Distribution Agreement

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

Haoran Cheng April 17, 2023

Aromatic compound metabolism and oxidative stress as potential mediators between ambient air pollution and hypertensive disorders of pregnancy--evidence from an MWAS study on Atlanta African American Maternal-Child Cohort

By

Haoran Cheng
Master of Public Health

Gangarosa Department of Environmental Health

Donghai Liang PhD, MPH
Committee Chair

Aromatic compound metabolism and oxidative stress as potential mediators between ambient air pollution and hypertensive disorders of pregnancy--evidence from an MWAS study on Atlanta African American Maternal-Child Cohort

By

Haoran Cheng

BS
Emory University
2022

Thesis Committee Chair: Donghai Liang

An abstract of
a thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Gangarosa Department of Environmental Health
2023

Abstract

Aromatic compound metabolism and oxidative stress as potential mediators between ambient air pollution and hypertensive disorders of pregnancy--evidence from an MWAS study on Atlanta African American Maternal-Child Cohort

By Haoran Cheng

Hypertensive disorders of pregnancies (HDP) are the most common medical disorder of pregnancy and a leading cause of maternal and infant morbidity and mortality. Previous studies have established an association between exposure to ambient air pollution and the occurrence of HDP, but the underlying biological mechanism is unclear, impeding targeted intention strategies.

329 individuals from the Atlanta African American (AA) Maternal-Child cohort were included in the study (68 with HDP including 42 gestational hypertension and 26 preeclampsia). We employed liquid chromatography-high-resolution mass spectrometry to conduct metabolomics profiling on serum samples collected between 8-14 weeks of gestation and developed a spatiotemporally resolved model to estimate exposure to three common ambient air pollutants ($PM_{2.5}$, NO_2, and O_3) during four critical exposure windows (1yr prior to conception, 1st trimester, 1m and 1w prior to blood draw). We investigated overlapping features and pathways using a Meet-in-the-Middle Approach (MITM) and High-Dimensional Mediation Analysis (HDMA).

13,980 and 11,106 metabolic features were extracted from HILIC and C18 chromatography columns. Several metabolites and pathways involved in oxidative stress and systemic inflammation are significantly associated with air pollutant exposures during critical exposure windows and HDP, including phenylalanine, indole, benzoate, LysoPC, degradation of the aromatic compound, and propionate metabolism. Biliverdin and porphyrin metabolism pathways are also associated with exposures and outcomes, indicative of potential hepatic impairment due to aromatic compound metabolism.

Our findings suggest a potentially critical role of various aromatic compounds, indicative of oxidative stress, in the pathophysiology underlying the association between air pollution and HDP.

Aromatic compound metabolism and oxidative stress as potential mediators between ambient air pollution and hypertensive disorders of pregnancy--evidence from an MWAS study on Atlanta African American Maternal-Child Cohort

By

Haoran Cheng

BS
Emory University
2022

Thesis Committee Chair: Donghai Liang PhD, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Gangarosa Department of Environmental Health
2023

Tables of Contents

Introduction

Hypertensive disorders of pregnancies (HDP), as a common type of pregnancy complication, is a leading cause of maternal and infant mortality, accounting for 14% of maternal death globally, with the highest burden in Africa[1]. HDP can be further categorified into 4 groups based on pathophysiology[2]. Chronic hypertension refers to hypertension onset prior to and continuously present during the pregnancy. Gestational hypertension (gHTN) is defined as new onset hypertension developed after 20 weeks of gestation, with systolic blood pressure greater than 140 mmHg and diastolic blood pressure greater than 90 mmHg. Preeclampsia (PE) is characterized by gestational hypertension and proteinuria. The last form is superimposed preeclampsia, with proteinuria as well as organ damage. The occurrence of HDP is closely related to the occurrence of several reproductive health problems, including intrauterine growth restrictions and preterm birth[2].

Air pollution has been recognized as a global public health risk factor, and influences different aspects of human health[3,4]. Common ambient air pollutants include ambient fine particulate matter ($PM_{2.5}$), nitrogen dioxide (NO_2), and Ozone (O_3). Pregnant people and newborns are among the most vulnerable population to air pollution[5]. Maternal exposure to environmental stressors has been linked to adverse pregnancy complications and birth outcomes.[5,60] Moreover, according to the Developmental Origins of Health Theory (DOHaD), these experiences during the early years of a child's life have a long-lasting impact throughout their lifespan, increasing their risk of developing chronic disease later in life[6,7].

Previous studies have established an association between exposure to ambient air pollution and the occurrence of HDP[8,9], but the specific biological mechanism was unclear. An in-depth understanding of the toxicity of various pollutant species is crucial for the development of interventions and policies to reduce the prevalence of HDP.

High-resolution metabolomics has emerged as a powerful biomonitoring tool to examine the biological pathways that are influenced by environmental exposures and disease pathogenesis. Previous studies have identified oxidative stress and inflammatory response pathway perturbations as well as changes in maternal metabolome as a result of air pollution exposure[10,11,12]. Similarly, there have also been ongoing efforts using metabolomics to predict the occurrence of HDP[13]. By identifying the overlapping metabolites and pathways linking ambient air pollution exposure and HDP will contribute to revealing the mechanisms underlying the toxicity of air pollution on HDP. However, few study has used metabolomics together with mediation analysis in environmental reproductive health research.[17,18,19] Historically, most environmental epidemiology study in the US predominantly focuses on Caucasian populations. At the same time, however, African Americans suffer disproportionately from ambient air pollution and other psychosocial stressors, making them particularly susceptible and in need of further research[14]. To address these critical knowledge gaps, we conducted the study using the Atlanta African American Maternal-Child Cohort, a cohort established in 2014 to analyze the impact of environmental exposures on maternal and child health[15,16,17,19,21]. We aimed to use high-resolution untargeted metabolomics, meet-in-

the-middle (MITM) approach, and high dimensional mediation analysis to identify metabolites and pathways that may potentially mediate the association between exposure to ambient air pollution and risk of HDP.

Method

Study population

In this study, we included participants from the Atlanta African Maternal-Child Cohort[15,16], a prospective cohort enrolling African American pregnant people in the Atlanta area at Emory University Hospital Midtown (Private) and Grady Memorial Hospital (Public). The inclusion criterion included ages between 18-40 years old, 8-14 weeks of gestation, singleton pregnancy, no chronic medical condition, and ability to communicate in English. Data collection includes a self-report questionnaire, information extraction for electronic medical records, and blood samples. For our analysis, we included 329 participants on whom we conducted metabolomics profiling. All samples are collected during the first prenatal visit and stored at -80C, until metabolomic profiling in 2020 to reduce the potential batch effect. This study was approved by Emory University Internal Review Board (IRB ID 1071) and signed informed consent was obtained from all study participants.

Air pollution exposure assessment

We used a previously validated ensemble model to estimate participants' exposure to 3 ambient air pollutant species, $PM_{2.5}$, NO_2, and O_3[20]. Specifically, the personal exposure to air

pollution was assessed based on the residential zip code of each participant. The ensemble model integrated multiple machine learning algorithms including neural network, random forest, and gradient boost, to yield pollutant concentrations with a spatial resolution of 1km x 1km. We selected 4 different exposure windows, including 1-year prior to conception, the first trimester of pregnancy,1 month, and 1 week prior to a blood draw to represent short (1tri, 1m, 1w) and long-term (1y) exposure.

Measure of HDP and covariates

2 types of HDP including gHTN and PE were extracted from the electronic medical record system. Specifically, participants were considered to have gHTN if they had new-onset hypertension (systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg at ≥20 weeks of gestation) in the absence of proteinuria or new signs of end-organ dysfunction, and PE if they met the same blood pressure criteria along with proteinuria or new signs of end-organ dysfunction. Covariates and confounding factors adjusted in the analysis are determined based on previous analysis and manifested using a directed acyclic graph (DAG). To better interpret the MITM results, we used the same set of covariates in the exposure-mediator and mediator-outcome model. Demographic information including maternal age and education level was obtained through a questionnaire. Infant sex, parity, gestational age, alcohol use, and other substance use (tobacco and marijuana) are extracted from the Electronic Health Record system. Maternal BMI was calculated using the weight

and height measured during their first visit, and the season of the conception of was calculated based on the time of the delivery in the system.

High-resolution metabolomics

Metabolomics profilings on non-fasting serum samples were conducted using high-resolution liquid chromatography coupled with mass spectrometry (HR-LCMS, Thermo Scientific™ Q-Exactive™ HF) using established protocols[16,21]. All samples, including actual study samples and quality control (QC) samples, were analyzed in triplicate in 2 chromatography columns, the hydrophilic interaction liquid chromatography (HILIC) with positive electrospray ionization (ESI) and C18 hydrophobic reversed-phase chromatography with negative ESI. Metabolic features including mass-to-charge ratio(m/z), retention time (rt), and relative intensity were extracted using the R package apLCMS and xMSanalyzer[22,23], averaged, and then log-transformed for the following analysis. We also calculated the relative standard deviation (RSD) for each feature in the QC sample and missingness in the study and QC samples as part of the quality control process. Features with missingness in study sample <90% or RSD >50% and QC missingness of <10% were excluded. In total, 11,269 and 9,565 metabolic features remained in the current analysis for the HILIC and C18 columns, respectively. Next, we utilized a second auxiliary feature that was correlated (with a Pearson's correlation greater than 0.5) to identify and separate missing values into two categories: those that were missing not at random (MNAR) and those that were missing at random (MAR). By using a correlated auxiliary feature, we assumed that we could gain an understanding of the

missing value patterns for a given feature by analyzing the non-missing observations of its auxiliary feature. The missing values of MNAR features were imputed by quantile regression imputation of left-censored data (QRILC), while those of MAR features were imputed by random forest (RF).[25]

MWAS Analysis and Meet-In-The-Middle Approach

We conducted MWASfor exposures and outcomes separately and then leveraged the MITM approach to identify the overlapping features associated with both exposures and outcomes[26]. Specifically, we use series of multiple linear regression (i.e., exposure-mediator) models and logistic regression (mediator-outcome) models to evaluate the association of metabolic features with exposures and outcomes, respectively, using the fowling equations:

$$\ln(Feature_j) = \beta_{0_j} + \beta_{1j}AP_i + \beta_{2j}Age + \beta_{3j}Education + \beta_{4j}Sex + \beta_{5j}BMI + \beta_{6j}Season + \beta_{7j}Parity + \beta_{8j}Alcohol + \beta_{9j}Tobaco\_Marijuana + \beta_{10j}GA\_Samp \quad \text{Eq. (1)}$$

$$Logit(P(HDP)) = \theta_{0j} + \theta_{1j}\ln(Feature_j) + \theta_{2j}Age + \theta_{3j}Education + \theta_{4j}Sex + \theta_{5j}BMI + \theta_{6j}Season + \theta_{7j}Parity + \theta_{8j}Alcohol + \theta_{9j}Tobaco\_Marijuana + \theta_{10j}GA\_Samp \quad \text{Eq. (2)}$$

