Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution

Donghai Liang, Emory University
Jennifer L. Moutinho, Georgia Institute of Technology
Rachel Golan, Ben Gurion University of the Negev
Tianwei Yu, Emory University
Chandresh N. Ladva, Emory University
Megan Niedzwiecki, Icahn School of Medicine at Mount Sinai
Douglas Walker, Emory University
Stefanie Ebelt Sarnat, Emory University
Howard Chang, Emory University
Roby Greenwald, Georgia State University

Only first 10 authors above; see publication for full author list.

Journal Title: Environment International
Volume: Volume 120
Publisher: Elsevier: Creative Commons Licenses | 2018-11-01, Pages 145-154
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1016/j.envint.2018.07.044
Permanent URL: https://pid.emory.edu/ark:/25593/tp0z2

Final published version: http://dx.doi.org/10.1016/j.envint.2018.07.044

Copyright information:
© 2018 Elsevier Ltd
This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Accessed August 11, 2019 8:56 PM EDT
Use of high-resolution metabolomics for the identification of metabolic signals associated with traffic-related air pollution

Donghai Liang\textsuperscript{a,}*, Jennifer L. Moutinho\textsuperscript{b}, Rachel Golan\textsuperscript{c}, Tianwei Yu\textsuperscript{d}, Chandresh N. Ladva\textsuperscript{e}, Megan Niedzwiecki\textsuperscript{f}, Douglas I. Walker\textsuperscript{a}, Stefanie Ebelt Sarnat\textsuperscript{a}, Howard H. Chang\textsuperscript{d}, Roby Greenwald\textsuperscript{h}, Dean P. Jones\textsuperscript{h}, Armistead G. Russell\textsuperscript{b}, Jeremy A. Sarnat\textsuperscript{g}

\textsuperscript{a} Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, USA
\textsuperscript{b} School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, USA
\textsuperscript{c} Department of Public Health, Ben-Gurion University of the Negev, Beer Sheva, Israel
\textsuperscript{d} Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, USA
\textsuperscript{e} Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, USA
\textsuperscript{f} Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, USA
\textsuperscript{g} Division of Environmental Health, Georgia State University School of Public Health, Atlanta, USA
\textsuperscript{h} Clinical Biomarkers Laboratory, Division of Pulmonary, Allergy, and Critical Care Medicine, School of Medicine, Emory University, Atlanta, USA

\begin{abstract}
\textbf{Background:} High-resolution metabolomics (HRM) is emerging as a sensitive tool for measuring environmental exposures and biological responses. The aim of this analysis is to assess the ability of high-resolution metabolomics (HRM) to reflect internal exposures to complex traffic-related air pollution mixtures.

\textbf{Methods:} We used untargeted HRM profiling to characterize plasma and saliva collected from participants in the Dorm Room Inhalation to Vehicle Emission (DRIVE) study to identify metabolic pathways associated with traffic emission exposures. We measured a suite of traffic-related pollutants at multiple ambient and indoor sites at varying distances from a major highway artery for 12 weeks in 2014. In parallel, 54 students living in dormitories near (20 m) or far (1.4 km) from the highway contributed plasma and saliva samples. Untargeted HRM profiling was completed for both plasma and saliva samples; metabolite and metabolic pathway alternations were evaluated using a metabolome-wide association study (MWAS) framework with pathway analyses.

\textbf{Results:} Weekly levels of traffic pollutants were significantly higher at the near dorm when compared to the far dorm (p < 0.05 for all pollutants). In total, 20,766 metabolic features were extracted from plasma samples and 29,013 from saliva samples. 45% of features were detected and shared in both plasma and saliva samples. 1291 unique metabolic features were significantly associated with at least one or more traffic indicator, including black carbon, carbon monoxide, nitrogen oxides and fine particulate matter (p < 0.05 for all significant features), after controlling for confounding and false discovery rate. Pathway analysis of metabolic features associated with traffic exposure indicated elicitation of inflammatory and oxidative stress related pathways, including leukotriene and vitamin E metabolism. We confirmed the chemical identities of 10 metabolites associated with traffic pollutants, including arginine, histidine, ω-linolenic acid, and hypoxanthine.

\textbf{Conclusions:} Using HRM, we identified and verified biological perturbations associated with primary traffic pollutant in panel-based setting with repeated measurement. Observed response was consistent with endogenous metabolic signaling related to oxidative stress, inflammation, and nucleic acid damage and repair. Collectively, the current findings provide support for the use of untargeted HRM in the development of metabolic biomarkers of traffic pollution exposure and response.
\end{abstract}

1. Introduction

Accurately assessing exposures to traffic-related air pollution (TRAP) is critical given the abundance of observational and controlled studies reporting associations between traffic sources and adverse health effects (Health Effects Institute, 2010). For primary traffic...
emission exposures, in particular, an added challenge lies in its chemical and physical heterogeneity, consisting of hundreds of different components. Beyond the task involving the external characterization of traffic pollution mixtures, measuring internal, biologically-relevant exposure and corresponding responses also remains challenging. In particular, the specific biological factors and pathways, as well as how they respond upon exposure of traffic pollution mixture remain poorly understood.

High resolution metabolomics (HRM) has emerged as a powerful platform to improve internal exposure estimation to complex environmental mixtures by providing identification and quantitation of thousands of metabolites in biological samples associated with endogenous and exogenous processes (Bundy et al., 2009; Lankadurai et al., 2013; Miller and Jones, 2014; Simpson and McKelvie, 2009; Viant, 2008). Although several analytical HRM platforms exist, we have utilized an approach based on liquid chromatography coupled with high-resolution mass spectrometry, which have the capability of identifying tens of thousands of metabolic features. We (Ladva et al., 2017), and others (Breitner et al., 2016; Martens et al., 2017; Menni et al., 2015; Surowiec et al., 2016; Vlaanderen et al., 2017), have begun to examine the capability of HRM to capture metabolic perturbations following exposures to ambient air pollutants. Currently, analytical and scientific uncertainties in HRM application, however, have limited its use for measuring the response to exposures from individual sources, such as TRAP (Hines et al., 2007; Morrison et al., 2007). It is unclear that whether HRM can identify specific metabolic indicators of primary traffic pollution with sufficient sensitivity and robust signals.

