	C00129	C5H12O7P2	M+Cl[-]	280.9741	183.1	0.182	1.57E-02	C18neg	3
		C00129	C5H12O7P2	M+Br[-]	324.9215	45.2	0.079	2.97E-01	C18neg	3
N-Glycan biosynthesis	CMP	C00055	C9H14N3O8P	M-H+O[-]	338.0362	49.5	0.019	8.01E-01	C18neg	3
		C00055	C9H14N3O8P	M+K-2H[-]	358.9951	67.8	0.121	1.09E-01	C18neg	3
	GDP	C00035	C10H15N5O11P2	M+Na-2H[-]	462.9889	38.1	0.134	7.53E-02	C18neg	3
	CTP	C00063	C9H16N3O14P3	M+ACN-H[-]	523.0008	68.6	0.083	2.73E-01	C18neg	3
Nucleotide Sugar Metabolism	dTTP	C00459	C10H17N2O14P3	M+2H[2+]	241.9996	178.9	0.233	1.86E-03	HILICpos	3
	4,6-Dideoxy-4-oxo-dTDP-D-glucose	C11907	C16H24N2O15P2	M[1+]	546.0627	103.8	0.211	5.03E-03	HILICpos	3
Phenylalanine metabolism	L-Phenylalanine	C00079	C9H11NO2	M-H[-]	164.0717	47.2	0.016	8.31E-01	C18neg	1
	3-(2-Hydroxyphenyl)propanoic acid	C01198	C9H10O3	M-H[-]	165.0558	177.7	0.106	1.63E-01	C18neg	1
	Hexadecanoate (n-C16:0)	C00249	C16H32O2	M+Cl[-]	291.2093	38.6	0.030	6.89E-01	C18neg	3
	CMP	C00055	C9H14N3O8P	M-H+O[-]	338.0362	49.5	0.019	8.01E-01	C18neg	3
		C00055	C9H14N3O8P	M+K-2H[-]	358.9951	67.8	0.121	1.09E-01	C18neg	3
	Cellobiose	C00185	C12H22O11	M-H+O[-]	357.1037	226	0.139	6.50E-02	C18neg	3
	CMP	C00055	C9H14N3O8P	M+K-2H[-]	358.9951	67.8	0.121	1.09E-01	C18neg	3
	1D-myo-Inositol 1,3,4-trisphosphate	C01243	C6H15O15P3	M-H2O-H[-]	400.9459	67.7	0.116	1.24E-01	C18neg	3
Phosphatidylinositol phosphate metabolism		C01243	C6H15O15P3	M+HCOO[-]	464.9585	59.5	0.064	4.02E-01	C18neg	3
	1-Phosphatidyl-1D-myo-inositol 3-phosphate	C04549	C11H18O16P2R2	M+K-2H[-]	505.9605	56.8	0.075	3.19E-01	C18neg	3
	1D-myo-Inositol 1,3,4,5,6-pentakisphosphate	C01284	C6H17O21P5	M+HCOO[-]	624.8963	55.8	0.089	2.38E-01	C18neg	3
		C01284	C6H17O21P5	M+CH3COO[-]	638.9116	55.8	0.089	2.43E-01	C18neg	3
	phosphatidylinositol-3,4,5-trisphosphate	C05981	C11H20O22P4R2	M+HCOO[-]	688.9671	69.6	0.071	3.52E-01	C18neg	3
		C05981	C11H20O22P4R2	M+CH3COO[-]	702.983	69.6	0.049	5.21E-01	C18neg	3
	Succinate	C00042	C4H6O4	M+Cl[-]	152.9959	36.5	0.099	1.92E-01	C18neg	3
Phytanic acid peroxisomal oxidation	Phytanate	C01607	C20H40O2	M-H+O[-]	327.2912	240.2	0.164	2.98E-02	C18neg	3
		C01607	C20H40O2	M+HCOO[-]	357.301	268.2	0.021	7.85E-01	C18neg	3
	4-Hydroxybutanoic acid	C00989	C4H8O3	M+H[1+]	105.0547	18.5	0.155	3.95E-02	HILICpos	3
Propanoate metabolism	D-Glyceraldehyde	C00577	C3H6O3	M+HCOONa[1+]	159.0262	108.9	0.062	4.15E-01	HILICpos	3
	Propinol adenylate	C05983	C13H18N5O8P	M+HCOONa[1+]	472.0838	83.8	0.173	2.13E-02	HILICpos	3
Tyrosine metabolism	Beta-Alanine	C00099	C3H7NO2	M-H[-]	88.0404	47	0.156	3.83E-02	C18neg	1

2-Methyl-3-oxopropanoate	C00349	C4H6O3	M(C13)-H[-]	102.0277	24.3	0.192	1.05E-02	C18neg	3
Sulfate	C00059	H2SO4	M+Cl[-]	132.9368	34.6	0.023	7.65E-01	C18neg	3
Phenylacetic acid	C00548	C8H8O2	M-H[-]	135.0452	252.4	0.192	1.08E-02	C18neg	1
Indole-56-quinone	C05579	C8H5NO2	M-H[-]	146.0248	180.6	0.019	7.98E-01	C18neg	3
2-Aminoacrylate	C02218	C3H5NO2	M+Br[-]	165.9509	25.2	0.228	2.33E-03	C18neg	3
Hippuric acid	C01586	C9H9NO3	M-H[-]	178.051	41.8	0.157	3.74E-02	C18neg	1
L-Glutamate	C00025	C5H9NO4	M+Cl37[-]	184.0194	36.6	0.021	7.86E-01	C18neg	3
2-Phenyl-13-propanediol monocarbamate	C16586	C10H13NO3	M-H[-]	194.0823	162.8	0.107	1.58E-01	C18neg	3
Phenylacetaldehyde	C00601	C8H8O	M+Br[-]	198.9756	45.3	0.069	3.61E-01	C18neg	3
cis-Aconitate	C00417	C6H6O6	M+Cl37[-]	210.9837	54.8	0.131	8.21E-02	C18neg	3
2-Carboxy-2,3-dihydro-5,6-dihydroxyindole	C05604	C9H9NO4	M+Na-2H[-]	215.0225	262	0.097	2.03E-01	C18neg	3
3,4-Dihydroxy-L-phenylalanine	C00355	C9H11NO4	M+Na-2H[-]	217.0369	39.2	0.171	2.36E-02	C18neg	3
Homovanillate	C05582	C9H10O4	M+ACN-H[-]	222.0767	39.4	0.023	7.62E-01	C18neg	3
Oxalosuccinate	C05379	C6H6O7	M+Cl37[-]	226.9778	286.7	0.142	5.93E-02	C18neg	3
	C05642	C13H16N2O4	M(C13)-H[-]	264.1066	37.7	0.119	1.16E-01	C18neg	3
Formyl-N-acetyl-5-methoxykynurenamine	C05642	C13H16N2O4	M+Cl[-]	299.0807	173.7	0.031	6.83E-01	C18neg	3
	C05642	C13H16N2O4	M+Cl37[-]	301.0771	281.5	0.055	4.66E-01	C18neg	3
Vitamin B5 - CoA biosynthesis from pantothenate	Pantetheine 4'-phosphate — C01134	C11H23N2O7PS	M(S34)-H[-]	359.0832	184.1	0.145	5.49E-02	C18neg	3
	CTP — C00063	C9H16N3O14P3	M+ACN-H[-]	523.0008	68.6	0.083	2.73E-01	C18neg	3
Vitamin K metabolism	Vitamin K1 epoxide — C05849	C31H46O3	M+K-2H[-]	502.2881	196.7	0.118	1.20E-01	C18neg	3
	C05849	C31H46O3	M+Br[-]	545.2683	186.1	0.134	7.53E-02	C18neg	3
	Phylloquinone — C02059	C31H46O2	M+Br[-]	529.2684	248.4	0.072	3.42E-01	C18neg	3

