ARTICLE



# Racial disparity in $PM_{2.5}$ species and DNA methylation in CARDIA: An epigenome-wide association study

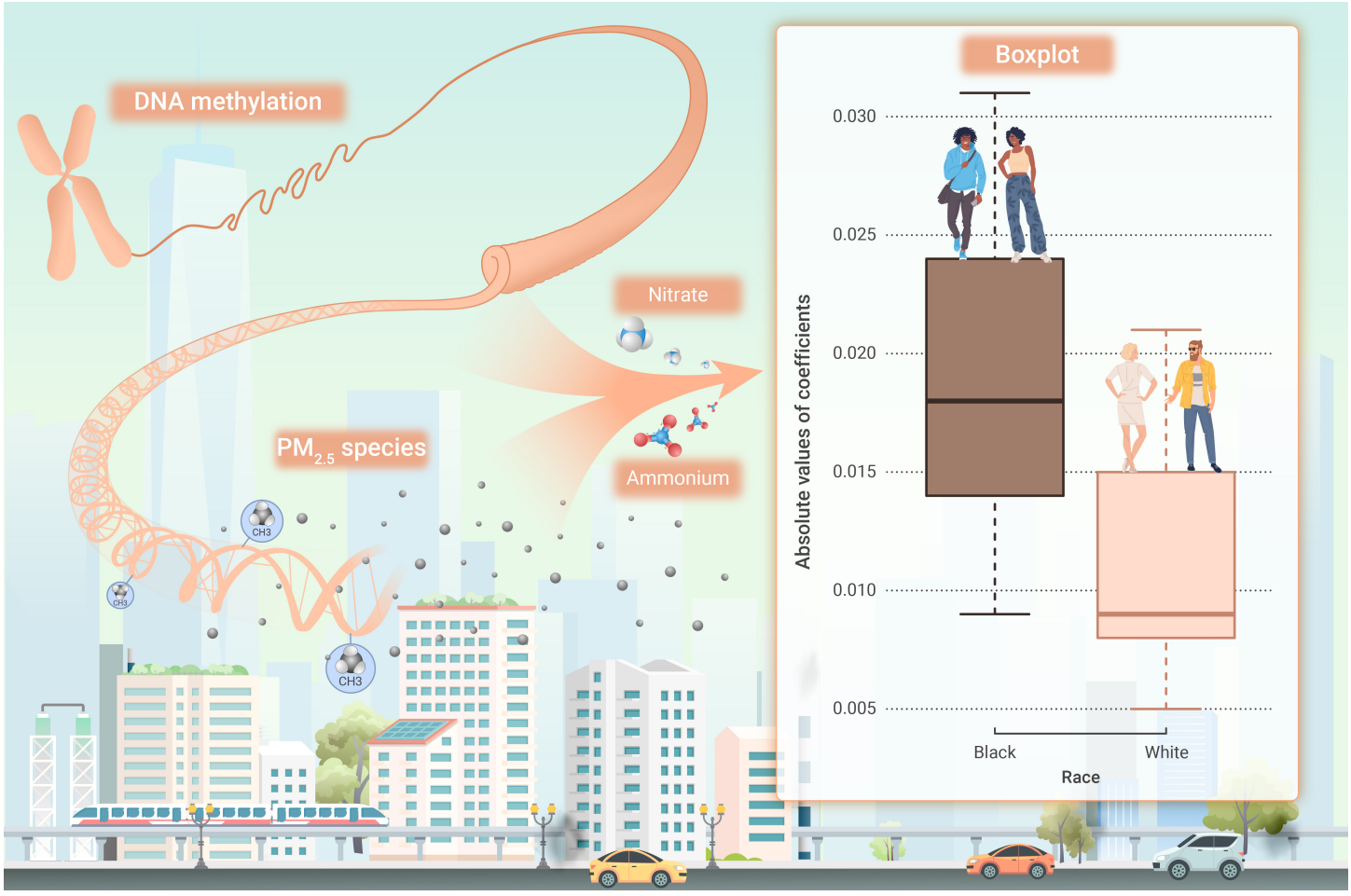
Jie Yao,[1] Yinan Zheng,[2] Jincheng Shen,[3] Brian T Joyce,[2] Ke Hao,[4] Ravi V Shah,[5] Venkatesh L Murthy,[6] Stephen Sidney,[7] Donald Lloyd-Jones,[2] Lifang Hou,[2,*] and Kai Zhang[8,*]

*Correspondence: l-hou@northwestern.edu (L.H.); kzhang9@albany.edu (K.Z.)

Received: October 31, 2023; Accepted: February 28, 2024; Published Online: April 1, 2024; https://doi.org/10.59717/j.xinn-med.2024.100061



## GRAPHICAL ABSTRACT




## PUBLIC SUMMARY

- There are statistically significant associations between DNA methylation level and $PM_{2.5}$ species.
- Black participants exhibited stronger associations compared to White participants.
- Gene targets had relevance to pathways to basal cancers, platelet activation, signaling, and aggregation.

ARTICLE

www.the-innovation.org/medicine

# Racial disparity in $PM_{2.5}$ species and DNA methylation in CARDIA: An epigenome-wide association study

Jie Yao,[1] Yinan Zheng,[2] Jincheng Shen,[3] Brian T Joyce,[2] Ke Hao,[4] Ravi V Shah,[5] Venkatesh L Murthy,[6] Stephen Sidney,[7] Donald Lloyd-Jones,[2] Lifang Hou,[2,*] and Kai Zhang[8,*]

[1]Department of Epidemiology and Biostatistics, School of Public Health, University at Albany, State University of New York, Rensselaer 12144, USA
[2]Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago 60611, USA
[3]Department of Population Health Sciences, University of Utah, Salt Lake City 84108, USA
[4]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York 10029, USA
[5]Vanderbilt Translational and Clinical Cardiovascular Research Center, Vanderbilt University Medical Center, Nashville 37232, USA
[6]Department of Medicine and Radiology, University of Michigan, Ann Arbor 48109, USA
[7]Division of Research, Kaiser Permanente Northern California, Oakland 94612, USA
[8]Department of Environmental Health Sciences, School of Public Health, University of Albany, State University of New York, Rensselaer 12144, USA
*Correspondence: l-hou@northwestern.edu (L.H.); kzhang9@albany.edu (K.Z.)

Received: October 31, 2023; Accepted: February 28, 2024; Published Online: April 1, 2024; https://doi.org/10.59717/j.xinn-med.2024.100061



Citation: Yao J., Zheng Y., Shen J., et al., (2024). Racial disparity in $PM_{2.5}$ species and DNA methylation in CARDIA: An epigenome-wide association study. The Innovation Medicine **2(2)**: 100061.

Minorities experience a disproportionate impact from $PM_{2.5}$ pollution, and biological mechanisms remain unclear. No studies have examined the epigenome-wide association between DNA methylation level and $PM_{2.5}$ species using longitudinal measurements of methylation among Blacks. This study aims to evaluate the relationship between exposure to $PM_{2.5}$ species and DNA methylation as well as assess possible race-specific effects. Genome-wide DNA methylation levels were profiled on 1,081 longitudinally followed participants (432 Black and 649 White adults) using the Infinium MethylationEPIC array. We examined the association between one-year average $PM_{2.5}$ species (black carbon, ammonium, nitrate, organic matter, sulfate, soil, and sea salt) exposure prior to exam year (Y)15 (2000-2001) and methylation at 841,639 CpG sites at Y15 and Y20 (2005-2006), respectively. In the association analysis of $PM_{2.5}$ species at Y15 with DNA methylation at Y20, methylation in 17 and 2 CpG sites, respectively, was significantly associated with nitrate and ammonium exposure. The magnitudes of coefficients of these significant CpG sites at Y20 increased by 2- to 48-fold compared to the cross-sectional association analysis of $PM_{2.5}$ species exposure prior to Y15 and DNA methylation at Y15. Black participants had stronger associations than White participants. Gene enrichment analysis indicated gene targets might be relevant to pathways including basal cancers, platelet activation, signaling, and aggregation. This study represents the first one to investigate the association between $PM_{2.5}$ species exposure and DNA methylation among Black and White adults using longitudinal measurement, highlighting the significance of considering race-specific epigenetic regulation in relation to air pollution-related health issues.

