Manuscript Details

Manuscript number	ENVINT_2018_2800_R1
Title	Methylome-wide association study provides evidence of particulate matter air pollution-associated DNA methylation
Article type	Research Paper

Abstract

Background: DNA methylation (DNAm) may contribute to processes that underlie associations between air pollution and poor health. Therefore, our objective was to evaluate associations between DNAm and ambient concentrations of particulate matter (PM) ≤2.5, ≤10, and 2.5-10 μm in diameter (PM2.5; PM10; PM2.5-10). Methods: We conducted a methylome-wide association study among twelve cohort- and race/ethnicity-stratified subpopulations from the Women's Health Initiative and the Atherosclerosis Risk in Communities study (n = 8,397; mean age: 61.5 years; 83% female; 45% African American; 9% Hispanic/Latino American). We averaged geocoded address-specific estimates of daily and monthly mean PM concentrations over 2, 7, 28, and 365 days and 1 and 12 months before exams at which we measured leukocyte DNAm in whole blood. We estimated subpopulation-specific, DNAm-PM associations at approximately 485,000 Cytosine-phosphate-Guanine (CpG) sites in multi-level, linear, mixed-effects models. We combined subpopulation- and site-specific estimates in fixed-effects, inverse variance-weighted meta-analyses, then for associations that exceeded methylome-wide significance and were not heterogeneous across subpopulations ($P < 1.0x10-7$; PCochran's $Q > 0.10$), we characterized associations using publicly accessible genomic databases and attempted replication in the Cooperative Health Research in the Region of Augsburg (KORA) study. Results: Analyses identified significant DNAm-PM associations at three CpG sites. Twenty-eight-day mean PM10 was positively associated with DNAm at cg19004594 (chromosome 20; MATN4; $P = 3.33x10-8$). One-month mean PM10 and PM2.5-10 were positively associated with DNAm at cg24102420 (chromosome 10; ARPP21; $P = 5.84x10-8$) and inversely associated with DNAm at cg12124767 (chromosome 7; CFTR; $P = 9.86x10-8$). The PM-sensitive CpG sites mapped to neurological, pulmonary, endocrine, and cardiovascular disease-related genes, but DNAm at those sites was not associated with gene expression in blood cells and did not replicate in KORA. Conclusions: Ambient PM concentrations were associated with DNAm at genomic regions potentially related to poor health among racially, ethnically and environmentally diverse populations of U.S. women and men. Further investigation is warranted to uncover mechanisms through which PM-induced epigenomic changes may cause disease.

Keywords	particulate matter; dna methylation; epigenetics; air pollution; epigenome-wide association study
Taxonomy	Air Pollution Health Impact, Environmental Epidemiology
Corresponding Author	Rahul Gondalia
Corresponding Author's Institution	University of North Carolina Department of Epidemiology
Order of Authors	Rahul Gondalia, Antoine Baldassari, Kathryn Holliday, Anne Justice, Raul Mendez Giraldez, James Stewart, Duanping Liao, Jeff Yanosky, Kasey Brennan, Stephanie Engel, Kristina Jordahl, Elizabeth Kennedy, Cavin Ward-Caviness, Kathrin Wolf, Josef Cyrys, Parveen Bhatti, Steve Horvath, Themistocles Assimes, James Pankow, Ellen Demerath, Weihua Guan, Myriam Fornage, Jan Bressler, Kari North, Karen Conneely, Yun Li, Lifang Hou, Andrea Baccarelli, Eric Whitsel
Suggested reviewers	Radhika Dhingra, Colleen Sitlani, Gregory Wellenius, John Hartman

Submission Files Included in this PDF

File Name [File Type]

PM-DNAm_CoverLetter_2019_03_27.docx [Cover Letter]

pm-mwas_response_2019_03_27.docx [Response to Reviewers]

PM-DNAm_2019_03_27.docx [Revised Manuscript with Changes Marked]

PM-DNAm_Highlights_2019_03_27.docx [Highlights]

PM-DNAm_abstract_2019_03_27_clean.docx [Abstract]

PM-DNAm_2019_03_27_clean.docx [Manuscript File]

manhattan_plots.pdf [Supplementary Material]

PM-DNAm_2019_03_27_suppl_clean.docx [Supplementary Material]

Submission Files Not Included in this PDF

File Name [File Type]

TableS3_2018_12_01.xlsx [Supplementary Material]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Research Data Related to this Submission

There are no linked research data sets for this submission. The following reason is given:
Data from the ARIC and WHI study are available on request: https://www2.cscc.unc.edu/aric/distribution-agreements. https://www.whi.org/researchers/SitePages/Write%20a%20Paper.aspx.

Title: Methylome-wide association study provides evidence of particulate matter air pollution-associated DNA methylation

Highlights

- DNA methylation (DNAm) may underlie processes linking air pollution and poor health
- We conducted a methylome-wide association study of 8,397 women and men with daily and monthly estimates of $PM_{2.5}$, PM_{10}, and $PM_{2.5-10}$ data
- Mid-duration PM_{10} and $PM_{2.5-10}$ were significantly associated with methylation at three CpG sites, but associations did not replicate in an independent population
- The three PM-sensitive CpG sites mapped to neurological, pulmonary, endocrine, and cardiovascular-disease related genes

Abstract

Background: DNA methylation (DNAm) may contribute to processes that underlie associations between air pollution and poor health. Therefore, our objective was to evaluate associations between DNAm and ambient concentrations of particulate matter (PM) ≤2.5, ≤10, and 2.5-10 μm in diameter ($PM_{2.5}$; PM_{10}; $PM_{2.5-10}$).

Methods: We conducted a methylome-wide association study among twelve cohort- and race/ethnicity-stratified subpopulations from the Women's Health Initiative and the Atherosclerosis Risk in Communities study (n = 8,397; mean age: 61.5 years; 83% female; 45% African American; 9% Hispanic/Latino American). We averaged geocoded address-specific estimates of daily and monthly mean PM concentrations over 2, 7, 28, and 365 days and 1 and 12 months before exams at which we measured leukocyte DNAm in whole blood. We estimated subpopulation-specific, DNAm-PM associations at approximately 485,000 Cytosine-phosphate-Guanine (CpG) sites in multi-level, linear, mixed-effects models. We combined subpopulation- and site-specific estimates in fixed-effects, inverse variance-weighted meta-analyses, then for associations that exceeded methylome-wide significance and were not heterogeneous across subpopulations ($P < 1.0\text{x}10^{-7}$; $P_{Cochran's\ Q} > 0.10$), we characterized associations using publicly accessible genomic databases and attempted replication in the Cooperative Health Research in the Region of Augsburg (KORA) study.

Results: Analyses identified significant DNAm-PM associations at three CpG sites. Twenty-eight-day mean PM_{10} was positively associated with DNAm at cg19004594 (chromosome 20; *MATN4; P* = $3.33\text{x}10^{-8}$). One-month mean PM_{10} and $PM_{2.5-10}$ were positively associated with DNAm at cg24102420 (chromosome 10; *ARPP21; P* = $5.84\text{x}10^{-8}$) and inversely associated with DNAm at cg12124767 (chromosome 7; *CFTR*; *P* = $9.86\text{x}10^{-8}$). The PM-sensitive CpG sites mapped to neurological, pulmonary, endocrine, and cardiovascular disease-related genes, but DNAm at those sites was not associated with gene expression in blood cells and did not replicate in KORA.

Conclusions: Ambient PM concentrations were associated with DNAm at genomic regions potentially related to poor health among racially, ethnically and environmentally diverse populations of U.S. women and men. Further investigation is warranted to uncover mechanisms through which PM-induced epigenomic changes may cause disease.

Title: Methylome-wide association study provides evidence of particulate matter air pollution-associated DNA methylation

Authors

Rahul Gondalia[a], Antoine Baldassari[a], Katelyn M Holliday[a,b], Anne E Justice[a,c], Raúl Méndez-Giráldez[a], James D Stewart[a], Duanping Liao[d], Jeff D Yanosky[d], Kasey JM Brennan[e], Stephanie M Engel[a], Kristina M Jordahl[f], Elizabeth Kennedy[g], Cavin K Ward-Caviness[h], Kathrin Wolf[i], Josef Cyrysj[i], Parveen Bhatti[f], Steve Horvath[j,k], Themistocles L Assimes[l], James S Pankow[m], Ellen W Demerath[m], Weihua Guan[n], Myriam Fornage[o], Jan Bressler[p] Kari E North[a,q], Karen N Conneely[r], Yun Li[s,t,u], Lifang Hou[v,w], Andrea A Baccarelli[e], Eric A Whitsel[a,x]

Affiliations

[a]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

[b]Department of Community and Family Medicine, Duke University School of Medicine, Durham, NC

[c]Geisinger Health System, Danville, PA, USA

[d]Division of Epidemiology, Department of Public Health Sciences, Pennsylvania State University College of Medicine, Hershey, PA, USA

[e]Laboratory of Environmental Epigenetics, Departments of Environmental Health Sciences and Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

[f]Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA

[g]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

[h]Environmental Public Health Division, National Health and Environmental Effects Research Laboratory, 104 Mason Farm Rd, Chapel Hill, NC 27514

[i]Institute of Epidemiology, Helmholtz Zentrum München, Ingolstaedter Landstrasse 1, Neuherberg, Germany 85764

[j]Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

[k]Biostatistics, School of Public Health, University of California Los Angeles, Los Angeles, CA, , USA

[l]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

[m]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, USA

[n]Division of Biostatistics, University of Minnesota, Minneapolis, MN, USA

[o]Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA
[p]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA
[q]Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC, USA
[r]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA
[s]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA
[t]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA
[u]Department of Computer Science, University of North Carolina, Chapel Hill, NC, USA
[v]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University Chicago, Evanston, IL, USA
[w]Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University Chicago, Evanston, IL, USA
[x]Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC, USA

Corresponding Author:
Rahul Gondalia, MPH
123 W. Franklin St.
Chapel Hill, North Carolina 27516, USA
Phone: 304-210-7823
Email**:** **rahgonda@unc.edu**

Abstract

Background: DNA methylation (DNAm) may contribute to processes that underlie associations between air pollution and poor health. Therefore, our objective was to evaluate associations between DNAm and ambient concentrations of particulate matter (PM) ≤2.5, ≤10, and 2.5-10 μm in diameter ($PM_{2.5}$; PM_{10}; $PM_{2.5-10}$).

Methods: We conducted a methylome-wide association study among twelve cohort- and race/ethnicity-stratified subpopulations from the Women's Health Initiative and the Atherosclerosis Risk in Communities study (n = 8,397; mean age: 61.5 years; 83% female; 45% African American; 9% Hispanic/Latino American). We averaged geocoded address-specific estimates of daily and monthly mean PM concentrations over 2, 7, 28, and 365 days and 1 and 12 months before exams at which we measured leukocyte DNAm in whole blood. We estimated subpopulation-specific, DNAm-PM associations at approximately 485,000 Cytosine-phosphate-Guanine (CpG) sites in multi-level, linear, mixed-effects models. We combined subpopulation- and site-specific estimates in fixed-effects, inverse variance-weighted meta-analyses, then for associations that exceeded methylome-wide significance and were not heterogeneous across subpopulations ($P < 1.0\text{x}10^{-7}$; $P_{Cochran's\ Q} > 0.10$), we characterized associations using publicly accessible genomic databases and attempted replication in the Cooperative Health Research in the Region of Augsburg (KORA) study.

Results: Analyses identified significant DNAm-PM associations at three CpG sites. Twenty-eight-day mean PM_{10} was positively associated with DNAm at cg19004594 (chromosome 20; *MATN4;* $P = 3.33\text{x}10^{-8}$). One-month mean PM_{10} and $PM_{2.5-10}$ were positively associated with DNAm at cg24102420 (chromosome 10; *ARPP21;* $P = 5.84\text{x}10^{-8}$) and inversely associated with DNAm at cg12124767 (chromosome 7; *CFTR*; $P = 9.86\text{x}10^{-8}$). The PM-sensitive CpG sites mapped to neurological, pulmonary, endocrine, and cardiovascular disease-related genes, but DNAm at those sites was not associated with gene expression in blood cells and did not replicate in KORA.

Conclusions: Ambient PM concentrations were associated with DNAm at genomic regions potentially related to poor health among racially, ethnically and environmentally diverse populations of U.S. women and men. Further investigation is warranted to uncover mechanisms through which PM-induced epigenomic changes may cause disease.

Keywords: particulate matter; DNA methylation; epigenetics; air pollution; epigenome-wide association study

Abbreviations: AA, African American; AV, annual visit; ARIC, Atherosclerosis Risk in Communities; AS311, Ancillary Study 311; AQS, United States Environmental Protection Agency Air Quality System; BAA23, Broad Agency Award 23; CI, confidence interval; CpG, Cytosine-phosphate-Guanine; CT, Clinical Trial; DNAm, deoxyribonucleic acid methylation; CVD, cardiovascular disease; EA, European American; eFORGE, Functional element Overlap analysis of Regions; EMPC, Epigenetic Mechanisms of PM-Mediated CVD Risk; FDR, false discovery rate; GTP, Grady Trauma Project; GWAS, genome-wide association study; HLA, Hispanic/Latino American; KORA, Cooperative Health Research in the Region Augsburg study; LLS, Long Life Study; LMM, linear mixed models; MESA, Multi-Ethnic Study of Atherosclerosis ; MICE, multiple imputation by chained equations; MWAS, methylome-wide association study; NAAQS, National Ambient Air Quality Standards; OS, Observational Study; PE, prediction error; PM_{10}, PM < 10 μm in diameter; $PM_{2.5}$, PM < 2.5 μm in diameter; $PM_{2.5\text{-}10}$, PM > 2.5 and < 10 μm in diameter; QQ, quantile-quantine; RMSS, root mean square standardized; SD, standard deviation; SE, standard error; SPE, standardized prediction error; WHI, Women's Health Initiative

1. Introduction

Ambient particulate matter (PM) air pollution is a modifiable exposure that has been consistently associated with morbidity and mortality (Cohen et al. 2017; Di et al. 2017; Miller et al. 2007) attributed to cardiovascular disease (Brook et al. 2004; Brook et al. 2010), respiratory disease (Dominici et al. 2006; Gan et al. 2013; Laumbach and Kipen 2012), and lung cancer (Pope et al. 2002; Raaschou-Nielsen et al. 2013). Despite the ubiquity of air pollution exposure and the continued population burden of PM (Cohen et al. 2017), the causal mechanisms underlying PM associations with poor health have not been adequately investigated.

One such mechanism could involve methylation of deoxyribonucleic acids (DNAm), conventionally measured at Cytosine-phosphate-Guanine (CpG) sites. DNAm is a heritable, but dynamic epigenetic modification that can influence gene expression without altering the DNA sequence (Clouaire and Stancheva 2008; Neidhart 2016) and may be central to mediation of PM-associated disease risk (Baccarelli et al. 2010; Bollati and Baccarelli 2010; Zhong et al. 2016). Indeed, PM exposure has been implicated in whole blood DNAm near candidate genes involved in inflammation, oxidative stress, coagulation and vasoconstriction (Bellavia et al. 2013; Chen et al. 2016; Chen et al. 2015; Tarantini et al. 2009; Tarantini et al. 2013), abnormalities of which have established associations with cardiovascular and respiratory disease. A few studies have agnostically evaluated DNAm associations with PM on a methylome-wide scale (de F.C. Lichtenfels et al. 2018; Panni et al. 2016; Plusquin et al. 2017), but none have done so in large, sociodemographically and environmentally diverse, well-characterized populations of adult women and men.

The present study therefore examined methylome-wide associations between DNAm and ambient concentrations of PM ≤ 2.5, ≤ 10, and 2.5-10 μm in diameter ($PM_{2.5}$, PM_{10}, and $PM_{2.5-10}$) within the Women's Health Initiative (WHI) and the Atherosclerosis Risk in Communities study (ARIC) cohorts, and their replication in subpopulations of the Cooperative Health Research in the Region Augsburg (KORA) study.