Supplemental Table 4-S8. Biweighted midcorrelations between OP-related CpGs and OP-related metabolomic features (top 20 pairs)
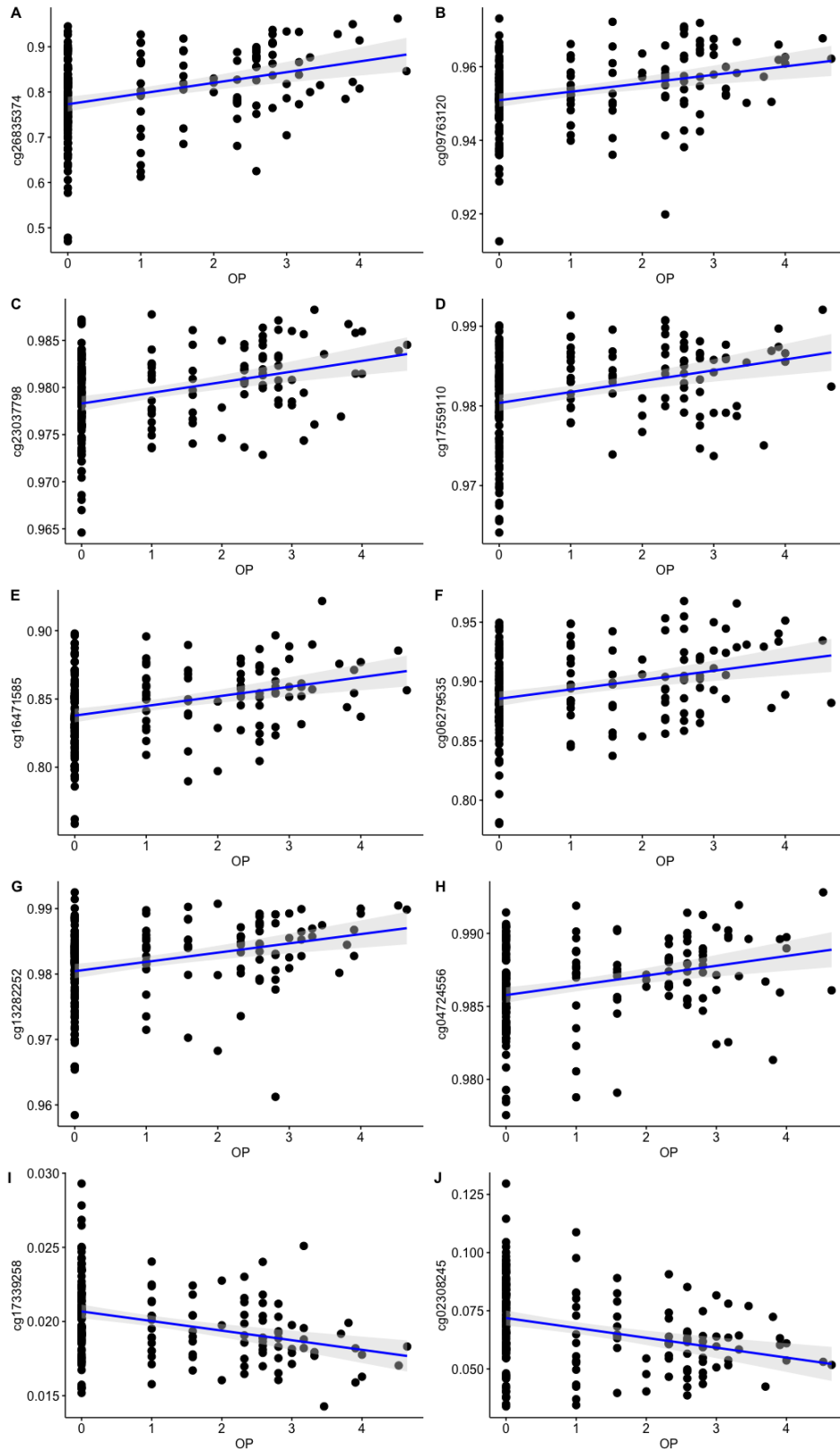
CpG	Metabolite	CpG-Metabolite_cor	CpG-Metabolite_pvalue	CpG-Metabolite_FDR	OP-CpG_coef	CpG_CHR	CpG_MAPINFO	CpG_gene_name	OP-Metabolite_coef	Metabolite_mz	Metabolite_time	Metabolite_Name
cg07385362	met_33_24_c18	0.354	1.48E-06	5.99E-02	0.260	16	1417710	UNKL	-0.200	374.3171	285.1	Docosanamide
cg24350213	met_14_0_c18	-0.282	1.46E-04	3.18E-01	-0.311	17	39943012	JUP	-0.192	102.0277	24.3	2-Methyl-3-oxopropanoate
cg04701182	met_43_0_c18	0.313	2.28E-05	2.18E-01	-0.266	10	69834353	HERC4	-0.192	135.0452	252.4	Phenylacetic acid
cg22016716	met_43_0_c18	-0.290	9.38E-05	3.05E-01	0.290	6	45152974	SUPT3H	-0.192	135.0452	252.4	Phenylacetic acid
cg16655166	met_43_0_c18	-0.302	4.73E-05	2.18E-01	0.263	2	8063199	LINC00298	-0.192	135.0452	252.4	Phenylacetic acid
cg16768320	met_43_0_c18	-0.289	1.01E-04	3.05E-01	0.264	17	57966691	TUBD1	-0.192	135.0452	252.4	Phenylacetic acid
cg13910460	met_43_0_c18	-0.281	1.60E-04	3.18E-01	0.274	3	112059676	CD200	-0.192	135.0452	252.4	Phenylacetic acid
cg01010012	met_10_46_c18	0.306	3.54E-05	2.18E-01	0.275	13	29933537	MTUS2	0.194	194.9059	114	
cg13699934	met_17_70_c18	-0.286	1.16E-04	3.13E-01	-0.307	9	93956018		-0.168	254.0256	47.3	Brassicanal B
cg18803171	met_17_70_c18	-0.290	9.50E-05	3.05E-01	-0.289	22	24093429	ZNF70	-0.168	254.0256	47.3	
cg10133171	met_17_70_c18	0.288	1.06E-04	3.05E-01	0.281	4	188534448	RP11-565A3.1	-0.168	254.0256	47.3	
cg11987576	met_17_70_c18	-0.281	1.57E-04	3.18E-01	-0.273	13	52378294	DHRS12	-0.168	254.0256	47.3	
cg15918732	met_21_24_c18	0.284	1.33E-04	3.18E-01	0.262	11	57364939	SERPING1	-0.182	280.9741	183.1	Isopentenyl diphosphate
cg14818343	met_27_60_c18	-0.301	4.86E-05	2.18E-01	0.262	11	114831791		-0.195	328.9842	54.1	2-[(5-Methylsulfinyl)-4-penten-2-ynylidene]-1,6-dioxaspiro[4.4]non-3-ene
cg02909136	met_44_65_c18	-0.302	4.55E-05	2.18E-01	-0.270	20	44941542		0.197	466.2742	213.9	N-Oleoyl phenylalanine
cg07701049	met_45_91_c18	-0.303	4.48E-05	2.18E-01	0.268	8	141559946	AGO2	-0.183	476.9149	52	[2,6-dihydroxy-3-(3,5,7-trihydroxy-4-oxo-4H-chromen-2-yl)phenyl]oxidanesulfonic acid
cg06394887	met_51_43_c18	0.312	2.45E-05	2.18E-01	-0.298	7	72937120	BAZ1B	-0.177	528.3206	202.5	
cg04121028	met_51_70_c18	-0.291	8.91E-05	3.05E-01	0.261	1	181622655	CACNA1E	-0.170	530.8885	57.1	
cg22721765	met_64_58_c18	-0.281	1.55E-04	3.18E-01	-0.287	5	149380233	HMGXB3	-0.228	696.3172	231.7	
cg17003212	met_64_58_c18	0.308	3.27E-05	2.18E-01	-0.262	10	97054967		-0.228	696.3172	231.7	

Supplemental Table 4-S9. Biweighted midcorrelations between OP-related CpGs and the metabolome (top 20 pairs)

| CpG	Metabolite	CpG-Metabolite_cor	CpG-Metabolite_p value	CpG-Metabolite_FDR	OP-CpG_coef	CpG_CHR	CpG_MAPINFO	CpG_gene_name	OP-Metabolite_coef	Metabolite_mz	Metabolite_time	Metabolite_Name
cg089	met_387									261.176		
01151	8_hilic	0.434	1.81E-09	1.19E-02	0.261	7	2172260	MAD1L1	0.221	1	215	
cg114	met_4_hi						12278532					
96747	lic	-0.367	5.24E-07	2.32E-01	0.285	3	8	PDIA5	-0.035	85.0523	109.8	
cg144	met_205									174.112		
15214	9_hilic	0.361	8.51E-07	2.80E-01	-0.265	11	2950403	PHLDA2	0.048	5	253.4	N-Acetylleucine
cg204	met_215									177.986		
97205	9_hilic	0.372	3.58E-07	2.32E-01	0.283	6	47754005	OPN5	0.038	9	219.1	
cg219	met_223									181.122		4-
15368	8_hilic	0.370	4.21E-07	2.32E-01	0.295	1	60166962	FGGY	-0.035	3	271.2	(Butoxymethyl)phenol
cg070	met_597											
19072	5_hilic	-0.371	3.88E-07	2.32E-01	0.274	6	27215328	PRSS16	-0.076	384.295	44.5	
cg276	met_640									416.120		
46848	4_hilic	0.369	4.49E-07	2.32E-01	-0.260	11	62389601	B3GAT3	-0.022	3	85.6	
cg133	met_806						14523532					
72862	7_hilic	0.367	5.52E-07	2.32E-01	0.267	8	8	MROH1	0.003	549.718	43.5	
cg147	met_870									598.724		
77817	5_hilic	-0.367	5.39E-07	2.32E-01	-0.266	13	80055594	NDFIP2;NDFIP2-AS1	-0.004	1	79.1	
cg128	met_889									611.409		
61193	3_hilic	-0.361	8.42E-07	2.80E-01	0.267	7	62693631	PHKG1P1;RP5-905H7.3	0.022	4	56.8	
cg070	met_926									634.433		
19072	3_hilic	-0.372	3.78E-07	2.32E-01	0.274	6	27215328	PRSS16	-0.027	5	54.3	
cg133	met_976									663.624		
72862	8_hilic	-0.380	1.91E-07	2.32E-01	0.267	8	8	MROH1	0.058	3	41	
							14523532			721.503		
cg133	met_106											
72862	77_hilic	0.361	8.48E-07	2.80E-01	0.267	8	8	MROH1	-0.034	1	46.8	PG(16:016:1(9Z))
cg266	met_115									818.562		PE(20:1(11Z)22:6(4Z7Z10Z13Z16Z19Z))
68919	62_hilic	0.378	2.25E-07	2.32E-01	0.261	6	35311534	PPARD	0.095	1	74.3	
cg098	met_157						11895984	B4GALT4;B4GALT4-AS1		238.937		
62076	7_c18	-0.371	4.14E-07	2.32E-01	-0.265	3	0		0.058	7	81	
cg000	met_157									238.937		
19137	7_c18	-0.367	5.65E-07	2.32E-01	-0.270	19	41103107	LTBP4	0.058	7	81	
cg089	met_167						14914661			245.175		
33615	2_c18	-0.365	6.49E-07	2.51E-01	0.274	1	8		-0.036	4	143.9	
cg162	met_196							C8orf58;CCAR		268.901		
66672	5_c18	0.390	9.06E-08	2.32E-01	0.318	8	22461575	2	0.032	3	107.1	
cg082	met_423							SLC4A1AP;SU				
39375	0_c18	-0.378	2.34E-07	2.32E-01	0.267	2	27887547	PT7L	-0.053	445.244	206	
cg270	met_622									654.905		
88191	5_c18	0.370	4.28E-07	2.32E-01	-0.262	18	46065262	CTIF	-0.027	9	59.1	

