Differential DNA methylation and PM$_{2.5}$ species in a 450K epigenome-wide association study

Lingzhen Dai, Amar Mehta, Irina Mordukhovich, Allan C. Just, Jincheng Shen, Lifang Hou, Petros Koutrakis, David Sparrow, Pantel S. Vokonas, Andrea A. Baccarelli & Joel D. Schwartz

To cite this article: Lingzhen Dai, Amar Mehta, Irina Mordukhovich, Allan C. Just, Jincheng Shen, Lifang Hou, Petros Koutrakis, David Sparrow, Pantel S. Vokonas, Andrea A. Baccarelli & Joel D. Schwartz (2017) Differential DNA methylation and PM$_{2.5}$ species in a 450K epigenome-wide association study, Epigenetics, 12:2, 139-148, DOI: 10.1080/15592294.2016.1271853

To link to this article: https://doi.org/10.1080/15592294.2016.1271853
ABSTRACT
Although there is growing evidence that exposure to ambient particulate matter is associated with global DNA methylation and gene-specific methylation, little is known regarding epigenome-wide changes in DNA methylation in relation to particles and, especially, particle components. Using the Illumina Infinium HumanMethylation450 BeadChip, we examined the relationship between one-year moving averages of PM$_{2.5}$ species (Al, Ca, Cu, Fe, K, Na, Ni, S, Si, V, and Zn) and DNA methylation at 484,613 CpG probes in a longitudinal cohort that included 646 subjects. Bonferroni correction was applied to adjust for multiple comparisons. Bioinformatics analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment was also performed. We observed 20 Bonferroni significant (P-value $< 9.4 \times 10^{-7}$) CpGs for Fe, 8 for Ni, and 1 for V. Particularly, analysis at Schlafen Family Member 11 (SLFN11) cg10911913 was positively associated with measured levels of all 3 species. The SLFN11 gene codes for an interferon-induced protein that inhibits retroviruses and sensitizes cancer cells to DNA-damaging agents. Bioinformatics analysis suggests that gene targets may be relevant to pathways including cancers, signal transduction, and cell growth and death. Ours is the first study to examine the epigenome-wide association between ambient particles species and DNA methylation. We found that long-term exposures to specific components of ambient particle pollution, especially particles emitted during oil combustion, were associated with methylation changes in genes relevant to immune responses. Our findings provide insight into potential biologic mechanisms on an epigenetic level.

Introduction

The associations between fine particulate matter (particles with an aerodynamic size of $\leq 2.5 \mu m$, PM$_{2.5}$) and health outcomes are well established. It is estimated that PM$_{2.5}$ leads to 3.15 million premature deaths per year on a global scale. Nationwide and multiple-community studies report that exposure to PM$_{2.5}$ is linked with an increased risk of hospital admissions and mortality in the US, and that certain PM$_{2.5}$ species may be more toxic than others. DNA methylation is an epigenetic process in which a methyl group is added to deoxycytosine bases to form deoxymethylcytosine. The literature to date indicates that DNA methylation may be an important pathway linking particulate matter (PM) to health outcomes. For example, changes in patterns of DNA methylation may be associated with processes leading to cardiovascular diseases. There is growing evidence that exposure to PM is associated with global DNA methylation and gene-specific methylation, especially for exposure to metal-rich PM in occupational settings. Yet, little is known regarding how PM exposure affects DNA methylation on a genome-wide level. Measurements of DNA methylation at cytosine-guanine dinucleotide (CpG) loci have recently become available; for instance, the Illumina Infinium HumanMethylation450 BeadChip (450K) measures methylation at more than 450,000 CpG sites. This makes an epigenome-wide association study of the relationship between particles and DNA methylation possible. Several studies have investigated the epigenome-wide association between smoking, an important PM source, and DNA methylation. Using the 450K platform and Bonferroni correction, Joubert et al. identified 26 CpGs with methylation changes in newborns related to maternal smoking during pregnancy in a large Norwegian birth cohort. Specifically, the authors documented changes at CpGs in cytochrome p450 1A1 (CYP1A1) and aryl-hydrocarbon receptor repressor (AhR), genes known to play a key role in the aryl hydrocarbon receptor (AhR) signaling pathway that mediates the detoxification of components of tobacco smoke. In a more recent study, the authors identified numerous loci associated with maternal smoking during pregnancy in a genome-wide consortium meta-analysis and found substantial persistence of such effects.

CONTACT Lingzhen Dai, ldai@mail.harvard.edu, Department of Environmental Health, Harvard T.H. Chan School of Public Health, Landmark Center 4th floor, 401 Park Drive, Boston, MA 02215, USA.