Where $\ln(Feature)$ refers to the log transformed intensity of metabolic feature j; AP_i is the averaged exposure of pollutant i in a specific window; Sex is the sex of the child born;

$Season$ is the season of conception; $Parity$ is the times of previous pregnancies; $Alcohol$ is the use of alcohol; $Tobaco\_Marijuana$ is the use of either tobacco or marijuana, GA_Samp is the gestational age at sampling; HDP denotes having gHTN or PE, and those without either condition was treated as a reference; Results were presented using Manhattan plots. All statistical analyses are conducted in R (Version 3.6)

High-Dimensional Mediation Analysis

In addition, we also employed the High-Dimensional Mediation Analysis (HDMA) method using the R package HIMA to uncover the potential mediators linking air pollution exposure to HDP[27]. This method is a complementary approach to previous studies that have developed a mediation analysis framework capable of handling multiple mediators and analyzing the indirect impact of each mediator[28]. In comparison to the MITM approach, HDMA can integrate several mediators into a single mediator-outcome model, allowing us to examine the relationship between specific indirect effects and mediators. HDMA extends the multiple mediator framework to the high-dimensional context by reducing the dimensionality of omics data. We used Benjamin-Hochberg (BH) corrected p-value to identify significant mediators. The detected features were annotated subsequently. Separate analyses were conducted for each column (HILIC positive ESI and C18 negative ESI).

Pathway enrichment and chemical annotation

To predict the biological function of the significant unknown features identified from MITM and MDMA, we conducted pathway enrichment analysis using *metapone*. Metapone is an innovative bioinformatics tool that predicts functional biological activities using untargeted metabolomic data extracted in both positive and negative ESI modes[29]. It combines the Small Molecule Pathway Database (SMPDB) and mummichog database to create its own pathway database. The tool takes in metabolic features that are putatively annotated with weights based on uncertainty in metabolite-feature matching. The significance of enriched biological pathways is then tested, taking into account the weight schema. This approach allows for more accurate predictions of biological pathway activity. To minimize false positive discovery, we only included those pathways associated with either pollutant exposure or HDP at a raw p-value < 0.05, with more than 3 metabolites enriched.

Finally, metabolic features that were significantly associated with exposure or outcome were annotated based on mz, rt, and extracted ion chromatographs (EIC) compared to the authentic chemical reference (confidence level 1).[23] All analyses were completed in R (version 3.6).

Results

A total of 329 individuals from the Atlanta AA cohort were included in the current analysis and their demographic characteristics are described in Table 1, stratified by no HDP, gHTN, and PE. Among all participants, 42 (12.77%) have gHTN and 26 (7.9 %) have PE. Those in gHTN and PE groups have less education beyond the college level compared to those without

HDP. They also have a higher proportion of the use of addictive substances, and higher BMI. The air pollution exposure concentration has been summarized in Table S1 and the correlation among different exposure windows are summarized in Figure S1. Of all 3 pollutant species, NO_2 has the highest correlation between 4 exposure windows. For O_3, there's an inverse relationship between long-term exposure and 3 short-term exposure windows.

Metabolome-wide association analysis

After quality assurance and control, 13,980 and 11,106 metabolic features were extracted from HILIC and C18 chromatography columns, respectively. The number of statistically significant (raw $p<0.05$ and FDR-corrected $q<0.2$) metabolic features associated with different air pollutant exposure windows or HDP outcomes were shown in Table S2, with Manhattan plots from each model in HILIC and C18 columns shown respectively in Figure S2 and S3. Overlapping features associated with at least one air pollution exposure and an adverse pregnancy complication outcome (raw $p<0.05$) have been summarized in Table S3. We were able to confirm a total of 15 features associated with at least one exposure or outcome with level 1 evidence from the 2 columns and the chemical identities of these features have been summarized in Table 2. Specifically, we identified Di (2-ethyl hexyl) phthalate (DEHP) as an overlapping feature that's associated with 1tri and 1m exposure to $PM_{2.5}$ and PE incidence. We have also identified various amino and derivatives, (phenylalanine, acetyl serine, glutamate, indole, cysteine, ketoleucine), lipids

(lysophosphatidylcholine or LysoPC), vitamins and cofactors (retinoate, biliverdin), and xenobiotics (benzoate, phthalic anhydride) associated with at least exposure or outcome.

High-dimensional mediation analysis

Details of the potential mediating features identified by HDMA for each model have been summarized in Table S4 and S5. In the HILIC column, HDMA has identified a total of 29 potential mediating features, among which 27 features are also associated identified by in the same Exposure-Mediator model in MITM. However, only 2 features (mz= 391.2842, rt=22; mz = 91.9894, rt = 37.5) are identified in the Mediator-Outcome model. 4 features identified by HDMA in the HILIC were not identified by MITM in either exposure-mediator or mediator- outcome mode. No feature was identified only in the mediator-outcome model. Similarly, in the C18 column, HDMA has identified a total of 24 potential mediating features, among which 20 features are also associated identified by in the same Exposure-Mediator model. 5 features are also identified in the Mediator-Outcome model, among which one (mz = 455.248, rt = 23.2) is only identified in the Mediator-Outcome model in MITM. 3 features identified by HDMA in the C18 column were not identified by MITM in either exposure-mediator or mediator-outcome mode. There is one feature recognized by both MITM and HDMA (mz= 391.2842, rt=22) as overlapping and we confirmed the identity of this feature to be DEHP with level 1 evidence. We were not able to match any other feature identified by HDMA with level 1 evidence.

Pathway Enrichment Analysis

A total of 68 pathways are identified associated with at least one exposure window or outcome. Pathways include amino acid (glycine, serine, tryptophan, tyrosine, leucine), lipid (sphingolipid), vitamin and cofactor (Vitamin B6, purine), energy (pyruvate and gluconeogenesis) and xenobiotics (xylene, aromatic compound degradation). 24 of all pathways identified are associated with 1y O_3 exposure. Overlapping features between exposure and outcomes include porphyrin metabolism associated with 1-m $PM_{2.5}$, 1-y O_3 exposure, and PE; propanoate metabolism associated 1-y and 1-tri NO_2 exposure and PE, degradation of aromatic compounds associated with all 4 exposure windows of NO_2 and gHTN. While not associated with any outcome, there were several common pathways associated with multiple pollutant species, or exposure windows including tryptophan metabolism common to 1y $PM_{2.5}$ exposure and 1tri NO_2 exposure, and glycine and serine metabolism common to 1y $PM_{2.5}$. and 1y O_3 exposure. Details of such pathways have been summarized in Table S6.

Discussion

In this study, we employed a parallel strategy of MITM and HDMA to investigate the metabolomics perturbation in the maternal blood serum associated with ambient air pollution exposure and HDP incidence in the Atlanta African American Maternal Child cohort. We also identified several metabolomics perturbations unique to each pollutant exposure and

health endpoint, related to systemic inflammation and oxidative stress. Our findings suggest a potentially critical role of various aromatic compounds, indicative of oxidative stress, in the pathophysiology underlying the association between air pollution and HDP.

Oxidative Stress as Underlying Pathophysiology

Phthalate and Oxidative Stress

Phthalates are a group of plasticizers used in hundreds of products and have very large consumption around the world. As of 2016, Di(2-ethylhexyl) phthalate (DEHP) was the most frequently used plasticizer world[30]. PMs, on the other hand, are a complex mixture of tiny solid and liquid particles composed of a variety of materials such as dust, soot, smoke, organic compounds, and metals. The presence of phthalates in $PM_{2.5}$ may be due to their use in consumer products such as plastics, cosmetics, and personal care products, which can release phthalates into the air as a result of product use or disposal.[31,32] The metabolism of phthalate in humans mainly takes place in the liver by various enzymes including the Cytochrome P450, and is eventually excreted from the urine or bile after enterohepatic circulation[33].

A previous MWAS study from the same cohort has confirmed metabolomics perturbations following phthalate exposure[34]. Specifically, the perturbed metabolites are tyramine, phenethylamine, and bilirubin, each corresponding to potential adverse effects including pre-term birth, oxidative stress, and neurotoxicity. Perturbed pathways include tyrosine

metabolism and porphyrin metabolism. In our analysis, bilirubin is enriched in porphyrin metabolism for both exposure and outcome models. Bilirubin is a yellow-orange pigment that is formed as a byproduct of the breakdown of hemoglobin in red blood cells. It is produced mainly in the liver and is excreted from the body through bile. In the normative aging study (NAS) focusing entirely on the elderly Caucasian population, bilirubin, and porphyrin metabolism pathway were found to be significantly associated with long-term (annual) $PM_{2.5}$ exposure[46]. While the population and exposure windows are different, our findings confirmed the association between $PM_{2.5}$ exposure and these metabolites and pathways.

While our analysis didn't directly identify bilirubin in any models, we did discover biliverdin, a green tetrapyrrole pigment right upstream of bilirubin in the hemoglobin metabolism pathway to be associated with gHTN. Biliverdin is also enriched in the porphyrin metabolism pathway for 1m $PM_{2.5}$ exposure. Comparing our results with previous studies published based on the cohort on different exposures, we found that bilirubin is significantly associated with Bisphenol A (BPA), another type of plasticizer, while biliverdin is significantly associated with nicotine exposure[18,19]. Phthalate, BPA, and tobacco are all xenobiotics with aromatic rings in their structure and are metabolized through liver cytochrome P450. A systemic review has confirmed that exposure to $PM_{2.5}$ is associated with increased enzymatic activity, a sign of liver damage[40]. The oxidation potential of PM and O_3 has been confirmed in several previous studies.[58,59] Bilirubin and biliverdin have several important roles in the body, including serving as antioxidants and markers of liver function[35,36]. Therefore, it's reasonable

to hypothesize the association of bilirubin and biliverdin with environmental exposures may indicate maternal hepatotoxicity and liver damage from excessive oxidative stress when metabolizing these aromatic compounds.

NO_2 and oxidative stress

While NO_2 exposure is not significantly associated with phthalate or porphyrin metabolism, all 4 exposure windows and gHTN incidence are significantly (raw $p<0.05$) associated with the degradation of aromatic compounds. 3 short-term exposure windows are all associated ($p<0.05$) with xylene degradation, among which 1tri and 1w exposure is significant even at $q<0.2$. Metabolites enriched in the aromatic compound metabolism pathway include benzaldehyde, acetophenone, xylene, and ethylbenzene. Previous studies have suggested exposure to NO_2 can affect the metabolism of aromatic compounds in the body through oxidative stress and changes in gene expression[39]. While Cytochrome P450 is mainly responsible for the metabolism of aromatic compounds, a significant portion of such aromatic compounds is metabolized by microorganisms, and other studies have also detected the metabolite from microorganism metabolism in human bodies.[41] Our results shed light on the interaction of host and microbiomes in xenobiotics metabolism, and a more flexible, interdisciplinary approach that combines blood metabolomics with, gut metabolomics or even microbiomics has the potential to generate more insightful results.