CpG	Metabolite	CpG-Metabolite_cor	CpG-Metabolite_pvalue	CpG-Metabolite_FDR	OP-CpG_coef	CpG_CHR	CpG_MAPINFO	CpG_gene_name	OP-Metabolite_coef	Metabolite_mz	Metabolite_time	Metabolite_Name
cg09973663	met_430_c18	-0.433	2.04E-09	4.84E-02	0.079	2	192115168	MYO1B	-0.192	135.0452	252.4	Phenylacetic acid
cg24365042	met_1620_c18	0.432	2.23E-09	4.84E-02	0.072	1	212780517	ATF3	-0.184	241.2147	219.1	
cg15590780	met_3172_c18	0.421	6.17E-09	8.91E-02	0.052	1	216596639	USH2A	-0.179	361.9381	34.8	
cg10933002	met_3324_c18	0.417	8.84E-09	9.58E-02	-0.006	11	78780983	TENM4	-0.200	374.3171	285.1	
cg25467493	met_1770_c18	-0.409	1.78E-08	1.20E-01	-0.064	2	55845064	RP11-554J4.1;SMEK2	-0.168	254.0256	47.3	
cg07540722	met_430_c18	0.405	2.45E-08	1.20E-01	-0.080	2	70056697	GMCL1	-0.192	135.0452	252.4	Phenylacetic acid
cg23055315	met_4323_c18	-0.403	2.92E-08	1.20E-01	-0.031	8	56792234	LYN	-0.168	452.4562	204.3	
cg19156483	met_3264_hilic	0.403	2.94E-08	1.20E-01	-0.007	2	164592571	FIGN	-0.248	230.1215	103.7	
cg24295424	met_6631_c18	0.402	3.27E-08	1.20E-01	-0.030	12	106530091	NUAK1	-0.169	746.8558	55.6	
cg11365617	met_1770_c18	-0.402	3.31E-08	1.20E-01	0.038	12	110174620	FAM222A;FAM222A-AS1	-0.168	254.0256	47.3	
cg23644795	met_1770_c18	-0.400	3.85E-08	1.20E-01	0.006	18	267711	RP11-705O1.8;THOC1	-0.168	254.0256	47.3	
cg00585621	met_1305_c18	-0.398	4.41E-08	1.20E-01	-0.006	4	100868748	DNAJB14	-0.214	216.9709	39.1	
cg22513528	met_5636_hilic	-0.398	4.47E-08	1.20E-01	-0.002	15	77861673	RP11-307C19.2	-0.233	362.2166	31	
cg26954751	met_3960_c18	-0.397	4.74E-08	1.20E-01	0.062	6	107506867	PDSS2	-0.183	423.2601	235.1	
cg19609284	met_430_c18	0.397	5.07E-08	1.20E-01	-0.111	5	72143978	TNPO1	-0.192	135.0452	252.4	Phenylacetic acid
cg10335659	met_1770_c18	-0.396	5.48E-08	1.20E-01	-0.110	2	99757356	C2orf15;TSGA10	-0.168	254.0256	47.3	
cg04072270	met_4172_c18	-0.396	5.50E-08	1.20E-01	0.041	2	15190874	NA	-0.170	439.718	41.4	
cg09515262	met_430_c18	-0.395	5.58E-08	1.20E-01	0.129	8	77927788	NA	-0.192	135.0452	252.4	Phenylacetic acid
cg16317457	met_11700_hilic	-0.395	5.62E-08	1.20E-01	0.060	11	123173103	NA	0.232	843.901	106.1	
cg15884430	met_2248_c18	-0.395	5.78E-08	1.20E-01	0.038	5	170819088	NPM1	0.169	289.9565	66.4	

Supplemental Table 4-S10. Biweighted midcorrelations between OP-related metabolomic features and the methylome (top 20 pairs)

Supplemental Figure 4-S1. Scatter plots showed top 10 CpGs correlated with OP exposure based on bicor p-value
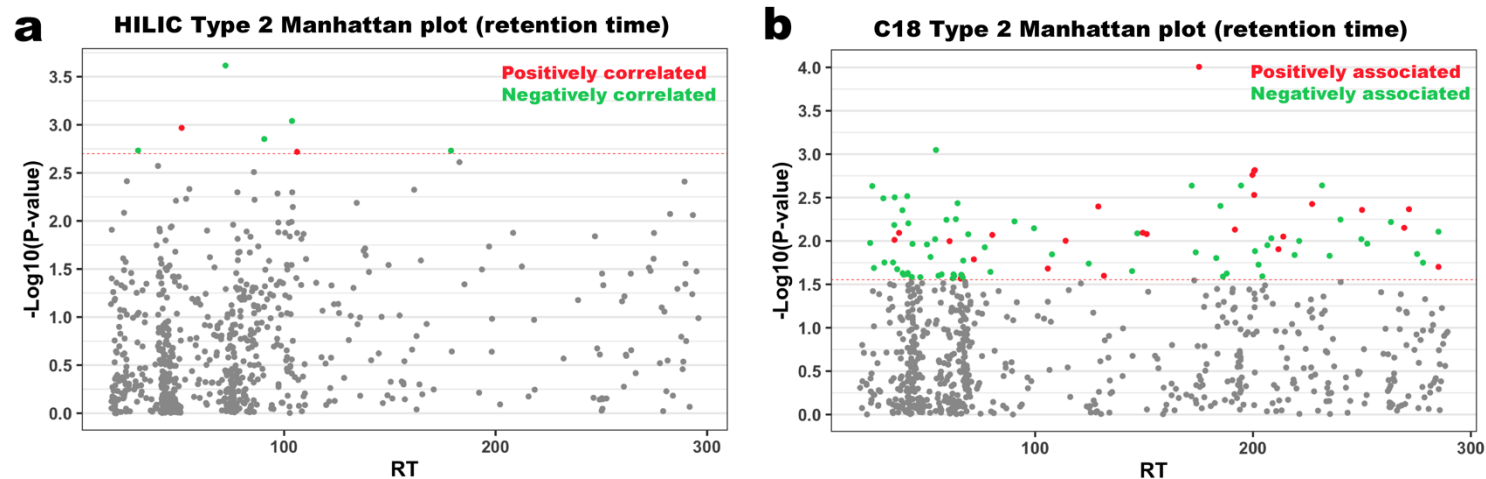
Supplemental Figure 4-S2. Identification of metabolic features associated with OP pesticides exposure. (**a**) Type 2 Manhattan plot for features in the HILIC column (positive ion mode), -log10(p-value) vs retention time. Only features selected by sparse PLS were included here. 7 m/z features were found with FDR < 0.2. Red dots represent the features that were positively correlated with OP, and the green dots represent the features that were negatively correlated with OP; (**b**) Type 2 Manhattan plot for features in the C18 column (negative ion mode), -log10(p-value) vs. retention time. 92 m/z features were found with FDR < 0.2
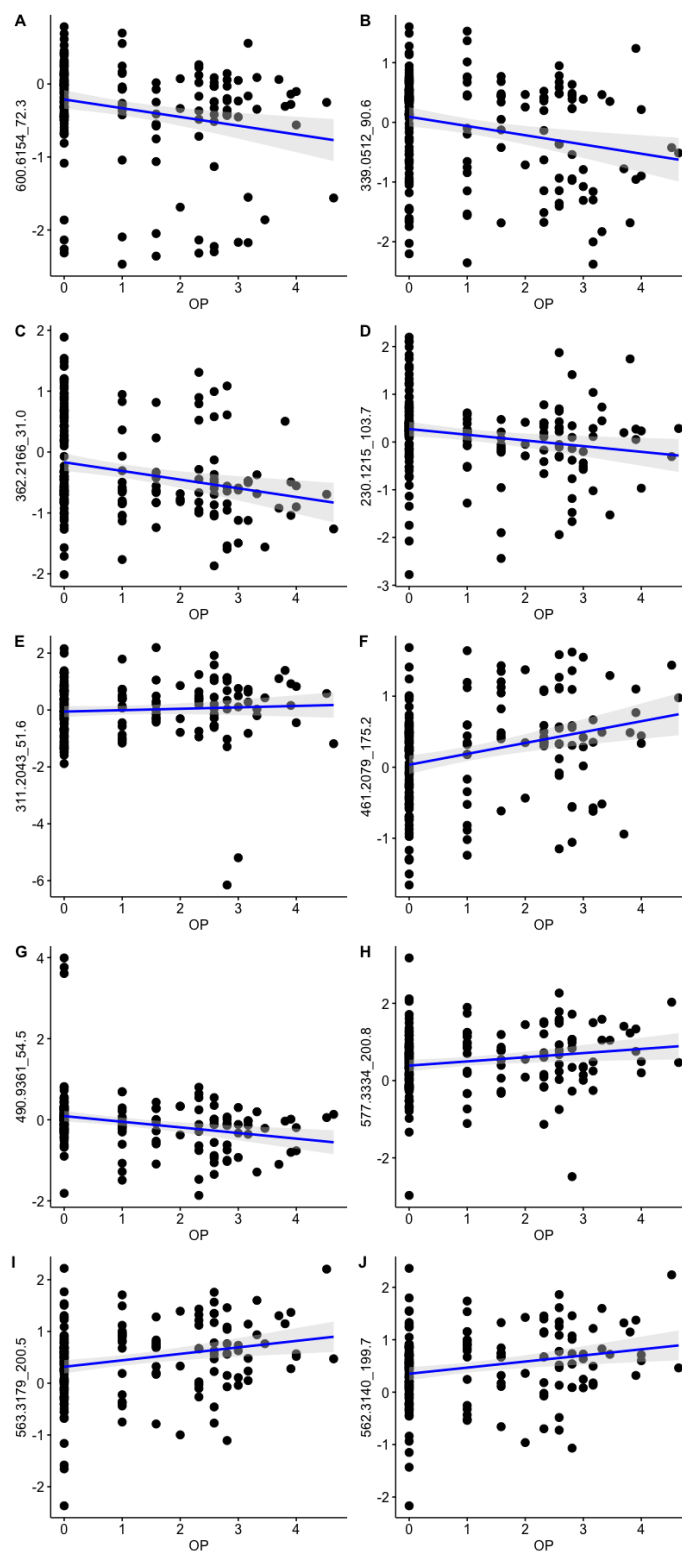
Supplemental Figure 4-S3. Scatter plots showed top 5 metabolites correlated with OP exposure based on bicor p-value from HILICpos and C18neg platforms, respectively.

5 Public Health Relevance and Expected Contributions

This dissertation investigated the metabolomic and epigenomic changes associated with traffic-related air pollution and chronic environmental pesticide exposures to better understand the underlying molecular mechanisms of harmful environmental exposures. In the first project, we found that metabolomic features and pathways associated with maternal air pollution exposure during pregnancy were primarily involved in oxidative stress and inflammatory responses that have been previously implicated in pregnancy complications and adverse outcomes. In the second project, we identified alterations in serum metabolome associated with chronic organophosphates (OPs), organochlorines (OCs), or pyrethroids (PYRs) exposure. Perturbed metabolomic pathways included mitochondrial energy metabolism, fatty acid and lipid metabolism, and amino acid metabolism. In addition, we found that the fatty acid beta-oxidation pathway is a common pathway shared across all three pesticide classes. In the third project, we demonstrated that chronic OP exposure could affect the glycosphingolipid biosynthesis pathway at both DNA methylation and metabolome level. We also showed that epigenome and metabolome exhibiting various distinct mechanisms in response to chronic OP exposure. For example, epigenomic signals were involved in intracellular membrane transport, cell adhesion, and carcinogenesis; while metabolomic responses included aromatic amino acids metabolism, oxidative stress, and energy metabolism.