ARTICLE HISTORY
Received 8 June 2016
Revised 26 November 2016
Accepted 7 December 2016

KEYWORDS
Ambient particulate matter; DNA methylation; epigenome-wide association study; metals; pathway enrichment

ABSTRACT
Although there is growing evidence that exposure to ambient particulate matter is associated with global DNA methylation and gene-specific methylation, little is known regarding epigenome-wide changes in DNA methylation in relation to particles and, especially, particle components. Using the Illumina Infinium HumanMethylation450 BeadChip, we examined the relationship between one-year moving averages of PM$_{2.5}$ species (Al, Ca, Cu, Fe, K, Na, Ni, S, Si, V, and Zn) and DNA methylation at 484,613 CpG probes in a longitudinal cohort that included 646 subjects. Bonferroni correction was applied to adjust for multiple comparisons. Bioinformatics analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment was also performed. We observed 20 Bonferroni significant (P-value $< 9.4 \times 10^{-7}$) CpGs for Fe, 8 for Ni, and 1 for V. Particularly, analysis at Schlafen Family Member 11 (SLFN11) cg10911913 was positively associated with measured levels of all 3 species. The SLFN11 gene codes for an interferon-induced protein that inhibits retroviruses and sensitizes cancer cells to DNA-damaging agents. Bioinformatics analysis suggests that gene targets may be relevant to pathways including cancers, signal transduction, and cell growth and death. Ours is the first study to examine the epigenome-wide association between ambient particles species and DNA methylation. We found that long-term exposures to specific components of ambient particle pollution, especially particles emitted during oil combustion, were associated with methylation changes in genes relevant to immune responses. Our findings provide insight into potential biologic mechanisms on an epigenetic level.

Introduction

The associations between fine particulate matter (particles with an aerodynamic size of $\leq 2.5 \mu m$, PM$_{2.5}$) and health outcomes are well established. It is estimated that PM$_{2.5}$ leads to 3.15 million premature deaths per year on a global scale. Nationwide and multiple-community studies report that exposure to PM$_{2.5}$ is linked with an increased risk of hospital admissions and mortality in the US, and that certain PM$_{2.5}$ species may be more toxic than others. DNA methylation is an epigenetic process in which a methyl group is added to deoxycytosine bases to form deoxymethylcytosine. The literature to date indicates that DNA methylation may be an important pathway linking particulate matter (PM) to health outcomes. For example, changes in patterns of DNA methylation may be associated with processes leading to cardiovascular diseases. There is growing evidence that exposure to PM is associated with global DNA methylation and gene-specific methylation, especially for exposure to metal-rich PM in occupational settings. Yet, little is known regarding how PM exposure affects DNA methylation on a genome-wide level. Measurements of DNA methylation at cytosine-guanine dinucleotide (CpG) loci have recently become available; for instance, the Illumina Infinium HumanMethylation450 BeadChip (450K) measures methylation at more than 450,000 CpG sites. This makes an epigenome-wide association study of the relationship between particles and DNA methylation possible. Several studies have investigated the epigenome-wide association between smoking, an important PM source, and DNA methylation. Using the 450K platform and Bonferroni correction, Joubert et al. identified 26 CpGs with methylation changes in newborns related to maternal smoking during pregnancy in a large Norwegian birth cohort. Specifically, the authors documented changes at CpGs in cytochrome p450 1A1 (CYP1A1) and aryl-hydrocarbon receptor repressor (AhR), genes known to play a key role in the aryl hydrocarbon receptor (AhR) signaling pathway that mediates the detoxification of components of tobacco smoke. In a more recent study, the authors identified numerous loci associated with maternal smoking during pregnancy in a genome-wide consortium meta-analysis and found substantial persistence of such effects.
in later childhood.\textsuperscript{18} Zeilinger et al.\textsuperscript{19} conducted an epigenome-wide study comparing methylation levels among current, former, and never smokers in a population-based panel, and found widespread differences in the degree of site-specific methylation as a function of tobacco smoking. Moreover, they observed that the most significant associations with smoking were for DNA methylation sites in \textit{AHRR}.

To date, very few epigenome-wide association studies have focused on exposures to ambient particles, with the exception of one recent study that identified several CpG sites associated with cumulative exposure to ambient particles among populations from Germany and the US.\textsuperscript{20} To our knowledge, no epigenome-wide association studies have investigated the toxicity of specific PM species. Such research would be of great importance because epigenome-wide association research based on DNA methylation microarrays could provide insight into genetic mechanisms underlying associations between air particles and health outcomes and thereby improve our understanding of human disease. To our knowledge, our study is the first to examine epigenome-wide associations between PM$_{2.5}$ species and DNA methylation using data from a large longitudinal cohort.

\section*{Methods}

\subsection*{Study population}

Study subjects were participants of the Normative Aging Study (NAS), a closed cohort that was established in 1963 by the Department of Veterans Affairs. This ongoing longitudinal study of aging enrolled 2,280 community-dwelling, healthy men living in the Greater Boston area.\textsuperscript{21} Participants were free of known chronic medical conditions at enrollment and have undergone examinations every 3 to 5 y on a continuous rolling basis. All participants provided written informed consent. The study was reviewed and approved by the Institutional Review Boards of all participating institutions.