## INTRODUCTION

$PM_{2.5}$ (particulate matter less than 2.5 micrometers in diameter) is associated with numerous adverse health effects, including respiratory diseases, cardiovascular disease, neurological disorders, and cancer.[1-4] There were 3.2 million deaths worldwide attributed to $PM_{2.5}$ in 2010, making it the 6$^{th}$ greatest premature death risk factor.[5] The U.S. Environmental Protection Agency (US EPA) reported a causal relationship between long-term $PM_{2.5}$ exposure and total mortality.[4]

$PM_{2.5}$ is a dynamic mixture of various pollutants, including ammonium, nitrate and sea salt. The different chemical compositions of $PM_{2.5}$ mass play a crucial role in shaping its association with mortality.[4,6,7] $PM_{2.5}$ mass alone is not enough to explain its adverse effects when its composition plays a key role in the association.[7,8] Moreover, recent studies suggest that exposure to $PM_{2.5}$ might disproportionate and adverse impacts on minorities. A nationwide study in 2019 reported that, for the association between long-term $PM_{2.5}$ exposure and mortality, the estimated hazard ratio was 1.05 (95% CI: 1.03, 1.09) in Black individuals and 1.02 (95% CI: 1.00, 1.05) in White counterparts.[4] It is likely that other factors related to social determinants of health along with race, may explain the observed differences.

The biological mechanisms by which $PM_{2.5}$ causes adverse health outcomes are not yet fully understood. The literature suggests DNA methylation (DNAm) may be an important mediator between $PM_{2.5}$ exposure and adverse health events.[9-12] DNAm involves the addition of a methyl group to the fifth carbon of cytosine residues,[13] which can be driven by environmental exposures and in turn alter gene expression and influence disease risks.[13] There is growing evidence indicating that DNAm is a potential mediator between air pollution and health outcomes.[14] For example, a previous review provided evidence suggesting that DNAm might contribute as a partial mediator in the association of air pollution with metabolic syndrome.[15] Thus, it is important to understand how $PM_{2.5}$ may affect DNAm and related gene function. Yet, there are still very few epigenome-wide association studies (EWAS) on $PM_{2.5}$ species and DNAm. EWAS is an approach to derive the association between epigenetic variation and a particular phenotype by examining a genome-wide set of epigenetic marks, such as DNAm.[16] Previous studies have revealed that methylated regions associated with $PM_{2.5}$ and related genes were associated with cardiovascular diseases, respiratory diseases, and cancer.[17-19] An EWAS study has been conducted to evaluate the $PM_{2.5}$-associated DNAm among 12 cohorts which contain African American, European American and Hispanic/Latino American from Women's Health Initiative (WHI) and the Atherosclerosis Risk in Communities study (ARIC).[20,21] They found that 2 CpG sites were significantly associated with $PM_{2.5}$. In a recent EWAS meta-analysis of the association between DNAm and prenatal $PM_{2.5}$ exposures in newborns and children, 14 CpG sites were reported to be related to the exposures to $PM_{2.5}$.[22] Panni et al. conducted an EWAS study of DNAm and air pollution exposure in 3 study populations consisting of White participants: Cooperative Health Research in the Region of Augsburg (KORA) F3, KORA F4 and the Normative Aging Study (NAS). 12 CpG sites were found associated with $PM_{2.5}$.[23] A recent EWAS study was performed using the NAS cohort in the Greater Boston area,[24] 29 significant CpGs were identified associated with $PM_{2.5}$ species after Bonferroni correction. Although some of these studies focused on the epigenome-wide association between $PM_{2.5}$ species and DNAm, to our knowledge, there have been no EWAS studies on specific $PM_{2.5}$ species and DNAm, and data regarding specific associations by race are lacking.

We sought to conduct an EWAS analysis to investigate DNAm changes associated with long-term exposure which was based on one-year average concentrations for $PM_{2.5}$ mass and species among Black and White adults in the Coronary Artery Risk Determinants in Young Adults (CARDIA) Study. We hypothesized that $PM_{2.5}$ species would be significantly associated with specific DNAm CpG sites and that these associations might differ for Black and White individuals.

## MATERIALS AND METHODS

### CARDIA

This study leverages a prospective multicenter cohort study – CARDIA -- with 5,115 young adults. It was designed to explore lifestyle and other potential factors that influence the progress of coronary heart disease risk factors.[25] Black or White women and men aged 18-30 years were enrolled

between 1985-1986 (study baseline; Year [Y] 0) with follow-up examinations at Y2 (1987-1988), Y5 (1990-1991), Y7 (1992-1993), Y10 (1995-1996), Y15 (2000-2001), Y20 (2005-2006), Y25 (2010-2011), Y30 (2015-2016), and Y35 (2021-2022). At baseline, participants were balanced on race, gender, age, and education at each of four centers in Birmingham, AL; Chicago, IL; Minneapolis, MN; and Oakland, CA. Details of the CARDIA study design have been previously published.[25,26] Participants are contacted at each visit for the ascertainment of information, e.g., self-identified race, gender, smoking status and the highest degree received, and physical measurements such as weight and skinfold fat were also used to collect the data at each visit.[25] Written informed consent was obtained from CARDIA participants at each exam, and institutional review boards have approved all CARDIA procedures.

### $PM_{2.5}$ exposure assessment

To evaluate long-term effects of air pollution on DNAm, we used one-year average concentrations for $PM_{2.5}$ and species (black carbon, ammonium, nitrate, organic matter, sulfate, soil, and sea salt).

The one-year average concentrations of exposures were estimated by calculating the means of the monthly values over the past year prior to the month when the participants underwent examination at Y15 according to geocoded residential addresses of CARDIA participants at Y15. The monthly $PM_{2.5}$ and species were extracted according to the grid cell that matched each participant's residential address from the V4.NA.02. The one-year average $PM_{2.5}$ mass and species exposures were calculated based on monthly $PM_{2.5}$ and species concentrations prior to each visit according to geocoded addresses. The concentrations of $PM_{2.5}$ and its species were estimated using the V4.NA.02 product, developed by the Dalhousie University Atmospheric Composition Analysis Group (DUACAG), with a spatial resolution of 1 km and monthly values.[27] This product was developed using a hybrid model based on the GEOS-Chem chemical transport model simulation and satellite observations of aerosol optical depth (AOD). A geographically weighted regression was applied to integrate these data and calibrated with ground-based observations. Through this model, the concentrations of $PM_{2.5}$ total mass, black carbon, ammonium, nitrate, organic matter, sulfate, soil, and sea salt were calculated. The estimated annual $PM_{2.5}$ concentrations are highly consistent with the ground-based observations ($R^2$ = 0.70).[27]

### Blood sample and DNA methylation processing

DNA was extracted from buffy coat fractions using the Puregene DNA extraction kit (Qiagen, Venlo, Netherlands) and stored at −70 °C. EZ DNA Methylation Kit (Zymo Research, Irvine, CA) was used to perform bisulfite conversion to the extracted genomic DNA. The DNA methylation data were measured using the Illumina Infinium MethylationEPIC BeadChip. The detailed methylation data processing pipeline used has been described previously.[28] Briefly, in the quality control step, we identified and excluded 6,209 CpGs with a detection rate <95% and 87 samples with >5% low-quality methylation measurements or extremely low bisulfite conversion probe intensity. After removing low-quality CpGs and samples, extreme outlier samples were excluded using Tukey's method based on total intensity values. The remaining samples underwent preprocessing with Enmix,[29] a model-based background correction method. Dye bias was corrected using RELIC,[30] and M or U intensities for Infinium I or II probes were separately quantile-normalized. Low-quality methylation values and extreme β-value outliers were set as missing. The final clean methylation working dataset contained 841,639 autosomal CpG sites (N=1042 from Y15, N=957 from Y20).