2. Methods

2.1. Study design and populations

The study included 8,397 consenting participants from subpopulations within the WHI and ARIC cohorts who had available peripheral blood leukocyte DNA.

The WHI is a multicenter prospective study of risk factors for cardiovascular disease (CVD), cancer, osteoporotic fractures, and other causes of morbidity and mortality among postmenopausal women (Anderson et al. 2003; NIH 1998). Between 1993 and 1998, women aged 50-79 years from forty WHI clinical centers throughout the United States (US) were enrolled in the Clinical Trials (CT) (n = 68,132) or Observational Study (OS) (n = 93,676). All WHI participants completed a screening visit (SV). CT participants also completed an annual visit (AV) at one, three, six, and nine years after randomization (AV1, AV3, AV6, AV9), and OS participants three years after enrollment (AV3). An additional visit of CT and OS participant subsets occurred between 2011 and 2012 (ranging from 14 to 19 years after enrollment) as part of the WHI Long Life Study (LLS) (Anderson and LaCroix).

For the current study, WHI participants were drawn from three ancillary studies: *Epigenetic Mechanisms of PM-Mediated CVD Risk* (WHI-EMPC) (Whitsel), *Broad Agency Announcement 23* (WHI-BAA23) (Assimes et al.) and *Ancillary Study 311* (WHI-AS311) (Jordahl et al. 2018). WHI-EMPC is a study of epigenetic mechanisms underlying associations between ambient PM air pollution and CVD within the WHI CT. From this population, DNAm was measured in 2,200 randomly selected participants (stage 1: SV, AV3, or AV6), remeasured in 200 participants at a second visit (stage 2: AV3 or AV6), and remeasured again in 43 participants at a third visit among those who participated in the WHI Long Life Study (stage 3: LLS), yielding 2,443 total observations. WHI-BAA23, also known as *Integrative Genomics and Risk of CHD and Related Phenotypes in the Women's Health Initiative*, is a case-control study of coronary heart disease within the WHI CT (n = 1,546) and OS (n = 442). By design, WHI-BAA23 oversampled African Americans and Hispanic/Latino Americans and required all participants to have undergone genome-wide genotyping and profiling of seven cardiovascular disease biomarkers. DNAm was measured in blood collected at the SV, before the incidence of coronary heart disease. WHI-AS311 is a matched case-control study of bladder cancer among women within the WHI CT (n = 405) and OS (n = 455). Bladder cancer cases were matched to controls based on enrollment year, age at enrollment, follow-up time, and DNAm extraction method. DNAm was measured in blood collected at the SV, before the incidence of bladder cancer.

ARIC is a community-based prospective study of atherosclerosis and its clinical outcomes in four US communities: Washington County, Maryland; Forsyth County, North Carolina; selected suburbs of Minneapolis, Minnesota; and Jackson, Mississippi (ARIC Investigators 1989). Enrollment in 1987-1989

(Visit 1) was followed by five subsequent visits (Visits 2-6) between 1990-2017. The present study included all 2,796 African Americans from Forsyth County or Jackson (ARIC-AA) with DNA and 1,139 European Americans from Forsyth County or Minneapolis (ARIC-EA) with cerebral magnetic resonance imaging data (Mosley et al. 2005), all at Visits 2 (1990-1992) or 3 (1993-1995).

Replication involved up to 2,176 participants from two studies of the population-based KORA cohort: F3 (n = 464) and F4 (n = 1,712). KORA F3 (2004-2005) and F4 (2006-2008) are follow-up studies of the KORA S3 and S4 cohort participants, including German nationals aged 25-74 years from Augsburg, Germany (Holle et al. 2005; Wichmann et al. 2005).

2.2. Particulate matter exposure estimation

The study focuses on three ambient particulate matter (PM) air pollutants, including two ($PM_{2.5}$ and PM_{10}) that are regulated under the Clean Air Act by the US Environmental Protection Agency (EPA) according to its National Ambient Air Quality Standards (NAAQS) (EPA 2017).

PM exposures were estimated at all geocoded WHI and ARIC participant addresses (Whitsel et al. 2006; Whitsel et al. 2004) in the contiguous US since the baseline examinations using two exposure modeling approaches, both based on US EPA Air Quality System (AQS) monitoring data for PM_{10} (since 1987) and $PM_{2.5}$ (since 1999). In the WHI, the median distance from geocoded participant addresses to PM_{10} and $PM_{2.5}$ EPA monitors was 7.8 and 7.6 kilometers. In ARIC, it was 4.8 and 7.2 kilometers. Geocoded address-specific daily mean PM_{10} concentrations ($\mu g/m^3$) were spatially estimated using national-scale, log-normal ordinary kriging. Exposure measurement error using kriging methods may yield misclassification and increase variance or bias associations (Alexeeff et al. 2014; Lee et al. 2012), therefore validity of the estimation was assessed, using standard cross-validation statistics: average prediction error (PE), standardized prediction error (SPE), root mean square standardized (RMSS), and standard error (SE). Observed values of PE and SPE near zero, RMSS near one, and RMS near SE have provided evidence of model validity (Liao et al. 2006; Liao et al. 2007).

Also, geocoded address-specific monthly mean concentrations ($\mu g/m^3$) were spatiotemporally estimated using generalized additive mixed models and geographic information system-based predictors. Because EPA AQS monitoring data for $PM_{2.5}$ were not widely available until 1999, spatiotemporal estimation also involved the log-transformed ratio of $PM_{2.5}$ to predicted PM_{10} between 1987 and 1999. A five- or ten-fold, out-of-sample cross-validation of the estimates in which the squared Pearson correlation between excluded monthly observations and model predictions (R^2 = 0.68-0.77) indicated that estimation models performed well (Yanosky et al. 2014).

Daily mean concentrations of PM_{10} were averaged over the 2-, 7-, 28-, and 365-day periods ending on (including) the examination day. Monthly mean concentrations of $PM_{2.5}$ and PM_{10} were

averaged over the 12-month period ending on (including) the calendar month of examination. Finally, coarse PM ($PM_{2.5-10}$) concentrations for each averaging duration were calculated as differences between PM_{10} and $PM_{2.5}$ concentrations.

2.3. DNA methylation

Peripheral blood leukocytes were isolated from visit-specific, fasting blood drawn from study participants. DNA was extracted from the peripheral blood leukocytes and then DNAm was measured on a methylome-wide scale at 485,577 CpG sites using the Illumina 450K Infinium Methylation BeadChip (Illumina Inc.; San Diego, CA, USA). Methylation was quantitatively represented by beta, the proportion of methylated cytosines over the sum of methylated and unmethylated cytosines across the same loci. The data from all studies were quality controlled (Table S1), Beta Mixture Quantile (BMIQ)-normalized to adjust for probe bias (Teschendorff et al. 2013), and in WHI-EMPC, ComBat-adjusted for stage and plate using empirical Bayes methods (Johnson et al. 2007). Otherwise, technical covariates (assay plate, chip, and row) were available to control for batch effects; and leukocyte proportions (CD8+ T cell, CD4+ T cell, B cell, natural killer cell, monocyte, and granulocyte) to account for leukocyte composition (Houseman et al. 2012). Among ARIC-AA participants, missing lymphocyte, monocyte, neutrophil, eosinophil, and basophil proportions were imputed based on measured proportions. Analyses excluded CpG sites at which DNAm distributions were multi-modal (Andrews et al. 2016) in at least one study.

2.4. Multiple imputation

To avoid potential for selection bias in complete-data analysis when data are missing at random (Hernan et al. 2004), multivariate imputation by chained equations (MICE) (Azur et al. 2011; Stuart et al. 2009) as implemented in SAS 9.3 (Cary, NC) was used to impute infrequently missing $PM_{2.5}$, PM_{10}, and $PM_{2.5-10}$ concentrations (missing range: 3.3%, 3.5%) and other covariates (missing range: 0%, 10.4%), excluding methylome-wide DNAm. Binary and categorical data were imputed using the logistic and discriminant functions whereas interval-scale data were imputed using predictive means matching with a k-nearest neighbor (k=5) approach.

2.5. Statistical analysis

All analyses were stratified by cohort and race/ethnicity (African-, European-, and Hispanic/Latino-American) and adjusted for age (years) at blood draw, education (high school education or lower, more than high school), smoking status (current, former, never), alcohol use (current, former, never), physical activity (metabolic equivalent of task [MET-hours/week]), body mass index (BMI, kg/m^2), neighborhood socioeconomic status (Roux et al. 2001), mean temperature (°C), mean dew point

(°C), mean barometric pressure (kPa), season, and methylation-related variables, which included ten principal components (PCs) for genetic ancestry (when available), leukocyte proportions, and technical covariates. Analyses additionally controlled for cohort-specific covariates, including binary sex (male, female) in ARIC; randomly assigned treatment group (CT subpopulations of WHI-AS311, WHI-BAA23, WHI-EMPC); case-control status (WHI-AS311, WHI-BAA23); and control matching criteria (WHI-AS311).

In each subpopulation, covariate-adjusted, multi-level, linear, mixed-effects models (LMMs) were used to estimate DNAm-PM associations. In WHI-EMPC, three-level, longitudinal models had a random intercept for examination at the participant level, a random intercept and slope for PM at the WHI center level, and a random intercept for chip, as given by

$$(1) \qquad DNAm_{ijk} = \beta_0 + \beta_1 PM_{ijk} + \beta_2 Z_{ijk} + b_{0k}^{C} + b_{1k}^{C} PM_{ijk} + b_{0jk}^{P} + b_{0ijk}^{E} + \varepsilon_{ijk}^{E}.$$

In WHI-BAA23 CT & OS, and WHI-AS311 CT & OS, two-level cross-sectional models had a random intercept and slope for PM at the WHI center level and a random intercept for plate and chip, as given by

$$(2) \qquad DNAm_{ik} = \beta_0 + \beta_1 PM_{ik} + \beta_2 Z_{ik} + b_{0k}^{C} + b_{1k}^{C} PM_{ik} + b_{0ik}^{E} + \varepsilon_{ik}^{E}.$$

In ARIC-AA and ARIC-EA, one-level cross-sectional models had a random intercept for plate and chip, as given by

$$(3) \qquad DNAm_{i} = \beta_0 + \beta_1 PM_{i} + \beta_2 Z_{i} + b_{0i}^{E} + \varepsilon_{i}^{E}.$$

Above, i, j and k denote the i^{th} examination of the j^{th} participant in the k^{th} center; $DNAm$ is the CpG site-specific beta value; β_0 is the intercept; PM is the 2-, 7-, 28-, 365-day, or 1- or 12-month mean of $PM_{2.5}$, PM_{10}, or $PM_{2.5\text{-}10}$; and Z is a vector of covariates. The terms $(b_0^C, b_1^C) \sim N(O,G)$ are a random intercept and a random slope for PM at the center level, $(b_0^P) \sim N(O,G)$ is a random intercept for examination at the participant level, $(b_0^E) \sim N(O,G)$ are random intercepts for technical covariates, and $\varepsilon^E \sim (O,\sigma^2)$ is the random error at the examination level. Measures of association (β_1) and their 95% confidence intervals ($\beta_1 \pm 1.96\ x\ standard\ error$) were reported as an absolute percentage change in DNAm per 10 μg/m³ increase in PM.

Given the focus on fixed effects, LMMs were fit with maximum likelihood using the MixedModels package (Bates 2017) in Julia v0.6 (Bezanson et al. 2017). Stratum-specific results were combined using fixed-effects, inverse-variance weighted meta-analysis. Homogeneity of associations was assessed using Cochran's Q test statistic (Cochran 1954). A $P_{Cochran's\ Q} < 0.10$ and Bonferroni-corrected

threshold of $P < 1 \times 10^{-7}$ (i.e. assuming 500,000 independent CpG tests) were used to identify significant CpG associations. The threshold of suggestive significance was $P < 1 \times 10^{-5}$.

Examination of stratified and meta-analyzed results included reviewing quantile-quantile (QQ) plots of the observed -$\log_{10}$-transformed *P* values for each CpG site against the expected values from a theoretical χ^2 distribution and estimating the associated genomic inflation factor (λ), where λ is defined as the ratio of the observed to expected median -$\log_{10}P$ values (Devlin et al. 2001).

2.6. Technical validation

In a random subset of 200 WHI-EMPC participants, bisulfite pyrosequencing was used to validate the Illumina 450K measures of DNAm at ten PM_{10}- or $PM_{2.5}$-sensitive CpG sites ($P < 1 \times 10^{-5}$). CpG sites with poor next generation sequencing data or situated in CpG-rich, repetitive element, or low sequence complexity regions of the genome were not candidates for pyrosequencing. Site-specific comparisons of DNAm measures were based on mean Illumina 450K minus bisulfite pyrosequencing differences (**Δ**), Pearson correlation coefficients (r), and Deming regression estimates of their intercepts (α) and slopes (β) (Cornbleet and Gochman 1979). When the two measures are nearly identical, **Δ**, r, α, and β approach values of 0, 1, 0, and 1, respectively.

2.7. Functional annotation

Published genotype-phenotype associations for variants annotated to or within 100 kilobases of genes containing statistically significant PM-sensitive CpG sites were identified in the National Human Genome Research Institute (NHGRI) Genome-Wide Association Study (GWAS) Catalog (Welter et al. 2014). Tissue-specific gene expression was assessed using the Genotype-Tissue Expression (GTEx) database (Lonsdale et al. 2013) and associations between DNAm and gene expression in human blood cells were obtained from a study of approximately 400,000 CpG sites and > 13,000 transcripts in the *Multi-Ethnic Study of Atherosclerosis* (MESA) and *Grady Trauma Project* (GTP) cohorts (Kennedy et al. 2018). PM-sensitive CpG sites ($P < 1 \times 10^{-5}$) were functionally characterized using experimentally derived Functional element Overlap analysis of ReGions from EWAS (eFORGE) v2.0 (Breeze et al. 2016) with data from the Encyclopedia of DNA elements (ENCODE) (Consortium 2012), Roadmap Epigenomics Project (Bernstein et al. 2010), and BLUEPRINT (Stunnenberg et al. 2016). Overlap of CpG site-specific PM sensitivity, histone modification, and DNase I hypersensitivity were evaluated in eFORGE with a false discovery rate (FDR) threshold of 0.05.

2.8. Replication

Significant CpG sites that were not heterogeneous across subpopulations ($P < 1.0\text{x}10^{-7}$; $P_{Cochran's\ Q} > 0.10$) underwent replication and meta-analyses in KORA F3 and F4. Pollutant- and averaging duration-specific replication thresholds were Bonferroni-corrected by dividing the conventional alpha level (0.05) by the number of CpG sites carried into replication.

3. Results

The study consisted of twelve ARIC and WHI subpopulations, collectively representing 8,397 participants, of whom 45.8% were African American, 8.4% were Hispanic/Latino American, and 83.0% were female (Table 1). Participants were on average 61.3 years of age and contributed methylation data at ≥ 461,014 CpG sites. One-month mean concentrations of PM_{10}, $PM_{2.5}$, and $PM_{2.5-10}$ were 20.9, 13.2, and 7.7 μg/m^3; varied by subpopulation and race/ethnicity (Tables 1 and S2); and did not exceed NAAQS in place at the time of data collection. Between-pollutant Pearson correlation coefficients depended on size fraction and averaging duration (Table 2). Overall, the median (range) was 0.35 (-0.14, 0.79) and among 2-, 7-, 28, and 365-day mean PM_{10} concentrations, it was 0.64 (0.43, 0.79). Correlations between PM_{10} and $PM_{2.5}$ concentrations were 0.73 and 0.64 when they were averaged over 1 and 12 months.