Panel-based designs have proven to be effective approaches of investigating TRAP health effects, given the ability to accurately measure both exposure and health parameters on an individual-level (Delfino et al., 2006; Delfino et al., 2008; McCreanor et al., 2007; Samat et al., 2012). However, the utility of this design for smaller environmental HRM applications is still unclear. Previous studies have either been conducted in cohorts numbering in the thousands, or in small panels exposed to extremely elevated concentrations of specific chemicals in occupational settings (Bundy et al., 2009; Morrison et al., 2007). From an analytical and statistical standpoint, therefore, questions remain concerning whether HRM sensitively captures statistically robust and biologically meaningful signals in panel-based environmental exposure protocols.

To address these research gaps and uncertainties, we conducted the Dorm Room Inhalation to Vehicle Emission (DRIVE) study, an intensive 12-week filed study that focused on assessing a complete emission-to-exposure pathway of primary traffic pollution. The present analysis utilized DRIVE study’s repeated biomonitoring (plasma and saliva) in a panel of 54 college students living in dormitories either near (20 m) or far (1.4 km) from the highway. We applied HRM to assess whether HRM is capable of detecting changes in specific metabolic profiles in plasma or saliva samples in relation to TRAP in panel-based setting. We then conducted comprehensive pathway analysis and chemical validation to identify specific metabolic patterns and to further investigate potentially biologically relevant indicators to primary traffic exposures. Finally, in examining the link between TRAPs and corresponding internal metabolic profiles, we aimed to identify metabolomics-based measurements that can be used to develop traffic exposure biomarkers. Such biomarkers may serve to inform more targeted regulations of traffic-related pollution with the ultimate goal of reducing the public health burden attributable to this air pollution source.

2. Methods

We conducted the DRIVE study to measure traditional and multi-pollutant traffic indicators along an emissions-to-dose exposure pathway. The centerpiece of this study was an intensive field sampling campaign that took place between September 8, 2014 and January 5, 2015. As part of this campaign, we measured a suite of traffic- and non-traffic-related pollutants at 6 outdoor and 2 indoor monitoring sites on the campus of the Georgia Institute of Technology (GIT) in Atlanta, GA, adjacent to the Downtown Connector (the Connector’), one of the most heavily trafficked highway arteries in the US (Fig. 1). While numerous smaller roadways surround the GIT campus, the Connector is the dominant mobile emissions source, with an annual average daily traffic...
(AADT) count of 305,365 vehicles during 2014, approximately 15 times that of the adjacent roads.

In addition to the field sampling, we recruited 54 GIT students living in one of two dormitories, located at different proximities from the highway, to participate in repeated personal biomonitoring. Both of the dormitories were buildings that did not permit smoking, with no predominant, known point sources of industrial pollution in the vicinity. Similarly, all participants in this protocol were non-smokers. The current analysis examined a subset of the measurements collected, specifically those related to traffic pollution indicators in these study participants and corresponding metabolic response. The study protocol and all aspects of human participants were approved and supervised by the Emory University Institutional Review Board.

2.1. Exposure assessment

Measurements included, but were not limited to, several, ubiquitous primary traffic-related pollutants using a range of sampling platforms. A detailed description of the DRIVE study design, as well as the sampling methods can be found elsewhere (Liang et al., 2018; Sarnat et al., 2018). For the current analysis, we used exposure measurements conducted at two student dormitories located at different proximities from the highway, including two outdoor sites and two indoor sites: ‘Near Dorm’, approximately 20 m west of the Connector, and ‘Far Dorm’, approximately 1.4 km west of the Connector (Fig. 1).

Sampling instrumentation located inside each dorm was identical, and utilized a three-way valve to alternate sampling between indoor and outdoor air. At each sampling location, we measured a suite of traditional primary traffic-related pollutants continuously including black carbon (BC), carbon monoxide (CO), nitric oxide (NO), nitrogen dioxide (NO2), and nitric oxides (NOX). In addition to the continuous and semi-continuous measurements, we conducted 48-hour quartz and Teflon filter-based measurements for particle mass. A total of 10 instruments were utilized, providing air pollutant concentration data at time scales from minutes to averages of over a week. All field instrumentation used to measure continuous pollutant concentrations were evaluated, refurbished if needed, and calibrated prior to field sampling. In order to compare concurrent pollutant measurements across the multiple sampling sites and ensure accurate concentrations during the sampling period, instruments measuring the same pollutant parameters were also co-located both before and after the sampling period and consistently calibrated throughout the 13-week intensive field sampling period. Calibration was done by varying the blend of pollutant gas from a cylinder of known concentration with a cylinder of zero air at given flow rates (Bios, DryCal). Instrument collocations were conducted, for continuous NO-NO2-NOX, CO, and integrated PM2.5 mass and reflectance over a multi-day period, both before and after field sampling to assess method precision and potential instrument offset.

2.2. Panel study recruitment

We conducted a nested panel study (from Sept 8 to Dec 15, 2014) as a component within the indoor and outdoor monitoring campaign. We recruited a cohort of students living in the same dormitories that housed the indoor Near Dorm and Far Dorm sampling instrumentation. Recruitment occurred on-site during a 3-week period to accommodate dorm move-in dates and the start of the fall school semester, and was conducted by researchers in accordance to pre-established protocols and the Institutional Review Board at GIT. Of the 28 and 38 students who signed consent forms to participate from the Near Dorm and Far Dorm, respectively, 26 from the Near Dorm and 31 from the Far Dorm were enrolled based on an assessment of the participant’s availability during the semester and likely compliance with the study protocol. During the 12-week sampling period, 2 participants from the Near Dorm and 1 from the Far Dorm dropped out of the study. No specific reasons were given for the attrition. In total, 54 students participated in the panel study.