Table 1. Characteristics of the CARDIA participants at Exam Year (Y) 15 (2000-2001; N=1081)

| Variables | | |
|---|---|---|
| Age (Years), mean (SD[a]) | | 40.2 (3.5) |
| Body mass index (Kg/m$^2$), mean (SD) | | 28.5 (6.2) |
| Education, N (%) | High school | 386 (35.7) |
| | Associated degree | 132 (12.2) |
| | Bachelor | 303 (28.0) |
| | Master | 123(11.4) |
| | Doctorate | 18 (1.7) |
| | Professional | 57 (5.3) |
| | Other | 62 (5.7) |
| Smoking status, N (%) | Never | 684 (63.3) |
| | Former | 193 (17.9) |
| | Current | 204 (18.9) |
| Sex, N (%) | Male | 540 (50.0) |
| | Female | 541 (50.0) |
| Race, N (%) | Black participants | 432 (40.0) |
| | White participants | 649 (60.0) |
| Study center, N (%) | Birmingham | 256 (23.7) |
| | Chicago | 235 (21.7) |
| | Minneapolis | 294 (27.2) |
| | Oakland | 296 (27.4) |

a: Standard deviation.

### Statistical analysis

Since the $PM_{2.5}$ species data are only available since Y15, we restricted our analysis to participants with related DNAm information at Y15 and Y20. There

Table 2. Summary of $PM_{2.5}$ mass and species and correlation coefficient at Exam Year (Y) 15 (2001-2002; N=1081)

| | | | Correlation coefficient | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Pollutants | Mean ± SD[a](μg/m$^3$) | IQR(μg/m$^3$) | $PM_{2.5}$ | Black carbon | Ammonium | Nitrate | Organic matter | Sulfate | Soil | Sea salt |
| $PM_{2.5}$ | 13.45 ± 3.25 | 6.06 | 1 | 0.65 | 0.69 | -0.03 | 0.41 | 0.78 | 0.73 | -0.21 |
| Black carbon | 0.95 ± 0.39 | 0.56 | | 1 | 0.26 | 0.09 | 0.79 | 0.32 | 0.47 | 0.20 |
| Ammonium | 1.34 ± 0.54 | 0.69 | | | 1 | 0.24 | -0.20 | 0.76 | 0.62 | -0.69 |
| Nitrate | 2.12 ± 1.02 | 1.78 | | | | 1 | -0.06 | -0.37 | -0.20 | 0.07 |
| Organic matter | 4.75 ± 2.25 | 3.42 | | | | | 1 | 0.02 | 0.18 | 0.60 |
| Sulfate | 2.96 ± 1.50 | 2.78 | | | | | | 1 | 0.79 | -0.58 |
| Soil | 0.67 ± 0.27 | 0.44 | | | | | | | 1 | -0.47 |
| Sea salt | 0.86 ± 0.86 | 0.61 | | | | | | | | 1 |

a: Standard deviation.

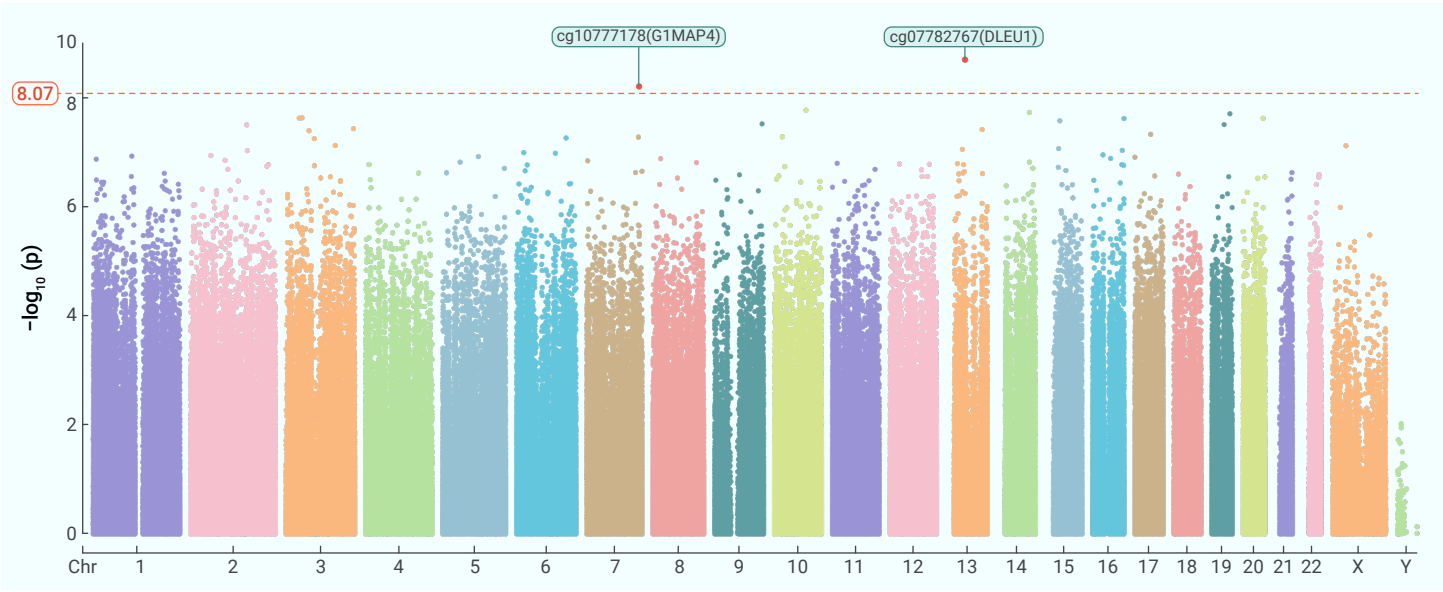



**Figure 1. Epigenome-wide associations of Ammonium species at Y15 (2000-2001) with DNA methylation at Y20 (2005-2006), displayed by Manhattan plots** The horizontal line represents the significance threshold with Bonferroni correction (P-value < 8.5e-09).

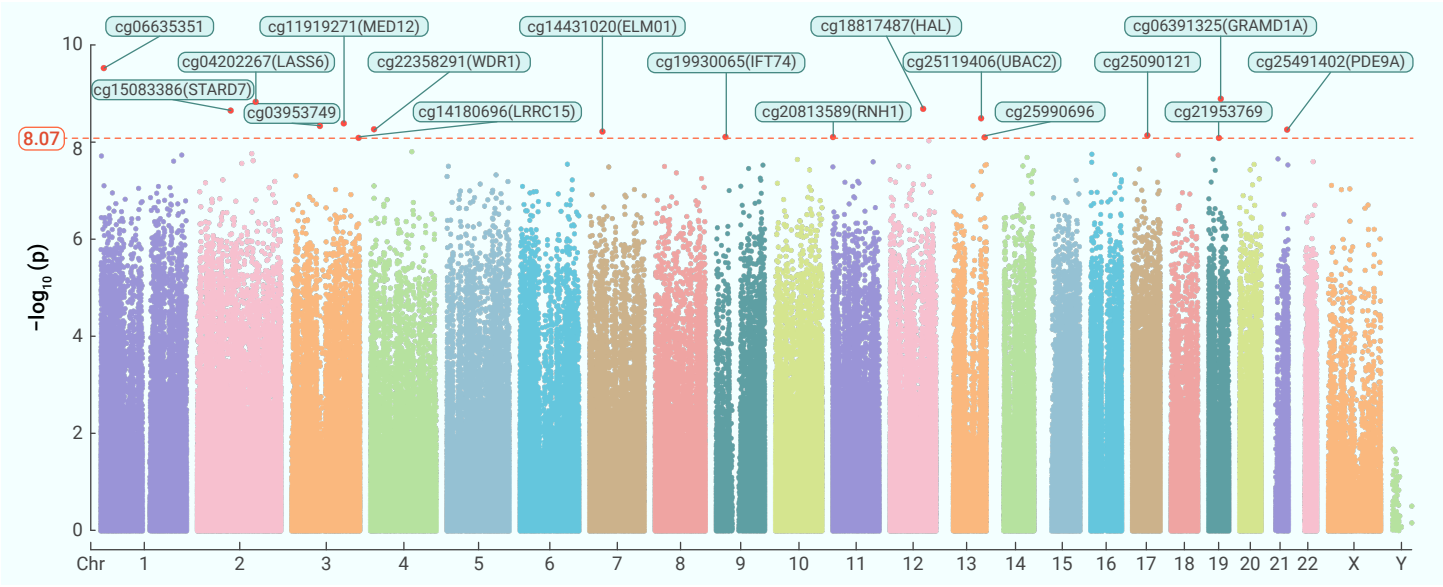



**Figure 2. Epigenome-wide associations of Nitrate species at Y15 (2000-2001) with DNA methylation at Y20 (2005-2006), displayed by Manhattan plots** The horizontal line represents the significance threshold with Bonferroni correction (P-value < 8.5e-09).

were 1,081 participants included in our EWAS analysis, of whom 850 have DNAm data at both Y15 and Y20; 158 and 73 participants have DNAm at Y15 only and Y20 only, respectively. We compared the characteristics of enrolled participants with those not included in our study at Y15 to check for potential selection bias.