Table 1

Characteristics of the study participants, by subpopulation

Subpopulation			Race / ethnicity	n	% female	Age, yrs $\bar{x}$ (SD)	Maximum CpGs	PM (μg/m^3), 1 mo $\bar{x}$ (SD)		
								PM_{10}	$PM_{2.5}$	$PM_{2.5-10}$
ARIC			AA	2,664	63%	56.6 (5.9)	463,431	20.5 (4.6)	13.2 (3.1)	7.3 (2.1)
			EA	1,100	58%	59.9 (5.4)	462,543	23.2 (5.3)	15.4 (4.3)	7.8 (3.5)
WHI	AS311	CT	EA	351	100%	64.7 (7.1)	461,136	19.8 (6.6)	11.9 (3.82)	7.9 (4.6)
		OS	EA	395	100%	66.2 (6.9)	461,136	19.9 (5.7)	12.0 (3.9)	7.9 (4.1)
	BAA23	CT	AA	371	100%	61.8 (6.3)	461,014	22.6 (6.2)	14.3 (4.2)	8.3 (3.8)
			EA	926	100%	67.8 (6.2)	461,014	19.7 (5.7)	11.7 (3.7)	8.0 (4.4)
			HLA	220	100%	60.7 (6.4)	461,014	21.4 (8.1)	10.3 (4.1)	11.1 (5.7)
		OS	AA	259	100%	62.8 (6.8)	461,014	22.3 (5.9)	14.0 (4.0)	8.3 (4.2)
			HLA	174	100%	62.8 (7.3)	461,014	23.0 (8.1)	11.0 (4.2)	11.9 (6.4)
	EMPC[a]		AA	553	100%	62.7 (6.9)	463,916	22.2 (6.2)	15.2 (5.1)	7.0 (4.7)
			EA	1,072	100%	64.6 (7.1)	463,916	19.4 (6.0)	13.0 (5.0)	6.4 (5.2)
			HLA	312	100%	61.5 (6.1)	463,916	21.9 (7.1)	12.8 (6.3)	9.1 (5.3)
All			AA (45.8%) HLA (8.4%) EA (45.8%)	8,397	83%	61.3 (7.4)	463,916	20.9 (5.8)	13.2 (4.3)	7.7 (4.0)

Abbreviations: AA, African American; ARIC, Atherosclerosis Risk in Communities; AS311, Ancillary Study 311; BAA23, Broad Agency Award 23; CpG, Cytosine-phosphate-Guanine; CT, Clinical Trial; EA, European American; EMPC, Epigenetic Mechanisms of PM-Mediated CVD Risk; HLA, Hispanic/Latino American; mo, month; OS, Observational Study; PM_{10}, PM < 10 μm in diameter; $PM_{2.5}$, PM < 2.5 μm in diameter; $PM_{2.5-10}$, PM > 2.5 and < 10 μm in diameter; SD, standard deviation; WHI, Women's Health Initiative; $\bar{x}$, mean

[a]At the 1st visit. Methylation data also were available among 185 & 43 WHI-EMPC participants @ the 2nd & 3rd visits

Table 2

Particulate matter concentration (μg/m³) means and Pearson correlations in the total population (n = 8,397)

		PM_{10} 2 d	PM_{10} 7 d	PM_{10} 28 d	PM_{10} 365 d	PM_{10} 1 mo	PM_{10} 12 mo	$PM_{2.5}$ 1 mo	$PM_{2.5}$ 12 mo	$PM_{2.5-10}$ 1 mo	$PM_{2.5-10}$ 12 mo
	$\bar{x}$	31.9	31.1	30.9	31.2	20.9	20.9	13.2	13.2	7.7	7.8
	(SD)	(12.1)	(9.2)	(7.1)	(5.1)	(5.8)	(4.0)	(4.3)	(3.0)	(4.0)	(3.1)
PM_{10}	2 d	1.00									
PM_{10}	7 d	0.74	1.00								
PM_{10}	28 d	0.58	0.79	1.00							
PM_{10}	365 d	0.43	0.56	0.70	1.00						
PM_{10}	1 mo	0.39	0.48	0.54	0.27	1.00					
PM_{10}	12 mo	0.15	0.18	0.24	0.35	0.62	1.00				
$PM_{2.5}$	1 mo	0.29	0.36	0.41	0.17	0.73	0.39	1.00			
$PM_{2.5}$	12 mo	0.11	0.12	0.15	0.23	0.40	0.64	0.66	1.00		
$PM_{2.5-10}$	1 mo	0.25	0.31	0.35	0.21	0.67	0.48	-0.02	-0.13	1.00	
$PM_{2.5-10}$	12 mo	0.08	0.12	0.17	0.23	0.41	0.67	-0.14	-0.14	0.74	1.00

Abbreviations: d, day; mo, month; PM, particulate matter; PM_{10}, PM < 10 μm in diameter; $PM_{2.5}$, PM < 2.5 μm in diameter; $PM_{2.5-10}$, PM > 2.5 and < 10 μm in diameter; SD, standard deviation; $\bar{x}$, mean

QQ plots (Fig. 1) based on the trans-ethnic, fixed-effects, inverse variance-weighted meta-analyses provided little evidence of inflation across pollutants and averaging durations: median (range) λ = 1.01, (0.89-1.07). Manhattan plots (Fig. 2) show three significant ($P < 1 \times 10^{-7}$) and 55 suggestively significant ($1 \times 10^{-5} < P < 1 \times 10^{-7}$) PM-sensitive CpG sites (Tables 3 and S3). The three significant CpG sites (cg19004594; cg24102420; cg12124767) were neither within ten base pairs of single nucleotide polymorphisms (minor allele frequency > 1%) nor previously identified as cross-reactive probes (Chen et al. 2013).

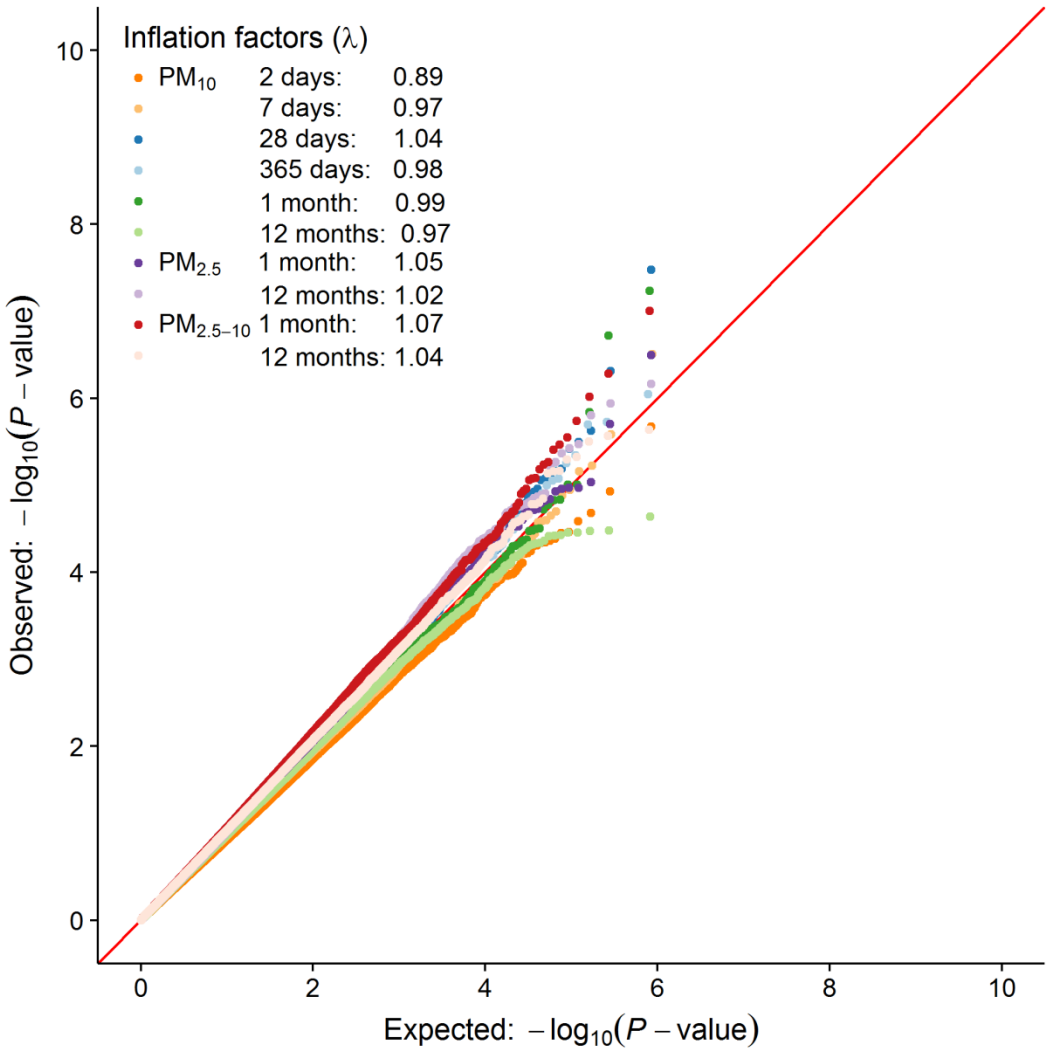

Fig. 1. Quantile-quantile (QQ) plot of observed vs. expected -$\log_{10}$ *p*-value of each CpG site from trans-ethnic, fixed-effects meta-analyses of 2-, 7-, 28-, and 365-day PM_{10} and 1- and 12-month PM_{10} and $PM_{2.5}$. The red diagonal line references the methylome-wide significance threshold ($P < 1.0\times10^{-7}$). Lambda (λ) is the inflation factor.

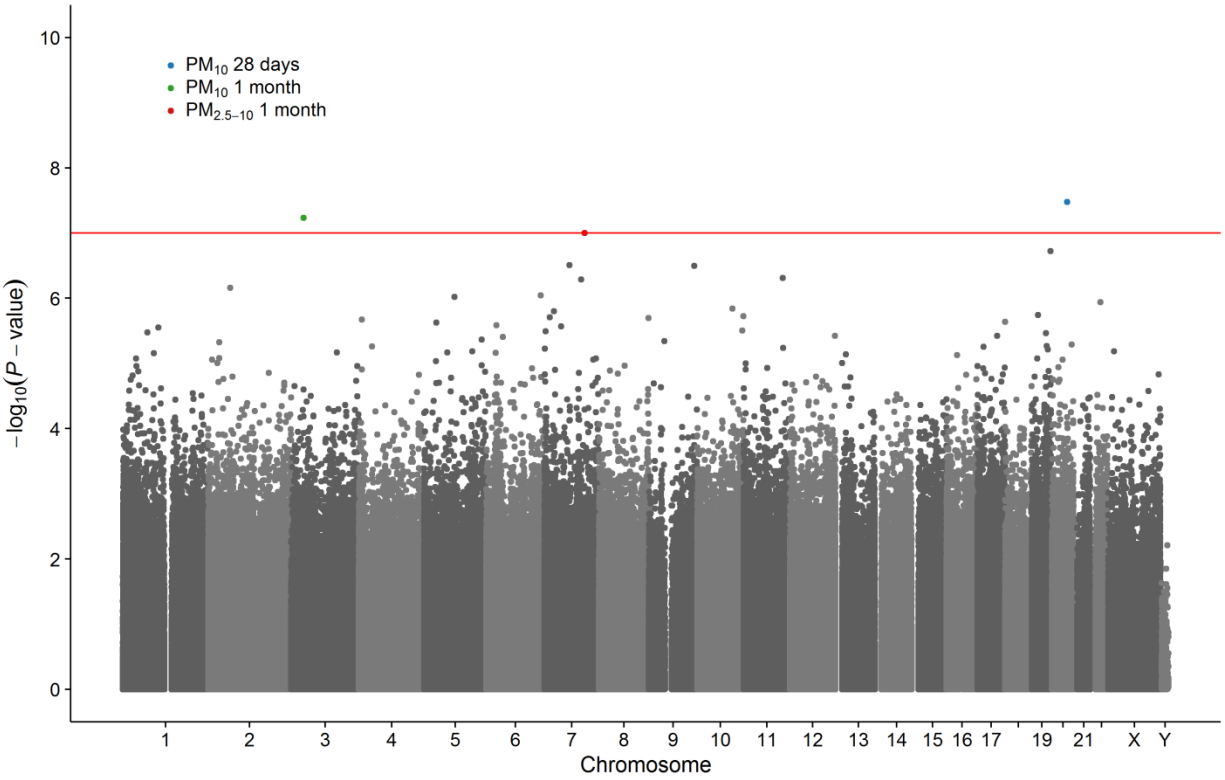

Fig. 2. Manhattan plot of $-\log_{10}$ *p*-value vs. chromosomal position of each CpG site from trans-ethnic, fixed-effects meta-analyses of 2-, 7-, 28-, and 365-day PM_{10} and 1- and 12-month PM_{10} and $PM_{2.5}$. The red line references the methylome-wide significance threshold ($P < 1.0 \times 10^{-7}$).

Table 3

Findings from trans-ethnic, fixed-effects meta-analyses ($P < 1 \times 10^{-7}$, $P_{Cochran's\ Q} > 0.10$).

Chr	Position[a]	CpG	Exposure	%Δ (95% CI)[b]	*P*	n_{obs}	Gene
20	43926884	cg19004594	PM_{10}, 28 d	0.3 (0.2, 0.4)	3.33×10^{-8}	8,622	*MATN4*
3	35785890	cg24102420	PM_{10}, 1 mo	-0.5 (-0.7, -0.3)	5.84×10^{-8}	8,575	*ARPP21 / MIR128-2*
7	117299297	cg12124767	$PM_{2.5-10}$, 1 mo	-0.5 (-0.7, -0.3)	9.96×10^{-8}	8,577	*CFTR*

Abbreviations: **Δ**, change; Chr, chromosome; CI, confidence interval; CpG, Cytosine-phosphate-Guanine; d, days; mo, month; PM_{10}, PM < 10 µm in diameter; $PM_{2.5}$, PM < 2.5 µm in diameter; $PM_{2.5-10}$, PM > 2.5 and < 10 µm in diameter

[a]Build 37

[b]Absolute percentage point per 10 µg/m³ increase in PM_{10}

On chromosome 20 within an exonic CpG island of *MATN4*, a 10 µg/m³ increase in 28-day mean PM_{10} was associated with a 0.3% (95% confidence interval [CI]: 0.2, 0.4) higher DNAm at cg19004594 ($P = 3.33 \times 10^{-8}$; Fig. 3A). On chromosome 3 intronic to *ARPP21*, a 10 µg/m³ increase in 1-month mean

PM_{10} was associated with a 0.5% (95% CI: 0.3, 0.7) lower DNAm at cg24102420 ($P = 5.84 \times 10^{-8}$; Fig. 3B). Cg24102420 is approximately 200 base pairs upstream from the transcriptional start site for microRNA 128-2 (*miR128-2*). On chromosome 7 intronic to *CFTR*, a 10 μg/m³ increase in 1-month mean $PM_{2.5-10}$ was associated with a 0.5% (95% CI: 0.3, 0.7) lower DNAm at cg12124767 ($P = 9.86 \times 10^{-8}$; Fig. 3C). Furthermore, PM associations with cg19004594, cg24102420, and cg12124767 were similar across race/ethnic strata (Fig. S1). Complete annotations for all PM-sensitive CpG sites ($P < 1 \times 10^{-7}$) are available in Excel Table S1.