Participants were given a baseline questionnaire detailing socio-demographic information, preliminary health, and typical time-activity patterns. Time-activity pattern data were collected through portable global positioning system (GPS) trackers that were distributed to a subset of participants (N = 6) each week. The collected GPS data aided in quantifying time spent in various microenvironments and proximity to traffic sources as potential modifiers of personal exposures.

2.3. High-resolution metabolomics

Metabolomics analysis was completed using up to four (monthly) venous blood and twelve (weekly) saliva samples collected from each of the 54 participants. In total, 175 plasma samples (average of 3.2 repeated samples per participant) and 621 2-ml vials of saliva (average of 11.5 repeated samples per participant) were collected. Metabolomics analyses were conducted on the monthly blood samples and saliva samples collected at the same time (collected on September 19th, October 24th, Nov 14th, and December 5th 2014) using established protocols (Go et al., 2015; Ladva et al., 2017). Each sample was treated with two volumes of ice-cold acetonitrile and analyzed using liquid chromatography with high-resolution mass spectrometry (LC-HRMS) techniques (Thermo Scientific™ Q-Exactive™ HF). Each sample was analyzed in triplicate. Two technical columns, hydrophilic interaction liquid chromatography (HILIC) with positive electrospray ionization (ESI) and C18 hydrophobic reversed-phase chromatography with negative ESI, were used to enhance the coverage of metabolic feature detection. Two quality control pooled reference plasma samples, which included NIST 1950 (Simon-Manso et al., 2013) and pooled human plasma purchased from Equitech Bio were included at the beginning and end of each analytical batch of 20 samples for normalization, control for background noise, batch evaluation, and post hoc quantification. Following instrument analyses of all samples, raw data files were converted to .mzML files using ProteoWizard and extracted using aplCMS with modifications by XMSanalyzer (Uppal et al., 2015; Yu et al., 2009). Detected signals (referred to as metabolic features) were uniquely defined by their mass-to-charge ratio (m/z), retention time and ion intensity. Only metabolic features detected in > 15% of plasma and saliva samples with median coefficient of variation (CV) among technical replicates < 30% and Pearson correlation > 0.7 were included in further analyses. Following quality assessment, replicate samples were averaged and averaged intensities were log2 transformed.

2.4. Data analysis

For the primary statistical analysis, we followed an untargeted Metabolome-Wide Association Study (MWAS) workflow, where metabolic profiles were analyzed without prior knowledge of their chemical identity. Linear mixed effect models were conducted to assess associations between metabolic feature intensity (i.e., relative concentration) and levels of traffic related pollutants, controlling for random subject effects. Since nested personal exposure assessment showed that students from both dorms tended to spend the majority of their time in or around their respective dormitories (57% for Near Dorm students; 61% for Far Dorm students), we selected measurements from these locations as a priori surrogates of actual personal pollutant exposures. We used the weekly average outdoor or indoor level of each of the six traditional single-species air pollutants (BC, CO, NO, NO2, NOX, and PM2.5) at the dormitory locations as the primary exposure indicator of interest. Models had the following form:

\[
Y_{ij} = \mu + \beta_0 + \beta_1_{Pollutant_{ij}} + \beta_2_{Dorm_{ij}} + \beta_3_{Age_{ij}} + \beta_4_{Gender_{ij}} + \beta_5_{BMI_{ij}} + \beta_6_{Race_{ij}} + \beta_7_{Movingdays_{ij}} + \beta_8_{Timepoints_{ij}} + \epsilon_{ij}
\]

(1)

where \(Y_{ij}\) refers to intensity (i.e., relative concentration) of metabolic
conducted for each metabolic feature from each of the two biosample
axis against the plots, which plot the retention time of each metabolic feature on the x-
axis for participant i on sampling date t. Separate models were
conducted for each metabolic feature from each of the two biosample
types and each of the two analytical configurations (plasma HILIC
column with positive ESI, plasma C18 column with negative ESI, saliva
HILIC column with positive ESI, and saliva C18 column with negative ESI). μ is the fixed-effect intercept and a random effect \( \theta_j \) is included to control for potential between participant variation. \( \text{Pollutant}_{itj} \) refers to the
weekly average outdoor or indoor level of the traffic related pol-
lutant k at the dorm location for participant i on biosampling date t. \( \text{Dorm}_{ij} \) refers to the dorm location for participant i, accounting for
potential differences in the non-traffic-pollutant-related factors among
the participants from the two dorms. Other covariates were included to control for potential between participant differences, including age,
gender (categorical), body mass index (BMI; continuous), and race
(categorical). We also controlled for \( \text{Movingdays}_{ip} \), the total number of
days between the biosample collection date and the date that partici-
pat i moved into the dorm; and \( \text{Timepoints}_{it} \), the time point order (i.e.,
month number for plasma or saliva) when the biosample was collected
from participant i on biosampling date t. \( \epsilon_{itj} \) represents residual random
normal error.

Hypothesis tests to identify differentially expressed features associated
with specific traffic related pollutant levels (by biometric column) were corrected for multiple comparisons using the Benjamini-
Hochberg false discovery rate (FDRBH) procedure at a 5% false positive
(i.e., Type I error) threshold. Results were presented using Manhattan
plots, which plot the retention time of each metabolic feature on the x-
axis against the \( - \log_{10}(p) \) for \( \beta_j \) from Eq. (1) on the y-axis. The sig-
nificant metabolic features (FDRBH < 0.05) associated with traffic-
related pollutants were then used to conduct the pathway enrichment
and metabolite annotation analyses.