Another piece of evidence we observed in support of NO_2's contribution to oxidative stress was the identification of several LysoPCs positively associated with 1 long-term and 2 short-term NO_2 exposure windows. LysoPC is a type of phospholipid that is found in cell membranes and plays a role in various physiological processes. One of its main functions is as a signaling molecule that regulates inflammation and immune responses in the body. An increase in LysoPC is positively associated with an increase in cardiovascular and neurodegenerative diseases, including hypertension.[53,54]

gHTN and PE pathophysiology

Molecules for Maternal Morbidity

While gHTN itself is not directly associated with increased risk for maternal and fetal complications, gHTN is associated with increased risk for PE, and PE is associated with various complications like hemolysis, elevated liver enzymes, low platelet count (HELLP) syndrome, fetal growth restriction (FGR).[37,56] While the pathophysiology of pre-eclampsia is not entirely clear yet, commonly considered factors include abnormal development and remodeling of the placenta, oxidative stress, and inflammation which are all associated with air pollution [37,56,57] Inadequate trophoblast extravillous trophoblast invasion and spiral artery remodeling reduce blood flow to the placenta, resulting in placental hypoxia and placental ischemia, and reperfusion injury, causing syncytiotrophoblast stress. [37,56] The same reason also gives rise to FGR, explaining why the 2 conditions mostly occur concurrently. Abnormal placental development leads to an imbalance in the production of certain proteins and

hormones, such as angiogenic factors, which can cause damage to the endothelial cells that line the blood vessels in the mother's body.

Previous studies have considered gHTN as a subclinical stage of PE from angiogenesis-related factors[38], and our results discovered similar results from a metabolite perspective. While There're no overlapping pathways or metabolites discovered between gHTN and PE, 2 metabolites identified as associated with gHTN, phenylalanine, and indole, both have an aromatic ring in their structure. Indole is formed from tryptophan metabolism, with kynurenine being another product. Kynurenine was found to be significantly positively associated with PE compared to gHTN in a previous study on the same cohort but with a smaller sample size.[39] In our analysis, we found indole to be positively associated with gHTN compared to those without HDP. Both tryptophan and phenylalanine are associated with cardiovascular disease risks. Tryptophan is known to exert regulatory effects on the development of atherosclerosis while plasma phenylalanine is positively related to hypertension. [51, 52]

Molecules for Child Morbidity

Amino acids play an important role in fetus growth and development. Most directly, the inability to metabolize phenylalanine is an inborn error of metabolism called phenylketonuria (PKU), causing intellectual disability and other serious health problems.[55] Amino acid imbalance during gestation may produce long-term morphological or functional changes in

offspring, such as an increased risk of developing hypertension in the later life of the child, known as the developmental origins of hypertension.[43,6,7] Perturbations of tryptophan metabolism are further associated with chronic kidney disease (CKD) in later life[44]. From here, we can see the double burden of aromatic compound metabolism in maternal and child health—both on the mother herself as well as the child. In our result, both short-term and long-term air pollution are associated with perturbations in various amino acid pathways, including tryptophan metabolism, tyrosine metabolism, aspartate, and asparagine metabolism, alanine and aspartate metabolism, many of which are aromatic amino acid pathways.

In addition to DEHP, another metabolite that is significantly associated with PE in our analysis is retinoate. In our analysis, retinoate is negatively associated with PE, which is contradictory to the previous study suggesting a positive association between retinoate and PE[42]. Retinoate is involved in the regulation of cell growth and differentiation, particularly in the development of various tissues and organs such as the eyes, skin, and central nervous system [47]. Given that PE is associated with fetal growth restriction and retinoate is an important messenger for growth, the negative relationship makes more sense. Further research on the association between retinoate and PE is needed.

Inflammation as Underlying Pathophysiology

Propanoate metabolism is another overlapping pathway shared by 1y, 1tri NO_2 exposure, and PE. Propanoate is a microbial-derived short-chain fatty acid that plays important roles in

energy metabolism, immune regulation, appetite, and food intake. In a different cohort study on gut metabolome, they identified the role of propionate and short-chain fatty acid in reducing the risk of preeclampsia by suppressing inflammation[43]. The discovery of propanoate metabolism also highlights the importance of integrating metabolomics with microbiomics to further investigate the molecular interactions of environmental perturbations on human health.

Analytic Strategies for Mediation Analysis

The HDMA is much more conservative compared to MITM when identifying overlapping features, given that it generates much fewer overlapping features for each model. It's quite surprising to see the big discrepancies in the overlap of MITM and HDMA results for exposure and outcome models. While FDR-corrected q of the features identified by HDMA are all smaller than 0.1 most of the FDR-corrected q in their MITM model is greater than 0.2, making these findings nonsignificant. The discrepancies here warrant that we should not overtly rely on p values to differentiate significant from nonsignificant findings.

Strengths & Limitations

Our analysis has several unique strengths. To begin with, it is one of the first studies that use untargeted high-resolution metabolomics to investigate perturbations associated with both environmental exposure and HDP. Our exposure assessment is conducted using a well-validated ensemble spatiotemporally resolved model and includes 3 pollutant species.

Additionally, we used an innovative parallel analysis strategy, leveraging 2 different mediation analysis approaches at the same time, and both discover phthalate as overlapping metabolite, making the result much more robust. To our knowledge, this is also the first study to discover porphyrin and aromatic compound metabolism a potentially important but underestimated pathway explaining the contribution of ambient pollution to the pathophysiology of HDP. What's more, we focused on African Americans, a population that has been historically underrepresented in biomedical and epidemiological research. Our cohort is one of the largest cohorts nowadays focusing on the reproductive outcome of African American populations and our results shed light on the persistent health disparities in the risk of HDP in the U.S. Finally, the workflow of the untargeted metabolomics profiling was well-established and has been shown to successfully analyze many non-fasting samples previously.[18,45]

This study also has several limitations. Firstly, while the total sample size is moderate, the sample size of those with PE is still relatively small (n= 26). This may give rise to relatively high variability of the results. Secondly, even though we're able to confirm phthalate as an overlapping metabolite, this result is not significant at FDR $q< 0.2$, suggesting the robustness of such results needs further validation. Given the aim of the study is hypothesis-generating, we lowered the threshold to include more positive results. However, we were only able to annotate a subset of all features identified, and that may leave out potentially important molecules. The association also does not imply a causal relationship given the cross-sectional

study design. Thirdly, similar to most air pollution epidemiological studies, ambient air pollution exposure was estimated based on the residential address of pregnant people, which did not consider daily mobility patterns and could be subject to exposure misclassification. Lastly, given this study specifically focuses on African American populations, the results might not be applicable to a more generalized population that is ethnically and socio-economically more heterogeneous.

This is a lot of variability and uncertainty in the composition of $PM_{2.5}$, making research on its health effects difficult. The study warrants further research to study the impact of ambient air pollutant species, particularly PM by breaking down their specific compositions. While Perfluoroalkyl Substances (PFAS) have long been known as one of the most ubiquitous substances in the environment and have been related to various health risks[61], the repetitive appearance of phthalate in our results suggests phthalate might become “the next PFAS” and further research on phthalate is urgently needed. At the same time, while there’s not a consensus yet on the way to perform high-dimensional mediation analysis as well as multi-omics cross talk, methodological innovations to come up with methods that are more flexible and robust would be greatly appreciated.

Conclusion

Our study confirms the great potential of untargeted high-resolution metabolomics in unraveling the mechanistic pathways linking environmental exposures and maternal and child health outcomes. The discoveries of various aromatic compounds and metabolic pathways suggest that it is crucial to consider pollutant compositions when studying the toxicity of air pollutants. The findings also suggest a potentially critical role of aromatic compounds indicative of oxidative stress, in the pathophysiology underlying the impact of air pollution on HDP. Further hypothesis testing research is warranted to replicate and validate these findings, using methods like the mixture analysis approach and multi-pollutant models.

Reference:

1. Jiang L, Tang K, Magee LA, et al. A global view of hypertensive disorders and diabetes mellitus during pregnancy. Nature Reviews Endocrinology 2022; 18(12): 760-75.
2. Metoki H, Iwama N, Hamada H, et al. Hypertensive disorders of pregnancy: definition, management, and out-of-office blood pressure measurement. Hypertension Research 2022; 45(8): 1298-309.
3. Boogaard H, Walker K, Cohen AJ. Air pollution: the emergence of a major global health risk factor. International Health 2019; 11(6): 417-21.
4. Manisalidis I, Stavropoulou E, Stavropoulos A, Bezirtzoglou E. Environmental and Health Impacts of Air Pollution: A Review. Frontiers in Public Health 2020; 8.
5. Li Z, Tang Y, Song X, Lazar L, Li Z, Zhao J. Impact of ambient PM2.5 on adverse birth outcome and potential molecular mechanism. Ecotoxicol Environ Saf. Mar 2019;169:248-254. doi:10.1016/j.ecoenv.2018.10.109
6. Hales CN, Barker DJ, Clark PM, et al. Fetal and infant growth and impaired glucose tolerance at age 64. Bmj 1991; 303(6809): 1019-22.
7. Barker DJ. The origins of the developmental origins theory. J Intern Med 2007; 261(5): 412-7.
8. Pedersen M, Stayner L, Slama R, et al. Ambient Air Pollution and Pregnancy-Induced Hypertensive Disorders. Hypertension 2014; 64(3): 494-500.