We restricted our analysis to study visits with DNA samples, which were collected from subjects starting in 1999. A comparison of visits with DNA samples and visits without showed no substantial differences with regard to body mass index, education, alcohol consumption, and smoking, although older participants were less likely to provide DNA samples (mean age 74.5 vs. 76.3 years; \textit{P}-value of \textit{t}-test in mean age < 0.05). In addition, since Caucasians comprise most of our study population (approximately 97%), we excluded non-Caucasians to reduce study heterogeneity that may be introduced by diverse genetic ancestry. We also excluded observations that did not have complete information on the exposures or covariates of interest. In the final analyses, samples from 646 subjects with 1,031 total visits were included. Of the 646 subjects, 314 presented for one visit, 279 presented for 2 visits, and 53 presented for 3 visits. The flowchart of our study participants is shown in Fig. 1.

\subsection*{DNA methylation}

DNA samples were collected between 1999 and 2013. We used the QIAamp DNA Blood Kit (Qiagen, CA, USA) to extract DNA from buffy coat, and performed bisulfite conversion with the EZ-96 DNA Methylation Kit (Zymo Research, CA, USA). To minimize batch effects, we randomized chips across plates and randomized samples based on a 2-stage age stratification so that age distributed similarly across chips and plates.

We measured DNA methylation of CpG probes using Illumina Infinium HumanMethylation450 BeadChip.\textsuperscript{22} Quality control analysis was performed, where samples with >1% of probes with a detection \textit{P}-value > 0.05 were removed, and probes with \textit{P}-value > 5% samples with a bad count < 3 or with >1% samples with a detection \textit{P}-value > 0.05 were removed. The remaining samples were preprocessed using the Illumina-type background correction\textsuperscript{23} and normalized with dye-bias\textsuperscript{24} and BMIQ3 adjustments,\textsuperscript{25} which were used to generate methylation values. Extreme outliers, defined by Tukey’s method (i.e., <25\textsuperscript{th} percentile – 3 \times IQR or > 75\textsuperscript{th} percentile + 3 \times IQR), were trimmed. As a result, 484,613 CpG probes were included in the working set.

We used the \textit{β} value to indicate methylation level at each CpG site. The \textit{β} value represents the ratio of the fluorescence of the methylated signal to the combined methylated and unmethylated signals; that is, \textit{β} = \text{intensity of the methylated signal (M)} / \text{[intensity of the unmethylated signal (U) + intensity of the methylated signal (M) + 100]}.\textsuperscript{22}

The methylation data has been uploaded to the database of Genotypes and Phenotypes (dbGaP; available at https://www.ncbi.nlm.nih.gov/gap) established by the National Institutes of Health, along with some other basic information (dbGaP accession phs000853.v1.p1).

\subsection*{Air pollution}

We measured daily concentrations of PM$_{2.5}$ mass and species at the stationary monitoring supersite at the Harvard University Countway Library, Boston, MA, USA. PM$_{2.5}$ was assessed by

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{flowchart.png}
\caption{Flowchart of study participants.}
\end{figure}
To estimate long-term effects of air pollution on DNA methylation, we calculated one-year moving average concentrations for PM$_{2.5}$ and the following 11 species: Al, Ca, Cu, Fe, K, Na, Ni, S, Si, V, and Zn. These species are representative of different sources of particles: Al, Ca, Fe, and Si are major components of road dust; Cu and Zn originate from traffic emissions; Ni and V are tracers of oil combustion; S is a regional pollutant that is associated with coal combustion and to a lesser extent with metallurgic activities; Na is a major component of sea salt particles; K originates from wood burning and soil.

Statistical analysis

For each specific CpG site, we applied a linear mixed-effects model as described below,

$$E(Y_{ij}) = \beta_0 + \beta_1 * \text{Spec}_{ij} + \beta_2 * \text{PM}_{2.5j} + \beta_3 * \text{Age}_{ij} + \ldots + \mu_i,$$

where $Y_{ij}$ is the methylation level of subject i at visit j, Spec$_{ij}$ and PM$_{2.5j}$ are one-year moving average species and PM$_{2.5}$ concentrations before visit j, Age$_{ij}$ is subject i’s age at visit j (this notation is relevant to all other covariates, since we used values assessed at each visit in our analyses), and $\mu_i$ is the random intercept that accounts for correlation within subject. The full formula can be found in the Supplemental Document.

In the analysis, we controlled for the following covariates a priori based on the relevant literature:\cite{6,19,26,27} total PM$_{2.5}$ mass, cell proportions (CD4$^+$ T lymphocytes, CD8$^+$ T lymphocytes, natural killer cells, B cells, monocytes) estimated by the Houseman et al. method,\cite{28} age, body mass index [BMI, computed as weight (in kilograms) divided by height (in square meters)], cigarette pack years, smoking status (never, ever), alcohol consumption (in kilograms divided by height (in square meters)), years of education, age and season (spring: March–May, summer: June–August, fall: September–November, winter: December–February). Potential batch effects were also considered, including plate, the position of the chip on plate, and the row and column position on the chip.