Over the past decade, numerous research efforts have pioneered various omics technologies, and brought us into an era where various "omes" can be readily characterized. The integration of omics approaches into environmental epidemiological studies can go a long way in helping researchers better trace environmental exposures and understand underlying biological mechanisms. Investigating the cumulative effects of environmental exposures across the lifespan or during specific vulnerable periods during development and capturing their multifactorial

137

etiology for health and disease has always been a major challenge for traditional epidemiological and toxicological studies. Our studies utilized a state-of-art GIS-based exposure assessment model linking macro-level population exposure to individual addresses, making it possible to study chronic, real-world levels of air pollution and pesticide exposures' contributions to health in human populations. We further mapped chronic exposures to the micro-level metabolome and epigenome, delineating a holistic picture of the deleterious effects of air pollution and pesticide exposure at the molecular level, thus potentially providing new insights into the pathogenesis of environmental exposure-related diseases.

Our studies identified altered molecular signals in response to chronic air pollution or pesticide exposures, some of which, if confirmed in future studies, can be used as easily obtainable biomarkers (by standard blood draw). Biomarkers that reflect chronic persistent adverse exposures are vital for developing targeted prevention efforts and to detect environmental damage at pre-clinical stage. Prevention and remediation could then be started for at-risk individuals before significant deleterious health effects leading to a variety of chronic diseases are detected.

Complementing traditional environmental epidemiology studies, our findings also allow regulatory agencies to understand the adverse effects of air pollution and pesticide exposures in human populations and thus support the decision-making process. For example, some pesticides may be considered safe in terms of acute toxicity or at low levels in animal models but when they are shown to alter a wide range of molecular pathways that are known to lead to adverse health outcomes in the long run, long term human studies should be conducted and these pesticides should be regulated, and fall under a 'pesticidovigilance' clause, or be removed from the market.

In summary, by combining large-scale population-based exposure data with multiple omics approaches, our studies facilitate a novel understanding of mechanisms through which the

exposure, the epigenome, and the metabolome interact and affect health, producing urgently needed information for translation and prevention, and eventually guide policy making.

6 References

Abeliovich A, Gitler AD. 2016. Defects in trafficking bridge parkinson's disease pathology and genetics. Nature 539:207-216.

Al-Gubory KH, Fowler PA, Garrel C. 2010. The roles of cellular reactive oxygen species, oxidative stress and antioxidants in pregnancy outcomes. The international journal of biochemistry & cell biology 42:1634-1650.

Alfano R, Chadeau-Hyam M, Ghantous A, Keski-Rahkonen P, Chatzi L, Perez AE, et al. 2020. A multi-omic analysis of birthweight in newborn cord blood reveals new underlying mechanisms related to cholesterol metabolism. Metabolism 110:154292.

Anderson SM, Naidoo RN, Ramkaran P, Phulukdaree A, Muttoo S, Asharam K, et al. 2018. The effect of nitric oxide pollution on oxidative stress in pregnant women living in durban, south africa. Arch Environ Contam Toxicol 74:228-239.

Androutsopoulos VP, Hernandez AF, Liesivuori J, Tsatsakis AM. 2013. A mechanistic overview of health associated effects of low levels of organochlorine and organophosphorous pesticides. Toxicology 307:89-94.

Anthonymuthu TS, Kenny EM, Lamade AM, Kagan VE, Bayir H. 2018. Oxidized phospholipid signaling in traumatic brain injury. Free Radic Biol Med 124:493-503.

Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. 2014. Minfi: A flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. Bioinformatics (Oxford, England) 30:1363-1369.

Bandres-Ciga S, Saez-Atienzar S, Bonet-Ponce L, Billingsley K, Vitale D, Blauwendraat C, et al. 2019. The endocytic membrane trafficking pathway plays a major role in the risk of parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 34:460-468.

Banerjee BD, Seth V, Ahmed RS. 2001. Pesticide-induced oxidative stress: Perspectives and trends. Rev Environ Health 16:1-40.

Banks CN, Lein PJ. 2012. A review of experimental evidence linking neurotoxic organophosphorus compounds and inflammation. Neurotoxicology 33:575-584.

Baraniuk JN, El-Amin S, Corey R, Rayhan R, Timbol C. 2013. Carnosine treatment for gulf war illness: A randomized controlled trial. Glob J Health Sci 5:69-81.

Barnham KJ, Masters CL, Bush AI. 2004. Neurodegenerative diseases and oxidative stress. Nature reviews Drug discovery 3:205-214.

Becerra TA, Wilhelm M, Olsen J, Cockburn M, Ritz B. 2013. Ambient air pollution and autism in los angeles county, california. Environ Health Perspect 121:380-386.

Beilina A, Rudenko IN, Kaganovich A, Civiero L, Chau H, Kalia SK, et al. 2014. Unbiased screen for interactors of leucine-rich repeat kinase 2 supports a common pathway for sporadic and familial parkinson disease. Proc Natl Acad Sci U S A 111:2626-2631.

Benson P. 1989. Caline4 - a dispersion model for predicting air pollutant concentrations near roadways.California Department of Transportation, CA.

Benson PE. 1984. Caline 4-a dispersion model for predictiong air pollutant concentrations near roadways.

Berlett BS, Stadtman ER. 1997. Protein oxidation in aging, disease, and oxidative stress. J Biol Chem 272:20313-20316.

Biagi G, De Rosa V, Pelusi G, Scagliarini G, Sani G, Coccheri S. 1990. Increased placental production of leukotriene b4 in gestational hypertension. Thromb Res 60:377-384.

Blair A, Ritz B, Wesseling C, Freeman LB. 2015. Pesticides and human health. Occup Environ Med 72:81-82.

Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008:P10008.

Bloszies CS, Fiehn O. 2018. Using untargeted metabolomics for detecting exposome compounds. Current Opinion in Toxicology 8:87-92.

Boldyrev AA, Dupin AM, AYa B, Babizhaev M, Severin SE. 1987. The antioxidative properties of carnosine, a natural histidine containing dipeptide. Biochemistry international 15:1105-1113.

Bonder MJ, Luijk R, Zhernakova DV, Moed M, Deelen P, Vermaat M, et al. 2017. Disease variants alter transcription factor levels and methylation of their binding sites. Nature Genetics 49:131-138.

Bonvallot N, Tremblay-Franco M, Chevrier C, Canlet C, Warembourg C, Cravedi J-P, et al. 2013. Metabolomics tools for describing complex pesticide exposure in pregnant women in brittany (france). Plos One 8:e64433.

Bonvallot N, Tremblay-Franco M, Chevrier C, Canlet C, Warembourg C, Cravedi JP, et al. 2013. Metabolomics tools for describing complex pesticide exposure in pregnant women in brittany (france). Plos One 8:e64433.

Bradberry SM, Cage SA, Proudfoot AT, Vale JA. 2005. Poisoning due to pyrethroids. Toxicological reviews 24:93-106.

Breeze CE, Paul DS, van Dongen J, Butcher LM, Ambrose JC, Barrett JE, et al. 2016. Eforge: A tool for identifying cell type-specific signal in epigenomic data. Cell Rep 17:2137-2150.

Broderick BM, Budd U, Misstear BD, Ceburnis D, Jennings SG. 2005. Validation of caline4 modelling for carbon monoxide concentrations under free-flowing and congested traffic conditions in ireland. International Journal of Environment and Pollution 24:104-113.

Busse WW. 1998. Leukotrienes and inflammation. Am J Respir Crit Care Med 157:S210-213.

California Air Resources Board. 2013. Emfac2011 - technical documentation. Sacramento, CA:California Air Resources Board.

Cardenas A, Sordillo JE, Rifas-Shiman SL, Chung W, Liang L, Coull BA, et al. 2019. The nasal methylome as a biomarker of asthma and airway inflammation in children. Nat Commun 10:3095.

Carrizo D, Chevallier OP, Woodside JV, Brennan SF, Cantwell MM, Cuskelly G, et al. 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.

Castellanos A, Andres A, Bernal L, Callejo G, Comes N, Gual A, et al. 2018. Pyrethroids inhibit k2p channels and activate sensory neurons: Basis of insecticide-induced paraesthesias. Pain 159:92.

Ch R, Singh AK, Pathak MK, Singh A, Kesavachandran CN, Bihari V, et al. 2019. Saliva and urine metabolic profiling reveals altered amino acid and energy metabolism in male farmers exposed to pesticides in madhya pradesh state, india. Chemosphere 226:636-644.

Chow CK. 1991. Vitamin e and oxidative stress. Free Radic Biol Med 11:215-232.

Clary T, Ritz B. 2003. Pancreatic cancer mortality and organochlorine pesticide exposure in california, 1989-1996. Am J Ind Med 43:306-313.

Cockburn M, Mills P, Zhang X, Zadnick J, Goldberg D, Ritz B. 2011. Prostate cancer and ambient pesticide exposure in agriculturally intensive areas in california. Am J Epidemiol 173:1280-1288.

Costa LG. 2006. Current issues in organophosphate toxicology. Clin Chim Acta 366:1-13.

Cunningham GC, Tompkinison DG. 1999. Cost and effectiveness of the california triple marker prenatal screening program. Genet Med 1:199-206.

Curwin BD, Hein MJ, Sanderson WT, Striley C, Heederik D, Kromhout H, et al. 2007. Urinary pesticide concentrations among children, mothers and fathers living in farm and non-farm households in iowa. Ann Occup Hyg 51:53-65.

Dadvand P, Figueras F, Basagana X, Beelen R, Martinez D, Cirach M, et al. 2013. Ambient air pollution and preeclampsia: A spatiotemporal analysis. Environ Health Perspect 121:1365-1371.

Daher N, Saliba NA, Shihadeh AL, Jaafar M, Baalbaki R, Shafer MM, et al. 2014. Oxidative potential and chemical speciation of size-resolved particulate matter (pm) at near-freeway and urban background sites in the greater beirut area. Science of The Total Environment 470-471:417-426.

DellaPenna D, Mène-Saffrané L. 2011. Chapter 5 - vitamin e. In: Advances in botanical research, Vol. 59, (Rébeillé F, Douce R, eds):Academic Press, 179-227.

143

Demur C, Métais B, Canlet C, Tremblay-Franco M, Gautier R, Blas-Y-Estrada F, et al. 2013. Dietary exposure to a low dose of pesticides alone or as a mixture: The biological metabolic fingerprint and impact on hematopoiesis. Toxicology 308:74-87.

Deng K, Zhang F, Tan Q, Huang Y, Song W, Rong Z, et al. 2019. Waveica: A novel algorithm to remove batch effects for large-scale untargeted metabolomics data based on wavelet analysis. Analytica Chimica Acta.