We applied linear regression models on each CpG site while setting DNAm levels as our dependent variable and the exposure as independent variable. We conducted models for exposure to $PM_{2.5}$ mass concentrations and species separately. Since the DNAm at Y15 may not immediately reflect the influence of the exposures at Y15, our primary model examined the longitudinal association by considering exposures at Y15 and DNAm at Y20. We also modeled cross-sectional exposures at Y15 with DNAm at Y15 to examine whether there were changes in associations over time. The model is described below,

$$E(Y) = \beta_0 + \beta_1 * Spceies + \beta_2 * PM_{2.5} + \cdots + \varepsilon, \quad (1)$$

where Y is the methylation level of observation at a selected visit (Y15 or Y20), and species and $PM_{2.5}$ represent the one-year average values of $PM_{2.5}$ species and $PM_{2.5}$ total mass prior to the examination date at Y15. ε is the error term with Gaussian distribution $N(0, \sigma^2)$.

We adjusted for the following covariates in our model: age, body mass index (BMI), race (Black and White participants), gender (male, female), center (Birmingham, Chicago, Minneapolis, Oakland), smoking status (Never, Former, Current), highest degree the participant earned (high school, associate, bachelor, master, doctorate, professional, or other degree). We also adjusted for DNAm-based blood subtype proportion ($CD4^+$ T lymphocytes, $CD8^+$ T lymphocytes, natural killer cells, B cells, monocytes, and granulocytes) estimated using Houseman's method,[31] and eight surrogate variables derived from intensity data for non-negative internal control probes using principal components analysis.[28,29] Also, in order to assess the effect differences of $PM_{2.5}$ species exposures on DNAm from Y15 to Y20, we compared the magnitude of the coefficients of the same CpG sites at Y15 and Y20 with Y15 exposure to evaluate the changes.

To validate the identified CpG, we performed bootstrap with 1000 replicates for the identified CpG sites. In each bootstrap replicate, we got $t_1$, $t_2$,

Table 3. Genome-wide associations of $PM_{2.5}$ species at Y15 (2000-2001) with DNA methylation at Y20 (2005-2006): CpGs with Bonferroni-corrected statistical significance (P-value < 8.5e-09)

| Species | CpG | Chr[a] | Gene | Location | Coef[b] | SE[c] | P-value | Bootstrap P-value[d] |
|---|---|---|---|---|---|---|---|---|
| Ammonium | cg07782767 | 13 | DLEU1 | Body | 0.019 | 0.003 | 2.01E-09 | 0.001 |
| | cg10777178 | 7 | GIMAP4 | TSS200 | 0.025 | 0.004 | 6.24E-09 | 0.001 |
| Nitrate | cg06635351 | 1 | | | -0.020 | 0.003 | 2.96E-10 | 0.001 |
| | cg06391325 | 19 | GRAMD1A | Body | 0.008 | 0.001 | 1.27E-09 | 0.001 |
| | cg04202267 | 2 | LASS6 | Body | -0.011 | 0.002 | 1.47E-09 | 0.001 |
| | cg18817487 | 12 | HAL | TSS200 | -0.010 | 0.002 | 2.05E-09 | 0.001 |
| | cg15083386 | 2 | STARD7 | Body | -0.016 | 0.003 | 2.22E-09 | 0.001 |
| | cg25119406 | 13 | UBAC2 | Body | 0.012 | 0.002 | 3.20E-09 | 0.001 |
| | cg11919271 | 3 | MED12L; P2RY12 | Body; TSS1500 | -0.020 | 0.003 | 4.07E-09 | 0.001 |
| | cg03953749 | 3 | | | -0.017 | 0.003 | 4.61E-09 | 0.001 |
| | cg22358291 | 4 | WDR1 | Body | 0.013 | 0.002 | 5.38E-09 | 0.001 |
| | cg25491402 | 21 | PDE9A | Body;5'UTR | 0.010 | 0.002 | 5.47E-09 | 0.001 |
| | cg14431020 | 7 | ELMO1 | Body | -0.023 | 0.004 | 6.01E-09 | 0.001 |
| | cg25090121 | 17 | | | -0.009 | 0.002 | 7.20E-09 | 0.001 |
| | cg19930065 | 9 | IFT74 | Body | -0.010 | 0.002 | 7.73E-09 | 0.001 |
| | cg20813589 | 11 | RNH1 | 5'UTR | -0.016 | 0.003 | 7.79E-09 | 0.001 |
| | cg25990696 | 13 | | | -0.016 | 0.003 | 7.88E-09 | 0.001 |
| | cg14180696 | 3 | LRRC15 | 5'UTR | 0.011 | 0.002 | 8.09E-09 | 0.001 |
| | cg21953769 | 12 | | | 0.009 | 0.002 | 8.13E-09 | 0.001 |

a: Chromosome. b: Regression coefficients in linear models controlled for $PM_{2.5}$ mass, blood subtype proportion ($CD4^+$ T lymphocytes, $CD8^+$ T lymphocytes, natural killer cells, B cells, monocytes, granulocytes), age, BMI, race, gender, center, smoking status, education, and technical bias. c: Standard errors of the coefficients. d: To validate the identified CpG, we performed bootstrap with 1000 replicates for the identified CpG sites. The bootstrap p-value as p = (1+sum (t >= $t_0$))/(N+1), where t is the t-statistics at each replicate, $t_0$ is the t-statistics at original sample, and N is the replicate number. The statistical significance p-value level is 0.05.

$t_3$, ..., $t_n$, where ti is the t-statistic computed on the ith bootstrap sample, and where each bootstrap sample is sampled from the null distribution. Let $t_0$, which is the t-statistic we got in original sample, be the test statistic. We calculated the bootstrap p-value as p = (1+sum (t >= $t_0$))/(N+1), where N is the replicate number.

Given that our analysis involved a large number of statistical tests across many CpG site and seven $PM_{2.5}$ species, Bonferroni correction was used to account for multiple testing and p-value < 8.5e-09 (i.e., 0.05/ (841,639 *7)) was set as the study-wise significance level. Since our study included a biracial cohort with 40.0% (n = 432) Black participants, we further performed subgroup analyses for the $PM_{2.5}$-DNAm pairs identified from the primary analyses and assessed how these associations may differ between Black and White participants. We evaluated the significant CpG sites derived from the overall models by race groups and compared how the coefficients of these sites differed. R 4.1.0 was applied to perform all the analyses of this study.[32]

### Bioinformatics analysis

We matched unique gene symbols associated with the differentially methylated CpG sites to proximal genes based on the UCSC RefSeq annotation database. We further conducted pathway enrichment analysis on the annotated genes. Molecular function (MF) and biological process (BP) of Gene Ontology (GO) enrichment analysis, Reactome pathway analysis, KEGG pathway analysis, and Disease gene network were all used in our study to understand the biological function and relevance of those genes.[33-37] The enrichment p-values were calculated and adjusted using the Benjamini–Hochberg false discovery rate (BH-FDR) procedure.[38]

## RESULTS

Table 1 provides a descriptive summary of the 1,081 CARDIA participants included in this study at Y15; 540 (50.0%) were males, and 432 (40.0%) self-identified as Black participants. The average age at the baseline (Y0) was 25.2 years [standard deviation (SD) = 3.5], and 34.8% participants' education levels were less than or equal to high school. For smoking status, 684 participants (63.3%) never smoked, 193 (17.9%) were former smokers. Table S1 presents the race-stratified summary statistics of the 1,081 CARDIA participants included in this study at Y15. Table S2 shows characteristics of the 1,081 CARDIA participants stratified by study center at Y15.