Since overall type I error inflates as the number of statistical tests increases, we used Bonferroni correction to adjust for multiple comparisons. This method, which accounts for the number of CpG sites and exposures tested, has been widely used in similar studies.\cite{31,32,33} As a result, the adjusted level of significance in the current study was $P$-value $= 9.4 \times 10^{-9}$, i.e., 0.05 / (484,613 x 11). After identifying significant CpGs, we further checked whether they were non-unimodal in the dip test,\cite{34} were cross-reactive,\cite{35,36} or were located within 10 bases of SNPs relevant to European populations.\cite{37}

Our study population was a heavily smoking population. Since we controlled for smoking status and pack years in the analysis, we believe there was sufficient adjustment, but we understand that the presence of a large number of smokers may obscure or distort associations. To address this, we have performed subset analyses among nonsmokers and smokers.

All statistical analyses were performed using R 3.1.2\cite{37}.

Bioinformatics analysis

We performed Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis\cite{38} using the genes targets of CpGs, identified by the Illumina annotation assignment, with Benjamini-Hochberg false discovery rates (FDR)\cite{39} <1% and <5%, respectively. Bioinformatics analysis was performed using Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resources 6.7,\cite{31,32} Briefly, DAVID consists of an integrated biologic knowledgebase with wide annotation content coverage. It allowed us to identify the most relevant pathways associated with the gene lists we provided.

Results

Characteristics of study subjects are presented in Table 1. At baseline, the mean age was 74 y [standard deviation (SD) = 7]. The study population was well educated, with a mean of 15 y of education (SD = 3). Furthermore, 186 of the 646 (28.8%) individuals never smoked, and 131 (20.3%) had 2 or more drinks per day.

Table 2 shows the descriptive summary of concentrations of PM$_{2.5}$ mass and species. During the study period 1,999 – 2,013, the annual average PM$_{2.5}$ concentration was 10.5 (SD = 1.1) $\mu$g/m$^3$. Among the examined species, S accounted for the largest mass proportion (10%). The oil combustion tracers, Ni and V, accounted for 0.03% and 0.04% of PM$_{2.5}$ mass with mean concentrations of 0.0037 (SD = 0.0013) $\mu$g/m$^3$ and 0.0041 (SD = 0.0012) $\mu$g/m$^3$, respectively. We report a strong correlation between Ni and V ($r = 0.96$).

Table 3 presents CpGs that reached Bonferroni-corrected significance ($P$-value $< 9.4 \times 10^{-9}$) in relation to PM$_{2.5}$ species. After excluding CpGs that were non-unimodal in the dip test, were cross-reactive, or contained SNPs within 10 bases, we observed 20 significant CpGs for Fe, 8 for Ni, and 1 for V. Particularly, Schlafen Family Member 11 (SLFN11) cg10911913 was significant for all 3 metals. The ratio of methylation at this CpG site is expected to increase by 0.073, 0.031, and 0.044 per IQR increase in the one-year
moving averages of Fe, Ni, and V, respectively. We found strong seasonal patterns at this site for the 3 metals (Figure S1 in the Supplemental Material), with the highest methylation levels in the spring. We also found that methylation levels at CASZ1 cg16238819, CUECD2 cg06024834, ONECUT1 cg15446043, and FOXO4 cg12453500 were related to both Fe and Ni. To illustrate the epigenome-wide association between PM$_{2.5}$ species and methylation, we plotted the -log$_{10}$(P-value) from the mixed-effects models for all the 484,613 CpG sites in Fig. 2. The horizontal line represents the Bonferroni-corrected significance threshold. Subset analyses among nonsmokers and smokers showed consistent associations between PM species and methylation within smoking subgroups. For example, the coefficients for Fe at our most significant CpG (cg10911913) were 0.122 overall (Table S1 in the Supplemental Document), 0.131 among smokers and 0.086 among nonsmokers, and were statistically significant at a genome-wide level for both smoking and non-smoking subgroups. We observed the same trend for Ni and V.

Results of the bioinformatics analysis are shown in Fig. 3. We conducted KEGG pathway enrichment analysis