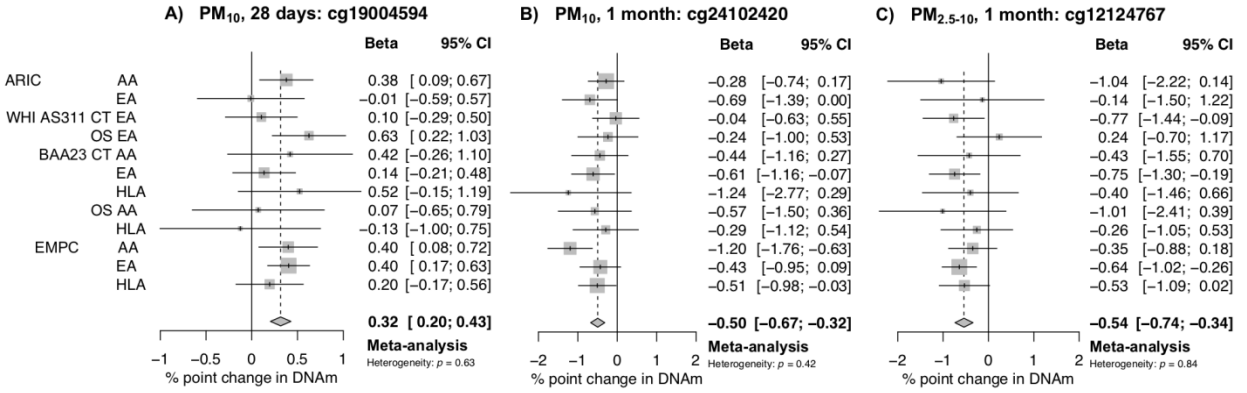

Fig. 3. Forest plots of PM-CpG associations (95% confidence intervals) for A) cg19004594, B) cg2410240, and C) cg12124767 with a 10 μg/m³ increase in PM by subpopulation and overall after fixed-effects meta-analysis.

3.1. Technical validation

Overall, bisulfite pyrosequencing and Illumina 450K-based DNAm measures were similar (Table S4). The medians (interdecile ranges) of **Δ**, r, α and β were: 0.01 (-0.06, 0.07), 0.73 (0.20, 0.83), 0.04 (-0.27, 0.24), and 0.98 (0.09, 1.62). Corresponding estimates (95% CIs) for cg24102420 were -0.04 (-0.04, -0.03), 0.79 (0.73, 0.83), -0.16 (-0.38, 0.07) and 1.13 (0.88, 1.39). Cg19004594 and cg12124767 were not pyrosequenced.

3.2. Functional annotation

MATN4 is highly expressed in the pancreas, reproductive tract, and skin (Fig. S2), but variants of this gene have not been significantly associated ($P < 5 \times 10^{-8}$) with any phenotypes in prior GWAS. *ARPP21* is primarily expressed in the brain (Fig. S3), is significantly associated with neuroticism and severe H1N1 influenza, and suggestively associated ($5 \times 10^{-8} < P < 5 \times 10^{-6}$) with entorhinal cortical thickness and childhood-onset asthma in prior GWAS. *CFTR* is expressed in various tissues, including the pancreas, colon, minor salivary gland, digestive tract, and lung (Fig. S4). *CFTR* polymorphisms are

associated with cystic fibrosis (CF), Barrett's esophagus / esophageal carcinoma, and coronary artery disease.

Differential methylation at cg19004594, cg24102420, or cg12124767 was not associated with gene expression in blood cells at any of the > 13,000 transcripts evaluated ($P > 10^{-5}$) in the MESA/GTP cohorts. Although genomic regions around PM-sensitive CpG sites were associated with tri-methylation of histone 3 at lysine 9 (H3K9me3) in natural killer cells, derived mesenchymal stem cells, the fetal adrenal gland, fetal lung fibroblasts, and foreskin fibroblasts (FDR < 0.05; Fig. 4), they were not associated with mono- or tri-methylation of histone 3 at lysine 4, 27, or 36 (H3K4me1, H3K4me3, H3K27me3, or H3K36me3) or DNase I hypersensitivity in any tissues catalogued by eFORGE.

3.3. Replication

The three statistically significant, non-heterogeneous PM-sensitive CpG sites (cg19004594; cg24102420; cg12124767) did not replicate in KORA F3 / F4 (Table S5).

Fig. 4. Enrichment of PM-sensitive CpG sites in regions overlapping H3K9me3 using Roadmap data.

4. Discussion

This methylome-wide association study (MWAS) discovered three CpG sites at which higher levels of monthly mean ambient particulate matter air pollution concentrations were associated with DNAm. The DNAm-PM associations at all three CpG sites were homogeneous across the twelve subpopulations and each site was annotated to a neurological, pulmonary, endocrine, or cardiovascular disease-related gene (*MATN4*, *ARPP21* or *CFTR*). Although a recent MWAS also implicated cigarette smoking in DNA methylation at *ARPP21* and *CFTR* (Joehanes et al. 2016)—two genes that may underlie epigenetically mediated responses to inhalable environmental exposures—the CpG sites discovered herein are in different regions of *ARPP21* and *CFTR*, suggesting varied responses to particulate exposures, and none of them were associated with gene expression of blood cells in MESA/GTP.

Methylation of cg19004594 (exon of *MATN4*) was positively associated with 28-day mean PM_{10} concentrations. *MATN4* encodes Matrilin 4, a von Willebrand factor A domain-containing protein, which contributes to cardiac remodeling (Barallobre-Barreiro et al. 2012) and inhibits the proliferation of hematopoietic stem cells at rest. Additionally, environmental stressors trigger expression of the *CXCL12*-encoded chemokine (SDF1) (Liberda et al. 2010) and activation of its G protein-coupled receptor (CXCR4), leading to inhibition of Matrilin 4 and subsequent expansion of hematopoietic stem cell pools (Uckelmann et al. 2016). SDF1-activated CXCR4 also inhibits beta-adrenergically activated calcium influx through myocardial L-type calcium ion channels (Pyo et al. 2006), a process that may affect PM_{10}-associated ventricular action potential and electrocardiographic QT interval duration (Gondalia et al. 2017). Methylation of *MATN4* may therefore underlie commonly observed hematological and electrocardiographic of effects of PM_{10}.

Methylation at cg24102420 (intron of *ARPP21*) was positively associated with 1-month mean PM_{10} concentrations. *ARPP21* encodes a neuronal cAMP-regulated phosphoprotein, a regulator of calmodulin signaling (RCS) that is highly enriched in medium spiny neurons within the basal ganglia, cerebral cortex, and other regions of the brain (Rakhilin et al. 2004), with dual evidence of expression in cardiac tissues (Kahr et al. 2011; Kirchhof et al. 2011; Mathar et al. 2013). Variants of *ARPP21* have been associated with entorhinal cortical thickness (Furney et al. 2010). Calmodulin signaling (O'Day et al. 2015), entorhinal cortical thickness (Velayudhan et al. 2013), and PM air pollution (Cacciottolo et al. 2017) are all associated with Alzheimer's disease progression, suggesting a potential epigenetic mechanism of PM_{10}-related neuropathology.

Indeed, *ARPP21* and *miR128-2*, a microRNA within *ARPP21*, are both regulators of dendritic growth (Rehfeld et al. 2018). In a study of rats, exposure to ammonium sulfate, a major component of $PM_{2.5}$, was associated with diminished dendritic complexity in hippocampal neurons (Cheng et al. 2017).

Additionally, *miR128* expression in peripheral blood of steel plant workers increased with increases in PM exposure, as was confirmed by an *in vitro* study of PM-treated pulmonary tissue (Bollati et al. 2015). Additional roles of *miR128* include the inhibition of *ABCA1* and *ABCG1*, adenosine triphosphate-binding cassette (ABC) transporter genes also involved in homeostasis of cholesterol (Adlakha et al. 2013), an established risk factor for stroke, myocardial infarction, and other common forms of cardiovascular disease.

Methylation at cg12124767 (intron of *CFTR*) was inversely associated with 1-month mean $PM_{2.5-10}$ concentrations. *CFTR* encodes a transmembrane conductance regulator; specifically, an ABC transporter of chloride and thiocyanate ions. The *CFTR*-encoded ABC transporter controls fluid secretion and absorption in epithelial tissues (Saint-Criq and Gray 2017). Its most common mutation impairs folding and trafficking of the encoded protein in pulmonary and pancreatic epithelia, causing CF and CF-related diabetes (Brennan et al. 2004). However, cigarette smoke and chronic inflammation also reduce *CFTR* chloride channel function (Rasmussen et al. 2014), a hypothesized molecular pathway underlying the development of chronic obstructive pulmonary disease (Rab et al. 2013). Furthermore, CFTR chloride channel currents in the myocardium shorten action potential and QT interval duration (Duan 2013). Their activation by cAMP protein kinase A (PKA), protein kinase C (PKC), or extracellular adenosine triphosphate (ATP) through purinergic receptors (al-Awqati 1995; Duan 2013) can be arrhythmogenic (Cacciapuoti et al. 1991; Engler and Yellon 1996; Leonard et al. 2017; Najeed et al. 2002; Yamazaki and Hume 1997). Hypomethylation of *CFTR* at this site therefore highlights another epigenetic mechanism that may underlie PM_{10}-related pulmonary and electrocardiographic manifestations of disease.

While the putative mechanisms described above are biologically plausible, analyses on which they are based are limited by their reliance on DNAm derived from leukocytes. Although other (e.g. heart, lung, nervous) tissues may be more appropriate for studying the role of DNAm on human disease, their collection is highly invasive (McCullough et al. 2017; Zhong et al. 2016); as such, leukocytes extracted from peripheral blood are widely used surrogate tissues (Zhong et al. 2016) with demonstrated consistency of DNAm patterns across relevant tissues types (Byun et al. 2009; Fan and Zhang 2009; Ma et al. 2014). Still, DNAm at cg19004594, cg24102420, cg12124767 was not associated with gene expression of blood cells in GTP/MESA (Kennedy et al. 2018). Unlike DNAm patterns though, gene expression is highly variable by tissue type (Aguet et al. 2017), and *MATN4*, *ARPP21* and *CFTR* are primarily expressed in other tissues.

The inability to replicate associations in KORA F3 and F4 participants is noteworthy. Although independent from the discovery populations, KORA represents a population of white, European men and women living in Augsburg, Germany, one distinct from that of the environmentally diverse, multi-racial/ethnic U.S. populations in the discovery. In addition, PM composition in ARIC and WHI (1990-

2012) may differ from that in Augsburg during KORA F3 and F4 (2004-2006). Furthermore, PM concentrations in KORA were measured at community monitors, while those in WHI and ARIC were spatially or spatiotemporally estimated at participant geocoded addresses from monitoring networks in the 48 contiguous US states.

DNAm associations with $PM_{2.5}$ – potentially the driver for PM-associated disease (Brook et al. 2010) – were not detected in this study. Inability to do so may be due to lower power to detect $PM_{2.5}$ versus PM_{10} associations with DNAm given lower-variance $PM_{2.5}$ exposure estimates, lack of short-duration $PM_{2.5}$ data before 1999 when EPA AQS started monitoring it, and / or induction of $PM_{2.5}$ health effects that are not epigenetically mediated.

The analyses also were limited by predominantly cross-sectional data, high multiple testing burden, small effect sizes, and residual need for functional characterization. However, repeated measures of PM and DNAm over time were leveraged in WHI-EMPC to increase statistical power. Among-pollutant correlations also were moderate in this context, so the multiple comparisons made were not strictly independent. Similarly, the Bonferroni-corrected threshold used herein ($P < 1 \times 10^{-7}$) is conservative because of methylome-wide correlations among CpG sites (Saffari et al. 2018; Tsai et al. 2012), decreasing the likelihood of false positives. Moreover, observed effect sizes were consistent with those seen in other epigenetic studies of particulate matter exposure (de F. C. Lichtenfels et al. 2018; Panni et al. 2016; Plusquin et al. 2017) and smoking (Joehanes et al. 2016). Further investigation is nonetheless needed to determine the clinical impact of CpG-specific changes in methylation although functional validation of epigenetic associations was outside the scope of presently funded work. Still, this is a well-powered study of geographically diverse, multi-racial/ethnic populations of women and men with methylome-wide DNAm and geocoded address-specific PM data, that leveraged multivariate imputation to minimize selection-related biases otherwise known to affect epidemiologic associations in complete data analyses.

5. Conclusions

Findings from this large, racially/ethnically and environmentally diverse methylome-wide association study of women and men in EPA regions 1-10 suggest that ambient particulate matter air pollution affects DNAm at regions of the genome potentially related to neurological, pulmonary, endocrine, and cardiovascular disease. Although the discovered associations are biologically plausible, functional characterization in relevant tissues or animal models remain necessary to validate associations and elucidate putative epigenetic mechanisms of PM-associated disease.

500

Acknowledgements

The Atherosclerosis Risk in Communities study has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I and HHSN268201700005I). The authors thank the staff and participants of the ARIC study for their important contributions. Funding was also supported by 5RC2HL102419 and R01NS087541. Data from the ARIC study are available on request at https://www2.cscc.unc.edu/aric/distribution-agreements.

The WHI program is funded by the NHLBI, U.S. Department of Health and Human Services, through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. WHI-AS311 was supported by American Cancer Society award 125299-RSG-13-100-01-CCE. All contributors to WHI science are listed at https://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf. Data from the WHI are available on request at https://www.whi.org/researchers/SitePages/Write%20a%20Paper.aspx.

The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

This work was supported by NIEHS grant R01-ES020836 (LH, AB, EAW), NHLBI contract HHSN268201100046C (KC), NIEHS grant R01-ES017794 (EAW), NHLBI National Research Service Award T32-HL007055 (RG), NIEHS National Research Service Award T32-ES007018 (KH), and NCI grant R25-CA094880 (KJ).

Conflicts of interest

No authors have declared a potential conflicts of interest.

References

Adlakha, Y.K.; Khanna, S.; Singh, R.; Singh, V.P.; Agrawal, A.; Saini, N. Pro-apoptotic miRNA-128-2 modulates ABCA1, ABCG1 and RXRα expression and cholesterol homeostasis. Cell Death & Disease 2013;4:e780