In this paper, we consider complete case analyses. There were 4034 CARDIA participants at Y15 not enrolled in the study due to DNAm data not being measured. Table S3 in the supplement presents the characteristics of the CARDIA participants at Y15 not included in our study. We found that the participants without DNAm measures were similar at Y15 across most of the demographic variables. However, there were some differences in race and sex: 2205 individuals (54.7%) were Black participants, and 1787 (44.3%) were males.

Table 2 summarizes $PM_{2.5}$ mass and species and correlation statistics at Y15. The one-year average of $PM_{2.5}$ mass exposure at Y15 was 13.45 $\mu g/m^3$ (SD=3.25). Among all the correlations between $PM_{2.5}$ mass and species which ranged from -0.69 to 0.79, the correlation between black carbon and organic matter was the strongest (r=0.79), followed by sulfate and $PM_{2.5}$ mass, sulfate and ammonium.

Table 3 shows the top hits of epigenome-wide associations between $PM_{2.5}$ species at Y15 and DNAm at Y20 (Y15-Y20 model). After Bonferroni adjustment, there were two significant CpG sites associated with ammonium expo-

Table 4. The estimated coefficients of significant CpG sites for overall model in subgroup of Black and White participants

| | | | | | Black participants | | | White participants | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Species | CpG | Chr[a] | Gene | Location | Coef[b] | SE* | P-value | Coef[b] | SE[c] | P-value |
| Ammonium | cg07782767 | 13 | DLEU1 | Body | 0.028 | 0.006 | 9.98E-07 | 0.013 | 0.004 | 5.49E-04 |
| | cg10777178 | 7 | GIMAP4 | TSS200 | 0.031 | 0.007 | 1.28E-05 | 0.020 | 0.005 | 1.83E-04 |
| Nitrate | cg06635351 | 1 | | | -0.027 | 0.006 | 9.79E-07 | -0.017 | 0.004 | 1.38E-05 |
| | cg06391325 | 19 | GRAMD1A | Body | 0.009 | 0.002 | 1.42E-05 | 0.007 | 0.002 | 3.45E-06 |
| | cg04202267 | 2 | LASS6 | Body | -0.014 | 0.003 | 1.24E-05 | -0.008 | 0.002 | 7.59E-05 |
| | cg18817487 | 12 | HAL | TSS200 | -0.014 | 0.003 | 1.44E-06 | -0.008 | 0.002 | 2.90E-04 |
| | cg15083386 | 2 | STARD7 | Body | -0.018 | 0.004 | 6.95E-05 | -0.015 | 0.003 | 5.67E-06 |
| | cg25119406 | 13 | UBAC2 | Body | 0.018 | 0.004 | 1.21E-06 | 0.009 | 0.002 | 4.77E-04 |
| | cg11919271 | 3 | MED12L; P2RY12 | Body; TSS1500 | -0.025 | 0.005 | 5.05E-06 | -0.018 | 0.004 | 5.64E-05 |
| | cg03953749 | 3 | | | -0.023 | 0.005 | 3.69E-06 | -0.015 | 0.004 | 3.28E-05 |
| | cg22358291 | 4 | WDR1 | Body | 0.020 | 0.004 | 1.22E-07 | 0.009 | 0.003 | 8.92E-04 |
| | cg25491402 | 21 | PDE9A | Body;5'UTR | 0.015 | 0.003 | 2.85E-07 | 0.007 | 0.002 | 3.97E-04 |
| | cg14431020 | 7 | ELMO1 | Body | -0.027 | 0.006 | 3.76E-05 | -0.021 | 0.005 | 1.20E-05 |
| | cg25090121 | 17 | | | -0.013 | 0.003 | 3.13E-07 | -0.005 | 0.002 | 6.56E-03 |
| | cg19930065 | 9 | IFT74 | Body | -0.012 | 0.003 | 3.32E-05 | -0.009 | 0.002 | 4.83E-05 |
| | cg20813589 | 11 | RNH1 | 5'UTR | -0.023 | 0.004 | 3.82E-07 | -0.011 | 0.004 | 2.37E-03 |
| | cg25990696 | 13 | | | -0.023 | 0.004 | 4.23E-07 | -0.012 | 0.003 | 6.28E-04 |
| | cg14180696 | 3 | LRRC15 | 5'UTR | 0.015 | 0.003 | 7.45E-07 | 0.008 | 0.002 | 5.94E-04 |
| | cg21953769 | 12 | | | 0.013 | 0.003 | 4.17E-07 | 0.006 | 0.002 | 9.37E-04 |

a: Chromosome. b: Regression coefficients in linear models controlled for $PM_{2.5}$ mass, blood subtype proportion ($CD4^+$ T lymphocytes, $CD8^+$ T lymphocytes, natural killer cells, B cells, monocytes, granulocytes), age, BMI, race, gender, center, smoking status, education, and technical bias. c: Standard errors of the coefficients.

sure and 17 significant CpG sites for nitrate exposure. For other species, there were no significant CpG sites found. There was no significant CpG site in the cross-sectional model of $PM_{2.5}$ species at Y15 and DNAm at Y15 (Y15-Y15 model, data not shown).

Among the significant sites identified in the longitudinal Y15-Y20 model, we compared the magnitude of these associations to the corresponding ones from the Y15-Y15 model and found that the coefficients of significant CpG sites of the Y15-Y20 model increased by 2- to 48-fold (Table S4 in the Supplemental Document). Figures 1-2 provide the distribution of 841,639 CpG sites of the epigenome-wide association for ammonium and nitrate. The points above the horizontal line indicate the significant CpG sites.

Table 4 shows results for the longitudinal Y15-Y20 model stratified by race, and Table S5 presents the results of the interaction test of the EWAS model with an interaction term between $PM_{2.5}$ species and race; Black participants were the reference group in those significant CpG sites. We found that several interaction terms were significant which indicated the race difference in those sites was significant. For example, the significant interaction term of site cg07782767 for ammonium was -0.011 (SE=0.004); We conducted a further stratified model by race, revealing that the coefficient of the same site is 0.028 (SE=0.006) in Black participants and 0.013 (SE=0.004) in White participants. The negative values of significant interaction term indicated that the magnitudes of the coefficients of the relation between $PM_{2.5}$ species and DNAm were statistically lower among White participants compared to Black participants. The situation is similar for nitrate. The significant interaction term of site cg07782767 for nitrate was -0.006 (SE=0.004); Our further stratified model by race shows that the coefficient of the same site is 0.015 (SE=0.003) in Black participants and 0.007 (SE=0.002) in White participants.

Figures 3-4 signify the potential pathways involved in the DNAm associations. We annotated the differentially methylated CpG sites to genes based on the UCSC RefSeq annotation database. All the significant CpG sites for nitrate and ammonium were included in the pathway enrichment analysis. The results of the disease gene network pathway analysis for nitrate showed that the identified pathways were associated with basal cell cancer, cerebral aneurysm, and diseases of blood vessels.[39] The results of Reactome pathway analysis identified the pathways linked to platelet activation, signaling, aggregation and guanosine 3',5'-cyclic monophosphate effect, and these pathways were related to heart attack and stroke.[40]

## DISCUSSION

To the best of our knowledge, this study represents the first examination of the effects of long-term exposure to $PM_{2.5}$ species on DNAm specifically in Black participants using repeated DNAm measurements. We evaluated epigenome-wide associations between $PM_{2.5}$ species and DNAm in a large biracial cohort of young adults. We found that there were two CpG sites significantly associated with ammonium exposure and 17 sites with nitrate exposure. Moreover, the subgroup analyses of the associations between $PM_{2.5}$ species and these top DNAm identified suggested that the associations were stronger among Black participants than White participants, and the interaction test of $PM_{2.5}$ species and race indicated that race differences were significant among 3 CpG sites. The 17 CpG sites identified were associated with gene pathways of basal cancers, platelet activation, signaling and aggregation, and cGMP effect related to cognitive disorders.