<p>| Table 3. Genome-wide associations of one-year moving averages of PM$_{2.5}$ species with DNA methylation: CpGs with Bonferroni-corrected statistical significance (P-value &lt; 9.4 × 10$^{-9}$). |
|---|---|---|---|---|---|---|---|---|---|---|---|---|</p>
<table>
<thead>
<tr>
<th>Species</th>
<th>Chr$^a$</th>
<th>Gene</th>
<th>CpG</th>
<th>Location</th>
<th>Estimate$^b$</th>
<th>Coef$^c$</th>
<th>SE$^d$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td>1</td>
<td>DHR53</td>
<td>ch.1.471396R</td>
<td>Body</td>
<td>0.053</td>
<td>0.089</td>
<td>0.015</td>
<td>2.37E-09</td>
</tr>
<tr>
<td>1</td>
<td>CASZ1</td>
<td>cg16238819</td>
<td>5’UTR</td>
<td>0.076</td>
<td>0.126</td>
<td>0.021</td>
<td>6.94E-09</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SLC39A7;RKB8</td>
<td>cg24319547</td>
<td>5’UTR;TSS1500</td>
<td>0.031</td>
<td>0.052</td>
<td>0.009</td>
<td>2.20E-09</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>KIAA1949</td>
<td>cg08052856</td>
<td>5’UTR</td>
<td>0.072</td>
<td>0.120</td>
<td>0.020</td>
<td>2.58E-09</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>CDCL3</td>
<td>cg22787480</td>
<td>TSS200</td>
<td>0.068</td>
<td>0.113</td>
<td>0.019</td>
<td>8.30E-09</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>FBXO925</td>
<td>ch.8.742998</td>
<td>TSS200</td>
<td>0.062</td>
<td>0.103</td>
<td>0.017</td>
<td>4.40E-09</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>RNF112</td>
<td>cg06024834</td>
<td>TSS200</td>
<td>0.039</td>
<td>0.065</td>
<td>0.011</td>
<td>1.98E-09</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>RPS25;TRAPPC4</td>
<td>cg05223675</td>
<td>TSS1500</td>
<td>0.077</td>
<td>0.128</td>
<td>0.021</td>
<td>6.97E-09</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>SMEK1</td>
<td>cg01561368</td>
<td>TSS1500</td>
<td>0.024</td>
<td>0.040</td>
<td>0.007</td>
<td>6.20E-09</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ONECUT1</td>
<td>cg15446043</td>
<td>TSS200</td>
<td>0.146</td>
<td>0.243</td>
<td>0.039</td>
<td>1.37E-09</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>TPM1</td>
<td>cg12302829</td>
<td>TSS200</td>
<td>0.039</td>
<td>0.065</td>
<td>0.011</td>
<td>7.22E-09</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>SLFN11</td>
<td>cg10917952</td>
<td>TSS200</td>
<td>0.073</td>
<td>0.122</td>
<td>0.018</td>
<td>1.84E-10</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>CHT1</td>
<td>cg26525457</td>
<td>Body</td>
<td>-0.076</td>
<td>-0.126</td>
<td>0.020</td>
<td>1.26E-09</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>RNF138</td>
<td>cg16078856</td>
<td>Body</td>
<td>0.065</td>
<td>0.108</td>
<td>0.018</td>
<td>2.55E-09</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>MYPOP</td>
<td>cg03766703</td>
<td>3’UTR</td>
<td>-0.119</td>
<td>-0.198</td>
<td>0.033</td>
<td>6.91E-09</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>REEP6;PCSK1</td>
<td>cg1970591</td>
<td>1stExon;TS1500</td>
<td>0.222</td>
<td>0.370</td>
<td>0.063</td>
<td>8.10E-09</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>C21orf57</td>
<td>cg10434947</td>
<td>3’UTR</td>
<td>0.043</td>
<td>0.072</td>
<td>0.012</td>
<td>4.32E-09</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>CYB11;LOC4565851</td>
<td>cg11002378</td>
<td>5’UTR;TSS1500</td>
<td>0.050</td>
<td>0.084</td>
<td>0.013</td>
<td>4.71E-10</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>FOXO4</td>
<td>cg12453500</td>
<td>TSS1500</td>
<td>0.059</td>
<td>0.098</td>
<td>0.014</td>
<td>2.89E-11</td>
<td></td>
</tr>
<tr>
<td>Ni</td>
<td>1</td>
<td>CASZ1</td>
<td>cg16238819</td>
<td>5’UTR</td>
<td>0.032</td>
<td>0.102</td>
<td>0.017</td>
<td>8.25E-09</td>
</tr>
<tr>
<td>10</td>
<td>CUECD2</td>
<td>cg06024834</td>
<td>TSS200</td>
<td>0.017</td>
<td>0.056</td>
<td>0.009</td>
<td>1.10E-09</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ONECUT1</td>
<td>cg15446043</td>
<td>TSS200</td>
<td>0.059</td>
<td>0.195</td>
<td>0.030</td>
<td>8.31E-09</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>WDR90</td>
<td>cg0234477</td>
<td>TSS200</td>
<td>0.035</td>
<td>0.118</td>
<td>0.019</td>
<td>4.61E-09</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>SLFN11</td>
<td>cg10917952</td>
<td>TSS200</td>
<td>0.031</td>
<td>0.104</td>
<td>0.016</td>
<td>7.25E-11</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>RNF138</td>
<td>cg16078856</td>
<td>Body</td>
<td>0.029</td>
<td>0.073</td>
<td>0.014</td>
<td>1.91E-10</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>SEC1;DBP</td>
<td>cg13402773</td>
<td>TSS1500;TSS200</td>
<td>0.056</td>
<td>0.185</td>
<td>0.031</td>
<td>6.22E-09</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>FOXO4</td>
<td>cg12453500</td>
<td>TSS1500</td>
<td>0.022</td>
<td>0.073</td>
<td>0.012</td>
<td>2.10E-09</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>17</td>
<td>SLFN11</td>
<td>cg10917952</td>
<td>TSS200</td>
<td>0.044</td>
<td>0.103</td>
<td>0.017</td>
<td>6.81E-10</td>
</tr>
</tbody>
</table>

$^a$Chromosome.