Dick CAJ, Singh P, Daniels M, Evansky P, Becker S, Gilmour MI. 2003. Murine pulmonary inflammatory responses following instillation of size-fractionated ambient particulate matter. Journal of Toxicology and Environmental Health, Part A 66:2193-2207.

Ding G, Cui C, Chen L, Gao Y, Zhou Y, Shi R, et al. 2015. Prenatal exposure to pyrethroid insecticides and birth outcomes in rural northern china. J Expo Sci Environ Epidemiol 25:264-270.

Dong K. 2007. Insect sodium channels and insecticide resistance. Invertebrate Neuroscience 7:17.

Du L, Wang H, Xu W, Zeng Y, Hou Y, Zhang Y, et al. 2013. Application of ultraperformance liquid chromatography/mass spectrometry–based metabonomic techniques to analyze the joint toxic action of long-term low-level exposure to a mixture of organophosphate pesticides on rat urine profile. Toxicological Sciences 134:195-206.

Du L, Li S, Qi L, Hou Y, Zeng Y, Xu W, et al. 2014. Metabonomic analysis of the joint toxic action of long-term low-level exposure to a mixture of four organophosphate pesticides in rat plasma. Mol Biosyst 10:1153-1161.

Duhig K, Chappell LC, Shennan AH. 2016. Oxidative stress in pregnancy and reproduction. Obstet Med 9:113-116.

EPA C. 2017. Summary of pesticide use report data - 2017. California Department of Pesticide Regulation.

Eze IC, Hemkens LG, Bucher HC, Hoffmann B, Schindler C, Kunzli N, et al. 2015. Association between ambient air pollution and diabetes mellitus in europe and north america: Systematic review and meta-analysis. Environ Health Persp 123:381-389.

Farkhondeh T, Mehrpour O, Forouzanfar F, Roshanravan B, Samarghandian S. 2020. Oxidative stress and mitochondrial dysfunction in organophosphate pesticide-induced neurotoxicity and its amelioration: A review. Environmental Science and Pollution Research 27:24799-24814.

Fenske RA, Lu CS, Curl CL, Shirai JH, Kissel JC. 2005. Biologic monitoring to characterize organophosphorus pesticide exposure among children and workers: An analysis of recent studies in washington state. Environ Health Persp 113:1651-1657.

Ferreira JD, Couto AC, Pombo-de-Oliveira MS, Koifman S, Leukemia BCSGoIA. 2013. In utero pesticide exposure and leukemia in brazilian children< 2 years of age. Environ Health Persp 121:269-275.

Fillmore N, Mori J, Lopaschuk G. 2014. Mitochondrial fatty acid oxidation alterations in heart failure, ischaemic heart disease and diabetic cardiomyopathy. British journal of pharmacology 171:2080-2090.

Foley JE. 1992. Rationale and application of fatty acid oxidation inhibitors in treatment of diabetes mellitus. Diabetes Care 15:773-784.

Ford-Hutchinson AW, Bray MA, Doig MV, Shipley ME, Smith MJH. 1980. Leukotriene b, a potent chemokinetic and aggregating substance released from polymorphonuclear leukocytes. Nature 286:264.

Foreman WT, Majewski MS, Goolsby DA, Wiebe FW, Coupe RH. 2000. Pesticides in the atmosphere of the mississippi river valley, part ii--air. The Science of the total environment 248:213-226.

Fortin J-P, Triche TJ, Jr., Hansen KD. 2017. Preprocessing, normalization and integration of the illumina humanmethylationepic array with minfi. Bioinformatics (Oxford, England) 33:558-560.

Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. 2014. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol 15:503.

Frye RE, Rose S, Chacko J, Wynne R, Bennuri SC, Slattery JC, et al. 2016. Modulation of mitochondrial function by the microbiome metabolite propionic acid in autism and control cell lines. Translational Psychiatry 6:e927-e927.

Funk CD. 2001. Prostaglandins and leukotrienes: Advances in eicosanoid biology. Science (New York, NY) 294:1871-1875.

Furlong MA, Barr DB, Wolff MS, Engel SM. 2017. Prenatal exposure to pyrethroid pesticides and childhood behavior and executive functioning. Neurotoxicology 62:231-238.

Furlong MA, Paul KC, Yan Q, Chuang YH, Cockburn MG, Bronstein JM, et al. 2020. An epigenome-wide association study of ambient pyrethroid pesticide exposures in california's central valley. Int J Hyg Environ Health 229:113569.

Gangemi S, Gofita E, Costa C, Teodoro M, Briguglio G, Nikitovic D, et al. 2016. Occupational and environmental exposure to pesticides and cytokine pathways in chronic diseases (review). Int J Mol Med 38:1012-1020.

Garcia-Ruiz C, Morales A, Fernández-Checa JC. 2015. Glycosphingolipids and cell death: One aim, many ways. Apoptosis 20:607-620.

Garcia-Sevillano MA, Garcia-Barrera T, Gomez-Ariza JL. 2015. Environmental metabolomics: Biological markers for metal toxicity. Electrophoresis 36:2348-2365.

Ghio AJ, Carraway MS, Madden MC. 2012. Composition of air pollution particles and oxidative stress in cells, tissues, and living systems. Journal of Toxicology and Environmental Health, Part B 15:1-21.

Glinianaia SV, Rankin J, Bell R, Pless-Mulloli T, Howel D. 2004. Particulate air pollution and fetal health: A systematic review of the epidemiologic evidence. Epidemiology 15:36-45.

Go YM, Walker DI, Liang Y, Uppal K, Soltow QA, Tran V, et al. 2015. Reference standardization for mass spectrometry and high-resolution metabolomics applications to exposome research. Toxicol Sci 148:531-543.

Go YM, Jones DP. 2016. Exposure memory and lung regeneration. Ann Am Thorac Soc 13 Suppl 2:S452-S461.

Goldberg DW, Wilson JP, Knoblock CA, Ritz B, Cockburn MG. 2008. An effective and efficient approach for manually improving geocoded data. Int J Health Geogr 7:60.

Gu F, Derkach A, Freedman ND, Landi MT, Albanes D, Weinstein SJ, et al. 2016. Cigarette smoking behaviour and blood metabolomics. Int J Epidemiol 45:1421-1432.

Gu J, Su F, Hong P, Zhang Q, Zhao M. 2019. (1)h nmr-based metabolomic analysis of nine organophosphate flame retardants metabolic disturbance in hep g2 cell line. The Science of the total environment 665:162-170.

Guerra R, Vera-Aguilar E, Uribe-Ramirez M, Gookin G, Camacho J, Osornio-Vargas AR, et al. 2013. Exposure to inhaled particulate matter activates early markers of oxidative stress, inflammation and unfolded protein response in rat striatum. Toxicology Letters 222:146-154.

Guxens M, Garcia-Esteban R, Giorgis-Allemand L, Forns J, Badaloni C, Ballester F, et al. 2014. Air pollution during pregnancy and childhood cognitive and psychomotor development: Six european birth cohorts. Epidemiology 25:636-647.

Hanke W, Romitti P, Fuortes L, Sobala W, Mikulski M. 2003. The use of pesticides in a polish rural population and its effect on birth weight. International archives of occupational and environmental health 76:614-620.

Happo MS, Uski O, Jalava PI, Kelz J, Brunner T, Hakulinen P, et al. 2013. Pulmonary inflammation and tissue damage in the mouse lung after exposure to pm samples from biomass heating appliances of old and modern technologies. Science of The Total Environment 443:256-266.

Heck JE, Wu J, Lombardi C, Qiu J, Meyers TJ, Wilhelm M, et al. 2013. Childhood cancer and traffic-related air pollution exposure in pregnancy and early life. Environ Health Perspect 121:1385-1391.

Ho SM, Johnson A, Tarapore P, Janakiram V, Zhang X, Leung YK. 2012. Environmental epigenetics and its implication on disease risk and health outcomes. ILAR J 53:289-305.

Hoggatt KJ, Flores M, Solorio R, Wilhelm M, Ritz B. 2012. The "latina epidemiologic paradox" revisited: The role of birthplace and acculturation in predicting infant low birth weight for latinas in los angeles, ca. J Immigr Minor Health 14:875-884.

Horrobin DF. 1992. Nutritional and medical importance of gamma-linolenic acid. Prog Lipid Res 31:163-194.

Horvath S, Levine AJ. 2015. Hiv-1 infection accelerates age according to the epigenetic clock. J Infect Dis 212:1563-1573.

Horvath S, Ritz BR. 2015. Increased epigenetic age and granulocyte counts in the blood of parkinson's disease patients. Aging (Albany NY) 7:1130-1142.

Hossain MM, Richardson JR. 2011. Mechanism of pyrethroid pesticide–induced apoptosis: Role of calpain and the er stress pathway. Toxicological Sciences 122:512-525.

Hou L, Zhang X, Wang D, Baccarelli A. 2012. Environmental chemical exposures and human epigenetics. Int J Epidemiol 41:79-105.

Hougaard KS, Jensen KA, Nordly P, Taxvig C, Vogel U, Saber AT, et al. 2008. Effects of prenatal exposure to diesel exhaust particles on postnatal development, behavior, genotoxicity and inflammation in mice. Part Fibre Toxicol 5.

Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. Bmc Bioinformatics 13.

Hunn BHM, Cragg SJ, Bolam JP, Spillantini M-G, Wade-Martins R. 2015. Impaired intracellular trafficking defines early parkinson's disease. Trends Neurosci 38:178-188.

Jarrell ZR, Smith MR, Hu X, Orr M, Liu KH, Quyyumi AA, et al. 2020. Plasma acylcarnitine levels increase with healthy aging. Aging 12:13555-13570.

Jayaraj R, Megha P, Sreedev P. 2016. Organochlorine pesticides, their toxic effects on living organisms and their fate in the environment. Interdiscip Toxicol 9:90-100.

Jellali R, Gilard F, Pandolfi V, Legendre A, Fleury MJ, Paullier P, et al. 2018. Metabolomics-on-a-chip approach to study hepatotoxicity of ddt, permethrin and their mixtures. J Appl Toxicol 38:1121-1134.

Jenner P. 2003. Oxidative stress in parkinson's disease. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society 53:S26-S38.

Jeong A, Fiorito G, Keski-Rahkonen P, Imboden M, Kiss A, Robinot N, et al. 2018. Perturbation of metabolic pathways mediates the association of air pollutants with asthma and cardiovascular diseases. Environ Int 119:334-345.

Jin L, Godri Pollitt KJ, Liew Z, Rosen Vollmar AK, Vasiliou V, Johnson CH, et al. 2021. Use of untargeted metabolomics to explore the air pollution-related disease continuum. Current Environmental Health Reports.