9. Nobles CJ, Williams A, Ouidir M, Sherman S, Mendola P. Differential Effect of Ambient Air Pollution Exposure on Risk of Gestational Hypertension and Preeclampsia. Hypertension 2019; 74(2): 384-90.
10. Yan Q, Liew Z, Uppal K, et al. Maternal serum metabolome and traffic-related air pollution exposure in pregnancy. Environ Int. Sep 2019;130:104872. doi:10.1016/j.envint.2019.05.066
11. Liang D, Li Z, Vlaanderen J, Tang Z, Vermeulen R, Sarnat J. Systematic Review on Untargeted Metabolomics Application in Air Pollution Health Research: Current Progress, Analytical Challenges, and Future Direction. 2020:
12. Liang D, Moutinho JL, Golan R, et al. Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution. Environ Int. Nov 2018;120:145-154. doi:10.1016/j.envint.2018.07.044
13. Mayrink J, Leite DF, Nobrega GM, Costa ML, Cecatti JG. Prediction of pregnancy-related hypertensive disorders using metabolomics: a systematic review. BMJ Open 2022; 12(4): e054697.
14. Perry MJ, Arrington S, Freisthler MS, et al. Pervasive structural racism in environmental epidemiology. Environmental Health 2021; 20(1): 119.
15. Corwin EJ, Hogue CJ, Pearce B, et al. Protocol for the Emory University African American Vaginal, Oral, and Gut Microbiome in Pregnancy Cohort Study. BMC Pregnancy Childbirth. Jun 1 2017;17(1):161. doi:10.1186/s12884-017-1357-x
16. Brennan PA, Dunlop AL, Smith AK, Kramer M, Mulle J, Corwin EJ. Protocol for the Emory University African American maternal stress and infant gut microbiome cohort study. BMC Pediatr. Jul 22 2019;19(1):246. doi:10.1186/s12887-019-1630-4
17. Tan Y, Barr DB, Ryan PB, et al. High-resolution metabolomics of exposure to tobacco smoke during pregnancy and adverse birth outcomes in the Atlanta African American maternal-child cohort. Environ Pollut. Jan 1 2022;292(Pt A):118361. doi:10.1016/j.envpol.2021.118361
18. Gaskins AJ, Tang Z, Hood RB, et al. Periconception air pollution, metabolomic biomarkers, and fertility among women undergoing assisted reproduction. Environ Int. Oct 2021;155:106666. doi:10.1016/j.envint.2021.106666
19. Tchen R, Tan Y, Boyd Barr D, et al. Use of high-resolution metabolomics to assess the biological perturbations associated with maternal exposure to Bisphenol A and Bisphenol F among pregnant African American women. Environment International 2022; 169: 107530.
20. Di Q, Amini H, Shi L, et al. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. Environ Int. Sep 2019;130:104909. doi:10.1016/j.envint.2019.104909
21. Chang CJ, Barr DB, Ryan PB, et al. Per- and polyfluoroalkyl substance (PFAS) exposure, maternal metabolomic perturbation, and fetal growth in African American women: A meet-in-the-middle approach. Environ Int. Jan 2022;158:106964. doi:10.1016/j.envint.2021.106964

22. Yu T, Park Y, Johnson JM, Jones DP. apLCMS--adaptive processing of high-resolution LC/MS data. Bioinformatics. Aug 1 2009;25(15):1930-6. doi:10.1093/bioinformatics/btp291
23. Uppal K, Soltow QA, Strobel FH, et al. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. BMC Bioinformatics. Jan 16 2013;14:15. doi:10.1186/1471-2105-14-15
24. Mustillo S, Kwon S. Auxiliary Variables in Multiple Imputation When Data Are Missing Not at Random. The Journal of Mathematical Sociology 2015; 39(2): 73-91.
25. Wei R, Wang J, Su M, et al. Missing Value Imputation Approach for Mass Spectrometry-based Metabolomics Data. Scientific Reports 2018; 8(1): 663.
26. Chadeau-Hyam M, Athersuch TJ, Keun HC, et al. Meeting-in-the-middle using metabolic profiling - a strategy for the identification of intermediate biomarkers in cohort studies. Biomarkers 2011; 16(1): 83-8.
27. Perera C, Zhang H, Zheng Y, et al. HIMA2: high-dimensional mediation analysis and its application in epigenome-wide DNA methylation data. BMC Bioinformatics 2022; 23(1): 296.
28. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behavior Research Methods 2008; 40(3): 879-91.
29. Tian L, Li Z, Ma G, et al. Metapone: a Bioconductor package for joint pathway testing for untargeted metabolomics data. Bioinformatics 2022; 38(14): 3662-4.
30. Benjamin S, Masai E, Kamimura N, Takahashi K, Anderson RC, Faisal PA. Phthalates impact human health: Epidemiological evidences and plausible mechanism of action. J Hazard Mater 2017; 340: 360-83.
31. Szewczyńska M, Dobrzyńska E, Pośniak M. Determination of phthalates in particulate matter and gaseous phase emitted in indoor air of offices. Environ Sci Pollut Res Int 2021; 28(42): 59319-27.
32. Li X, Zhang W, Lv J, et al. Distribution, source apportionment, and health risk assessment of phthalate esters in indoor dust samples across China. Environmental Sciences Europe 2021; 33(1): 19.
33. Radke EG, Galizia A, Thayer KA, Cooper GS. Phthalate exposure and metabolic effects: a systematic review of the human epidemiological evidence. Environment International 2019; 132: 104768.
34. Zhang X, Barr DB, Dunlop AL, et al. Assessment of metabolic perturbations associated with exposure to phthalates among pregnant African American women. Science of The Total Environment 2022; 818: 151689.
35. Chowdhury NR, Li Y, Chowdhury JR. Disorders of Bilirubin Metabolism. The Liver; 2020: 229-44.
36. Ziberna L, Martelanc M, Franko M, Passamonti S. Bilirubin is an Endogenous Antioxidant in Human Vascular Endothelial Cells. Scientific Reports 2016; 6(1): 29240.
37. Dimitriadis E, Rolnik DL, Zhou W, et al. Pre-eclampsia. Nature Reviews Disease Primers 2023; 9(1): 8.

38. Hirashima C, Ohkuchi A, Takahashi K, et al. Gestational hypertension as a subclinical preeclampsia in view of serum levels of angiogenesis-related factors. Hypertension Research 2011; 34(2): 212-7.
39. Kim HJ, Min JY, Seo YS, Min KB. Association of Ambient Air Pollution with Increased Liver Enzymes in Korean Adults. Int J Environ Res Public Health 2019; 16(7).
40. Pan L, Sui J, Xu Y, et al. Effect of Fine Particulate Matter Exposure on Liver Enzymes: A Systematic Review and Meta-Analysis. International Journal of Environmental Research and Public Health 2023; 20(4): 2803.
41. Diallo AF, Lockwood MB, Maki KA, et al. Metabolic Profiling of Blood and Urine for Exploring the Functional Role of the Microbiota in Human Health. Biol Res Nurs 2020; 22(4): 449-57.
42. Ferranti EP, Frediani JK, Mitchell R, et al. Early Pregnancy Serum Metabolite Profiles Associated with Hypertensive Disorders of Pregnancy in African American Women: A Pilot Study. Journal of Pregnancy 2020; 2020: 1515321.
43. Hsu CN, Tain YL. Amino Acids and Developmental Origins of Hypertension. Nutrients 2020; 12(6).
44. Hsu CN, Tain YL. Developmental Programming and Reprogramming of Hypertension and Kidney Disease: Impact of Tryptophan Metabolism. Int J Mol Sci 2020; 21(22).
45. Jin J, Gao L, Zou X, et al. Gut Dysbiosis Promotes Preeclampsia by Regulating Macrophages and Trophoblasts. Circ Res 2022; 131(6): 492-506.
46. Nassan FL, Kelly RS, Kosheleva A, et al. Metabolomic signatures of the long-term exposure to air pollution and temperature. Environ Health 2021; 20(1): 3.
47. Das BC, Thapa P, Karki R, et al. Retinoic acid signaling pathways in development and diseases. Bioorg Med Chem 2014; 22(2): 673-83.
48. Go YM, Walker DI, Liang Y, et al. Reference Standardization for Mass Spectrometry and High-resolution Metabolomics Applications to Exposome Research. Toxicol Sci 2015; 148(2): 531-43.
49. Blanchard O, Deguen S, Kihal-Talantikite W, François R, Zmirou-Navier D. Does residential mobility during pregnancy induce exposure misclassification for air pollution? Environ Health 2018; 17(1): 72.
50. Singh U, Hernandez KM, Aronow BJ, Wurtele ES. African Americans and European Americans exhibit distinct gene expression patterns across tissues and tumors associated with immunologic functions and environmental exposures. Scientific Reports 2021; 11(1): 9905.
51. Teymoori F, Asghari G, Mirmiran P, Azizi F. High dietary intake of aromatic amino acids increases risk of hypertension. Journal of the American Society of Hypertension 2018; 12(1): 25-33.
52. Nitz K, Lacy M, Atzler D. Amino Acids and Their Metabolism in Atherosclerosis. Arterioscler Thromb Vasc Biol 2019; 39(3): 319-30.
53. Law SH, Chan ML, Marathe GK, Parveen F, Chen CH, Ke LY. An Updated Review of Lysophosphatidylcholine Metabolism in Human Diseases. Int J Mol Sci 2019; 20(5).
54. Kim M, Jung S, Kim SY, Lee SH, Lee JH. Prehypertension-associated elevation in

circulating lysophosphatidlycholines, Lp-PLA2 activity, and oxidative stress. PLoS One 2014; 9(5): e96735.
55. Al Hafid N, Christodoulou J. Phenylketonuria: a review of current and future treatments. Transl Pediatr 2015; 4(4): 304-17.
56. Burton GJ, Redman CW, Roberts JM, Moffett A. Pre-eclampsia: pathophysiology and clinical implications. BMJ 2019; 366: l2381.
57. Saenen ND, Martens DS, Neven KY, et al. Air pollution-induced placental alterations: an interplay of oxidative stress, epigenetics, and the aging phenotype? Clinical Epigenetics 2019; 11(1): 124.
58. Tang Z, Sarnat JA, Weber RJ, et al. The Oxidative Potential of Fine Particulate Matter and Biological Perturbations in Human Plasma and Saliva Metabolome. Environmental Science & Technology 2022; 56(11): 7350-61.
59. Miller DB, Karoly ED, Jones JC, et al. Inhaled ozone (O3)-induces changes in serum metabolomic and liver transcriptomic profiles in rats. Toxicol Appl Pharmacol 2015; 286(2): 65-79.
60. Mitku AA, Zewotir T, North D, et al. Impact of ambient air pollution exposure during pregnancy on adverse birth outcomes: generalized structural equation modeling approach. BMC Public Health 2023; 23(1): 45.
61. Bommarito PA, Ferguson KK, Meeker JD, McElrath TF, Cantonwine DE. Maternal Levels of Perfluoroalkyl Substances (PFAS) during Early Pregnancy in Relation to Preeclampsia Subtypes and Biomarkers of Preeclampsia Risk. Environmental Health Perspectives 2021; 129(10): 107004.

Tables and Figures

Table 1. Selected population characteristics by birth outcomes among the subjects enrolled in Atlanta African American Maternal-Child Cohort study, 2014-2018 (N = 329).