$^b$Estimate of coefficient in methylation level per IQR increase in the species concentration.

$^c$Regression coefficient in the mixed-effects models, with adjustment for PM$_{2.5}$ mass, Houseman estimated cell proportions (CD4$^+$ T lymphocytes, CD8$^+$ T lymphocytes, natural killer cells, B cells, monocytes), age, BMI, cigarette pack years, smoking status, alcohol consumption, education, year, season, and batch effects (plate, row and column position on chip).

$^d$Standard error for regression coefficient.
for gene targets of CpGs with FDR <1% and <5%, respectively. Pathways whose enrichment FDRs were less than 5% were considered significant. In the analysis for gene targets of CpGs with FDR <1%, we identified 9 significant KEGG pathways. Three of them (i.e., “pathways in cancer,” “non-small cell lung cancer,” and “glioma”) belong to the “Cancers” network. Results for gene targets of CpGs with FDR <5% suggest that those genes may also be linked to other cancers (e.g., myeloid leukemia) and pathways in the immune system. Other identified pathways include the diabetes and mitogen-activated protein kinases (MAPK) signaling pathways; the latter is involved in cellular functions including cell proliferation, differentiation, and migration.

**Discussion**

In this study, we investigated the relationship between PM<sub>2.5</sub> species and epigenome-wide DNA methylation in a large longitudinal cohort of elderly men. To the best of our knowledge, this is the first study that examined epigenome-wide associations between ambient PM<sub>2.5</sub> components and DNA methylation. We adjusted for multiple tests for the 11 PM components and 484,613 CpGs analyzed. We observed that 20 CpGs met the conservative Bonferroni-corrected level of significance (P-value < 9.4 x 10<sup>-9</sup>) in the analysis for Fe, 8 for Ni, and 1 for V. Fe mainly originates from crustal sources and the iron and steel industry; Ni and V are by-products of oil combustion, whereas
Ni is also associated with the iron and steel industry. The strong correlations between these species are shown in Table 2; the correlation coefficients were 0.93 for Fe-Ni, 0.86 for Fe-V, and 0.96 for Ni-V. In addition, we found that all 3 metals were associated with methylation levels at cg10911913 within the SLFN11 gene. Our bioinformatics analysis suggests that gene targets of CpGs with FDR < 1% may be relevant to pathways including cancers, signal transduction, and cell growth and death.

To date, only a limited number of research studies have examined the association of ambient particles with DNA methylation. Most of them used surrogates for global DNA methylation, such as methylation at repetitive elements, or focused on methylation in candidate genes. The main shortcomings of these studies are: 1) methylation levels measured by surrogates do not consider the location in the genome and the patterns of methylation at surrogates may not well represent the patterns of gene-specific methylation, and 2) the use of candidate genes can only reveal methylation changes in particular genes selected a priori and may miss the genes that exhibit the most profound effects.

![KEGG pathway enrichment analysis](image)

Figure 3. KEGG pathway enrichment analysis for (A) gene targets of CpGs with FDR < 1%, and (B) gene targets of CpGs with FDR < 5%. Pathways whose enrichment FDR was less than 5% were considered significant and are indicated by colored chunks.

One previous study regarding PM exposure and DNA methylation reported that exposures to 90-day moving averages of black carbon and sulfate particles concentrations were associated with hypomethylation of the long interspersed nucleotide element-1 (LINE-1) and the short interspersed nucleotide element Alu. Baccarelli et al. found decreased LINE-1 methylation in association with 7-day moving average black carbon and PM2.5 levels, whereas no association was observed for Alu methylation. A controlled human exposure experiment observed hypomethylation of Alu and the toll-like receptor 4 (TLR4) gene among 15 healthy adults exposed to fine concentrated ambient particles (CAPs) for 130 min. Recently, a longitudinal study reported a relationship between decreased methylation at 3 CpG loci located in the nitric oxide synthase 2, inducible (NOS2A) gene and short-term exposures to organic carbon, elemental carbon, NO3⁻, and NH4⁺ among 28 patients with chronic obstructive pulmonary disease.

A more recent epigenome-wide study by Panni et al. on short- and mid-term PM2.5 exposure among 3 independent cohorts—the US NAS cohort and the German KORA F3 and F4 cohorts—identified 12 CpG sites at a Bonferroni adjusted
significance level. Methylation at 9 out of these 12 CpGs showed increases related to PM exposure. This trend was consistent with our study: we found that 27 out of the 29 associations across all species displayed increased methylation levels. However, our study did not identify the same CpGs. A possible explanation for this difference is that the previous study evaluated short- or medium-term effects of PM, whereas we focused on long-term exposures. In addition, the results of the previous study were from a meta-analysis of 3 cohorts, including our study population. The 2 German cohorts had a much larger sample size than the NAS cohort (2299 vs. 657 participants), and the Panni et al. study reported heterogeneity at some CpG sites across the cohorts. Hence, their findings were more likely to reveal methylation changes within the German population and cannot be generalized to the US NAS population.