Johnson JM, Yu T, Strobel FH, Jones DP. 2010. A practical approach to detect unique metabolic patterns for personalized medicine. The Analyst 135:2864.

Johnson WE, Li C, Rabinovic A. 2007. Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics 8:118-127.

Jones DP. 2016. Sequencing the exposome: A call to action. Toxicol Rep 3:29-45.

Jones DP, Walker DI, Uppal K, Rohrbeck P, Mallon TM, Go YM. 2016. Metabolic pathways and networks associated with tobacco use in military personnel. J Occup Environ Med 58:S111-S116.

Kaaja R, Tikkanen MJ, Viinikka L, Ylikorkala O. 1995. Serum-lipoproteins, insulin, and urinary prostanoid metabolites in normal and hypertensive pregnant-women. Obstet Gynecol 85:353-356.

Kampa M, Castanas E. 2008. Human health effects of air pollution. Environ Pollut 151:362-367.

Kang GA, Bronstein JM, Masterman DL, Redelings M, Crum JA, Ritz B. 2005. Clinical characteristics in early parkinson's disease in a central california population-based study. Mov Disord 20:1133-1142.

Karami-Mohajeri S, Abdollahi M. 2011. Toxic influence of organophosphate, carbamate, and organochlorine pesticides on cellular metabolism of lipids, proteins, and carbohydrates: A systematic review. Hum Exp Toxicol 30:1119-1140.

Katsouyanni K. 2003. Ambient air pollution and health. Br Med Bull 68:143-156.

Kegley S, Hill B, Orme S, Choi A. 2000. Pan pesticide database. Pesticide Action Network, North America (San.

Kelly FJ. 2003. Oxidative stress: Its role in air pollution and adverse health effects. Occup Environ Med 60:612-616.

Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. 2015. Integrative analysis of 111 reference human epigenomes. Nature 518:317-330.

Ladva CN, Golan R, Liang D, Greenwald R, Walker DI, Uppal K, et al. 2018. Particulate metal exposures induce plasma metabolome changes in a commuter panel study. Plos One 13:e0203468.

Lai CH, Lee CN, Bai KJ, Yang YL, Chuang KJ, Wu SM, et al. 2016. Protein oxidation and degradation caused by particulate matter. Sci Rep 6:33727.

Lane RK, Hilsabeck T, Rea SL. 2015. The role of mitochondrial dysfunction in age-related diseases. Biochimica et Biophysica Acta (BBA)-Bioenergetics 1847:1387-1400.

Langfelder P, Horvath S. 2008. Wgcna: An r package for weighted correlation network analysis. Bmc Bioinformatics 9.

Langfelder P, Horvath S. 2012. Fast r functions for robust correlations and hierarchical clustering. J Stat Softw 46.

Lavigne E, Burnett RT, Stieb DM, Evans GJ, Godri Pollitt KJ, Chen H, et al. 2018. Fine particulate air pollution and adverse birth outcomes: Effect modification by regional nonvolatile oxidative potential. Environ Health Perspect 126:077012.

Le Cao KA, Rossouw D, Robert-Granie C, Besse P. 2008. A sparse pls for variable selection when integrating omics data. Stat Appl Genet Mol Biol 7:Article 35.

Le Cao KA, Gonzalez I, Dejean S. 2009. Integromics: An r package to unravel relationships between two omics datasets. Bioinformatics 25:2855-2856.

Lee KY, Wong CK, Chuang KJ, Bien MY, Cao JJ, Han YM, et al. 2014. Methionine oxidation in albumin by fine haze particulate matter: An in vitro and in vivo study. Journal of hazardous materials 274:384-391.

Leung MCK, Meyer JN. 2019. Mitochondria as a target of organophosphate and carbamate pesticides: Revisiting common mechanisms of action with new approach methodologies. Reproductive Toxicology 89:83-92.

Levitin J, Harkonen J, Kukkonen J, Nikmo J. 2005. Evaluation of the caline4 and car-fmi models against measurements near a major road. Atmospheric Environment 39:4439-4452.

Li F, Wang L, Ji C, Wu H, Zhao J, Tang J. 2017. Toxicological effects of tris (2-chloropropyl) phosphate in human hepatic cells. Chemosphere 187:88-96.

Li H, Cai J, Chen R, Zhao Z, Ying Z, Wang L, et al. 2017. Particulate matter exposure and stress hormone levels: A randomized, double-blind, crossover trial of air purification. Circulation 136:618-627.

Li M, Wang J, Lu Z, Wei D, Yang M, Kong L. 2014. Nmr-based metabolomics approach to study the toxicity of lambda-cyhalothrin to goldfish (carassius auratus). Aquatic Toxicology 146:82-92.

Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, et al. 2013. Predicting network activity from high throughput metabolomics. Plos Comput Biol 9:e1003123.

Li S, Dunlop AL, Jones DP, Corwin EJ. 2016. High-resolution metabolomics: Review of the field and implications for nursing science and the study of preterm birth. Biol Res Nurs 18:12-22.

Li S, Todor A, Luo R. 2016. Blood transcriptomics and metabolomics for personalized medicine. Computational and Structural Biotechnology Journal 14:1-7.

Li S, Sullivan NL, Rouphael N, Yu T, Banton S, Maddur MS, et al. 2017. Metabolic phenotypes of response to vaccination in humans. Cell 169:862-877 e817.

Li S, Cirillo P, Hu X, Tran V, Krigbaum N, Yu S, et al. 2019. Understanding mixed environmental exposures using metabolomics via a hierarchical community network model in a cohort of california women in 1960's. Reproductive Toxicology.

Li W, Wilker EH, Dorans KS, Rice MB, Schwartz J, Coull BA, et al. 2016. Short-term exposure to air pollution and biomarkers of oxidative stress: The framingham heart study. J Am Heart Assoc 5.

Liang D, Moutinho JL, Golan R, Yu T, Ladva CN, Niedzwiecki M, et al. 2018. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environ Int 120:145-154.

Lie PPY, Nixon RA. 2019. Lysosome trafficking and signaling in health and neurodegenerative diseases. Neurobiology of disease 122:94-105.

Lin MT, Beal MF. 2006. Mitochondrial dysfunction and oxidative stress in neurodegenerative diseases. Nature 443:787-795.

Liu KH, Walker DI, Uppal K, Tran V, Rohrbeck P, Mallon TM, et al. 2016. High-resolution metabolomics assessment of military personnel: Evaluating analytical strategies for chemical detection. J Occup Environ Med 58:S53-61.

Liu KH, Nellis M, Uppal K, Ma C, Tran V, Liang Y, et al. 2020. Reference standardization for quantification and harmonization of large-scale metabolomics. Analytical Chemistry 92:8836-8844.

Liu Y, Hou Y, Wang G, Zheng X, Hao H. 2020. Gut microbial metabolites of aromatic amino acids as signals in host-microbe interplay. Trends in Endocrinology & Metabolism 31:818-834.

Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al. 2013. The carcinogenicity of outdoor air pollution. Lancet Oncol 14:1262-1263.

Louis P, Flint HJ. 2017. Formation of propionate and butyrate by the human colonic microbiota. Environmental Microbiology 19:29-41.

Lu C, Fenske RA, Simcox NJ, Kalman D. 2000. Pesticide exposure of children in an agricultural community: Evidence of household proximity to farmland and take home exposure pathways. Environ Res 84:290-302.

Lukaszewicz-Hussain A, Moniuszko-Jakoniuk J. 2005. A low dose of chlorfenvinphos affects hepatic enzymes in serum and antioxidant enzymes in erythrocytes and liver of rats. Polish Journal of Environmental Studies 14.

Maitre L, Robinson O, Martinez D, Toledano MB, Ibarluzea J, Marina LS, et al. 2018. Urine metabolic signatures of multiple environmental pollutants in pregnant women: An exposome approach. Environ Sci Technol 52:13469-13480.

Mallozzi M, Bordi G, Garo C, Caserta D. 2016. The effect of maternal exposure to endocrine disrupting chemicals on fetal and neonatal development: A review on the major concerns. Birth Defects Research Part C: Embryo Today: Reviews 108:224-242.

Marmur A, Mamane Y. 2003. Comparison and evaluation of several mobile-source and line-source models in israel. Transportation Research Part D-Transport and Environment 8:249-265.

Martens DS, Gouveia S, Madhloum N, Janssen BG, Plusquin M, Vanpoucke C, et al. 2017. Neonatal cord blood oxylipins and exposure to particulate matter in the early-life environment: An environage birth cohort study. Environ Health Perspect 125:691-698.

Mattingly CJ, Rosenstein MC, Colby GT, Forrest JN, Jr., Boyer JL. 2006. The comparative toxicogenomics database (ctd): A resource for comparative toxicological studies. J Exp Zool A Comp Exp Biol 305:689-692.

McGeer PL, McGeer EG. 2004. Inflammation and neurodegeneration in parkinson's disease. Parkinsonism & related disorders 10:S3-S7.

Menni C, Metrustry SJ, Mohney RP, Beevers S, Barratt B, Spector TD, et al. 2015. Circulating levels of antioxidant vitamins correlate with better lung function and reduced exposure to ambient pollution. Am J Respir Crit Care Med 191:1203-1207.

Meredith G, Totterdell S, Beales M, Meshul C. 2009. Impaired glutamate homeostasis and programmed cell death in a chronic mptp mouse model of parkinson's disease. Experimental neurology 219:334-340.

Mihaylova MM, Shaw RJ. 2011. The ampk signalling pathway coordinates cell growth, autophagy and metabolism. Nature Cell Biology 13:1016-1023.

Nagiah S, Phulukdaree A, Naidoo D, Ramcharan K, Naidoo RN, Moodley D, et al. 2015. Oxidative stress and air pollution exposure during pregnancy: A molecular assessment. Hum Exp Toxicol 34:838-847.

Narayan S, Liew Z, Paul K, Lee PC, Sinsheimer JS, Bronstein JM, et al. 2013. Household organophosphorus pesticide use and parkinson's disease. Int J Epidemiol 42:1476-1485.

Nassan FL, Kelly RS, Kosheleva A, Koutrakis P, Vokonas PS, Lasky-Su JA, et al. 2021. Metabolomic signatures of the long-term exposure to air pollution and temperature. Environ Health-Glob 20.

Nicolson GL. 2007. Metabolic syndrome and mitochondrial function: Molecular replacement and antioxidant supplements to prevent membrane peroxidation and restore mitochondrial function. Journal of cellular biochemistry 100:1352-1369.