Our EWAS analysis supports the association between $PM_{2.5}$ species and epigenome-wide DNAm. Furthermore, by leveraging bioinformatics analysis, we have gained valuable insights into the gene-level pathways that establish a potential link between exposure to $PM_{2.5}$ and its associated health outcomes. Both analyses provide robust evidence highlighting the significant role of DNAm in mediating the path between exposure to $PM_{2.5}$ and adverse

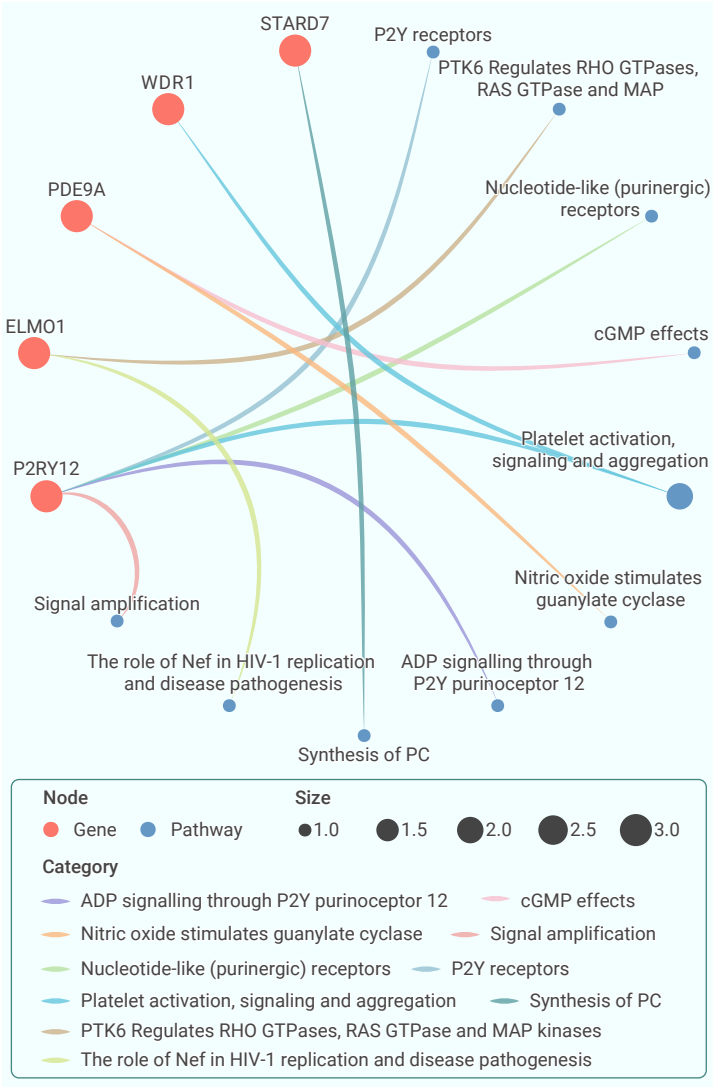



**Figure 3. Reactome pathway enrichment analysis for significant gene targets of CpGs in the overall model.**

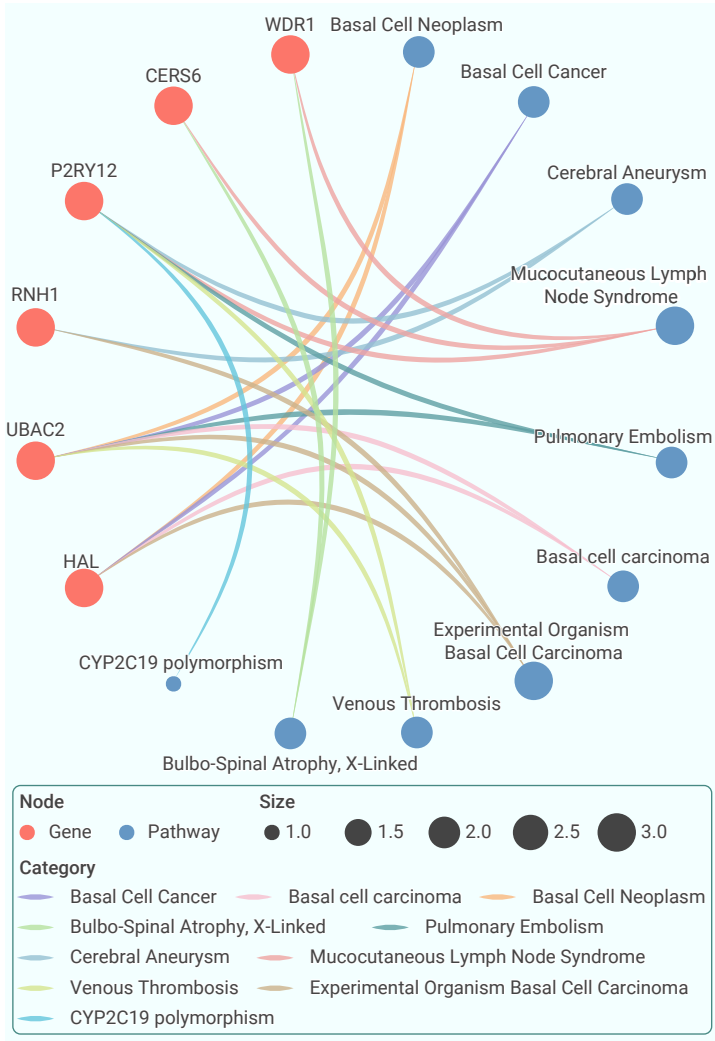



**Figure 4. Disease gene network pathway enrichment analysis for significant gene targets of CpGs in the overall model.**

health outcomes.

To our knowledge, to date, there have been only limited studies focused on the epigenome-wide association between $PM_{2.5}$ species and DNAm. A recent EWAS on DNAm and $PM_{2.5}$ species in a 450K platform by Dai et al. indicated that 29 CpG sites were identified at a Bonferroni adjusted significance level.[24] DNAm at *Schlafen Family Member 11* (*SLFN11*) cg10911913 was positively associated $PM_{2.5}$ species Fe, Ni, and V. Nevertheless, our study did not share the same CpG sites with theirs. The reason we have different results may be caused by the different $PM_{2.5}$ species we considered and the differences in the targeted populations. Compared to their study that used the following 11 species: Al, Ca, Cu, Fe, K, Na, Ni, S, Si, V, and Zn, we used black carbon, ammonium, nitrate, organic matter, sulfate, soil, and sea salt. Compared to that investigation in the Normative Aging Study (NAS), our CARDIA cohort also has a larger sample size (1,081 vs. 657). In addition, their study primarily consisted of elderly White men,[24] where we have measurements among Black participants. Moreover, the participants of our study were from four different urban sites in the US, where NAS only has the participants in Greater Boston Area.

One of the findings of our study was the positive association between the methylation at *GRAMD1A* (cg06391325) with the concentration of the $PM_{2.5}$ species nitrate. *GRAM Domain Containing 1A* (*GRAMD1A*) gene is localized to the endoplasmic reticulum.[41] The role of *GRAMD* family genes is to transport the phosphatidylinositol monophosphate from plasma membrane to the endoplasmic reticulum. During the transport process, *GRAMD1A* enrichment occurs and connect to the autophagy inhibitors autogramin-1 and 2, which are required for autophagosome biogenesis.[42,43] Current research suggests that expansion of hepatocellular carcinoma (HCC) stem cells and hepatocellular carcinoma tumor growth are promoted by overexpression of *GRAMD1A*.[44] Our finding of a positive epigenome-wide association between nitrate $PM_{2.5}$ exposure and DNAm suggests that higher exposure to nitrate may induce hypermethylation of CpGs within *GRAMD1A*. The hypothesis is consistent with the recent research showing an association between concentration of $PM_{2.5}$ air pollution and hepatocellular carcinoma incidence in the United States.[45]

We also observed that *UBAC2* (cg25119406) was positively associated with the exposure of nitrate. This gene can negatively regulate the Wnt signaling pathway and retrograde protein transport,[45] acting upstream of or within protein localization to endoplasmic reticulum.[46] These pathways may facilitate venous thrombosis and pulmonary embolism.[47]

In addition, we also found that exposure to nitrate is positively associated with the methylation of cg25491402 in gene *PDE9A*. The protein encoded by this gene catalyzes the hydrolysis of cAMP and cGMP to their corresponding monophosphates, and plays a role in signal transduction by regulating the intracellular concentration of these cyclic nucleotides, which has been shown to link with cognitive disorders.[48]