	No (N=261)	**gHTN (N=42)**	**Preeclampsia (N=26)**
Age			
Mean (SD)	25.2 (4.74)	23.5 (4.90)	24.6 (4.70)
Median [Min, Max]	24.0 [18.0, 39.0]	22.0 [18.0, 40.0]	24.0 [18.0, 33.0]
Highest education			
Less than high school	39 (14.9%)	10 (23.8%)	5 (19.2%)
High school	91 (34.9%)	18 (42.9%)	14 (53.8%)
Some college or more	131 (50.2%)	14 (33.3%)	7 (26.9%)
Sex			
Male	132 (50.6%)	19 (45.2%)	13 (50.0%)
Female	129 (49.4%)	23 (54.8%)	13 (50.0%)
Season of conception			
fall	62 (23.8%)	5 (11.9%)	5 (19.2%)
spring	69 (26.4%)	10 (23.8%)	4 (15.4%)
summer	78 (29.9%)	18 (42.9%)	10 (38.5%)
winter	52 (19.9%)	9 (21.4%)	7 (26.9%)
Parity			
Nulliparity	109 (41.8%)	22 (52.4%)	17 (65.4%)
Primiparity	74 (28.4%)	10 (23.8%)	6 (23.1%)
Multiparity	78 (29.9%)	10 (23.8%)	3 (11.5%)
Tobacco or Marijuana Use			
No	156 (59.8%)	20 (47.6%)	12 (46.2%)
Yes	105 (40.2%)	22 (52.4%)	14 (53.8%)
Alcohol use			
No	240 (92.0%)	35 (83.3%)	21 (80.8%)
Yes	21 (8.0%)	7 (16.7%)	5 (19.2%)
BMI			
Mean (SD)	27.7 (7.15)	31.2 (7.45)	33.0 (10.2)
Median [Min, Max]	25.6 [17.1, 54.1]	30.6 [20.5, 46.9]	32.1 [18.7, 51.4]
GA at sampling			
Mean (SD)	11.6 (2.18)	11.0 (2.21)	10.5 (1.99)
Median [Min, Max]	11.9 [6.00, 17.9]	10.9 [7.43, 17.0]	10.5 [6.86, 14.9]

Table 2. Chemical identity of the metabolites significantly associated with air pollution exposure or HDP outcomes (raw p<0.05).

mz	Time	Compound Name	Preferred Adduct	Model
HILIC Column				
137.0458	40.7	HYPOXANTHINE	M+H	pm25_1y
391.2842	22	BIS(2-ETHYLHEXYL)PHTHALATE	M+H	pm25_1tri, pm25_1m, PE
504.3067	28.6	LYSOPE(20:3)	M+H	pm25_1m
522.3558	31	LYSOPC(18:1)	M+H	no2_1y, no2_1tri, no2_1m
480.344	30.3	LYSOPC(O/P-16:1)	M+H	no2_1tri
524.3691	29.9	LYSOPC(18:0)	M+H	no2_1tri, no2_1m
149.0232	22.3	PHTHALIC ANHYDRIDE	M+H	no2_1m
468.3082	31.6	LYSOPC(14:0)	M+H	no2_1m
195.0877	28.9	CAFFEINE	M+H	o3_1y
284.2947	22.5	HEXADECANOL	M+ACN+H	o3_1tri
241.0312	186.9	L-CYSTINE/CYSTINE	M+H	o3_1tri
343.1241	95.3	SUCROSE	M+H, M-H2O+H	o3_1m
343.1241	95.3	MELIBIOSE	M+H	o3_1m
301.2162	79.5	RETINOATE	M+H	PE
166.0863	40.5	L-PHENYLALANINE	M+H	ghtn
148.0604	63.8	L-GLUTAMIC ACID	M+H	ghtn
148.0604	63.8	N-METHYL-D-ASPARTIC ACID	M+H	ghtn
148.0604	63.8	ACETYLSERINE	M+H	ghtn
118.0652	32.9	INDOLE	M+H	ghtn
C18 Column				
162.0196	26.6	ACETYLCYSTEINE	M-H, 2M-3H	pm25_1y
167.0348	194.8	VANILLIC ACID	M-H	pm25_1tri
167.0348	194.8	HYDROXYMANDELIC ACID	M-H	pm25_1tri
301.2387	24.2	RAC-GLYCEROL 1-MYRISTATE	M-H	pm25_1m
121.0295	20.3	BENZOATE	M-H	no2_1y
129.0558	21.1	KETOLEUCINE/KETOISOLEUCINE	M-H	no2_1w
581.2347	221	BILIVERDIN	M-H	ghtn

Table S1. Statistics of 3 pollutants ($PM_{2.5}$, NO_2, O_3) exposure for the four exposure windows among 329 pregnant people in the Atlanta African American Maternal-Child Cohort, 2014-2018.

Pollutant	Window	Q1	Q3	IQR	Median	Mean
$PM_{2.5}$*	1-yr	9.813	10.951	1.138	10.389	10.339
	1-tri	9.136	11.736	2.6	10.306	10.571
	1-w	8.185	12.037	3.852	9.936	10.37
	1-m	8.66	11.726	3.066	10.122	10.48
NO_2#	1-yr	20.56	27.215	6.655	24.318	23.776
	1-tri	18.754	26.712	7.958	23.558	22.774
	1-w	17.791	28.818	11.027	23.051	23.114
	1-m	18.635	27.338	8.703	23.357	22.668
O_3#	1-yr	36.92	38.44	1.52	37.59	37.79
	1-tri	36.12	45.89	9.77	42.66	40.42
	1-w	32.13	47.22	15.09	39.8	39.62
	1-m	32.94	45.8	12.86	41.31	39.85

*: The unit for $PM_{2.5}$ is

#: The unit for NO_2 and O_3 are ppb

Table S2. Number of statistically significant metabolic features associated with at least one pollutant exposure window or HDP outcome under different cutoff p-values, the Atlanta African American Maternal-Child cohort (2014-2016)

Exposure		HILIC+		C18	
	Exposure Window	FDR 0.2	RAW p<0.05	FDR 0.2	RAW p<0.05
$PM_{2.5}$	1-yr	2	929	0	555
	1-tri	6	891	0	573
	1-w	20	428	19	383
	1-m	1	570	0	562
NO_2	1-yr	0	573	0	480
	1-tri	0	659	0	512
	1-w	0	553	0	351
	1-m	0	647	0	563
O_3	1-yr	5	813	5	578
	1-tri	9	950	2	713
	1-w	0	549	0	477
	1-m	0	525	0	431
HDP	gHTN	0	756	0	413
	PE	0	605	0	561

Note: Both Benjamini-Hochberg false discovery rate (FDR) procedure and raw p-value were used to identify a reasonable number of significant metabolic features.

Abbreviations:

gHTN: gestational hypertension, PE: preeclampsia

[a] Those without hypertensive disorders of pregnancy are used as a reference group to compare with gHTN and PE.

Table S3. Number of statistically significant metabolic features associated with at least one pollutant exposure window and one HDP outcome under raw p-values <0.05 , the Atlanta African American Maternal-Child cohort (2014-2016)

Exposure	Exposure Window	HILIC		C18	
		gHTN	PE	gHTN	PE
$PM_{2.5}$	1-yr	54	53	27	39
	1-tri	50	53	31	40
	1-w	25	24	12	21
	1-m	27	33	13	37
NO_2	1-yr	56	26	20	39
	1-tri	62	26	24	25
	1-w	41	23	17	21
	1-m	53	26	24	27
O_3	1-m	59	38	27	48
	1-yr	52	52	35	64
	1-w	45	23	25	19
	1-m	41	31	21	19

Abbreviations:

gHTN: gestational hypertension, PE: preeclampsia

[a] Those without hypertensive disorders of pregnancy are used as a reference group to compare with gHTN and PE.

Table S4. Model statistics of metabolic features in the HILIC column identified by HDMA as mediating features between different exposure and outcome among 329 pregnant people in the Atlanta African American Maternal-Child Cohort, 2014-2018.[c]

mz	rt	Indirect Effect	BH.FDR	model
589.1548	87.9	-0.066	0.038	gHTN ~ pm25_1y
170.9852	55.8	-0.028	0.005	gHTN ~ pm25_1w
1069.7144	30.2	0.036	0.028	gHTN ~ pm25_1m
193.5697	83	-0.235	0.006	PE~ pm25_1y
456.1238	61.5	0.146	0.053	
600.1729	69.9	0.065	0.030	PE ~ pm25_1tri
175.96	77.8	0.138	0.015	
193.8944	111.5	0.110	0.068	
228.904	126.2	0.133	0.015	
302.1385	23.4	0.133	0.015	PE ~ pm25_1m
391.2842[b]	22	0.087	0.061	
415.357	21.1	-0.103	0.021	
741.9338	59.7	-0.163	0.026	
861.5489	32.6	0.008	0.033	gHTN ~ no2_1y
861.5489	32.6	0.000	0.079	gHTN ~ no2_1w"
861.5489	32.6	0.002	0.051	gHTN ~ no2_1m
91.9894[b]	37.5	-0.031	0.028	PE ~ no2_1tri
687.6077	46.5	0.029	0.028	
111.0918	238.5	0.002	0.029	PE ~ no2_1w
91.9894	37.5	-0.042	0.005	
161.0961	21.2	-0.041	0.005	PE ~ no2_1m
518.7669	51.9	0.017	0.045	
236.8852	116.6	-0.081	0.046	gHTN ~ o3_1y
606.316	23.1	0.087	0.046	
284.9951	186.6	-0.025	0.010	gHTN ~ o3_1tri
788.2795	46.4	0.018	0.020	
161.9053	43.4	0.003	0.004	gHTN ~ o3_1m
392.8422	92.2	-0.049	0.062	PE ~ o3_1y
415.357	21.1	-0.002	0.074	PE ~ o3_1m

Table S5. Model statistics of metabolic features in the C18 column identified by HDMA as mediating features between different exposure and outcome among 329 pregnant people in the Atlanta African American Maternal-Child Cohort, 2014-2018.[c]

mz	rt	Indirect Effect	BH.FDR	model
278.1593	23.4	0.027	0.039	gHTN ~ pm25_1tri
145.0698	176.9	-0.010	0.013	gHTN ~ pm25_1w
472.1447	201.1	0.045	0.008	gHTN ~ pm25_1m
708.3493	233.8	-0.040	0.008	
175.0989	282.5	-0.346	0.028	
414.2011	224.4	-0.452	0.009	
429.298	285.1	0.616	0.009	
447.1342	222.9	-0.593	0.011	PE ~ pm25_1y
455.2476[b]	23.2[a]	0.340	0.009	
461.3483	263.4	0.387	0.023	
573.3494[b]	41.2	-0.460	0.009	
557.4575[a]	269.3	0.018	0.004	PE~ pm25_1tri
721.9594	41.1	0.047	0.041	PE ~ pm25_1m
803.9812[b]	46.9	-0.070	0.033	
490.4301	30.6	0.002	0.034	gHTN ~ no2_1tri
786.8853	30.6	0.004	0.075	gHTN ~ no2_1w
490.4301	30.6	0.000	0.046	gHTN ~ no2_1m
397.0174[a]	20.6	-0.022	0.013	PE ~ no2_1y
572.6175[b]	23.1	-0.017	0.005	PE ~ no2_1w
432.1745[a]	199.7	0.079	0.034	gHTN ~ o3_1y
485.1568	209.7	0.003	0.063	gHTN ~ o3_1m
776.0182[b]	45.7	0.056	0.001	PE ~ o3_1tri
213.113	227.8	-0.004	0.174	PE ~ o3_1w
803.9812[b]	46.9	-0.003	0.019	PE ~ o3_1m