Niedzwiecki MM, Walker DI, Vermeulen R, Chadeau-Hyam M, Jones DP, Miller GW. 2019. The exposome: Molecules to populations. Annu Rev Pharmacol Toxicol 59:107-127.

Niu YC, Feng RN, Hou Y, Li K, Kang Z, Wang J, et al. 2012. Histidine and arginine are associated with inflammation and oxidative stress in obese women. Br J Nutr 108:57-61.

Noctor G, Lelarge-Trouverie C, Mhamdi A. 2015. The metabolomics of oxidative stress. Phytochemistry 112:33-53.

Ogburn PL, Jr., Williams PP, Johnson SB, Holman RT. 1984. Serum arachidonic acid levels in normal and preeclamptic pregnancies. Am J Obstet Gynecol 148:5-9.

Orta OR, Wesselink AK, Bethea TN, Henn BC, Sjödin A, Wegienka G, et al. 2020. Correlates of organochlorine pesticide plasma concentrations among reproductive-aged black women. Environ Res 184:109352.

Padayatty SJ, Katz A, Wang Y, Eck P, Kwon O, Lee JH, et al. 2003. Vitamin c as an antioxidant: Evaluation of its role in disease prevention. J Am Coll Nutr 22:18-35.

Parrón T, Requena M, Hernández AF, Alarcón R. 2011. Association between environmental exposure to pesticides and neurodegenerative diseases. Toxicology and applied pharmacology 256:379-385.

Patel S, Pandey AK, Bajpayee M, Parmar D, Dhawan A. 2006. Cypermethrin-induced DNA damage in organs and tissues of the mouse: Evidence from the comet assay. Mutation Research/Genetic Toxicology and Environmental Mutagenesis 607:176-183.

Patterson PH. 2009. Immune involvement in schizophrenia and autism: Etiology, pathology and animal models. Behav Brain Res 204:313-321.

Paul KC, Sinsheimer JS, Rhodes SL, Cockburn M, Bronstein J, Ritz B. 2016. Organophosphate pesticide exposures, nitric oxide synthase gene variants, and gene-pesticide interactions in a case-control study of parkinson's disease, california (USA). Environ Health Perspect 124:570-577.

Paul KC, Chuang YH, Cockburn M, Bronstein JM, Horvath S, Ritz B. 2018a. Organophosphate pesticide exposure and differential genome-wide DNA methylation. The Science of the total environment 645:1135-1143.

Paul KC, Jerrett M, Ritz B. 2018b. Type 2 diabetes mellitus and alzheimer's disease: Overlapping biologic mechanisms and environmental risk factors. Curr Environ Health Rep 5:44-58.

Peden DB. 1999. 37 - controlled exposures of asthmatics to air pollutants. In: Air pollution and health, (Holgate ST, Samet JM, Koren HS, Maynard RL, eds). London:Academic Press, 865-880.

Peden DB. 2001. Air pollution in asthma: Effect of pollutants on airway inflammation. Ann Allergy Asthma Immunol 87:12-17.

Pedersen M, Giorgis-Allemand L, Bernard C, Aguilera I, Andersen AM, Ballester F, et al. 2013. Ambient air pollution and low birthweight: A european cohort study (escape). Lancet Respir Med 1:695-704.

Pedersen M, Stayner L, Slama R, Sorensen M, Figueras F, Nieuwenhuijsen MJ, et al. 2014. Ambient air pollution and pregnancy-induced hypertensive disorders: A systematic review and meta-analysis. Hypertension 64:494-500.

Peng K-Y, Watt MJ, Rensen S, Greve JW, Huynh K, Jayawardana KS, et al. 2018. Mitochondrial dysfunction-related lipid changes occur in nonalcoholic fatty liver disease progression. Journal of lipid research 59:1977-1986.

Peter Stein T, Scholl TO, Schluter MD, Leskiw MJ, Chen X, Spur BW, et al. 2008. Oxidative stress early in pregnancy and pregnancy outcome. Free Radic Res 42:841-848.

Petersen AK, Zeilinger S, Kastenmuller G, Romisch-Margl W, Brugger M, Peters A, et al. 2014. Epigenetics meets metabolomics: An epigenome-wide association study with blood serum metabolic traits. Hum Mol Genet 23:534-545.

Pisoschi AM, Pop A. 2015. The role of antioxidants in the chemistry of oxidative stress: A review. Eur J Med Chem 97:55-74.

Powell WS, Rokach J, Khanapure SP, Manna S, Hashefi M, Gravel S, et al. 1996. Effects of metabolites of leukotriene b4 on human neutrophil migration and cytosolic calcium levels. Journal of Pharmacology and Experimental Therapeutics 276:728-736.

Priyadarshini M, A Kamal M, H Greig N, Realef M, M Abuzenadah A, GA Chaudhary A, et al. 2012. Alzheimer's disease and type 2 diabetes: Exploring the association to obesity and tyrosine hydroxylase. CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders) 11:482-489.

Reuter SE, Evans AM. 2012. Carnitine and acylcarnitines. Clinical pharmacokinetics 51:553-572.

Risom L, Moller P, Loft S. 2005. Oxidative stress-induced DNA damage by particulate air pollution. Mutat Res 592:119-137.

Ritz B, Yu F. 1999. The effect of ambient carbon monoxide on low birth weight among children born in southern california between 1989 and 1993. Environ Health Perspect 107:17-25.

Ritz B, Yu F, Chapa G, Fruin S. 2000. Effect of air pollution on preterm birth among children born in southern california between 1989 and 1993. Epidemiology 11:502-511.

Ritz B, Wilhelm M, Hoggatt KJ, Ghosh JK. 2007. Ambient air pollution and preterm birth in the environment and pregnancy outcomes study at the university of california, los angeles. Am J Epidemiol 166:1045-1052.

Ritz BR, Paul KC, Bronstein JM. 2016. Of pesticides and men: A california story of genes and environment in parkinson's disease. Curr Environ Health Rep 3:40-52.

Romero R, Espinoza J, Goncalves LF, Kusanovic JP, Friel L, Hassan S. 2007. The role of inflammation and infection in preterm birth. Semin Reprod Med 25:21-39.

Rull RA, Ritz B. 2003. Historical pesticide exposure in california using pesticide use reports and land-use surveys: An assessment of misclassification error and bias. Environ Health Persp 111:1582-1589.

Rull RP, Ritz B, Shaw GM. 2006. Neural tube defects and maternal residential proximity to agricultural pesticide applications. Am J Epidemiol 163:743-753.

Rusiecki JA, Beane Freeman LE, Bonner MR, Alexander M, Chen L, Andreotti G, et al. 2017. High pesticide exposure events and DNA methylation among pesticide applicators in the agricultural health study. Environmental and molecular mutagenesis 58:19-29.

Saccenti E, Hoefsloot HCJ, Smilde AK, Westerhuis JA, Hendriks MMWB. 2014. Reflections on univariate and multivariate analysis of metabolomics data. Metabolomics 10:361-374.

Saillenfait AM, Ndiaye D, Sabate JP. 2015. Pyrethroids: Exposure and health effects--an update. Int J Hyg Environ Health 218:281-292.

Salihovic S, Ganna A, Fall T, Broeckling CD, Prenni JE, van Bavel B, et al. 2016. The metabolic fingerprint of p,p'-dde and hcb exposure in humans. Environ Int 88:60-66.

Sanborn M, Kerr KJ, Sanin LH, Cole DC, Bassil KL, Vakil C. 2007. Non-cancer health effects of pesticides: Systematic review and implications for family doctors. Can Fam Physician 53:1712-1720.

Sato H, Taketomi Y, Murakami M. 2016. Metabolic regulation by secreted phospholipase a2. Inflamm Regen 36:7.

Sato R, Goldstein JL, Brown MS. 1993. Replacement of serine-871 of hamster 3-hydroxy-3-methylglutaryl-coa reductase prevents phosphorylation by amp-activated kinase and blocks inhibition of sterol synthesis induced by atp depletion. Proceedings of the National Academy of Sciences 90:9261-9265.

Schell LM, Gallo MV, Denham M, Ravenscroft J. 2006. Effects of pollution on human growth and development: An introduction. J Physiol Anthropol 25:103-112.

Schrimpe-Rutledge AC, Codreanu SG, Sherrod SD, McLean JA. 2016. Untargeted metabolomics strategies-challenges and emerging directions. J Am Soc Mass Spectrom 27:1897-1905.

Shelton JF, Geraghty EM, Tancredi DJ, Delwiche LD, Schmidt RJ, Ritz B, et al. 2014. Neurodevelopmental disorders and prenatal residential proximity to agricultural pesticides: The charge study. Environ Health Perspect 122:1103-1109.

Shetewy A, Shimada-Takaura K, Warner D, Jong CJ, Mehdi A-BA, Alexeyev M, et al. 2016. Mitochondrial defects associated with β-alanine toxicity: Relevance to hyper-beta-alaninemia. Mol Cell Biochem 416:11-22.

Simon-Manso Y, Lowenthal MS, Kilpatrick LE, Sampson ML, Telu KH, Rudnick PA, 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 85:11725-11731.

Soane L, Kahraman S, Kristian T, Fiskum G. 2007. Mechanisms of impaired mitochondrial energy metabolism in acute and chronic neurodegenerative disorders. J Neurosci Res 85:3407-3415.

Sonsalla P, Albers D, Zeevalk G. 1998. Role of glutamate in neurodegeneration of dopamine neurons in several animal models of parkinsonism. Amino Acids 14:69-74.

Stieb DM, Chen L, Eshoul M, Judek S. 2012. Ambient air pollution, birth weight and preterm birth: A systematic review and meta-analysis. Environ Res 117:100-111.

Stingone JA, Louis GMB, Nakayama SF, Vermeulen RCH, Kwok RK, Cui YX, et al. 2017. Toward greater implementation of the exposome research paradigm within environmental epidemiology. Annual Review of Public Health, Vol 38 38:315-327.

Suades-Gonzalez E, Gascon M, Guxens M, Sunyer J. 2015. Air pollution and neuropsychological development: A review of the latest evidence. Endocrinology 156:3473-3482.

Sultana Z, Maiti K, Aitken J, Morris J, Dedman L, Smith R. 2017. Oxidative stress, placental ageing-related pathologies and adverse pregnancy outcomes. Am J Reprod Immunol 77.

Sultatos LG. 1994. Mammalian toxicology of organophosphorus pesticides. J Toxicol Environ Health 43:271-289.

Surowiec I, Karimpour M, Gouveia-Figueira S, Wu J, Unosson J, Bosson JA, et al. 2016. Multi-platform metabolomics assays for human lung lavage fluids in an air pollution exposure study. Anal Bioanal Chem 408:4751-4764.