The present study is one of the largest investigations to date of epigenome-wide association between the $PM_{2.5}$ species and DNAm, which suggests that exposure to air pollution can induce change of methylation of certain genes which are related to inflammation, cancer, and biosynthesis in cells.[49] Our study has several strengths. First, we conducted a comparative analysis of the coefficients of epigenome-wide association between DNAm at Y20 and

DNAm at Y15, which provided insights into the dynamics of these associations. Second, we included seven $PM_{2.5}$ species in the study, which allowed us to investigate the species-specific associations between the exposures and DNAm. Third, the cohort we chose included participants from four geographically distinct urban areas across the US. The exposure to $PM_{2.5}$ mass and species can be influenced by measurement sites.[20] We were able to evaluate these associations across a wide range of $PM_{2.5}$ concentrations and species, which strengthened the robustness and generalizability of our findings. Fourth, the DNAm profiles were obtained during middle age, which is a crucial stage of life where DNAm may reflect cumulative environmental exposures from the past, as well as when the risks of certain diseases begin to emerge. Finally, since we included different race ethnicity groups in our analysis, we were able assess associations between DNAm and $PM_{2.5}$ species in diverse sociodemographic groups, and shed light on disparity in air pollution-related health issues. Since race is a social construct, it is highly likely that the disparities observed in this study may reflect other environmental exposures or factors related to social determinants of health that are also associated with race.

While our findings provide insights into the potential associations between $PM_{2.5}$ species and methylation patterns, it is important to acknowledge some limitations. Though we conducted EWAS for Y15 exposure to Y15 methylation and Y15 exposure to Y20 methylation, we cannot definitively establish a cause-and-effect relationship between $PM_{2.5}$ species and DNAm changes. $PM_{2.5}$ species data were available only at Y15, which restricted our ability to assess the time lag required for the observed effects on methylation to manifest. In addition, though we found that the associations were stronger among Black participants, there were only a few CpG sites that had significant interaction terms for $PM_{2.5}$ species and race. Further research on racial differences in associations between $PM_{2.5}$ species and DNAm is needed on larger cohorts, and should focus on other environmental exposures that may differ by race, socioeconomic status, or other factors. Moreover, this is only one level of gene regulation, whereas other forms of "epigenetic" regulation, like non-coding RNA, are likely to be relevant. In addition, the gene annotation here is mainly based on proximity in the genomic locations, although DNAm influences mostly local genes in nearby genomic regions, assessing the long-range regulation will be of interest for future studies to provide more comprehensive understanding of the associations found in this study.

## CONCLUSION

In summary, this study provides important evidence for the epigenome-wide association between $PM_{2.5}$ species and DNAm in a longitudinal cohort, which highlights the importance of epigenetic regulation in air pollution-related health problems. Future studies may help further elucidate the role of specific DNAm markers, together with other epigenetic mechanisms such as histone acetylation, for $PM_{2.5}$ induced adverse health effects.

## REFERENCES


1. Ostro, B., Roth, L., Malig, B., et al. (2009). The effects of fine particle components on respiratory hospital admissions in children. Environ. Health Perspect. **117**: 475–480. DOI: 10.1289/ehp.11848.
2. Hayes, R.B., Lim, C., Zhang, Y., et al. (2020). $PM_{2.5}$ air pollution and cause-specific cardiovascular disease mortality. Int. J. Epidemiol. **49**: 25-35. DOI: 10.1093/ije/dyz114.
3. Fu, P., Guo, X., Cheung, F.M.H., et al. (2019). The association between $PM_{2.5}$ exposure and neurological disorders: A systematic review and meta-analysis. Sci. Total Environ. **655**: 1240-1248. DOI: 10.1016/j.scitotenv.2018.11.218.
4. U.S. EPA. (2019). Supplement to the 2019 integrated science assessment for particulate matter. U.S. Environmental Protection Agency https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=354490.
5. Lozano, R., Naghavi, M., Foreman, K., et al. (2012). Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. Lancet **380**: 2095–2128. DOI: 10.1016/S0140-6736(12)61728-0.
6. Zhou, J., Ito, K., Lall, R., et al. (2011). Time-series analysis of mortality effects of fine particulate matter components in Detroit and Seattle. Environ. Health Perspect. **119**: 461–466. DOI: 10.1289/ehp.1002613.
7. Franklin, M., Koutrakis, P., and Schwartz, P. (2008). The role of particle composition on the association between $PM_{2.5}$ and mortality. Epidemiology **19**: 680-689. DOI: 10.1097/ede.0b013e3181812bb7.
8. Zanobetti, A., Franklin, M., Koutrakis, P., et al. (2009). Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environ. Health **8**: 58. DOI: 10.1186/1476-069X-8-58.
9. Madrigano, J., Baccarelli, A., Mittleman, M.A., et al. (2011). Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. Environ. Health Perspect. **119**: 977–982. DOI: 10.1289/ehp.1002773.
10. Chen, R., Qiao, L., Li, H., et al. (2015). Fine particulate matter constituents, nitric oxide synthase DNA methylation and exhaled nitric oxide. Environ. Sci. Technol. **49**: 11859–11865. DOI: 10.1021/acs.est.5b02527.
11. Chen, H., Li, S., Liu, J., et al. (2004). Chronic inorganic arsenic exposure induces hepatic global and individual gene hypomethylation: Implications for arsenic hepatocarcinogenesis. Carcinogenesis **25**: 1779–1786. DOI: 10.1093/carcin/bgh161.
12. Belinsky, S.A., Palmisano, W.A., Gilliland, F.D., et al. (2002). Aberrant promoter methylation in bronchial epithelium and sputum from current and former smokers. Cancer Res. **62**: 2370–2377.
13. Moen, E.L., Mariani, C.J., Zullow, H., et al. (2015). New themes in the biological functions of 5-methylcytosine and 5-hydroxymethylcytosine. Immunol. Rev. **263**: 36–49. DOI: 10.1111/imr.12242.
14. Rider, C.F., and Carlsten, C. (2019). Air pollution and DNA methylation: Effects of exposure in humans. Clin. Epigenetics **11**: 131. DOI: 10.1186/s13148-019-0713-2.
15. Poursafa, P., Kamali, Z., Fraszczyk, E., et al. (2022). DNA methylation: A potential mediator between air pollution and metabolic syndrome. Clin. Epigenetics **14**: 82. DOI: 10.1186/s13148-022-01301-y.
16. Rakyan, V.K., Down, T.A., Balding, D.J., et al. (2011). Epigenome-wide association studies for common human diseases. Nat. Rev. Genet. **12**: 529–541. DOI: 10.1038/nrg3000.
17. Plusquin, M., Guida, F., Polidoro, S., et al. (2017). DNA methylation and exposure to ambient air pollution in two prospective cohorts. Environ. Int. **108**: 127–136. DOI: 10.1016/j.envint.2017.08.006.
18. Lee, M.K., Xu, C.J., Carnes, M.U., et al. (2019). Genome-wide DNA methylation and long-term ambient air pollution exposure in Korean adults. Clin. Epigenetics **11**: 37. DOI: 10.1186/s13148-019-0635-z.
19. Callahan, C.L., Bonner, M.R., Nie, J., et al. (2018). Lifetime exposure to ambient air pollution and methylation of tumor suppressor genes in breast tumors. Environ. Res. **161**: 418–424. DOI: 10.1016/j.envres.2017.11.040.
20. Gondalia, R., Baldassari, A., Holliday, K.M., et al. (2019). Methylome-wide association study provides evidence of particulate matter air pollution-associated DNA methylation. Environ. Int. **132**: 104723. DOI: 10.1016/j.envint.2019.03.071.
21. Ho, S.M. (2010). Environmental epigenetics of asthma: An update. J. Allergy Clin. Immunol. **126**: 453–465. DOI: 10.1016/j.jaci.2010.07.030.
22. Gruzieva, O., Xu, C.J., Yousefi, P., et al. (2019). Prenatal particulate air pollution and DNA methylation in newborns: An epigenome-wide meta-analysis. Environ. Health Perspect. **127**: 57012. DOI: 10.1289/EHP4522.
23. Panni, T., Mehta, A.J., Schwartz, J.D., et al. (2016). Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. Environ. Health Perspect. **124**: 983–990. DOI: 10.1289/ehp.1509966.
24. Dai, L., Mehta, A., Mordukhovich, I., et al. (2017). Differential DNA methylation and $PM_{2.5}$ species in a 450K epigenome-wide association study. Epigenetics **12**: 139-148. DOI: 10.1080/15592294.2016.1271853.
25. Friedman, G.D., Cutter, G.R., Donahue, R.P., et al. (1988). CARDIA: Study design, recruitment, and some characteristics of the examined subjects. J. Clin. Epidemiol. **41**: 1105–1116. DOI: 10.1016/0895-4356(88)90080-7.
26. Cutter, G.R., Burke, G.L., Dyer, A.R., et al. (1991). Cardiovascular risk factors in young adults. The CARDIA baseline monograph. Control. Clin. Trials **12**: 1S-77S. DOI: 10.1016/0197-2456(91)90002-4.
27. van, D.A., Martin, R.V., Li, C., et al. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environ. Sci. Technol. **53**: 2595–2611. DOI: 10.1021/acs.est.8b06392.
28. Zheng, Y., Joyce, B.T., Hwang, S.J., et al. (2022). Association of cardiovascular health through young adulthood with genome-wide DNA methylation patterns in midlife: The CARDIA study. Circulation **146**: 94–109. DOI: 10.1161/CIRCULATIONAHA.121.055484.
29. Xu, Z., Niu, L., Li, L., et al. (2016). Enmix: A novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic Acids Res. **44**: e20. DOI: 10.1093/nar/gkv907.
30. Xu, Z., Langie, S.A., De, B.P., et al. (2017). RELIC: A novel dye-bias correction method for Illumina Methylation BeadChip. BMC Genomics **18**: 4. DOI: 10.1186/s12864-016-3426-3.
31. Houseman, E.A., Accomando, W.P., Koestler, D.C., et al. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics **13**: 86. DOI: 10.1186/1471-2105-13-86.
32. R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.
33. Wu, T., Hu, E., Xu, S., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation **2**: 100141. DOI: 10.1016/j.xinn.2021.100141.
34. Kanehisa, M., Goto, S., Kawashima, S., et al. (2002). The KEGG databases at