Aguet, F.; Brown, A.A.; Castel, S.E.; Davis, J.R.; He, Y.; Jo, B.; Mohammadi, P.; Park, Y.; Parsana, P.; Segrè, A.V.; Strober, B.J.; Zappala, Z.; Cummings, B.B.; Gelfand, E.T.; Hadley, K.; Huang, K.H.; Lek, M.; Li, X.; Nedzel, J.L.; Nguyen, D.Y.; Noble, M.S.; Sullivan, T.J.; Tukiainen, T.; MacArthur, D.G.; Getz, G.; Addington, A.; Guan, P.; Koester, S.; Little, A.R.; Lockhart, N.C.; Moore, H.M.; Rao, A.; Struewing, J.P.; Volpi, S.; Brigham, L.E.; Hasz, R.; Hunter, M.; Johns, C.; Johnson, M.; Kopen, G.; Leinweber, W.F.; Lonsdale, J.T.; McDonald, A.; Mestichelli, B.; Myer, K.; Roe, B.; Salvatore, M.; Shad, S.; Thomas, J.A.; Walters, G.; Washington, M.; Wheeler, J.; Bridge, J.; Foster, B.A.; Gillard, B.M.; Karasik, E.; Kumar, R.; Miklos, M.; Moser, M.T.; Jewell, S.D.; Montroy, R.G.; Rohrer, D.C.; Valley, D.; Mash, D.C.; Davis, D.A.; Sobin, L.; Barcus, M.E.; Branton, P.A.; Abell, N.S.; Balliu, B.; Delaneau, O.; Frésard, L.; Gamazon, E.R.; Garrido-Martín, D.; Gewirtz, A.D.H.; Gliner, G.; Gloudemans, M.J.; Han, B.; He, A.Z.; Hormozdiari, F.; Li, X.; Liu, B.; Kang, E.Y.; McDowell, I.C.; Ongen, H.; Palowitch, J.J.; Peterson, C.B.; Quon, G.; Ripke, S.; Saha, A.; Shabalin, A.A.; Shimko, T.C.; Sul, J.H.; Teran, N.A.; Tsang, E.K.; Zhang, H.; Zhou, Y.-H.; Bustamante, C.D.; Cox, N.J.; Guigó, R.; Kellis, M.; McCarthy, M.I.; Conrad, D.F.; Eskin, E.; Li, G.; Nobel, A.B.; Sabatti, C.; Stranger, B.E.; Wen, X.; Wright, F.A.; Ardlie, K.G.; Dermitzakis, E.T.; Lappalainen, T.; Aguet, F.; Ardlie, K.G.; Cummings, B.B.; Gelfand, E.T.; Getz, G.; Hadley, K.; Handsaker, R.E.; Huang, K.H.; Kashin, S.; Karczewski, K.J.; Lek, M.; Li, X.; MacArthur, D.G.; Nedzel, J.L.; Nguyen, D.T.; Noble, M.S.; Segrè, A.V.; Trowbridge, C.A.; Tukiainen, T.; Abell, N.S.; Balliu, B.; Barshir, R.; Basha, O.; Battle, A.; Bogu, G.K.; Brown, A.; Brown, C.D.; Castel, S.E.; Chen, L.S.; Chiang, C.; Conrad, D.F.; Cox, N.J.; Damani, F.N.; Davis, J.R.; Delaneau, O.; Dermitzakis, E.T.; Engelhardt, B.E.; Eskin, E.; Ferreira, P.G.; Frésard, L.; Gamazon, E.R.; Garrido-Martín, D.; Gewirtz, A.D.H.; Gliner, G.; Gloudemans, M.J.; Guigo, R.; Hall, I.M.; Han, B.; He, Y.; Hormozdiari, F.; Howald, C.; Kyung Im, H.; Jo, B.; Yong Kang, E.; Kim, Y.; Kim-Hellmuth, S.; Lappalainen, T.; Li, G.; Li, X.; Liu, B.; Mangul, S.; McCarthy, M.I.; McDowell, I.C.; Mohammadi, P.; Monlong, J.; Montgomery, S.B.; Muñoz-Aguirre, M.; Ndungu, A.W.; Nicolae, D.L.; Nobel, A.B.; Oliva, M.; Ongen, H.; Palowitch, J.J.; Panousis, N.; Papasaikas, P.; Park, Y.; Parsana, P.; Payne, A.J.; Peterson, C.B.; Quan, J.; Reverter, F.; Sabatti, C.; Saha, A.; Sammeth, M.; Scott, A.J.; Shabalin, A.A.; Sodaei, R.; Stephens, M.; Stranger, B.E.; Strober, B.J.; Sul, J.H.; Tsang, E.K.; Urbut, S.; van de Bunt, M.; Wang, G.; Wen, X.; Wright, F.A.; Xi, H.S.; Yeger-Lotem, E.; Zappala, Z.; Zaugg, J.B.; Zhou, Y.-H.; Akey, J.M.; Bates, D.; Chan, J.; Chen, L.S.; Claussnitzer, M.; Demanelis, K.; Diegel, M.; Doherty, J.A.; Feinberg, A.P.; Fernando, M.S.; Halow, J.; Hansen, K.D.; Haugen, E.; Hickey, P.F.; Hou, L.; Jasmine, F.; Jian, R.; Jiang, L.; Johnson, A.; Kaul, R.; Kellis, M.; Kibriya, M.G.; Lee, K.; Billy Li, J.; Li, Q.; Li, X.; Lin, J.; Lin, S.; Linder, S.; Linke, C.; Liu, Y.; Maurano, M.T.; Molinie, B.; Montgomery, S.B.; Nelson, J.; Neri, F.J.; Oliva, M.; Park, Y.; Pierce, B.L.; Rinaldi, N.J.; Rizzardi, L.F.; Sandstrom, R.; Skol, A.; Smith, K.S.; Snyder, M.P.; Stamatoyannopoulos, J.; Stranger, B.E.; Tang, H.; Tsang, E.K.; Wang, L.; Wang, M.; Van Wittenberghe, N.; Wu, F.; Zhang, R.; Nierras, C.R.; Branton, P.A.; Carithers, L.J.; Guan, P.; Moore, H.M.; Rao, A.; Vaught, J.B.; Gould, S.E.; Lockart, N.C.; Martin, C.; Struewing, J.P.; Volpi, S.; Addington, A.M.; Koester, S.E.; Little, A.R.; Brigham, L.E.; Hasz, R.; Hunter, M.; Johns, C.; Johnson, M.; Kopen, G.; Leinweber, W.F.; Lonsdale, J.T.; McDonald, A.; Mestichelli, B.; Myer, K.; Roe, B.; Salvatore, M.; Shad, S.; Thomas, J.A.; Walters, G.; Washington, M.; Wheeler, J.; Bridge, J.; Foster, B.A.; Gillard, B.M.; Karasik, E.; Kumar, R.; Miklos, M.; Moser, M.T.; Jewell, S.D.; Montroy, R.G.; Rohrer, D.C.; Valley, D.R.; Davis, D.A.; Mash, D.C.; Undale, A.H.; Smith, A.M.; Tabor, D.E.; Roche, N.V.; McLean, J.A.; Vatanian, N.; Robinson, K.L.; Sobin, L.; Barcus,

M.E.; Valentino, K.M.; Qi, L.; Hunter, S.; Hariharan, P.; Singh, S.; Um, K.S.; Matose, T.; Tomaszewski, M.M.; Barker, L.K.; Mosavel, M.; Siminoff, L.A.; Traino, H.M.; Flicek, P.; Juettemann, T.; Ruffier, M.; Sheppard, D.; Taylor, K.; Trevanion, S.J.; Zerbino, D.R.; Craft, B.; Goldman, M.; Haeussler, M.; Kent, W.J.; Lee, C.M.; Paten, B.; Rosenbloom, K.R.; Vivian, J.; Zhu, J. Genetic effects on gene expression across human tissues. Nature 2017;550:204

al-Awqati, Q. Regulation of ion channels by ABC transporters that secrete ATP. Science 1995;269:805-806

Alexeeff, S.E.; Schwartz, J.; Kloog, I.; Chudnovsky, A.; Koutrakis, P.; Coull, B.A. Consequences of kriging and land use regression for PM2.5 predictions in epidemiologic analyses: insights into spatial variability using high-resolution satellite data. Journal Of Exposure Science And Environmental Epidemiology 2014;25:138

Anderson, G.L.; LaCroix, A. W64 - Long Life Study (Long Life Study).

Anderson, G.L.; Manson, J.; Wallace, R.; Lund, B.; Hall, D.; Davis, S.; Shumaker, S.; Wang, C.-Y.; Stein, E.; Prentice, R.L. Implementation of the Women's Health Initiative study design. Annals of epidemiology 2003;13:S5-S17

Andrews, S.V.; Ladd-Acosta, C.; Feinberg, A.P.; Hansen, K.D.; Fallin, M.D. “Gap hunting” to characterize clustered probe signals in Illumina methylation array data. Epigenetics & Chromatin 2016;9:56

ARIC. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. Am J Epidemiol 1989;129:687-702

Assimes, T.; Tsao, P.; Absher, D.; Horvath, S. BA23 - Integrative genomics and risk of CHD and related phenotypes in the Women’s Health Initiative.

Azur, M.J.; Stuart, E.A.; Frangakis, C.; Leaf, P.J. Multiple imputation by chained equations: what is it and how does it work? International journal of methods in psychiatric research 2011;20:40-49

Baccarelli, A.; Rienstra, M.; Benjamin, E.J. Cardiovascular Epigenetics. Circulation: Cardiovascular Genetics 2010;3:567

Barallobre-Barreiro, J.; Didangelos, A.; Schoendube, F.A.; Drozdov, I.; Yin, X.; Fernández-Caggiano, M.; Willeit, P.; Puntmann, V.O.; Aldama-López, G.; Shah, A.M.; Doménech, N.; Mayr, M. Proteomics Analysis of Cardiac Extracellular Matrix Remodeling in a Porcine Model of Ischemia/Reperfusion Injury. Circulation 2012;125:789-802

Bates, D. Mixed-effects models in Julia. GitHub; 2017

Bellavia, A.; Urch, B.; Speck, M.; Brook, R.D.; Scott, J.A.; Albetti, B.; Behbod, B.; North, M.; Valeri, L.; Bertazzi, P.A.; Silverman, F.; Gold, D.; A. Baccarelli, A. DNA Hypomethylation, Ambient Particulate Matter, and Increased Blood Pressure: Findings From Controlled Human Exposure Experiments. Journal of the American Heart Association 2013;2

Bernstein, B.E.; Stamatoyannopoulos, J.A.; Costello, J.F.; Ren, B.; Milosavljevic, A.; Meissner, A.; Kellis, M.; Marra, M.A.; Beaudet, A.L.; Ecker, J.R.; Farnham, P.J.; Hirst, M.; Lander, E.S.;

Mikkelsen, T.S.; Thomson, J.A. The NIH Roadmap Epigenomics Mapping Consortium. Nature biotechnology 2010;28:1045-1048

Bezanson, J.; Edelman, A.; Karpinski, S.; Shah, V.B. Julia: A Fresh Approach to Numerical Computing. SIAM Review 2017;59:65-98

Bollati, V.; Angelici, L.; Rizzo, G.; Pergoli, L.; Rota, F.; Hoxha, M.; Nordio, F.; Bonzini, M.; Tarantini, L.; Cantone, L.; Pesatori, A.C.; Apostoli, P.; Baccarelli, A.A.; Bertazzi, P.A. Microvesicle-associated microRNA expression is altered upon particulate matter exposure in healthy workers and in A549 cells. Journal of Applied Toxicology 2015;35:59-67

Bollati, V.; Baccarelli, A. Environmental epigenetics. Heredity 2010;105:105-112

Breeze, Charles E.; Paul, Dirk S.; van Dongen, J.; Butcher, Lee M.; Ambrose, John C.; Barrett, James E.; Lowe, R.; Rakyan, Vardhman K.; Iotchkova, V.; Frontini, M.; Downes, K.; Ouwehand, Willem H.; Laperle, J.; Jacques, P.-É.; Bourque, G.; Bergmann, Anke K.; Siebert, R.; Vellenga, E.; Saeed, S.; Matarese, F.; Martens, Joost H.A.; Stunnenberg, Hendrik G.; Teschendorff, Andrew E.; Herrero, J.; Birney, E.; Dunham, I.; Beck, S. eFORGE: A Tool for Identifying Cell Type-Specific Signal in Epigenomic Data. Cell Reports 2016;17:2137-2150

Brennan, A.L.; Geddes, D.M.; Gyi, K.M.; Baker, E.H. Clinical importance of cystic fibrosis-related diabetes. Journal of Cystic Fibrosis 2004;3:209-222

Brook, R.D.; Franklin, B.; Cascio, W.; Hong, Y.; Howard, G.; Lipsett, M.; Luepker, R.; Mittleman, M.; Samet, J.; Smith, S.C.; Tager, I. Air Pollution and Cardiovascular Disease. Circulation 2004;109:2655

Brook, R.D.; Rajagopalan, S.; Pope, C.A.; Brook, J.R.; Bhatnagar, A.; Diez-Roux, A.V.; Holguin, F.; Hong, Y.; Luepker, R.V.; Mittleman, M.A.; Peters, A.; Siscovick, D.; Smith, S.C.; Whitsel, L.; Kaufman, J.D. Particulate Matter Air Pollution and Cardiovascular Disease. Circulation 2010;121:2331

Byun, H.-M.; Siegmund, K.D.; Pan, F.; Weisenberger, D.J.; Kanel, G.; Laird, P.W.; Yang, A.S. Epigenetic profiling of somatic tissues from human autopsy specimens identifies tissue- and individual-specific DNA methylation patterns. Human molecular genetics 2009;18:4808-4817

Cacciapuoti, F.; Spiezia, R.; Bianchi, U.; Lama, D.; D'Avino, M.; Varricchio, M. Effectiveness of glibenclamide on myocardial ischemic ventricular arrhythmias in non-insulin-dependent diabetes mellitus. The American Journal of Cardiology 1991;67:843-847

Cacciottolo, M.; Wang, X.; Driscoll, I.; Woodward, N.; Saffari, A.; Reyes, J.; Serre, M.L.; Vizuete, W.; Sioutas, C.; Morgan, T.E.; Gatz, M.; Chui, H.C.; Shumaker, S.A.; Resnick, S.M.; Espeland, M.A.; Finch, C.E.; Chen, J.C. Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. Translational Psychiatry 2017;7:e1022

Chen, R.; Meng, X.; Zhao, A.; Wang, C.; Yang, C.; Li, H.; Cai, J.; Zhao, Z.; Kan, H. DNA hypomethylation and its mediation in the effects of fine particulate air pollution on cardiovascular biomarkers: A randomized crossover trial. Environment International 2016;94:614-619

Chen, R.; Qiao, L.; Li, H.; Zhao, Y.; Zhang, Y.; Xu, W.; Wang, C.; Wang, H.; Zhao, Z.; Xu, X.; Hu, H.; Kan, H. Fine Particulate Matter Constituents, Nitric Oxide Synthase DNA Methylation and Exhaled Nitric Oxide. Environmental Science & Technology 2015;49:11859-11865

Chen, Y.a.; Lemire, M.; Choufani, S.; Butcher, D.T.; Grafodatskaya, D.; Zanke, B.W.; Gallinger, S.; Hudson, T.J.; Weksberg, R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013;8

Cheng, L.; Lau, W.K.W.; Fung, T.K.H.; Lau, B.W.M.; Chau, B.K.H.; Liang, Y.; Wang, Z.; So, K.F.; Wang, T.; Chan, C.C.H.; Lee, T.M.C. PM2.5 Exposure Suppresses Dendritic Maturation in Subgranular Zone in Aged Rats. Neurotoxicity Research 2017;32:50-57

Clouaire, T.; Stancheva, I. Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cellular and Molecular Life Sciences 2008;65:1509-1522

Cochran, W.G. The Combination of Estimates from Different Experiments. Biometrics 1954;10:101-129

Cohen, A.J.; Brauer, M.; Burnett, R.; Anderson, H.R.; Frostad, J.; Estep, K.; Balakrishnan, K.; Brunekreef, B.; Dandona, L.; Dandona, R.; Feigin, V.; Freedman, G.; Hubbell, B.; Jobling, A.; Kan, H.; Knibbs, L.; Liu, Y.; Martin, R.; Morawska, L.; Pope, C.A., III; Shin, H.; Straif, K.; Shaddick, G.; Thomas, M.; van Dingenen, R.; van Donkelaar, A.; Vos, T.; Murray, C.J.L.; Forouzanfar, M.H. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The Lancet 2017;389:1907-1918

Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74

Cornbleet, P.J.; Gochman, N. Incorrect least-squares regression coefficients in method-comparison analysis. Clinical Chemistry 1979;25:432-438

de F. C. Lichtenfels, A.J.; van der Plaat, D.A.; de Jong, K.; van Diemen, C.C.; Postma, D.S.; Nedeljkovic, I.; van Duijn, C.M.; Amin, N.; la Bastide-van Gemert, S.; de Vries, M.; Ward-Caviness, C.K.; Wolf, K.; Waldenberger, M.; Peters, A.; Stolk, R.P.; Brunekreef, B.; Boezen, H.M.; Vonk, J.M. Long-term Air Pollution Exposure, Genome-wide DNA Methylation and Lung Function in the LifeLines Cohort Study. Environ Health Perspect 2018;126:027004

Devlin, B.; Roeder, K.; Wasserman, L. Genomic control, a new approach to genetic-based association studies. Theoretical population biology 2001;60:155-166

Di, Q.; Wang, Y.; Zanobetti, A.; Wang, Y.; Koutrakis, P.; Choirat, C.; Dominici, F.; Schwartz, J.D. Air Pollution and Mortality in the Medicare Population. New England Journal of Medicine 2017;376:2513-2522

Dominici, F.; Peng, R.D.; Bell, M.L.; Pham, L.; McDermott, A.; Zeger, S.L. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. JAMA 2006;295

Duan, D.D. Phenomics of Cardiac Chloride Channels. Comprehensive Physiology; 2013

Engler, R.L.; Yellon, D.M. Sulfonylurea KATP Blockade in Type II Diabetes and Preconditioning in Cardiovascular Disease. Time for Reconsideration 1996;94:2297-2301

EPA. What are the Air Quality Standards for PM? : U.S. Environmental Protection Agency; 2017