Table 1
Traffic pollutant levels and baseline demographic data of the study participants.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall</th>
<th>Near Dorm</th>
<th>Far Dorm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic pollutant levels*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC (µg/m³), mean (SD)</td>
<td>0.88 (0.90)</td>
<td>0.78 (0.60)</td>
<td></td>
</tr>
<tr>
<td>CO (ppb), mean (SD)</td>
<td>343 (122)</td>
<td>209 (132)</td>
<td></td>
</tr>
<tr>
<td>NO (ppb), mean (SD)</td>
<td>15.9 (14.9)</td>
<td>12.2 (16.0)</td>
<td></td>
</tr>
<tr>
<td>NO₂ (ppb), mean (SD)</td>
<td>23.3 (9.6)</td>
<td>21.4 (4.7)</td>
<td></td>
</tr>
<tr>
<td>NO₃ (ppb), mean (SD)</td>
<td>39.2 (22.4)</td>
<td>33.7 (19.3)</td>
<td></td>
</tr>
<tr>
<td>PM₁₀ (µg/m³), mean (SD)</td>
<td>11.1 (5.5)</td>
<td>11.0 (5.9)</td>
<td></td>
</tr>
<tr>
<td>Demographic characteristics</td>
<td>n = 51</td>
<td>n = 23</td>
<td>n = 28</td>
</tr>
<tr>
<td>Age (SD)</td>
<td>19.3 (0.9)</td>
<td>19.2 (0.9)</td>
<td>19.4 (0.8)</td>
</tr>
<tr>
<td>BMI (SD)</td>
<td>23.3 (1.0)</td>
<td>22.7 (3.1)</td>
<td>23.9 (2.9)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>24 (47)</td>
<td>11 (48)</td>
<td>13 (46)</td>
</tr>
<tr>
<td>Male</td>
<td>27 (53)</td>
<td>12 (52)</td>
<td>15 (54)</td>
</tr>
<tr>
<td>Academic year, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freshman</td>
<td>29 (57)</td>
<td>16 (70)</td>
<td>13 (46)</td>
</tr>
<tr>
<td>Sophomore</td>
<td>14 (27)</td>
<td>2 (9)</td>
<td>12 (43)</td>
</tr>
<tr>
<td>Junior</td>
<td>7 (14)</td>
<td>4 (17)</td>
<td>3 (11)</td>
</tr>
<tr>
<td>Senior</td>
<td>1 (2)</td>
<td>1 (4)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Days in Dorm prior to first plasma collection (SD)</td>
<td>69 (119)</td>
<td>66 (161)</td>
<td>55 (67)</td>
</tr>
<tr>
<td>Time spent outdoors, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 1 h</td>
<td>4 (8)</td>
<td>2 (9)</td>
<td>2 (7)</td>
</tr>
<tr>
<td>1–2 h</td>
<td>21 (42)</td>
<td>11 (48)</td>
<td>10 (37)</td>
</tr>
<tr>
<td>3–4 h</td>
<td>18 (36)</td>
<td>8 (35)</td>
<td>10 (37)</td>
</tr>
<tr>
<td>5 h or more</td>
<td>7 (14)</td>
<td>2 (8)</td>
<td>5 (19)</td>
</tr>
<tr>
<td>Time spent in vehicle, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 1 h</td>
<td>20 (40)</td>
<td>8 (35)</td>
<td>12 (44)</td>
</tr>
<tr>
<td>1–2 h</td>
<td>27 (54)</td>
<td>14 (61)</td>
<td>13 (48)</td>
</tr>
<tr>
<td>3–4 h</td>
<td>2 (4)</td>
<td>1 (4)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>5 h or more</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (4)</td>
</tr>
</tbody>
</table>

* 24-hour average outdoor levels.

2.5. Metabolic pathway enrichment analysis and metabolite annotation

Pathway identification and module analysis were performed for
metabolic features meeting the 5% FDR threshold using mumichog (v.
1.0.5), a novel bioinformatics platform that infers and categorizes
functional biological activity directly from mass spectrometry output,
without prior metabolite validation (Amorim et al., 2013; Li et al.,
2013). Mumichog analyses were conducted separately for each set of
significant features from each of the 24 indoor/outdoor TRAP linear
mixed models in the HILIC and C18 columns. An adjusted \( p \) for each
pathway was calculated from resampling the reference input file in
mumichog using a gamma distribution, which penalizes pathways
with fewer reference hits, and assigning greater significance to path-
ways with more reference hits (Li et al., 2013). We classified pathways
with adjusted \( p < 0.05 \) for at least three of the TRAPs models, and with
at least four features from the experimental data matched with pathway
metabolites. To further minimize the possibility of false positive dis-
covery, candidate pathways were re-run using a subset of 6 most
common forms out of the 16 standard adduct forms in mumichog (for
HILIC positive ion mode, only the following adducts were considered:
M⁺, M + H⁺, M-H₂O + H⁺, M + Na⁺, M + K⁺, M + 2H₂⁺, and M(C13) + 2H⁺; for C18 negative ion mode, only the following adducts were considered:
M-H⁻, M + Cl⁻, M + ACN-H⁻, M + HOOC⁻, M(C13)-H⁻, M-H₂O-H⁻, and M + Na-2H⁻). We presented
final results in a metabolic-pathway-TRAPs heat map, with
each cell of the heat map representing a statistical association between
each of the metabolic pathways and each of the corresponding indoor/
outdoor TRAPs.

The metabolic features that were significantly associated with
TRAPs (FDRBH < 0.05) and also significantly enriched in a relevant
pathway \( (p < 0.05) \) using mumichog were annotated by matching mass
m/z value for adducts commonly formed to the METLIN (https://
metlin.scripps.edu/index.php), ChemSpider (http://www.chemspider.
com/), Human Metabolome Database (HMDB), and Kyoto Encyclopedia
of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/
pathway.html) databases, using a mass error threshold of 10 ppm.
Tentative matches were further screened on their retention time, iso-
tope patterns, and peak quality by manually examining the extracted
ion chromatogram (EIC). A select number of annotated metabolites
were then confirmed by comparison of accurate mass m/z, retention
time and ion dissociation patterns to authentic reference standards
analyzed using the identical method and instrument parameters.