The most consistent finding of our study was the significant increase in methylation at SLFN11 cg10911913 that was associated with one-year moving average concentrations of the PM$_{2.5}$ species Fe, Ni, and V. Schlafen Family Member (SLFN) genes encode a family of proteins that have been implicated in the regulation of cell growth and T cell development. SLFN11, a member of the SLFN family, potently and specifically abrogates the production of retroviruses, e.g., human immunodeficiency virus (HIV-1). Using cells with endogenously high and low SLFN11 expression and siRNA-mediated silencing, Zoppoli et al. found that SLFN11 causally determines cell death and cell cycle arrest in response to DNA-damaging agents in cancer cells from different tissues of origin. SLFN11 cg10911913 is located in a region whose function is promoter associated. Our results showed that methylation at this CpG site was positively associated with one-year moving averages of the PM$_{2.5}$ species Fe, Ni, and V. Our analysis suggests that exposure to ambient particles, especially those from oil combustion, was associated with hypermethylation of CpGs within SLFN11, which might lead to its downregulation. This hypothesis is consistent with recent literature on air-pollution-induced activation of viral sequences integrated in the human genome. For example, a study of truck drivers and office workers in Beijing, China showed a significant association of exposure to elemental carbon—taken as a marker of traffic particle exposure—with altered expression of viral microRNAs, possibly reflecting reactivation of latent retroviral sequences integrated in the host DNA.

We also observed that methylation at one cut homeobox 1 (ONECUT1) cg15446043 was positively associated with Fe and Ni exposures. ONECUT1 can inhibit hepatitis B virus gene expression and DNA replication via both transcriptional and post-transcriptional mechanisms.

In addition, we found that long-term exposure to Fe was linked to methylation in other genes expressed in immune cells, including cytohesin 1 (CYTH1) and ribosomal protein S6 kinase (RPS6KB2). CYTH1 regulates adhesion and transendothelial migration of monocytes, T lymphocytes, and dendritic cells; it is also a key regulator of neutrophil adhesion to endothelial cells and to components of the extracellular matrix. RPS6KB2 plays an important role in neutrophilic differentiation and neutrophilic proliferation in human promyelocytic leukemia (HL-60) cells. Hence, methylation changes at pertinent CpG sites within these 2 genes may affect immune function.

Interestingly, we found that Fe was associated with methylation levels of CpGs in genes that regulate the cell cycle, such as CDC5L cg22877480 and RPS6KB2 cg26457823. Cell division cycle 5-like (CDC5L) is a crucial regulator of cell cycle G2/M progression and a component of pre-mRNA processing; hence, previous research has suggested it as a target for cancer therapy. RPS6KB2 is also highly active in the G2 and M phases.

The DAVID bioinformatics analysis also indicated that Fe and Ni were associated with type II diabetes mellitus, insulin signaling pathways, and the MAPK pathway. This is consistent with previous work. For instance, long-term PM$_{2.5}$ exposure resulted in systemic inflammation, increased visceral adiposity, and insulin resistance among mice on a high fat diet. In mice consuming normal diets, PM exposure was a risk factor for the development of type II diabetes, as it induced insulin resistance and impaired glucose tolerance. Similarly, in humans, long-term exposure to particulate matter was linked to increased risk of type II diabetes in a large cohort study. We have also reported that particulate air pollution was associated with changes in methylation in the MAPK pathway within the NAS cohort, but were unable to analyze metal components as in this study.

Earlier, we found that particles from oil combustion were associated with markers of inflammation and endothelial dysfunction, and with increased blood pressure in the NAS population.

The current study may indicate potential mechanisms of PM-related health effects on an epigenetic level: specifically, a proinflammatory stimulus such as air pollution may modify methylation in genes relevant to immune responses, inducing inflammation and endothelial dysfunction, thereby influencing cardiovascular health. Our study has several strengths. First, this is the first study that examined the effects of ambient particulate matter species on epigenome-wide DNA methylation. Previous studies often estimated DNA methylation associated with ambient PM using surrogates for global methylation or gene-specific methylation at candidate genes. An epigenome-wide study like ours offers a novel perspective on such an association, and implicates potential biologic mechanisms on an epigenetic level. Second, we used a stringent Bonferroni correction threshold to adjust the level of significance, which accounted both for the 11 PM$_{2.5}$ species and for the number of CpG sites analyzed. Although we observed a small number of significant CpGs due to the strict correction, the probability of false positives is expected to remain appropriate. Third, the study population was geographically stable and well followed since enrollment in 1,963.