Fan, S.; Zhang, X. CpG island methylation pattern in different human tissues and its correlation with gene expression. Biochemical and Biophysical Research Communications 2009;383:421-425

Furney, S.J.; Simmons, A.; Breen, G.; Pedroso, I.; Lunnon, K.; Proitsi, P.; Hodges, A.; Powell, J.; Wahlund, L.O.; Kloszewska, I.; Mecocci, P.; Soininen, H.; Tsolaki, M.; Vellas, B.; Spenger, C.; Lathrop, M.; Shen, L.; Kim, S.; Saykin, A.J.; Weiner, M.W.; Lovestone, S. Genome-wide association with MRI atrophy measures as a quantitative trait locus for Alzheimer's disease. Molecular Psychiatry 2010;16:1130

Gan, W.Q.; FitzGerald, J.M.; Carlsten, C.; Sadatsafavi, M.; Brauer, M. Associations of Ambient Air Pollution with Chronic Obstructive Pulmonary Disease Hospitalization and Mortality. American Journal of Respiratory and Critical Care Medicine 2013;187:721-727

Gondalia, R.; Avery, C.L.; Napier, M.D.; Mendez-Giraldez, R.; Stewart, J.D.; Sitlani, C.M.; Li, Y.; Wilhelmsen, K.C.; Duan, Q.; Roach, J.; North, K.E.; Reiner, A.P.; Zhang, Z.M.; Tinker, L.F.; Yanosky, J.D.; Liao, D.; Whitsel, E.A. Genome-wide Association Study of Susceptibility to Particulate Matter-Associated QT Prolongation. Environ Health Perspect 2017;125:067002

Hernan, M.A.; Hernandez-Diaz, S.; Robins, J.M. A structural approach to selection bias. Epidemiology 2004;15:615-625

Holle, R.; Happich, M.; Löwel, H.; Wichmann, H.-E.; Group, n.f.t.M.K.S. KORA-a research platform for population based health research. Das Gesundheitswesen 2005;67:19-25

Houseman, E.A.; Accomando, W.P.; Koestler, D.C.; Christensen, B.C.; Marsit, C.J.; Nelson, H.H.; Wiencke, J.K.; Kelsey, K.T. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC bioinformatics 2012;13:1

Joehanes, R.; Just, A.C.; Marioni, R.E.; Pilling, L.C.; Reynolds, L.M.; Mandaviya, P.R.; Guan, W.; Xu, T.; Elks, C.E.; Aslibekyan, S.; Moreno-Macias, H.; Smith, J.A.; Brody, J.A.; Dhingra, R.; Yousefi, P.; Pankow, J.S.; Kunze, S.; Shah, S.; McRae, A.F.; Lohman, K.; Sha, J.; Absher, D.M.; Ferrucci, L.; Zhao, W.; Demerath, E.W.; Bressler, J.; Grove, M.L.; Huan, T.; Liu, C.; Mendelson, M.M.; Yao, C.; Kiel, D.P.; Peters, A.; Wang-Sattler, R.; Visscher, P.M.; Wray, N.R.; Starr, J.M.; Ding, J.; Rodriguez, C.J.; Wareham, N.J.; Irvin, M.R.; Zhi, D.; Barrdahl, M.; Vineis, P.; Ambatipudi, S.; Uitterlinden, A.G.; Hofman, A.; Schwartz, J.; Colicino, E.; Hou, L.; Vokonas, P.S.; Hernandez, D.G.; Singleton, A.B.; Bandinelli, S.; Turner, S.T.; Ware, E.B.; Smith, A.K.; Klengel, T.; Binder, E.B.; Psaty, B.M.; Taylor, K.D.; Gharib, S.A.; Swenson, B.R.; Liang, L.; DeMeo, D.L.; Connor, G.T.; Herceg, Z.; Ressler, K.J.; Conneely, K.N.; Sotoodehnia, N.; Kardia, S.L.R.; Melzer, D.; Baccarelli, A.A.; van Meurs, J.B.J.; Romieu, I.; Arnett, D.K.; Ong, K.K.; Liu, Y.; Waldenberger, M.; Deary, I.J.; Fornage, M.; Levy, D.; London, S.J. Epigenetic Signatures of Cigarette Smoking. Circulation: Cardiovascular Genetics 2016;

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

Jordahl, K.M.; Randolph, T.W.; Song, X.; Sather, C.L.; Tinker, L.F.; Phipps, A.I.; Kelsey, K.T.; White, E.; Bhatti, P. Genome-Wide DNA Methylation in Prediagnostic Blood and Bladder Cancer Risk in the Women's Health Initiative. Cancer Epidemiology, Biomarkers & Prevention 2018;27:689

Kahr, P.C.; Piccini, I.; Fabritz, L.; Greber, B.; Schöler, H.; Scheld, H.H.; Hoffmeier, A.; Brown, N.A.; Kirchhof, P. Systematic Analysis of Gene Expression Differences between Left and Right Atria in Different Mouse Strains and in Human Atrial Tissue. PloS one 2011;6:e26389

Kennedy, E.M.; Goehring, G.N.; Nichols, M.H.; Robins, C.; Mehta, D.; Klengel, T.; Eskin, E.; Smith, A.K.; Conneely, K.N. An integrated -omics analysis of the epigenetic landscape of gene expression in human blood cells. BMC Genomics 2018;19:476

Kirchhof, P.; Kahr, P.C.; Kaese, S.; Piccini, I.; Vokshi, I.; Scheld, H.-H.; Rotering, H.; Fortmueller, L.; Laakmann, S.; Verheule, S.; Schotten, U.; Fabritz, L.; Brown, N.A. PITX2c Is Expressed in the Adult Left Atrium, and Reducing Pitx2c Expression Promotes Atrial Fibrillation Inducibility and Complex Changes in Gene Expression. Circulation: Cardiovascular Genetics 2011;4:123-133

Laumbach, R.J.; Kipen, H.M. Respiratory health effects of air pollution: Update on biomass smoke and traffic pollution. Journal of Allergy and Clinical Immunology 2012;129:3-11

Lee, S.-J.; Serre Marc, L.; van Donkelaar, A.; Martin Randall, V.; Burnett Richard, T.; Jerrett, M. Comparison of Geostatistical Interpolation and Remote Sensing Techniques for Estimating Long-Term Exposure to Ambient PM2.5 Concentrations across the Continental United States. Environmental Health Perspectives 2012;120:1727-1732

Leonard, C.E.; Hennessy, S.; Han, X.; Siscovick, D.S.; Flory, J.H.; Deo, R. Pro- and Antiarrhythmic Actions of Sulfonylureas: Mechanistic and Clinical Evidence. Trends in Endocrinology & Metabolism 2017;28:561-586

Liao, D.; Peuquet, D.J.; Duan, Y.; Whitsel, E.A.; Dou, J.; Smith, R.L.; Lin, H.-M.; Chen, J.-C.; Heiss, G. GIS approaches for the estimation of residential-level ambient PM concentrations. Environmental health perspectives 2006:1374-1380

Liao, D.; Peuquet, D.J.; Lin, H.-M.; Duan, Y.; Whitsel, E.A.; Smith, R.L.; Heiss, G. National Kriging Exposure Estimation: Liao et al. Respond. Environmental Health Perspectives 2007;115:A338-A339

Liberda, E.N.; Cuevas, A.K.; Gillespie, P.A.; Grunig, G.; Qu, Q.; Chen, L.C. Exposure to inhaled nickel nanoparticles causes a reduction in number and function of bone marrow endothelial progenitor cells. Inhalation Toxicology 2010;22:95-99

Lonsdale, J.; Thomas, J.; Salvatore, M.; Phillips, R.; Lo, E.; Shad, S.; Hasz, R.; Walters, G.; Garcia, F.; Young, N.; Foster, B.; Moser, M.; Karasik, E.; Gillard, B.; Ramsey, K.; Sullivan, S.; Bridge, J.; Magazine, H.; Syron, J.; Fleming, J.; Siminoff, L.; Traino, H.; Mosavel, M.; Barker, L.; Jewell, S.; Rohrer, D.; Maxim, D.; Filkins, D.; Harbach, P.; Cortadillo, E.; Berghuis, B.; Turner, L.; Hudson, E.; Feenstra, K.; Sobin, L.; Robb, J.; Branton, P.; Korzeniewski, G.; Shive, C.; Tabor, D.; Qi, L.; Groch, K.; Nampally, S.; Buia, S.; Zimmerman, A.; Smith, A.; Burges, R.; Robinson, K.; Valentino, K.; Bradbury, D.; Cosentino, M.; Diaz-Mayoral, N.; Kennedy, M.; Engel, T.; Williams, P.; Erickson, K.; Ardlie, K.; Winckler, W.; Getz, G.; DeLuca, D.; MacArthur, D.; Kellis, M.; Thomson, A.; Young, T.; Gelfand, E.; Donovan, M.; Meng, Y.; Grant, G.; Mash, D.; Marcus, Y.; Basile, M.; Liu, J.; Zhu, J.; Tu, Z.; Cox, N.J.; Nicolae, D.L.; Gamazon, E.R.; Im, H.K.; Konkashbaev, A.; Pritchard, J.; Stevens, M.; Flutre, T.; Wen, X.; Dermitzakis, E.T.; Lappalainen, T.; Guigo, R.; Monlong, J.; Sammeth, M.; Koller, D.; Battle, A.; Mostafavi, S.; McCarthy, M.; Rivas, M.; Maller, J.; Rusyn, I.; Nobel, A.; Wright, F.; Shabalin, A.; Feolo, M.; Sharopova, N.; Sturcke, A.; Paschal, J.; Anderson, J.M.; Wilder, E.L.; Derr, L.K.; Green, E.D.;

Struewing, J.P.; Temple, G.; Volpi, S.; Boyer, J.T.; Thomson, E.J.; Guyer, M.S.; Ng, C.; Abdallah, A.; Colantuoni, D.; Insel, T.R.; Koester, S.E.; Little, A.R.; Bender, P.K.; Lehner, T.; Yao, Y.; Compton, C.C.; Vaught, J.B.; Sawyer, S.; Lockhart, N.C.; Demchok, J.; Moore, H.F. The Genotype-Tissue Expression (GTEx) project. Nature genetics 2013;45:580

Ma, B.; Wilker, E.H.; Willis-Owen, S.A.G.; Byun, H.-M.; Wong, K.C.C.; Motta, V.; Baccarelli, A.A.; Schwartz, J.; Cookson, W.O.C.M.; Khabbaz, K.; Mittleman, M.A.; Moffatt, M.F.; Liang, L. Predicting DNA methylation level across human tissues. Nucleic Acids Research 2014;42:3515-3528

Mathar, I.; Kecskes, M.; Van Der Mieren, G.; Jacobs, G.; Uhl, S.; Camacho Londoño, J.E.; Flockerzi, V.; Voets, T.; Freichel, M.; Nilius, B.; Herijgers, P.; Vennekens, R. Increased ß-Adrenergic Inotropy in Ventricular Myocardium from Trpm4^{-/-} Mice. Circulation Research 2013;

McCullough, S.D.; Dhingra, R.; Fortin, M.C.; Diaz-Sanchez, D. Air Pollution and the Epigenome: A Model Relationship for the Exploration of Toxicoepigenetics. Current Opinion in Toxicology 2017;

Miller , K.A.; Siscovick , D.S.; Sheppard , L.; Shepherd , K.; Sullivan , J.H.; Anderson , G.L.; Kaufman , J.D. Long-Term Exposure to Air Pollution and Incidence of Cardiovascular Events in Women. New England Journal of Medicine 2007;356:447-458

Mosley, T.H.; Knopman, D.S.; Catellier, D.J.; Bryan, N.; Hutchinson, R.G.; Grothues, C.A.; Folsom, A.R.; Cooper, L.S.; Burke, G.L.; Liao, D.; Szklo, M. Cerebral MRI findings and cognitive functioning. Neurology 2005;64:2056

Najeed, S.A.; Khan, I.A.; Molnar, J.; Somberg, J.C. Differential effect of glyburide glibenclamide and metformin on qt dispersion: a potential adenosine triphosphate sensitive k+ channel effect. American Journal of Cardiology 2002;90:1103-1106

Neidhart, M. DNA methylation and complex human disease ed^eds. Amsterdam: Elsevier; 2016

NIH. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials 1998;19:61-109

O'Day, D.H.; Eshak, K.; Myre, M.A. Calmodulin Binding Proteins and Alzheimer's Disease. Journal of Alzheimer's Disease 2015;46:553-569

Panni, T.; Mehta, A.J.; Schwartz, J.D.; Baccarelli, A.A.; Just, A.C.; Wolf, K. A 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 2016;

Plusquin, M.; Guida, F.; Polidoro, S.; Vermeulen, R.; Raaschou-Nielsen, O.; Campanella, G.; Hoek, G.; Kyrtopoulos, S.A.; Georgiadis, P.; Naccarati, A.; Sacerdote, C.; Krogh, V.; Bas Bueno-de-Mesquita, H.; Monique Verschuren, W.M.; Sayols-Baixeras, S.; Panni, T.; Peters, A.; Hebels, D.G.A.J.; Kleinjans, J.; Vineis, P.; Chadeau-Hyam, M. DNA methylation and exposure to ambient air pollution in two prospective cohorts. Environment International 2017;108:127-136

Pope, C.A., 3rd; Burnett, R.T.; Thun, M.J.; Calle, E.E.; Krewski, D.; Ito, K.; Thurston, G.D. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Jama 2002;287:1132-1141

Pyo, R.T.; Sui, J.; Dhume, A.; Palomeque, J.; Blaxall, B.C.; Diaz, G.; Tunstead, J.; Logothetis, D.E.; Hajjar, R.J.; Schecter, A.D. CXCR4 modulates contractility in adult cardiac myocytes. Journal of Molecular and Cellular Cardiology 2006;41:834-844

Raaschou-Nielsen, O.; Andersen, Z.J.; Beelen, R.; Samoli, E.; Stafoggia, M.; Weinmayr, G.; Hoffmann, B.; Fischer, P.; Nieuwenhuijsen, M.J.; Brunekreef, B.; Xun, W.W.; Katsouyanni, K.; Dimakopoulou, K.; Sommar, J.; Forsberg, B.; Modig, L.; Oudin, A.; Oftedal, B.; Schwarze, P.E.; Nafstad, P.; De Faire, U.; Pedersen, N.L.; Östenson, C.-G.; Fratiglioni, L.; Penell, J.; Korek, M.; Pershagen, G.; Eriksen, K.T.; Sørensen, M.; Tjønneland, A.; Ellermann, T.; Eeftens, M.; Peeters, P.H.; Meliefste, K.; Wang, M.; Bueno-de-Mesquita, B.; Key, T.J.; de Hoogh, K.; Concin, H.; Nagel, G.; Vilier, A.; Grioni, S.; Krogh, V.; Tsai, M.-Y.; Ricceri, F.; Sacerdote, C.; Galassi, C.; Migliore, E.; Ranzi, A.; Cesaroni, G.; Badaloni, C.; Forastiere, F.; Tamayo, I.; Amiano, P.; Dorronsoro, M.; Trichopoulou, A.; Bamia, C.; Vineis, P.; Hoek, G. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). The Lancet Oncology 2013;14:813-822

Rab, A.; Rowe, S.M.; Raju, S.V.; Bebok, Z.; Matalon, S.; Collawn, J.F. Cigarette smoke and CFTR: implications in the pathogenesis of COPD. American Journal of Physiology-Lung Cellular and Molecular Physiology 2013;305:L530-L541

Rakhilin, S.V.; Olson, P.A.; Nishi, A.; Starkova, N.N.; Fienberg, A.A.; Nairn, A.C.; Surmeier, D.J.; Greengard, P. A Network of Control Mediated by Regulator of Calcium/Calmodulin-Dependent Signaling. Science 2004;306:698-701

Rasmussen, J.E.; Sheridan, J.T.; Polk, W.; Davies, C.M.; Tarran, R. Cigarette Smoke-induced Ca2+ Release Leads to Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Dysfunction. Journal of Biological Chemistry 2014;289:7671-7681