3. Results

We collected 175 plasma and cell samples and 204 saliva samples
during four time points, spaced approximately 21–35 days apart from
each other, throughout the study from the 54 DRIVE participants
providing blood and saliva specimens. With the exception of indoor
PM₂.₅, mean indoor and outdoor levels for all the pollutants were significantly
higher in the Near Dorm compared to the Far Dorm during the entire
study period, indicative of a substantial contrast in the potential ex-
posure level to TRAPs among the two groups of study participants (t-
test: \( p < 0.05 \) (Tables 1 and S1). Baseline information collected from
each participant at the commencement of sampling showed generally
similar demographic characteristics among participants in both dorms,
although there was a greater relative number of sophomore (i.e., 2nd
year) students living in the Far Dorm compared to the Near Dorm
(42.9% vs 8.7%, respectively) (Table 1). The total number of days
the participants spent in their respective dorms prior to the first plasma
collection were comparable in both dorms (86 vs 55 days, \( p = 0.87 \)).

Among the 54 participants, continuous GPS data were collected from
43 participants; 21 participants from the Near Dorm and 22 from
the Far Dorm, over the course of 12 sampling weeks. As expected, there
was a clear bimodal distribution between participants from the two
dorms of time spent in closer proximity to the highway (Fig. S2),
supporting an observation that much of the students' time was spent within or near their respective dorms.

After data quality filtering, we detected 20,766 metabolic features in plasma samples (13,419 in HILIC plasma column and 7347 in C18 plasma column) and 29,013 in saliva samples (21,313 in HILIC plasma column and 7700 in C18 plasma column). The median CV across the triplicate for each metabolic feature was 24.8%, indicating good overall data quality. Comparison of the two chromatography columns showed 6667 (HILIC) and 2812 (C18) were detected in both plasma and saliva samples (Fig. S1). Spearman correlation coefficients between these shared metabolic feature intensities detected in both plasma and saliva were 0.53 for HILIC and 0.58 for C18 (across all four sampling time points: p < 0.0001, for all pairwise correlations). Correlations were stronger among the metabolic features with CVs < 10% (0.64 for HILIC; 0.68 for C18). In addition, > 45% of the detected metabolic features were found to have m/z matches (< 10 ppm) in either the Human Metabolome Database (HMDB) or the USEPA's Mobile Air Toxics database.

We conducted and analyzed 48 sets of MWAS models (12 indoor/outdoor TRAPs among metabolic features in plasma and saliva, with each analyzed using two chromatography columns). In total, 847 and 444 unique metabolic features were associated with at least one or more of the indoor or outdoor TRAPs using HILIC- and C18-plasma columns, respectively (FDR_{BH} < 5%). There were a similar number of significant metabolic features found in both columns for saliva samples (HILIC-saliva = 1320 features; C18-saliva = 399 features) (Table 2 and Fig. S3a–d).

We further screened each of the metabolic features associated with TRAP for spectrum peak quality and purity by manual examination of their respective extracted ion chromatograms (EICs). By including only those spectra with unambiguous EIC peaks (63% to 95%), we reduced the possibility of false positive discovery. Using mummichog, we examined whether the features that were significantly associated with TRAPs co-occurred as enriched metabolites within specific metabolic pathways.

The results indicated that 24 metabolic pathways were associated with at least three or more TRAPs in plasma samples (Fig. 2). A similar number of metabolic pathways (N = 23) were identified in saliva samples (Fig. 3). Eight pathways consistently appeared to be associated with varying indoor or outdoor TRAPs models for both plasma and saliva samples and were detected using both chromatography columns. These pathways are predominantly associated with xenobiotic-mediated oxidative stress and acute inflammatory response, such as leukotriene metabolism, vitamin E metabolism, and cytochrome P450 (Chow, 1991; Dahlén et al., 1981; Gonzalez, 2005; Henderson, 1994; Morgan, 1997; Singh et al., 2005). The leukotriene metabolism and vitamin E pathway were most strongly associated with a majority of the TRAPs. Based on the strength and consistency of the mummichog results in showing these TRAP-related associations, we focused on annotating the metabolic features within these pathways, with the aim of validating the targeted metabolic signals.

Using identified metabolite data available in KEGG, ChemSpider, HMDB, and a curated database of air toxics published by the US Environmental Protection Agency, we matched 17 and 11 unique metabolic features within the leukotriene pathway and vitamin E metabolism, respectively (Tables S2a–S3d). The metabolic features matching these pathways were significantly associated with multiple indoor or outdoor TRAPs and exhibited relatively pure spectrum peaks with common adduct forms.

Metabolic annotations from the leukotriene pathway included leukotriene B4 (LTB4), leukotriene D4 (LTD4) and leukotriene E4 (LTE4), as well as their metabolic precursors, 12-oxo-LTB4, 20-OH-LTB4 and 20-oxo-LTE4. Additional metabolites from this pathway included indicators of lipid oxidation, such as 6-trans-LTB4, 12-oxo-20-dihydroxy-LTB4 and 18-carboxy-dinor-LTE4, and leukotriene related metabolites, such as 10,11-dihydro-12R-hydroxy-LTC4, and 12-oxo-c-LTB3. Tentatively matched metabolites in the vitamin E pathway included dehydrogenation precursors of tocotrienols and tocopherols, which maintain primary vitamin E antioxidant activities. Notably, most of the features that matched pro-inflammatory metabolites were positively associated with the TRAPs in the MWAS models; while significant and negative associations included metabolites consistent with antioxidant processes.

Finally, we matched the samples with authentic reference standards, verified by tandem mass spectrometry, to confirm the chemical identity of metabolic features that were both associated with the measured TRAPs as well as those enriched within TRAP exposure-relevant metabolic pathways. In total, we confirmed the chemical identities of 10 metabolites (Table 3), 9 of which were indicative of endogenous metabolic signals, including arginine, histidine, γ-linolenic acid, and hypoxanthine. As described above, we observed negative associations (β < 0) between these anti-inflammatory metabolites with levels of TRAPs, while positive association (β > 0) between TRAP and corresponding oxidative or pro-inflammatory metabolites. Given their robust associations with TRAP indicators and substantial biological plausibility, we believe these endogenous metabolites have the potential to be developed further as source-specific biomarkers reflecting TRAP exposures or associated health responses. Future in vivo and in vitro studies are warranted, however, to replicate these findings and confirm the sensitivity and specificity of these promising biomarkers.