GenomeNet. Nucleic Acids Res. **30**: 42–46. DOI: 10.1093/nar/30.1.42.
35. Carbon, S., Ireland, A., Mungall, C.J., et al. (2009). AmiGO: Online access to ontology and annotation data. Bioinformatics **25**: 288–289. DOI: 10.1093/bioinformatics/btn615.
36. Fabregat, A., Korninger, F., Viteri, G., et al. (2018). Reactome graph database: Efficient access to complex pathway data. PloS Comput. Biol. **14**: e1005968. DOI: 10.1371/journal.pcbi.1005968.
37. Queralt-Rosinach, N., Piñero, J., Bravo, À., et al. (2016). DisGeNET-RDF: Harnessing the innovative power of the Semantic Web to explore the genetic basis of diseases. Bioinformatics **32**: 2236–2238. DOI: 10.1093/bioinformatics/btw214.
38. Subramanian, A., Tamayo, P., Mootha, V.K., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. **102**: 15545–15550. DOI: 10.1073/pnas.0506580102.
39. Estevez, B., and Du, X. (2017). New concepts and mechanisms of platelet activation signaling. Physiology (Bethesda) **32**: 162–177. DOI: 10.1152/physiol.00020.2016.
40. Münzel, T., Feil, R., Mülsch, A., et al. (2003). Physiology and pathophysiology of vascular signaling controlled by guanosine 3',5'-cyclic monophosphate-dependent protein kinase [corrected]. Circulation **108**: 2172–2183. DOI: 10.1161/01.CIR.0000094403.78467.C3.
41. Besprozvannaya, M., Dickson, E., Li, H., et al. (2018). GRAM domain proteins specialize functionally distinct ER-PM contact sites in human cells. Elife **7**: e31019. DOI: 10.7554/eLife.31019.
42. Wu, Y.W., and Waldmann, H. (2019). Toward the role of cholesterol and cholesterol transfer protein in autophagosome biogenesis. Autophagy **15**: 2167–2168. DOI: 10.1080/15548627.2019.1666595.
43. Laraia, L., Friese, A., Corkery, D.P., et al. (2019). The cholesterol transfer protein GRAMD1A regulates autophagosome biogenesis. Nat. Chem. Biol. **15**: 710–720. DOI: 10.1038/s41589-019-0307-5.
44. Fu, B., Meng, W., Zhao, H., et al. (2016). GRAM domain-containing protein 1A (GRAMD1A) promotes the expansion of hepatocellular carcinoma stem cell and hepatocellular carcinoma growth through STAT5. Sci. Rep. **6**: 31963. DOI: 10.1038/srep31963.
45. VoPham, T., Bertrand, K.A., Tamimi, R.M., et al. (2018). Ambient $PM_{2.5}$ air pollution exposure and hepatocellular carcinoma incidence in the United States. Cancer Causes Control **29**: 563-572. DOI: 10.1007/s10552-018-1036-x.
46. Gu, C., Zhao, K., Zhou, N., et al. (2020). UBAC2 promotes bladder cancer proliferation through BCRC-3/miRNA-182-5p/p27 axis. Cell Death Dis. **11**: 733. DOI: 10.1038/s41419-020-02935-7.
47. Wijk, S.C., Prabhala, P., Michaliková, B., et al. (2021). Human primary airway basal cells display a continuum of molecular phases from health to disease in chronic obstructive pulmonary disease. Am. J. Respir. Cell Mol. Biol. **65**: 103–113. DOI: 10.1165/rcmb.2020-0464OC.
48. Harms, J.F., Menniti, F.S., and Schmidt, C.J. (2019). Phosphodiesterase 9A in brain regulates cGMP signaling independent of nitric-oxide. Front. Neurosci. **13**: 837. DOI: 10.3389/fnins.2019.00837.
49. Rudez, G., Janssen, N.A., Kilinc, E., et al. (2009). Effects of ambient air pollution on hemostasis and inflammation. Environ. Health Perspect. **117**: 995–1001. DOI: 10.1289/ehp.0800437.

## FUNDING AND ACKNOWLEDGMENTS

The Coronary Artery Risk Development in Young Adults Study (CARDIA) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the University of Alabama at Birmingham (75N92023D00002 & 75N92023D00005), Northwestern University (75N92023D00004), University of Minnesota (75N92023D00006), and Kaiser Foundation Research Institute (75N92023D00003). This work was also supported by the National Institute on Aging (R01AG081244 and R01AG069120), the American Heart Association grant (19TPA34830085; PI, KZ), and the Empire Innovation Program (EIP) of the State University of New York (PI, KZ). This manuscript has been reviewed by CARDIA for scientific content. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## AUTHOR CONTRIBUTIONS

J.Y., Y.Z., L.H., and K.Z.: Conceptualization. J.Y., Y.Z., J.S., K.H., L.H., and K.Z.: Methodology. J.Y., Y.Z., and K.Z.: Data curation. J.Y.: Formal analysis. J.Y.: Visualization. J.Y.: Writing - Original Draft. J.Y.: Writing - Review & Editing. Y.Z., J.S., K.H., R.S., V.M., S.S., and D.L.: Critical review & editing. L.H., and K.Z.: Funding acquistion. K.Z.: Supervision. All authors have read and approved the final draft of the paper.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## ETHICAL STATEMENT AND PATIENT CONSENT

This study (No: 21X041) was reviewed and approved by the Institutional Review Board at the University at Albany, State University of New York. Written informed consent was obtained from CARDIA participants at each exam.

## DATA AND CODE AVAILABILITY

Data and materials used for this analysis are available on request to the CARDIA study at https://www.cardia.dopm.uab.edu, and data are also publicly available at the Biologic Specimen and Data Repository Information Coordinating Center (https://biolincc.nhlbi.nih.gov/studies/cardia/).

## SUPPLEMENTAL INFORMATION

It can be found online at https://doi.org/10.59717/j.xinn-med.2024.100061

## LEAD CONTACT WEBSITE

Kai Zhang: https://www.albany.edu/sph/faculty/kai-zhang