[a]: This feature is not identified by either the exposure-mediator model or mediator-outcome model in MITM

[b]: This feature is identified by the same mediator-outcome model in MITM

[c]: Unless otherwise specified, all features are identified by the same exposure-mediator model in MITM

Table S6. Metapone output of pathway analysis of Exposure-mediator, and Mediator-Outcome Models

Model	Pathway Name	p_value	significant metabolites	total metabolites	adjust.p
pm25_1y	glycine and serine metabolism	0.032	4.796	56	0.407
	amino sugar metabolism	0.029	3.509	46	0.407
	purine metabolism	0.002	8.945	108	0.284
	tryptophan metabolism	0.044	5.677	94	0.407
	vitamin b6 metabolism	0.009	3.498	30	0.407
pm25_1tri	tyrosine metabolism	0.038	6.163	133	0.647
pm25_1w	bile acid biosynthesis	0.006	3.773	79	0.204
pm25_1m	porphyrin metabolism	0.005	3.942	47	0.600
	abc transporters	0.049	4.410	73	0.636
	c21-steroid hormone biosynthesis and metabolism	0.007	5.192	77	0.600
no2_1y	degradation of aromatic compounds	0.047	4.338	93	0.450
	gluconeogenesis	0.005	3.385	35	0.181*
	propanoate metabolism	0.01	3.921	64	0.253
	pyruvate metabolism	0.031	3.207	59	0.386
	warburg effect	0.024	3.308	60	0.357
no2_1tri	2-oxocarboxylic acid metabolism	0.004	4.178	49	0.166*
	degradation of aromatic compounds	<0.001	7.473	93	<0.001*
	propanoate metabolism	0.035	3.565	64	0.443
	tryptophan metabolism	0.017	5.324	94	0.374
	valine, leucine, and isoleucine degradation	0.021	3.192	67	0.407
	xylene degradation	<0.001	3.667	22	<0.001*
no2_1w	degradation of aromatic compounds	0.019	4.589	93	0.842
	xylene degradation	0.001	3.180	22	0.100*
no2_1m	degradation of aromatic compounds	0.011	6.153	93	0.322
	pyruvate metabolism	0.04	3.420	59	0.624
	xylene degradation	0.003	3.471	22	0.276
	bile secretion	0.011	5.039	80	0.322
o3_1y	alanine and aspartate metabolism	0.005	4.692	48	0.127*
	alanine, aspartate and glutamate metabolism	0.001	3.169	25	0.101*
	glutamate metabolism	0.007	3.779	50	0.148*
	glycine and serine metabolism	0.002	5.959	56	0.101*
	glycine, serine and threonine metabolism	<0.001	5.476	41	<0.001*

	glycine, serine, alanine and threonine metabolism	0.013	6.780	86	0.219
	glyoxylate and dicarboxylate metabolism	0.004	4.104	45	0.127*
	amino sugar metabolism	0.032	3.307	46	0.225
	methionine and cysteine metabolism	0.049	5.768	81	0.258
	methionine metabolism	0.023	3.921	43	0.225
	phosphotransferase system (pts)	0.012	3.029	28	0.217
	porphyrin metabolism	0.003	3.982	47	0.127*
	protein digestion and absorption	0.009	4.269	43	0.175*
	pyrimidine metabolism	0.027	7.126	95	0.225
	abc transporters	0.027	5.315	73	0.225
	selenoamino acid metabolism	0.023	3.356	34	0.225
	urea cycle/amino group metabolism	0.029	4.779	71	0.225
	valine, leucine, and isoleucine degradation	0.03	3.220	67	0.225
	aspartate and asparagine metabolism	0.037	5.371	82	0.231
	biosynthesis of amino acids	0.035	6.366	92	0.227
	biosynthesis of antibiotics	0.041	9.480	207	0.231
	carbon metabolism	0.005	5.821	87	0.127*
	central carbon metabolism in cancer	0.029	3.164	36	0.225
	cyanoamino acid metabolism	0.026	3.058	29	0.225
o3_1w	sphingolipid metabolism	0.014	3.649	48	0.788
gHTN	degradation of aromatic compounds	<0.001	7.343	93	<0.001*
	carbon metabolism	0.008	4.936	87	<0.001*
	chloroalkane and chloroalkene degradation	0.011	3.130	25	0.298
PE	porphyrin and chlorophyll metabolism	0.005	3.278	40	0.322
	porphyrin metabolism	0.008	3.308	47	0.322
	propanoate metabolism	0.018	3.810	64	0.517

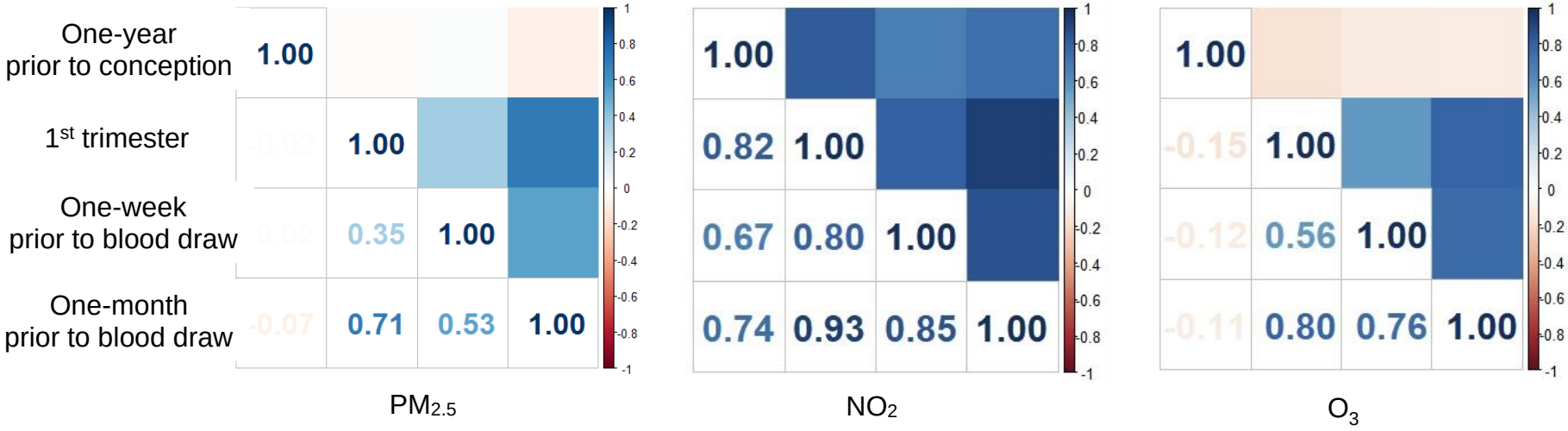
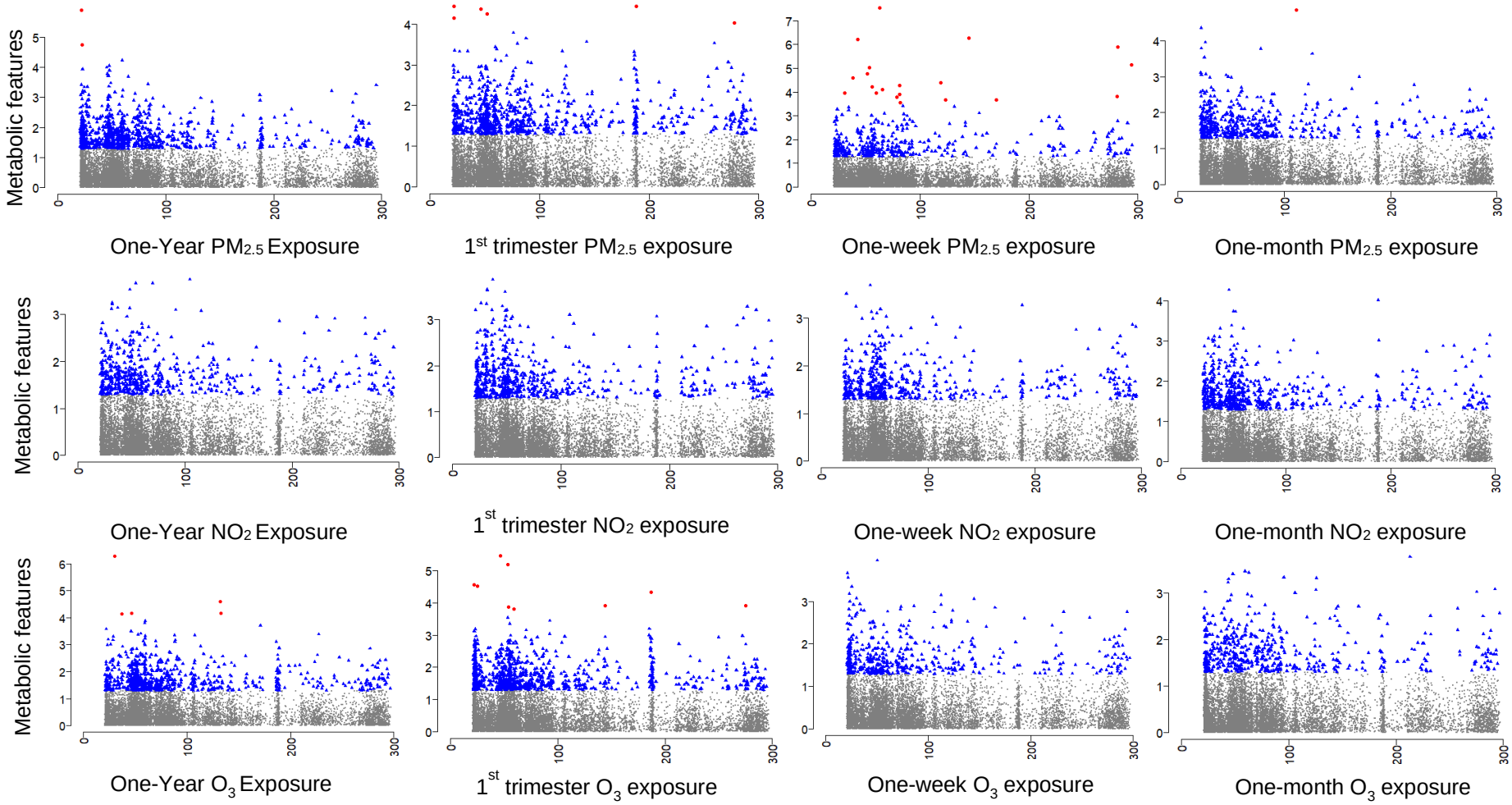