4. Discussion

Untargeted omics-based methods provide a promising platform for measuring exposure and response to thousands of xenobiotic agents (Ellis et al., 2012; Park et al., 2012; Walker et al., 2016). In the DRIVE study, we successfully demonstrated the ability of HRM to assess changes in specific metabolic profiles associated with realistic environmental exposures in plasma and saliva samples measured repeatedly in a relatively small, healthy, young adult panel. To our knowledge, this study constituted the single largest prospective longitudinal assessment examining perturbations in human metabolome due to traffic related pollutant exposures.

Table 2

| Biomarkers & technical column | Total number of features extracted | BC | CO | NO | NO2 | NOX | PM2.5 | Number of unique features
|-----------------------------|----------------------------------|----|----|----|-----|-----|-------|---------------------|
| HILIC-plasma                | 13,419                           | 182| 174| 138| 143| 198| 178  | 153  | 168  | 124 | 132 | 214 | 203 | 847
| C18-plasma                  | 7347                             | 90 | 102| 76 | 77 | 73 | 83   | 101  | 109  | 88  | 88  | 101 | 119 | 444
| HILIC-saliva                | 21,313                           | 312| 336| 185| 186| 257| 263  | 235  | 234  | 238 | 246 | 371 | 359 | 1320
| C18-saliva                  | 7700                             | 94 | 102| 67 | 67 | 70 | 90   | 86   | 77   | 68  | 75  | 99  | 105 | 399

Acronym: TRAP, traffic-related pollutant; FDR_{BH}, false discover rate correction using the Benjamini-Hochberg procedure; BC, black carbon; CO, carbon monoxide; NO, nitric oxide; NO2, nitrogen dioxide; NOX, nitrogen oxide; PM2.5, fine particulate matter.

* Number of unique metabolic features that were statistically significantly (FDR_{BH} < 0.05) associated with at least one or more TRPs.
For C18 negative ion mode, only the following adducts were considered: M-H(C13) + 2H[2+].

For HILIC positive ion mode, only the following adducts were considered: M[1+], M + H[1+], M-H2O + H[1+], M + Na[1+], M + K[1+], M + 2H[2+], and M + Na-2H[2+].

Among the most pronounced and consistent findings from the MWAS mixed effect modeling and metabolic pathway enrichment analyses was that cumulative exposure to elevated levels of TRAPs over a multi-month period was associated with perturbations in several key biological pathways (Figs. 2 and 3). Of added importance was that the identification of the specific elicited pathways were robust to the various indoor and outdoor TRAP models, both biological sample types, and in both technical columns. Moreover, a majority of the identified pathways are consistent with TRAP-related acute oxidative stress and inflammatory response (Baja et al., 2010; Chuang et al., 2007; Kelly, 2003; Lodovici and Bigagli, 2011; Nel et al., 2001). Among these pathways, those involved in leukotriene and vitamin E metabolism, in particular, consistently showed the strongest associations with TRAPs in both saliva and plasma samples.

Leukotrienes, synthesized from arachidonic acid by arachidonate 5-lipoxygenase, are a family of active eicosanoid inflammatory mediators formed by leukocytes, mastocytoma cells, macrophages, and other tissues and cells in response to immunological and non-immunological stimuli (Hammarström, 1983). Actively involved in asthmatic and allergic reactions and sustaining inflammatory reactions, leukotrienes are powerful biological signals, the overproduction of which is the major cause of inflammation in asthma and allergic rhinitis (Nelson et al., 2008; Salmon and Higgs, 1987). In particular, the dihydroxy fatty acid LTB4, synthesized by leukotriene A4 (LTA4) hydrolase from LTA4, is a potent chemoattractant and proinflammatory mediator that recruits cells from the immune system to produce inflammation and induces the release of lysosomal enzymes by neutrophils (Cotran et al., 1999; Devchand et al., 1996; Ford-Hutchinson et al., 1980). Numerous in vitro and in vivo studies have shown that LTB4 increases the generation of reactive oxygen species (ROS) and promotes leukocyte adherence (Biselli et al., 1996; Salas et al., 1999; Steiner et al., 2001). Moreover, LTB4 levels in exhaled breath condensate have been linked to elevated NO2 (Hüls et al., 2017; Vossoughi et al., 2014), ozone (Alfaro et al., 2007), and cigarette smoking (Carpinegranato et al., 2003).

Among the most pronounced and consistent findings from the MWAS mixed effect modeling and metabolic pathway enrichment analyses was that cumulative exposure to elevated levels of TRAPs over a multi-month period was associated with perturbations in several key biological pathways (Figs. 2 and 3). Of added importance was that the identification of the specific elicited pathways were robust to the various indoor and outdoor TRAP models, both biological sample types, and in both technical columns. Moreover, a majority of the identified pathways are consistent with TRAP-related acute oxidative stress and inflammatory response (Baja et al., 2010; Chuang et al., 2007; Kelly, 2003; Lodovici and Bigagli, 2011; Nel et al., 2001). Among these pathways, those involved in leukotriene and vitamin E metabolism, in particular, consistently showed the strongest associations with TRAPs in both saliva and plasma samples.
the propagation of reactive oxygen species, protects lipids and prevents the oxidation of polyunsaturated fatty acids (Choe and Min, 2009; Traber and Stevens, 2011; Whitney and Rolfes, 2012). Previously, nutritional research has revealed the protective role of vitamin E in preventing or minimizing free-radical damage associated with specific diseases and lifestyle patterns, including protection against ambient PM2.5 or ozone induced inflammatory response and oxidative stress (Bo et al., 2016; Menzel, 1979; Packer, 1991; Salvi, 2007). In our study, 11 metabolic features were matched to vitamin E metabolites and the intensities of these antioxidants decreased with higher TRAP exposures (Appendix Table S3a–d).