On the other hand, this study has several notable limitations. One major limitation is that we were unable to perform a replication analysis due to a lack of appropriate comparison cohorts. The German KORA cohorts had PM$_{2.5}$ mass data but did not have measurements on PM species. Since the chemical composition of air pollutants in Germany may differ from that in the US, we were not able to impute species data from PM$_{2.5}$ measurements. We do not have information on gene expression, and thus we cannot confirm the expected functional
implications of the identified methylation changes. Further research on gene expression in relation to PM exposures is needed. Additionally, we are unable to rule out residual confounding. The bioinformatics analysis of pathways used annotations of methylation sites within genes and thus genes that include more sites on the 450K array, such as those involved in cancer, are more likely to arise in pathway results. Last but not least, there exists a generalizability issue because our study population consisted almost entirely of elderly Caucasian men. One should consider factors like age, sex, and race when generalizing our findings to other populations.

In summary, we examined the association of long-term exposure to PM$_{2.5}$ species with differential DNA methylation in a 450K epigenome-wide association study using a large longitudinal cohort. Our study added evidence on the association of ambient particulate exposure with epigenome-wide DNA methylation. It also suggests differing relative toxicities of PM species, which is especially important given the US. National Research Council’s recommendation to consider determining the differential risk of PM$_{2.5}$ species a research priority. Additionally, we are unable to rule out residual confounding. The bioinformatics analysis of pathways used annotations of methylation sites within genes and thus genes that include more sites on the 450K array, such as those involved in cancer, are more likely to arise in pathway results. Last but not least, there exists a generalizability issue because our study population consisted almost entirely of elderly Caucasian men. One should consider factors like age, sex, and race when generalizing our findings to other populations.

In summary, we examined the association of long-term exposure to PM$_{2.5}$ species with differential DNA methylation in a 450K epigenome-wide association study using a large longitudinal cohort. Our study added evidence on the association of ambient particulate exposure with epigenome-wide DNA methylation. It also suggests differing relative toxicities of PM species, which is especially important given the US. National Research Council’s recommendation to consider determining the differential risk of PM$_{2.5}$ species a research priority. Additionally, we are unable to rule out residual confounding. The bioinformatics analysis of pathways used annotations of methylation sites within genes and thus genes that include more sites on the 450K array, such as those involved in cancer, are more likely to arise in pathway results. Last but not least, there exists a generalizability issue because our study population consisted almost entirely of elderly Caucasian men. One should consider factors like age, sex, and race when generalizing our findings to other populations.

In summary, we examined the association of long-term exposure to PM$_{2.5}$ species with differential DNA methylation in a 450K epigenome-wide association study using a large longitudinal cohort. Our study added evidence on the association of ambient particulate exposure with epigenome-wide DNA methylation. It also suggests differing relative toxicities of PM species, which is especially important given the US. National Research Council’s recommendation to consider determining the differential risk of PM$_{2.5}$ species a research priority. Additionally, we are unable to rule out residual confounding. The bioinformatics analysis of pathways used annotations of methylation sites within genes and thus genes that include more sites on the 450K array, such as those involved in cancer, are more likely to arise in pathway results. Last but not least, there exists a generalizability issue because our study population consisted almost entirely of elderly Caucasian men. One should consider factors like age, sex, and race when generalizing our findings to other populations.

Disclosures of potential conflicts of interest

No potential conflicts of interest were disclosed.

Funding

This work was supported by the National Institute of Environmental Health Sciences ES000002, ES021733 and ES015172–01 and the US. Environmental Protection Agency RD-83587201. A.J. was supported by the National Institute of Environmental Health Sciences K99 ES023450. D.S. was supported by a VA Research Career Scientist Award. The VA Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Centers of the US. Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and Information Center, Boston, Massachusetts.

Disclaimers

The contents are solely the responsibility of the grantee and do not necessarily represent the official views of the funders. Further, the funders do not endorse the purchase of any commercial products or services mentioned in the publication.

ORCID

Allan C. Just http://orcid.org/0000-0003-4312-5957
Andrea A. Baccarelli http://orcid.org/0000-0002-3436-0640

References

16. Kile ML, Fang S, Baccarelli AA, Tarantini L, Cavallari J, Christiani DC. A panel study of occupational exposure to fine particulate matter and changes in DNA methylation over a single workday and years


Triche TJ, Jr., Weibensener DJ, Van Den Berg D, Laird PW, Sieg- mund KD. Low-level processing of Illumina Infinium DNA Methyla- tion BeadArrays. Nucleic Acids Res 2013; 41:e90; PMID:23476028; http://dx.doi.org/10.1093/nar/gkt090


Schwarz DA, Katayama CD, Hedrick SM. Schlafen, a new family of growth regulatory genes that affect thymocyte development. Immunity 1998; 9:657-68; PMID:9846487; http://dx.doi.org/10.1016/S1074- 7613(00)80663-9


El azreq MA, Bourgoin SG. Cytohesin-1 regulates human blood neutrophil adhesion to endothelial cells through beta2 integrin activation. Mol Immunol 2011; 48:1408-16; PMID:21511340; http://dx.doi.org/10.1016/j.molimm.2011.03.018