Figure S1. The correlations among different air pollutant exposures for the four exposure windows

A. Exposure-mediator metabolome-wide associations



B. Mediator-outcome metabolome-wide associations

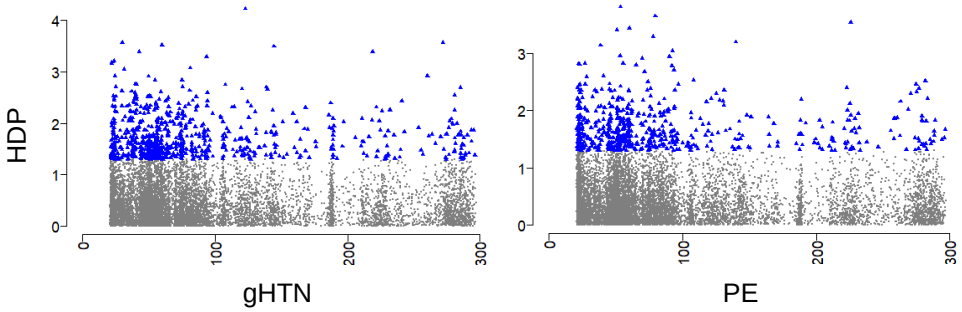
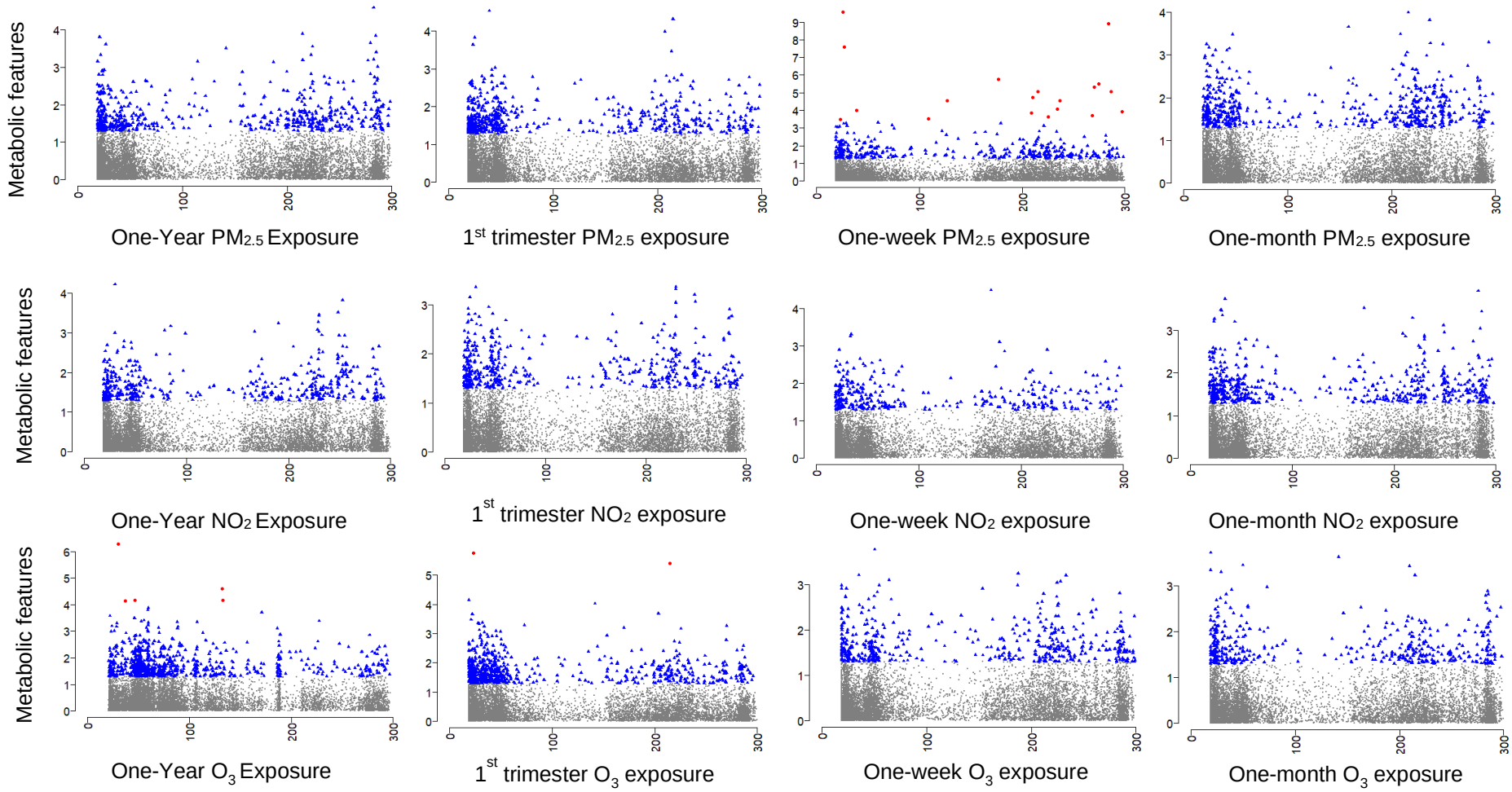


Figure S2. Manhattan plots of metabolome-wide association analysis in the HILIC column. A. Associations between air pollution exposures and changes in intensities of metabolic features; B. Associations between changes in intensities of metabolic features and gHTN or PE. X-axis denotes the retention time (in seconds) of the metabolic features, and Y-axis denotes the negative $\log_{10}$ of *p*-values. Red dots indicated significant associations at $FDR_{B\text{-}H} < 0.2$, and blue indicated associations at raw *p*-values < 0.05. Abbreviations: HILIC, hydrophilic interaction liquid chromatography; $PM_{2.5}$, fine particulate matter; NO_2, nitrogen dioxide, O_3, ozone, gHTN, gestational hypertension; PE, preeclampsia; $FDR_{B\text{-}H}$, Benjamini-Hochberg adjusted *p*-values.