The results of TRAP MWAS and confirmation with MSMS support the metabolic pathway enrichment findings. Of the 10 metabolites validated by comparison to authentic standards, nine were endogenous.

Fig. 3. Metabolic pathways associated with ≥3 traffic-related air pollution in saliva. Cells were shaded according to the strength (i.e. p-value) of the association between each of metabolic pathways and significant features (FDR<0.05) that were associated with each indoor/outdoor single traffic pollutant indicator. Pathways are ordered according to the total number of the significant pathway-traffic pollutant associations (p < 0.05) in the HILIC column (positive ion mode) and the C18 column (negative ion mode).

For C18 negative ion mode, only the following adducts were considered: M-H[−], M + Cl[−], M + ACN-H[−], M + HCOO[−], M(C13)-H[−], M-H2O-H[−], and M + Na-2H[−].

^Total number of metabolites within the specific metabolic pathway.

#Number of metabolic features in the samples with m/z matched to the metabolites within the specific metabolic pathway.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Chemical identity* of the metabolic features significantly associated with TRAP (FDR&lt;0.05) in the DRIVE study.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>m/z</strong></td>
<td><strong>RT (s)</strong></td>
</tr>
<tr>
<td>137.0463</td>
<td>44.2</td>
</tr>
<tr>
<td>279.2324</td>
<td>26.0</td>
</tr>
<tr>
<td>348.0709</td>
<td>147.6</td>
</tr>
<tr>
<td>391.2848</td>
<td>25.5</td>
</tr>
<tr>
<td>112.0511</td>
<td>50.0</td>
</tr>
<tr>
<td>116.0711</td>
<td>112.0</td>
</tr>
<tr>
<td>156.0773</td>
<td>103.0</td>
</tr>
<tr>
<td>89.0239</td>
<td>22.2</td>
</tr>
<tr>
<td>173.1039</td>
<td>23.3</td>
</tr>
<tr>
<td>223.0719</td>
<td>24.6</td>
</tr>
<tr>
<td>223.0719</td>
<td>24.6</td>
</tr>
</tbody>
</table>

Acronym: m/z, mass to charge ratio; RT, retention time; TRAP, traffic-related air pollutant; BC, black carbon; CO, carbon monoxide; NO, nitric oxide; NO2, nitrogen dioxide; NOx, nitrogen oxide; PM2.5, fine particulate matter.

* Chemical identification on the candidate metabolic features 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.
metabolites involved in oxidative stress, acute inflammatory response, and DNA damage and repair processes. The findings also provide preliminary support for these validated metabolites as potential novel biomarkers of TRAP exposures.

Our results involving arginine, specifically, warrant specific attention. Arginine is an α-aminio acid that produces NO endogenously in the airways via NO synthase (Silkoff et al., 2000), and has been shown to be depleted following air pollution-induced airway hyperresponsiveness, due to the augmentation of arginase (North et al., 2011). Correspondingly, in our samples, we identified plasma arginine and found it to be inversely associated with measured BC and NO levels. Another metabolite we identified in the DRIVE study was histidine, a semi-essential amino acid. Histidine is the precursor to histamine (Tabor, 1954), a vital and well-known inflammatory agent in immune responses, including the airway hyperresponsiveness (Hospers et al., 2000; Liu et al., 1990). Previously in a study examining the serum amino acid profiles in obese and non-obese women, both histidine and arginine were found to be negatively associated with inflammation and oxidative stress (Niu et al., 2012). Consistently in the DRIVE samples, we observed a significant negative association between histidine and outdoor PM$_{2.5}$ level. We also confirmed the identity of metabolite matching hypoxanthine, a naturally occurring purine derivative that protects against oxidant-induced cell injury by inhibiting activation of nuclear poly (ADP-ribose) polymerase (Durlacz et al., 1980; Szabó, 1998; Szabó and Dawson, 1998). In a recent study, Vlaanderen et al. (2017) also identified hypoxanthine as associated with exposure to ambient air pollutants following 18 h of exposure (Vlaanderen et al., 2017). In our study, hypoxanthine was positively associated with both indoor and outdoor CO levels. Finally, we identified γ-linolenic acid (GLA) in plasma, which is an essential inhibitor in the biosynthesis of LTB4 (Horrobin, 1992; Mancuso et al., 1997). Here, too, the negative observed relationship between GLA and PM$_{2.5}$ levels is as expected, indicative of its inverse relationship with LTB4 expression. Collectively, these confirmed results provide a small, yet promising advance in the development of specific markers for TRAP exposure. Future work should focus on replicating the present findings, as well as in identifying metabolite biomarkers expressed across multiple biometrics, rather than metabolites enriched in a single matrix.

Of all the extracted metabolic features, roughly 45% were detected in both plasma and saliva samples, matched by m/z ratio. Correlation coefficients among the shared features ranged from 0.5–0.9, which is consistent with our previous findings of moderate-to-strong correlation between saliva and plasma that we observed in a pilot study comparing metabolic profiles in plasma, saliva, and exhaled breath condensate (Ladva et al., 2017). Moreover, 12 significant metabolic pathways elucidated in mumichog were shared by both saliva and plasma. Taken together, these results support saliva samples as an alternative sensitive, and less invasive, biometric for metabolomics analysis.

Although somewhat unexpected, we also identified bis(2-ethylhexyl) phthalate (DEHP) using tandem mass spectrometry, an exogenous air pollutant identified in the EPA air toxic compounds database. DEHP was positively associated with multiple TRAPs within the MWAS results, including BC, CO, and NO. DEHP is commonly found indoors as a ubiquitous chemical plasticizer component, which may also be found from the emissions of volatile organic compounds from the interior materials within new vehicles, according to a recent study (Faber et al., 2013). A less well-known source of DEHP, however, is road dust (Omar et al., 2007; Wang et al., 2013), which may explain its association with TRAP exposures in this study.