Rehfeld, F.; Maticzka, D.; Grosser, S.; Knauff, P.; Eravci, M.; Vida, I.; Backofen, R.; Wulczyn, F.G. The RNA-binding protein ARPP21 controls dendritic branching by functionally opposing the miRNA it hosts. Nature Communications 2018;9:1235

Roux , A.V.D.; Merkin , S.S.; Arnett , D.; Chambless , L.; Massing , M.; Nieto , F.J.; Sorlie , P.; Szklo , M.; Tyroler , H.A.; Watson , R.L. Neighborhood of Residence and Incidence of Coronary Heart Disease. New England Journal of Medicine 2001;345:99-106

Saffari, A.; Silver, M.J.; Zavattari, P.; Moi, L.; Columbano, A.; Meaburn, E.L.; Dudbridge, F. Estimation of a significance threshold for epigenome-wide association studies. Genetic Epidemiology 2018;42:20-33

Saint-Criq, V.; Gray, M.A. Role of CFTR in epithelial physiology. Cellular and Molecular Life Sciences 2017;74:93-115

Stuart, E.A.; Azur, M.; Frangakis, C.; Leaf, P. Multiple imputation with large data sets: a case study of the Children's Mental Health Initiative. Am J Epidemiol 2009;169:1133-1139

Stunnenberg, H.G.; Abrignani, S.; Adams, D.; de Almeida, M.; Altucci, L.; Amin, V.; Amit, I.; Antonarakis, S.E.; Aparicio, S.; Arima, T.; Arrigoni, L.; Arts, R.; Asnafi, V.; Esteller, M.; Bae, J.-B.; Bassler, K.; Beck, S.; Berkman, B.; Bernstein, B.E.; Bilenky, M.; Bird, A.; Bock, C.; Boehm, B.; Bourque, G.; Breeze, C.E.; Brors, B.; Bujold, D.; Burren, O.; Bussemakers, M.J.; Butterworth, A.; Campo, E.; Carrillo-de-Santa-Pau, E.; Chadwick, L.; Chan, K.M.; Chen, W.; Cheung, T.H.; Chiapperino, L.; Choi, N.H.; Chung, H.-R.; Clarke, L.; Connors, J.M.; Cronet, P.; Danesh, J.; Dermitzakis, M.; Drewes, G.; Durek, P.; Dyke, S.; Dylag, T.; Eaves, C.J.; Ebert, P.; Eils, R.; Eils, J.; Ennis, C.A.; Enver, T.; Feingold, E.A.; Felder, B.; Ferguson-Smith, A.; Fitzgibbon, J.; Flicek, P.; Foo, R.S.Y.; Fraser, P.; Frontini, M.; Furlong, E.; Gakkhar, S.; Gasparoni, N.; Gasparoni, G.; Geschwind, D.H.; Glažar, P.; Graf, T.; Grosveld, F.; Guan, X.-Y.; Guigo, R.; Gut, I.G.; Hamann, A.; Han, B.-G.; Harris, R.A.; Heath, S.; Helin, K.; Hengstler, J.G.; Heravi-Moussavi, A.; Herrup, K.; Hill, S.; Hilton, J.A.; Hitz, B.C.; Horsthemke, B.; Hu, M.; Hwang, J.-Y.; Ip, N.Y.; Ito, T.; Javierre, B.-M.; Jenko, S.; Jenuwein, T.; Joly, Y.; Jones, S.J.M.; Kanai, Y.; Kang, H.G.; Karsan, A.; Kiemer, A.K.; Kim, S.C.; Kim, B.-J.; Kim, H.-H.; Kimura, H.; Kinkley, S.; Klironomos, F.; Koh, I.-U.; Kostadima, M.; Kressler, C.; Kreuzhuber, R.; Kundaje, A.; Küppers, R.; Larabell, C.; Lasko, P.; Lathrop, M.; Lee, D.H.S.; Lee, S.; Lehrach, H.; Leitão, E.; Lengauer, T.; Lernmark, Å.; Leslie, R.D.; Leung, G.K.K.; Leung, D.; Loeffler, M.; Ma, Y.; Mai, A.; Manke, T.; Marcotte, E.R.; Marra, M.A.; Martens, J.H.A.; Martin-Subero, J.I.; Maschke, K.; Merten, C.; Milosavljevic, A.; Minucci, S.; Mitsuyama, T.; Moore, R.A.; Müller, F.; Mungall, A.J.; Netea, M.G.; Nordström, K.; Norstedt, I.; Okae, H.; Onuchic, V.; Ouellette, F.; Ouwehand, W.; Pagani, M.; Pancaldi, V.; Pap, T.; Pastinen, T.; Patel, R.; Paul, D.S.; Pazin, M.J.; Pelicci, P.G.; Phillips, A.G.; Polansky, J.; Porse, B.; Pospisilik, J.A.; Prabhakar, S.; Procaccini, D.C.; Radbruch, A.; Rajewsky, N.; Rakyan, V.; Reik, W.; Ren, B.; Richardson, D.; Richter, A.; Rico, D.; Roberts, D.J.; Rosenstiel, P.; Rothstein, M.; Salhab, A.; Sasaki, H.; Satterlee, J.S.; Sauer, S.; Schacht, C.; Schmidt, F.; Schmitz, G.; Schreiber, S.; Schröder, C.; Schübeler, D.; Schultze, J.L.; Schulyer, R.P.; Schulz, M.; Seifert, M.; Shirahige, K.; Siebert, R.; Sierocinski, T.; Siminoff, L.; Sinha, A.; Soranzo, N.; Spicuglia, S.; Spivakov, M.; Steidl, C.; Strattan, J.S.; Stratton, M.; Südbeck, P.; Sun, H.; Suzuki, N.; Suzuki, Y.; Tanay, A.; Torrents, D.; Tyson, F.L.; Ulas, T.; Ullrich, S.; Ushijima, T.; Valencia, A.; Vellenga, E.; Vingron, M.; Wallace, C.; Wallner, S.; Walter, J.; Wang, H.; Weber, S.; Weiler, N.; Weller, A.; Weng, A.; Wilder, S.; Wiseman, S.M.; Wu, A.R.; Wu, Z.; Xiong, J.; Yamashita, Y.; Yang, X.; Yap, D.Y.; Yip, K.Y.; Yip, S.; Yoo, J.-I.; Zerbino, D.; Zipprich, G.; Hirst, M. The International Human Epigenome Consortium: A Blueprint for Scientific Collaboration and Discovery. Cell 2016;167:1145-1149

Tarantini, L.; Bonzini, M.; Apostoli, P.; Pegoraro, V.; Bollati, V.; Marinelli, B.; Cantone, L.; Rizzo, G.; Hou, L.; Schwartz, J.; Bertazzi, P.A.; Baccarelli, A. Effects of Particulate Matter on Genomic DNA Methylation Content and iNOS Promoter Methylation. Environmental Health Perspectives 2009;117:217-222

Tarantini, L.; Bonzini, M.; Tripodi, A.; Angelici, L.; Nordio, F.; Cantone, L.; Apostoli, P.; Bertazzi, P.A.; Baccarelli, A.A. Blood hypomethylation of inflammatory genes mediates the effects of metal-rich airborne pollutants on blood coagulation. Occup Environ Med 2013;70

Teschendorff, A.E.; Marabita, F.; Lechner, M.; Bartlett, T.; Tegner, J.; Gomez-Cabrero, D.; Beck, S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 2013;29:189-196

Tsai, P.-C.; Spector, T.D.; Bell, J.T. Using epigenome-wide association scans of DNA methylation in age-related complex human traits. Epigenomics 2012;4:511-526

Uckelmann, H.; Blaszkiewicz, S.; Nicolae, C.; Haas, S.; Schnell, A.; Wurzer, S.; Wagener, R.; Aszodi, A.; Essers, M.A.G. Extracellular matrix protein Matrilin-4 regulates stress-induced HSC proliferation via CXCR4. The Journal of Experimental Medicine 2016;

Velayudhan, L.; Proitsi, P.; Westman, E.; Muehlboeck, J.S.; Mecocci, P.; Vellas, B.; Tsolaki, M.; Kloszewska, I.; Soininen, H.; Spenger, C.; Hodges, A.; Powell, J.; Lovestone, S.; Simmons, A. Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. Journal of Alzheimer's disease : JAD 2013;33:755-766

Welter, D.; MacArthur, J.; Morales, J.; Burdett, T.; Hall, P.; Junkins, H.; Klemm, A.; Flicek, P.; Manolio, T.; Hindorff, L.; Parkinson, H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Research 2014;42:D1001-D1006

Whitsel, E.A. AS315 - Epigenetic Mechanisms of PM-mediated CVD Risk.

Whitsel, E.A.; Quibrera, P.M.; Smith, R.L.; Catellier, D.J.; Liao, D.; Henley, A.C.; Heiss, G. Accuracy of commercial geocoding: assessment and implications. Epidemiologic perspectives & innovations : EP+I 2006;3:8

Whitsel, E.A.; Rose, K.M.; Wood, J.L.; Henley, A.C.; Liao, D.; Heiss, G. Accuracy and repeatability of commercial geocoding. Am J Epidemiol 2004;160:1023-1029

Wichmann, H.; Gieger, C.; Illig, T.; Group, M.K.S. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany)) 2005;67:S26

Yamazaki, J.; Hume, J.R. Inhibitory Effects of Glibenclamide on Cystic Fibrosis Transmembrane Regulator, Swelling-Activated, and Ca2+-Activated Cl− Channels in Mammalian Cardiac Myocytes. Circulation Research 1997;81:101-109

Yanosky, J.D.; Paciorek, C.J.; Laden, F.; Hart, J.E.; Puett, R.C.; Liao, D.; Suh, H.H. Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors. Environmental health : a global access science source 2014;13:63

Zhong, J.; Agha, G.; Baccarelli, A.A. The Role of DNA Methylation in Cardiovascular Risk and Disease. Methodological Aspects, Study Design, and Data Analysis for Epidemiological Studies 2016;118:119-131

Title: Methylome-wide association study provides evidence of particulate matter air pollution-associated DNA methylation

Authors

Rahul Gondalia[a], Antoine Baldassari[a], Katelyn M Holliday[a,b], Anne E Justice[a,c], Raúl Méndez-Giráldez[a], James D Stewart[a], Duanping Liao[d], Jeff D Yanosky[d], Kasey JM Brennan[e], Stephanie M Engel[a], Kristina M Jordahl[f], Elizabeth Kennedy[g], Cavin K Ward-Caviness[h], Kathrin Wolf[i], Josef Cyrysj[i], Parveen Bhatti[f], Steve Horvath[j,k], Themistocles L Assimes[l], James S Pankow[m], Ellen W Demerath[m], Weihua Guan[n], Myriam Fornage[o], Jan Bressler[p] Kari E North[a,q], Karen N Conneely[r], Yun Li[s,t,u], Lifang Hou[v,w], Andrea A Baccarelli[e], Eric A Whitsel[a,x]

Affiliations

[a]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

[b]Department of Community and Family Medicine, Duke University School of Medicine, Durham, NC

[c]Geisinger Health System, Danville, PA, USA

[d]Division of Epidemiology, Department of Public Health Sciences, Pennsylvania State University College of Medicine, Hershey, PA, USA

[e]Laboratory of Environmental Epigenetics, Departments of Environmental Health Sciences and Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

[f]Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA

[g]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

[h]Environmental Public Health Division, National Health and Environmental Effects Research Laboratory, 104 Mason Farm Rd, Chapel Hill, NC 27514

[i]Institute of Epidemiology, Helmholtz Zentrum München, Ingolstaedter Landstrasse 1, Neuherberg, Germany 85764

[j]Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

[k]Biostatistics, School of Public Health, University of California Los Angeles, Los Angeles, CA, , USA

[l]Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

[m]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN, USA

[n]Division of Biostatistics, University of Minnesota, Minneapolis, MN, USA

[o]Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA
[p]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA
[q]Carolina Center for Genome Sciences, University of North Carolina, Chapel Hill, NC, USA
[r]Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA
[s]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA
[t]Department of Biostatistics, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA
[u]Department of Computer Science, University of North Carolina, Chapel Hill, NC, USA
[v]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University Chicago, Evanston, IL, USA
[w]Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University Chicago, Evanston, IL, USA
[x]Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC, USA

Corresponding Author:
Rahul Gondalia, MPH
123 W. Franklin St.
Chapel Hill, North Carolina 27516, USA
Phone: 304-210-7823
Email**:** **rahgonda@unc.edu**

Supplementary text

The Cooperative Health Research in the Region of Augsburg (KORA) study is a population-based cohort from the region of Augsburg, Southern Germany. Replication analyses involved data from the F3 (n = 3,006; 2004-2005) and F4 (n = 3,080; 2006-2008) follow-up studies of the KORA S3 and S4 participants (Rückert et al. 2011; Wichmann et al. 2005).

DNA methylation was analyzed from whole blood samples in 500 (F3) and 1799 (F4) participants using the Infinium HumanMethylation450 BeadChip Array (Illumina). Probes with signals from less than three functional beads, a detection *P* value > 0.05 in > 1% of samples, or covered single nucleotide polymorphisms (minor allele frequency in Europeans > 5%) were excluded. Sample exclusions included participants with a detection *P* value > 0.05 for > 1% of probes and those with a gender mismatch. DNAm measures were Beta Mixture Quantile (BMIQ)-normalized to adjust for probe bias (Teschendorff et al. 2013). DNAm at three CpG sites was analyzed: cg19004594, cg24102420, and cg12124767. Analyses controlled for technical variation by adjusting for CD4 T-cells, plasmablasts, natural killer cells, CD8 naive T-cells, monocytes, granulocytes, and a linear combination of CD8, CD45RA, and CD28 T-cells (Horvath 2013). Analyses also controlled for plate and batch effects using 20 principal components calculated from the control probes. Moreover, analyses controlled for demographic and clinical variables collected via standardized questionnaires at each visit, as well as meteorological variables: age, sex, years of education, smoking status (current regular, current irregular, former, never), alcohol consumption (alcohol usage, no alcohol usage), physical activity (active, inactive), body mass index (Rückert et al. 2011), mean temperature, mean barometric pressure, and mean relative humidity.

1 Rückert, I.-M. *et al.* Association between Markers of Fatty Liver Disease and Impaired Glucose Regulation in Men and Women from the General Population: The KORA-F4-Study. *PLoS ONE* **6**, e22932, doi:10.1371/journal.pone.0022932 (2011).

2 Horvath, S. DNA methylation age of human tissues and cell types. *Genome biology* **14**, R115 (2013).

3 Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98**, 288-295, doi:10.1016/j.ygeno.2011.07.007 (2011).

4 Smyth, G. K. in *Bioinformatics and computational biology solutions using R and Bioconductor* 397-420 (Springer, 2005).