A. Exposure-mediator metabolome-wide associations



B. Mediator-outcome metabolome-wide associations

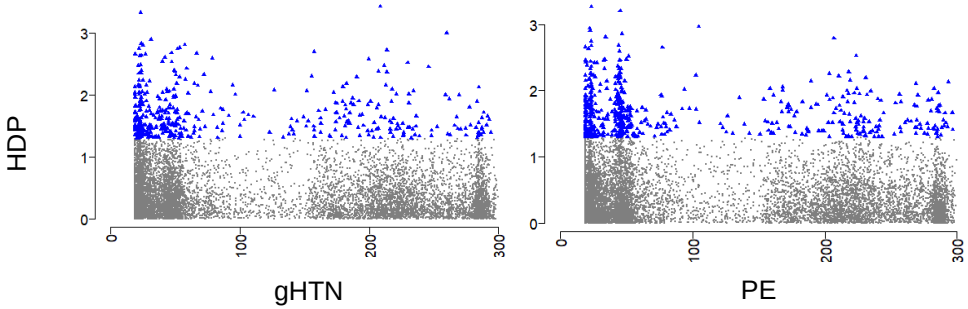
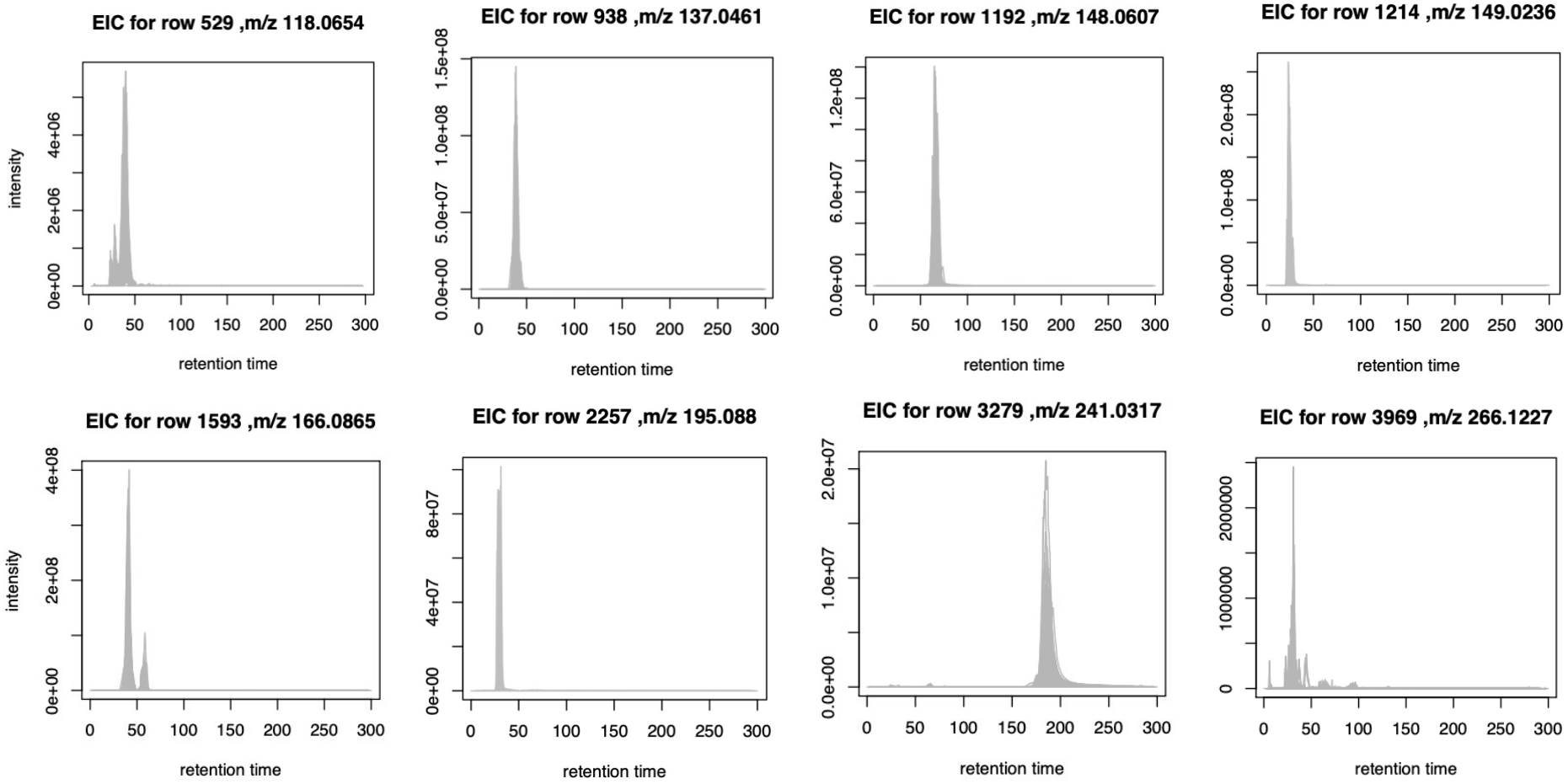
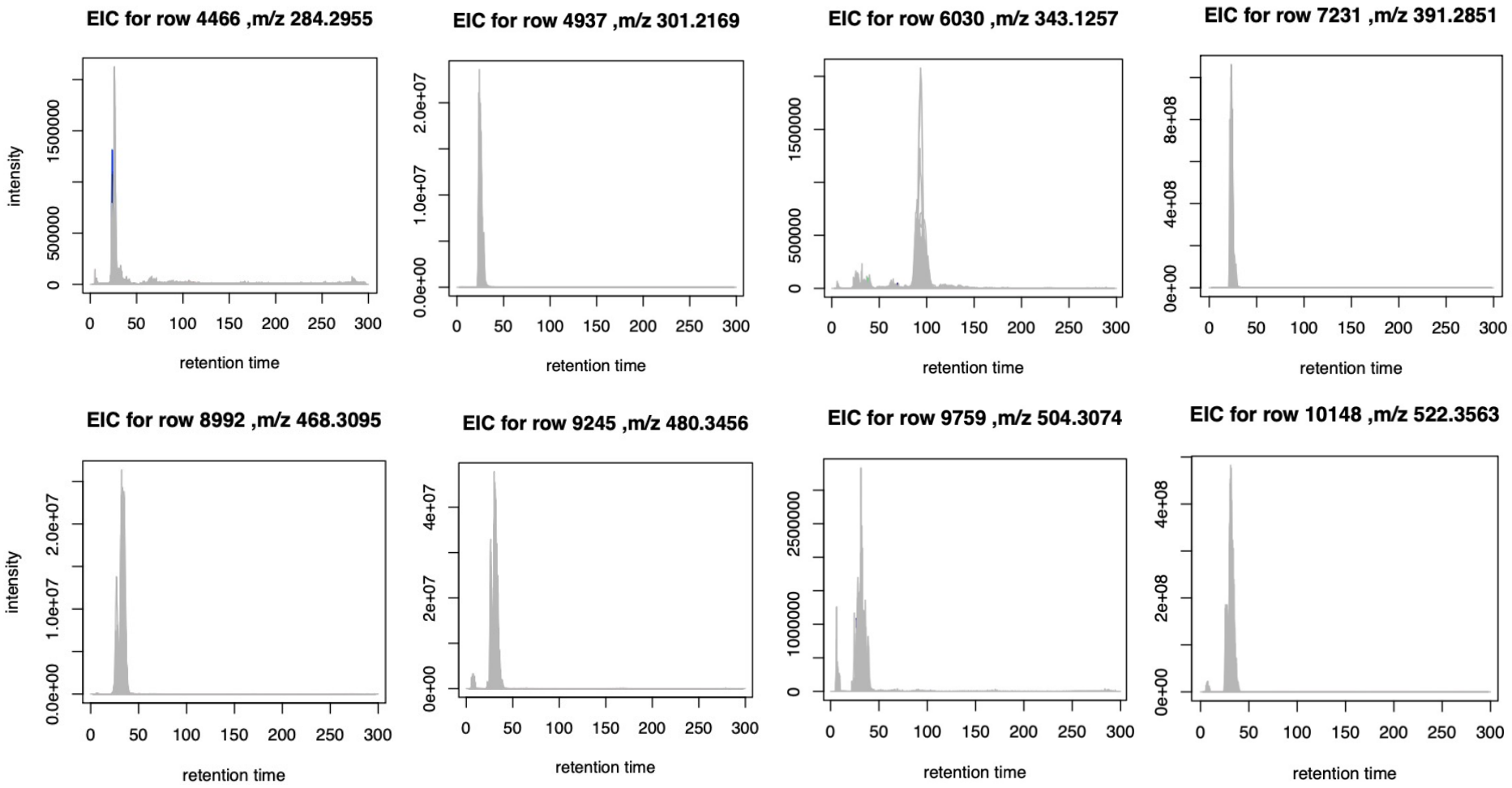


Figure S3. Manhattan plots of metabolome-wide association analysis in the C18 column. A. Associations between air pollution exposures and changes in intensities of metabolic features; B. Associations between changes in intensities of metabolic features and gHTN or PE. X-axis denotes the retention time (in seconds) of the metabolic features, and Y-axis denotes the negative $\log_{10}$ of *p*-values. Red dots indicated significant associations at $FDR_{B\text{-}H} < 0.2$, and blue indicated associations at raw *p*-values < 0.05. Abbreviations: C18, hydrophobic reversed-phase chromatography; $PM_{2.5}$, fine particulate matter; NO_2, nitrogen dioxide; O_3, Ozone; gHTN, gestational hypertension; PE, preeclampsia; $FDR_{B\text{-}H}$, Benjamini-Hochberg adjusted *p*-values


EIC for row 529 ,m/z 118.0654
EIC for row 938 ,m/z 137.0461
EIC for row 1192 ,m/z 148.0607
EIC for row 1214 ,m/z 149.0236
EIC for row 1593 ,m/z 166.0865
EIC for row 2257 ,m/z 195.088
EIC for row 3279 ,m/z 241.0317
EIC for row 3969 ,m/z 266.1227
intensity
retention time
0
50
100
150
200
250
300
0e+00
2e+06
4e+06
0.0e+00
5.0e+07
1.0e+08
1.5e+08
6.0e+07
1.2e+08
1.0e+08
2.0e+08
2e+08
4e+08
4e+07
8e+07
1.0e+07
2.0e+07
1000000
2000000


EIC for row 4466 ,m/z 284.2955
EIC for row 4937 ,m/z 301.2169
EIC for row 6030 ,m/z 343.1257
EIC for row 7231 ,m/z 391.2851
EIC for row 8992 ,m/z 468.3095
EIC for row 9245 ,m/z 480.3456
EIC for row 9759 ,m/z 504.3074
EIC for row 10148 ,m/z 522.3563
intensity
retention time

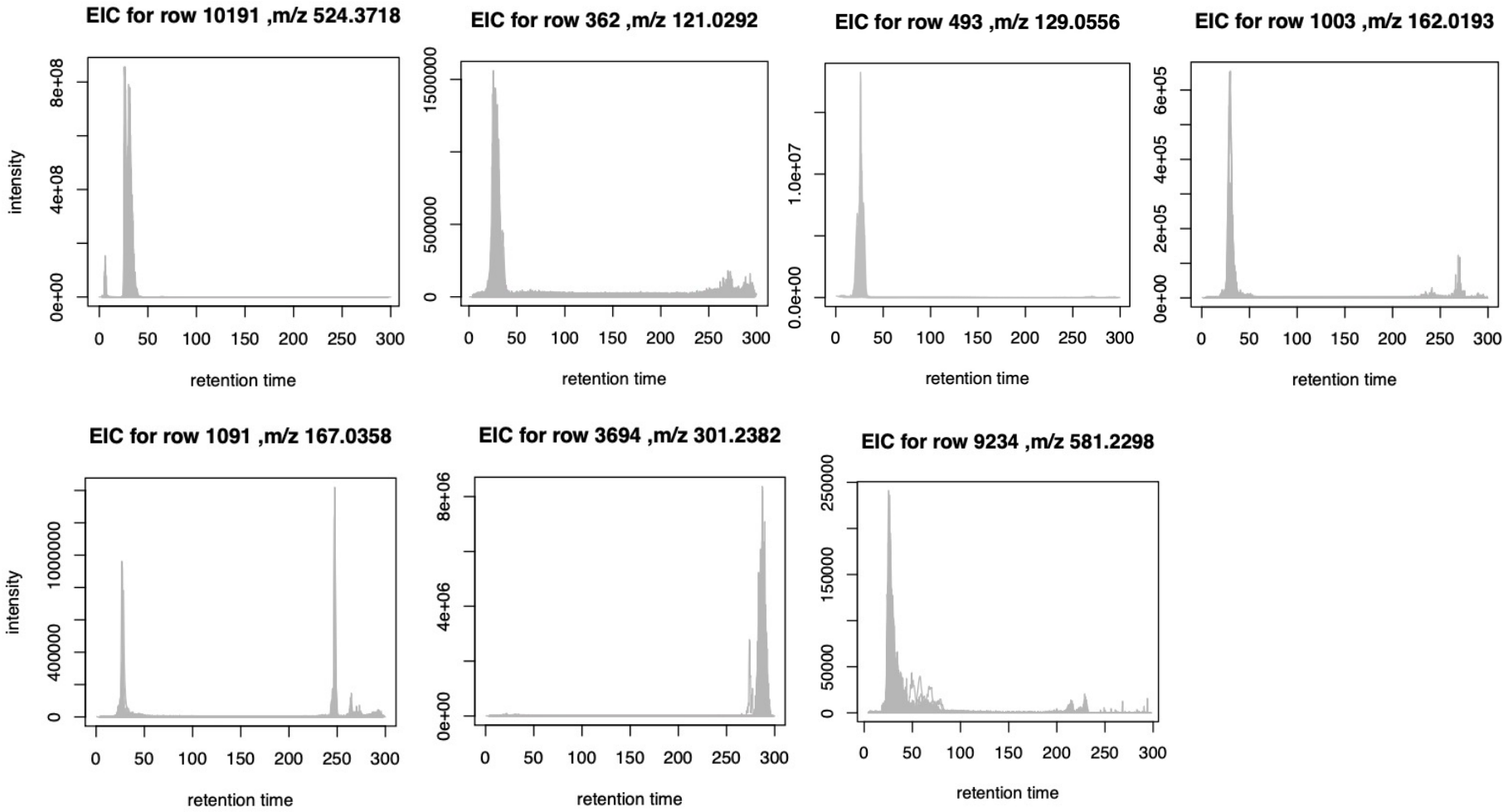


Figure S4. The extracted ion chromatograph of identified chemicals. The metabolites were considered to be acceptable for chemical identification that had one or multiple pure peaks.

Note: Pure peak refers to exhibiting clear gaussian peak shapes and signal-to-noise ratio above 3:1.