We used an MWAS approach where we independently modeled each of six single-pollutant measures as surrogates of exposure to primary traffic pollution, known to be a highly heterogeneous mixture. Thus, caution should be taken in interpreting the observed metabolic perturbations as indicative of causal associations with any individual TRAP indicator. To this point, we want to highlight the finding that 30%–40% of the metabolic features meeting the FDR threshold being associated with at least two or more individual traffic pollutant indicators.

Collectively, our study design and approach resolved statistically robust metabolic differences with changes in TRAP exposures. Despite this, limitations inherent in omics-based analyses and small panel designs deserve specific attention. Although we designed this study to recruit panels of students from both dorms that were demographically balanced, this was largely a convenience sample with some notable discrepancies between students in the two locations (Table 1). Students in the Near Dorm, on average, moved into their dorm rooms somewhat earlier than students in the Far Dorm, which could contribute to differences in exposure and associated metabolic phenotypes. Similarly, the Far Dorm panel included a greater fraction of second-, as compared to first-year students. These differences in the two panels, where differences in time spent on campus and exposure to either traffic or other unspecified environmental factors (e.g., diet, indoor pollutants) could contribute to the observed metabolomic differences. To account for this, the MWAS regression models included terms to address these between participant differences, including age, sex, BMI, race, and outdoor locations, as a means of controlling for non-trafficked environmental factors that might differ between these two subpanels (Walker et al., 2016). Results from sensitivity analyses of the MWAS output were consistent and robust to model specification and inclusion of these covariates.

There is also a possibility that findings from small panels may be unduly influenced by individual observations. Most of the 100 to 1000 unique metabolic features associated with at least one TRAP were detected in at least 38% of the participants, suggesting that the differences were not driven by extreme response in relatively few participants. Similarly, in highly multidimensional analyses, there may be an increased risk of false positives due to multiple comparisons and Type I errors. To address this, we applied several stringent criteria when conducting metabolic pathway enrichment analyses, including only using model findings with FDR$_{BH}$ < 5%, as well as restricting output to a subset of the six most abundant adducts in each ionization mode. Although previous studies using high-resolution metabolomics have identified biologically relevant results at FDR < 20% and when all 16 adduct forms are used when matching unknown features to metabolic databases (Chandler et al., 2016; Tebani et al., 2017), a more stringent threshold minimizes Type I error (false positives) at the increased risk of Type II error (false negatives). Since this was an exposure population without a formal control (i.e., unexposed) cohort, we balanced our findings towards increased stringency (smaller $p$) and decreased risk of false positives.

5. Conclusions

Recent advances in HRM support its use as a highly sensitive platform, capable of identifying thousands of small molecules from both endogenous and exogenous sources. We view the results from this study further support HRM as a powerful platform for elucidating biologically relevant pathways associated with exposures to key environmental pollutants and sources. Our results were consistent with the limited number of studies that reported perturbations in the human metabolome associated ambient air pollution and included changes in key oxidative stress and inflammatory pathways (Li et al., 2017; Surowiec et al., 2016; Vlaanderen et al., 2017). Most intriguingly, however, were results from our MWAS models, which point to the potential of HRM as a tool for biomarker discovery. Here, we identified and validated several metabolites in plasma and saliva that were directly associated with external traffic pollution measurements in our panel. Among these, arginine and histidine appeared to the most prevalent, accurately measured, and most strongly associated with corresponding traffic pollutant exposures. While highly intriguing, these findings warrant replication to solidify their use as novel, specific indicators of exposure and response to this critical source of urban air pollution.
Acknowledgement

Support for this project were provided through a contract with the Health Effects Institute (RFA #4942-RFA13-1/14-3). The field study conducted as part of this study benefitted greatly from the assistance of many students, staff, and faculty at both Georgia Tech and Emory. Special thanks go to C. Cornwell, K. Parada, S. Shim, K. Johnson and E. Yang for their tremendous help in conducting the field study. We want to thank Dr. R. Weber, Dr. V. Verma, and Ms. D. Gao for their measurements of oxidative potential of ambient fine particles via DTT assay. We are indebted to Dr. J. Schauer (U. Wisconsin) for loaning us several instruments to supplement our sampling network. We would also like to thank Dr. Seung-Hyun Cho from RTI, Inc. for her collaboration on this project. The Georgia EPD allowed us access to their roadside monitoring site and helped provide data from those monitors, and we particularly thank Ken Buckley for his assistance with this. The study used on the instrumentation assembled for field studies conducted as part of the Southeastern Center for Air Pollution and Epidemiology (SCAPE), which was funded by a US Environmental Protection Agency STAR grant R834799. The information in this document may not necessarily reflect the views of the Agency and no official endorsement should be inferred. We also like to thank Ms. V. Tran for conducting the LC/MS on the DRIVE biosamples. Dr. Golan gratefully acknowledges support by a post-doctoral fellowship from the Environment and Health Fund, Jerusalem, Israel. We acknowledge NSF for providing a fellowship to JM (DGE-1650044), and Dr. Russell made use of funds provided by a generous gift from Howard T. Tellepsen. Dr. Walker gratefully acknowledges the following grants supported by NIH (NIMH R01MH107205, NEIHS T32ES012870, and NIH S10OD18006). We acknowledge the support from the HERCULES exposure research center, supported by the National Institute of Environmental Health Sciences of the National Institutes of Health (P30ES019776). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We owe a debt of gratitude to the administrators at Georgia Tech for allowing us to conduct this study on campus and in their residence hall facilities.

Declaration of interests

None.

Funding

Support for this project was provided through a grant from the Health Effects Institute (RFA #4942-RFA13-1/14-3). The study used on the instrumentation assembled for field studies conducted as part of the Southeastern Center for Air Pollution and Epidemiology (SCAPE), which was funded by a US Environmental Protection Agency STAR grant R834799.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2018.07.044.

References