Swerdlow RH. 2011. Brain aging, alzheimer's disease, and mitochondria. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease 1812:1630-1639.

Tabrez S, R Jabir N, Shakil S, H Greig N, Alam Q, M Abuzenadah A, et al. 2012. A synopsis on the role of tyrosine hydroxylase in parkinson's disease. CNS & Neurological Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders) 11:395-409.

Tacconelli S, Capone ML, Patrignani P. 2010. Measurement of 8-iso-prostaglandin f2α in biological fluids as a measure of lipid peroxidation. In: Cyclooxygenases: Methods and protocols, (Ayoub SS, Flower RJ, Seed MP, eds). Totowa, NJ:Humana Press, 165-178.

Tam VC, Quehenberger O, Oshansky CM, Suen R, Armando AM, Treuting PM, et al. 2013. Lipidomic profiling of influenza infection identifies mediators that induce and resolve inflammation. Cell 154:213-227.

Terry Jr AV. 2012. Functional consequences of repeated organophosphate exposure: Potential non-cholinergic mechanisms. Pharmacology & therapeutics 134:355-365.

Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, et al. 2017. Champ: Updated methylation analysis pipeline for illumina beadchips. Bioinformatics 33:3982-3984.

Tiedje KE, Stevens K, Barnes S, Weaver DF. 2010. Beta-alanine as a small molecule neurotransmitter. Neurochem Int 57:177-188.

Torres-Altoro MI, Mathur BN, Drerup JM, Thomas R, Lovinger DM, O'Callaghan JP, et al. 2011. Organophosphates dysregulate dopamine signaling, glutamatergic neurotransmission, and induce neuronal injury markers in striatum. J Neurochem 119:303-313.

Triche TJ, Jr., Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. 2013. Low-level processing of illumina infinium DNA methylation beadarrays. Nucleic Acids Res 41:e90.

Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshirani R, et al. 2001. Missing value estimation methods for DNA microarrays. Bioinformatics 17:520-525.

Uchida K. 2003. Histidine and lysine as targets of oxidative modification. Amino Acids 25:249-257.

Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, Yu T, et al. 2013. Xmsanalyzer: Automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. Bmc Bioinformatics 14:15.

Uppal K, Walker DI, Liu K, Li S, Go YM, Jones DP. 2016. Computational metabolomics: A framework for the million metabolome. Chem Res Toxicol 29:1956-1975.

Uppal K, Walker DI, Jones DP. 2017. Xmsannotator: An r package for network-based annotation of high-resolution metabolomics data. Anal Chem 89:1063-1067.

Uppal K, Ma C, Go YM, Jones DP, Wren J. 2018. Xmwas: A data-driven integration and differential network analysis tool. Bioinformatics 34:701-702.

Valvi D, Walker DI, Inge T, Bartell SM, Jenkins T, Helmrath M, et al. 2020. Environmental chemical burden in metabolic tissues and systemic biological pathways in adolescent bariatric surgery patients: A pilot untargeted metabolomic approach. Environ Int 143:105957.

van der Plaat DA, de Jong K, de Vries M, van Diemen CC, Nedeljković I, Amin N, et al. 2018. Occupational exposure to pesticides is associated with differential DNA methylation. Occupational and environmental medicine 75:427-435.

van Veldhoven K, Kiss A, Keski-Rahkonen P, Robinot N, Scalbert A, Cullinan P, et al. 2018. Impact of short-term traffic-related air pollution on the metabolome - results from two metabolome-wide experimental studies. Environ Int 123:124-131.

Vlaanderen JJ, Janssen NA, Hoek G, Keski-Rahkonen P, Barupal DK, Cassee FR, et al. 2017. The impact of ambient air pollution on the human blood metabolome. Environ Res 156:341-348.

von Ehrenstein OS, Aralis H, Cockburn M, Ritz B. 2014. In utero exposure to toxic air pollutants and risk of childhood autism. Epidemiology 25:851-858.

Wainwright SL, Powell WS. 1991. Mechanism for the formation of dihydro metabolites of 12-hydroxyeicosanoids. Conversion of leukotriene b4 and 12-hydroxy-5,8,10,14-eicosatetraenoic acid to 12-oxo intermediates. J Biol Chem 266:20899-20906.

Walker DI, Pennell KD, Uppal K, Xia X, Hopke PK, Utell MJ, et al. 2016a. Pilot metabolome-wide association study of benzo(a)pyrene in serum from military personnel. J Occup Environ Med 58:S44-52.

Walker DI, Uppal K, Zhang L, Vermeulen R, Smith M, Hu W, et al. 2016b. High-resolution metabolomics of occupational exposure to trichloroethylene. Int J Epidemiol 45:1517-1527.

Walker DI, Lane KJ, Liu K, Uppal K, Patton AP, Durant JL, et al. 2018. Metabolomic assessment of exposure to near-highway ultrafine particles. J Expo Sci Environ Epidemiol.

Walker DI, Marder ME, Yano Y, Terrell M, Liang Y, Barr DB, et al. 2019a. Multigenerational metabolic profiling in the michigan pbb registry. Environ Res 172:182-193.

Walker DI, Perry-Walker K, Finnell RH, Pennell KD, Tran V, May RC, et al. 2019b. Metabolome-wide association study of anti-epileptic drug treatment during pregnancy. Toxicol Appl Pharmacol 363:122-130.

Wang A, Costello S, Cockburn M, Zhang X, Bronstein J, Ritz B. 2011. Parkinson's disease risk from ambient exposure to pesticides. Eur J Epidemiol 26:547-555.

Wang H-P, Liang Y-J, Long D-X, Chen J-X, Hou W-Y, Wu Y-J. 2009. Metabolic profiles of serum from rats after subchronic exposure to chlorpyrifos and carbaryl. Chemical Research in Toxicology 22:1026-1033.

Wang HP, Liang YJ, Zhang Q, Long DX, Li W, Li L, et al. 2011. Changes in metabolic profiles of urine from rats following chronic exposure to anticholinesterase pesticides. Pestic Biochem Phys 101:232-239.

Wang Z, Zheng Y, Zhao B, Zhang Y, Liu Z, Xu J, et al. 2015. Human metabolic responses to chronic environmental polycyclic aromatic hydrocarbon exposure by a metabolomic approach. J Proteome Res 14:2583-2593.

Ward MH, Lubin J, Giglierano J, Colt JS, Wolter C, Bekiroglu N, et al. 2006. Proximity to crops and residential exposure to agricultural herbicides in iowa. Environ Health Perspect 114:893-897.

Ward-Caviness CK, Breitner S, Wolf K, Cyrys J, Kastenmuller G, Wang-Sattler R, et al. 2016. Short-term no2 exposure is associated with long-chain fatty acids in prospective cohorts from augsburg, germany: Results from an analysis of 138 metabolites and three exposures. International Journal of Epidemiology 45:1528-1538.

Watanabe M, Suliman ME, Qureshi AR, Garcia-Lopez E, Barany P, Heimburger O, et al. 2008. Consequences of low plasma histidine in chronic kidney disease patients: Associations with inflammation, oxidative stress, and mortality. Am J Clin Nutr 87:1860-1866.

Weichenthal S, Moase C, Chan P. 2010. A review of pesticide exposure and cancer incidence in the agricultural health study cohort. Environ Health Persp 118:1117-1125.

Weisskopf M, Knekt P, O'reilly E, Lyytinen J, Reunanen A, Laden F, et al. 2010. Persistent organochlorine pesticides in serum and risk of parkinson disease. Neurology 74:1055-1061.

Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, et al. 2013. Hmdb 3.0--the human metabolome database in 2013. Nucleic Acids Res 41:D801-807.

Wold S, Sjostrom M, Eriksson L. 2001. Pls-regression: A basic tool of chemometrics. Chemometr Intell Lab 58:109-130.

Wu J, Ren C, Delfino RJ, Chung J, Wilhelm M, Ritz B. 2009. Association between local traffic-generated air pollution and preeclampsia and preterm delivery in the south coast air basin of california. Environ Health Perspect 117:1773-1779.

Wu J, Laurent O, Li L, Hu J, Kleeman M. 2016. Adverse reproductive health outcomes and exposure to gaseous and particulate-matter air pollution in pregnant women. Res Rep Health Eff Inst:1-58.

Xu MY, Wang P, Sun YJ, Wu YJ. 2019. Disruption of kidney metabolism in rats after subchronic combined exposure to low-dose cadmium and chlorpyrifos. Chem Res Toxicol 32:122-129.

Yan Q, Liew Z, Uppal K, Cui X, Ling C, Heck JE, et al. 2019. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ Int 130:104872.

Yan W, Yun Y, Ku T, Li G, Sang N. 2016. No2 inhalation promotes alzheimer's disease-like progression: Cyclooxygenase-2-derived prostaglandin e2 modulation and monoacylglycerol lipase inhibition-targeted medication. Sci Rep-Uk 6:22429.

Yang J, Sun X, Feng Z, Hao D, Wang M, Zhao X, et al. 2011. Metabolomic analysis of the toxic effects of chronic exposure to low-level dichlorvos on rats using ultra-performance liquid chromatography-mass spectrometry. Toxicol Lett 206:306-313.

Yang J, Wang H, Xu W, Hao D, Du L, Zhao X, et al. 2013. Metabolomic analysis of rat plasma following chronic low-dose exposure to dichlorvos. Hum Exp Toxicol 32:196-205.

Yang X, Zhang M, Lu T, Chen S, Sun X, Guan Y, et al. 2020. Metabolomics study and meta-analysis on the association between maternal pesticide exposome and birth outcomes. Environ Res 182:109087.

Yu GC, Wang LG, Han YY, He QY. 2012. Clusterprofiler: An r package for comparing biological themes among gene clusters. Omics-a Journal of Integrative Biology 16:284-287.

Yu T, Park Y, Johnson JM, Jones DP. 2009. Aplcms--adaptive processing of high-resolution lc/ms data. Bioinformatics 25:1930-1936.

Zhang T, De Waard AA, Wuhrer M, Spaapen RM. 2019. The role of glycosphingolipids in immune cell functions. Frontiers in Immunology 10.

Zhang X, Wallace AD, Du P, Kibbe WA, Jafari N, Xie H, et al. 2012. DNA methylation alterations in response to pesticide exposure in vitro. Environmental and molecular mutagenesis 53:542-549.

Zuluaga M, Robledo S, Osorio-Zuluaga GA, Yathe L, Gonzalez D, Taborda G. 2016. Metabolomics and pesticides: Systematic literature review using graph theory for analysis of references. Nova 14:121-138.