5 Teschendorff, A. E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189-196 (2013).

6 Pidsley, R. *et al.* A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC genomics* **14**, 293 (2013).

Table S1. Methylome-wide DNAm data exclusions in WHI and ARIC

Study	Sample Exclusions		Probe Exclusions				
	N after exclusions[a]	Detection p-value	n CpGs after exclusions[b]	Detection p-value	Y Chr	Bead Count	Non-CpG CH_3
WHI-EMPC[c]	1,937	> 0.01 in > 1%[d]	463,916	> 0.01 in > 10%[e]	Yes	No	No
WHI-BAA23	1,950	No	461,014	> 0.01 in > 10%[e]	Yes	No	Yes
WHI-AS311	746	No	461,136	> 0.01 in > 1%[e]	Yes	< 3 in > 10%[e]	Yes
ARIC-AA	2,664	> 0.01 in > 1%[d]	463,431	> 0.01 in > 1%[e]	No	< 3 in > 5%[e]	No
ARIC-EA	1,100	> 0.01 in > 1%[d]	462,543	> 0.01 in > 5%[e]	No	< 3 in >5%[e]	No

[a]Additional study-specific sample exclusions: gender mismatch or SNP discordance with previous genotyping, and / or outliers in principal component analysis

[b]Additional probe exclusion: CpG sites with multi-modal DNAm distributions in ≥ 1 study

[c]185 participants had a second and 43 had a third DNAm measure at a subsequent visit (n observations = 2,165)

[d]Of probes

[e]Of samples

Table S2. Mean concentrations (μg/m^3) of particulate matter (PM) by study

Study	Race / Ethnicity	PM_{10}						$PM_{2.5}$		$PM_{2.5-10}$	
		2 d $\bar{x}$(SD)	7 d $\bar{x}$(SD)	28 days $\bar{x}$(SD)	365 d $\bar{x}$(SD)	1 mo $\bar{x}$(SD)	12 mo $\bar{x}$(SD)	1 mo $\bar{x}$(SD)	12 mo $\bar{x}$(SD)	1 mo $\bar{x}$(SD)	12 mo $\bar{x}$(SD)
ARIC	AA	36.0 (12.3)	35.1 (9.1)	34.8 (6.3)	35.5 (3.3)	20.5 (4.6)	19.9 (1.69)	13.2 (3.1)	12.7 (1.3)	7.3 (2.1)	7.2 (0.8)
ARIC	EA	36.1 (11.5)	34.9 (8.2)	34.4 (5.8)	34.8 (3.0)	23.2 (5.3)	23.7 (2.4)	15.4 (4.3)	15.9 (2.1)	7.8 (3.5)	7.8 (1.4)
WHI-AS311[a]	EA	28.0 (11.0)	27.1 (7.9)	27.4 (6.5)	27.5 (4.1)	19.8 (6.6)	20.0 (4.8)	11.9 (3.82)	11.9 (2.7)	7.9 (4.6)	8.1 (3.8)
WHI-AS311[b]	EA	28.7 (11.1)	27.7 (8.9)	27.6 (6.6)	27.6 (4.2)	19.9 (5.7)	20.2 (4.5)	12.0 (3.9)	12.0 (2.6)	7.9 (4.1)	8.2 (3.5)
WHI-BAA23[a]	AA	28.2 (12.2)	27.0 (7.5)	27.8 (5.6)	28.3 (2.8)	22.6 (6.2)	22.3 (3.7)	14.3 (4.2)	14.1 (2.2)	8.3 (3.8)	8.2 (2.6)
WHI-BAA23[a]	EA	28.1 (10.7)	27.2 (8.4)	27.2 (6.4)	27.5 (4.0)	19.7 (5.7)	20.0 (4.5)	11.7 (3.7)	11.8 (2.5)	8.0 (4.4)	8.2 (3.7)
WHI-BAA23[a]	HLA	28.9 (10.4)	29.3 (8.3)	29.3 (6.8)	29.2 (4.1)	21.4 (8.1)	21.5 (5.9)	10.3 (4.1)	10.3 (3.0)	11.1 (5.7)	11.2 (4.5)
WHI-BAA23[b]	AA	28.8 (11.1)	28.8 (8.5)	28.1 (6.1)	28.1 (2.3)	22.3 (5.9)	22.6 (3.7)	14.0 (4.0)	14.1 (2.2)	8.3 (4.2)	8.5 (3.1)
WHI-BAA23[b]	HLA	30.2 (10.7)	29.3 (8.6)	29.9 (7.2)	30.0 (4.7)	23.0 (8.1)	23.1 (6.1)	11.0 (4.2)	10.9 (3.2)	11.9 (6.4)	12.2 (5.2)
WHI-EMPC[a,c]	AA	29.2 (11.2)	27.9 (7.3)	27.7 (5.5)	28.1 (3.0)	22.2 (6.2)	22.4 (4.3)	15.2 (5.1)	15.1 (3.8)	7.0 (4.7)	7.3 (3.4)
WHI-EMPC[a,c]	EA	28.3 (11.5)	27.3 (8.1)	27.2 (6.4)	27.5 (3.8)	19.4 (6.0)	19.8 (5.8)	13.0 (5.0)	12.9 (3.6)	6.4 (5.2)	6.8 (4.1)
WHI-EMPC[a,c]	HLA	28.5 (9.8)	28.4 (8.3)	28.3 (6.2)	28.3 (4.2)	21.9 (7.1)	22.3 (6.1)	12.8 (6.3)	12.9 (5.4)	9.1 (5.3)	9.4 (4.9)
All		31.9 (12.1)	31.1 (9.2)	30.9 (7.1)	31.2 (5.1)	20.9 (5.8)	20.9 (4.0)	13.2 (4.3)	13.2 (3.0)	7.7 (4.0)	7.8 (3.1)

Abbreviations: AA, African American; ARIC, Atherosclerosis Risk in Communities; AS311, Ancillary Study 311; BAA23OS, Broad Agency Award 23; CpG, Cytosine-phosphate-Guanine site; d, day; EA, European American; EMPC, Epigenetic Mechanisms of PM-Mediated CVD Risk; HLA, Hispanic/Latino American; mo, month; PM, particulate matter; SD, standard deviation; WHI, Women's Health Initiative

[a]WHI clinical trials participants

[b]WHI observational study participants

[c]Data from the first visit are presented for WHI-EMPC; 185 participants had a second and 43 had a third DNAm measure from a subsequent visit

Table S3. Findings from trans-ethnic, fixed-effects inverse variance-weighted meta-analyses ($P < 1 \times 10\text{-}5$, $P_{Cochran's\ Q} > 0.10$) with Illumina 450K Infinium Methylation

[Please see … Supplementary Table 3 uploaded separately due to large size]

Table S4. Comparison of DNA methylation measures from the Illumina 450K Infinium Methylation BeadChip versus bisulfite pyrosequencing

Chr	Position (B37)	CpG	Δ (95% CI)	r (95% CI)	α (95% CI)	β (95% CI)
4	8230847	cg01945624	0.07 (0.06, 0.07)	0.83 (0.78, 0.87)	0.17 (0.14, 0.20)	0.76 (0.68, 0.83)
2	64682236	cg01948201	0.02 (0.02, 0.02)	0.71 (0.63, 0.77)	0.04 (0.01, 0.08)	0.91 (0.78, 1.04)
22	30639730	cg07316313	-0.15 (-0.15, -0.14)	0.75 (0.68, 0.81)	-0.58 (-0.74, -0.42)	1.52 (1.33, 1.72)
9	132383003	cg09731694	0.08 (0.08, 0.08)	0.78 (0.72, 0.83)	0.03 (0.01, 0.06)	1.36 (1.21, 1.51)
8	144790656	cg09754549	-0.03 (-0.04, -0.02)	0.86 (0.82, 0.90)	-0.06 (-0.21, 0.09)	1.04 (0.85, 1.23)
6	159466542	cg16180082	0.04 (0.04, 0.05)	0.61 (0.52, 0.69)	-0.23 (-0.35, -0.12)	2.49 (1.90, 3.07)
7	15726411	cg18580296	0.00 (-0.01, 0.00)	0.22 (0.08, 0.34)	0.08 (0.08, 0.09)	0.09 (0.04, 0.15)
7	2968595	cg22989995	-0.05 (-0.06, -0.05)	0.04 (-0.10, 0.18)	0.82 (0.60, 1.04)	0.07 (-0.17, 0.30)
3	35785890	cg24102420	-0.04 (-0.04, -0.03)	0.79 (0.73, 0.83)	-0.16 (-0.38, 0.07)	1.13 (0.88, 1.39)
7	27225396	cg24988255	0.07 (0.07, 0.08)	0.40 (0.27, 0.51)	0.12 (0.09, 0.15)	0.55 (0.27, 0.82)

Abbreviations: B37, build 37; **Δ**, mean Illumina 450K minus bisulfite pyrosequencing difference in DNAm; Chr, chromosome; CI, confidence interval; CpG, cytosine-phosphate-guanine site; ICC, intra-class correlation; r, Pearson correlation coefficient

Table S5

Findings from Cooperative Health Research in the Region Augsburg study (KORA)

Chr	Position[a]	CpG	Exposure	%Δ (95% CI)[b]	*P*	n_{obs}	Gene
20	43926884	cg19004594	PM_{10}, 28 d	-0.1 (-0.3, 0.1)	0.42	2,168	*MATN4*
3	35785890	cg24102420	PM_{10}, 30 d	-0.2 (-0.6, 0.1)	0.13	2,176	*ARPP21 / MIR128-2*
7	117299297	cg12124767	$PM_{2.5-10}$, 30 d	0.4 (-0.2, 1.0)	0.21	2,036	*CFTR*

Abbreviations: **Δ**, change; Chr, chromosome; CI, confidence interval; CpG, Cytosine-phosphate-Guanine; d, days; PM_{10}, PM < 10 μm in diameter; $PM_{2.5}$, PM < 2.5 μm in diameter; $PM_{2.5-10}$, PM > 2.5 and < 10 μm in diameter

[a]Build 37

[b]Absolute percentage point per 10 μg/m^3 increase in PM_{10}

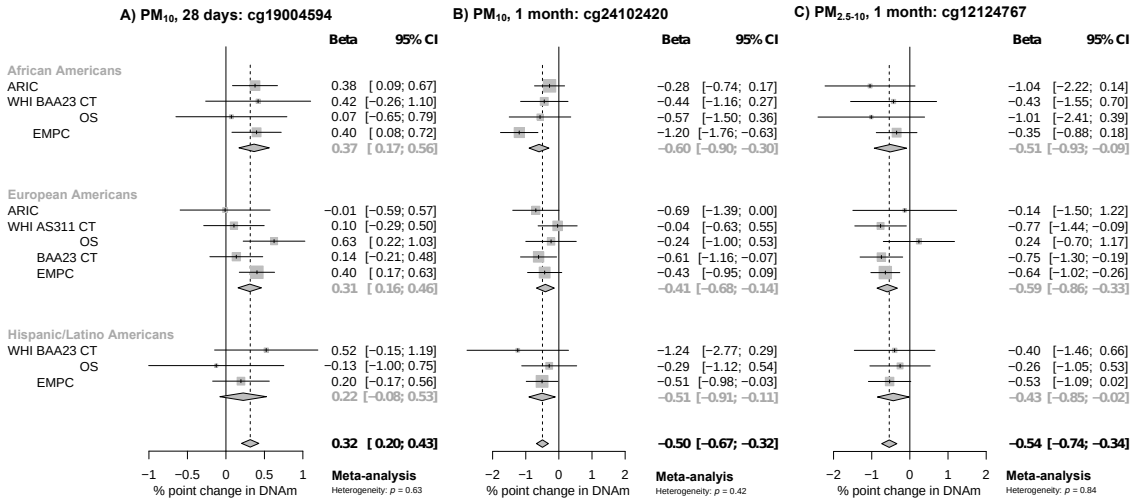

Figure S1. Forest plots of PM-CpG associations (95% confidence intervals) for A) cg19004594, B) cg2410240, and C) cg12124767 with a 10 μg/m³ increase in PM by subpopulation and by race/ethnicity and overall after fixed-effects meta-analysis.

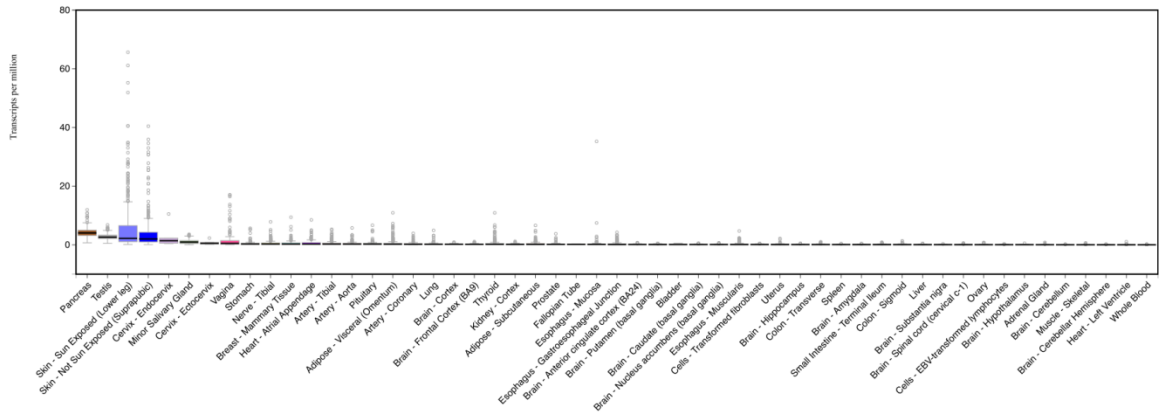

Figure S2. Gene expression for *MATN4*

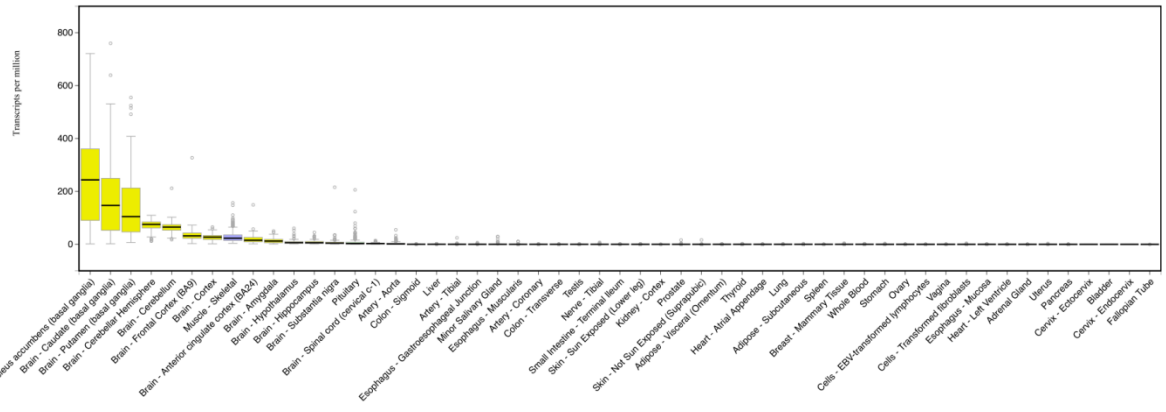

Figure S3. Gene expression for *ARPP21*

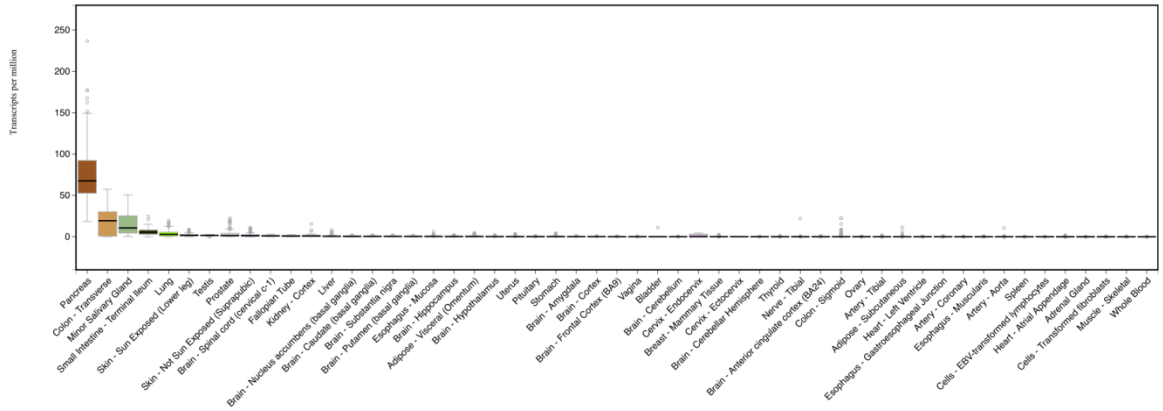

Figure S4. Gene expression for *CFTR*

Bibliography

Horvath S. 2013. DNA methylation age of human tissues and cell types. Genome biology 14:R115.

Rückert I-M, Heier M, Rathmann W, Baumeister SE, Döring A, Meisinger C. 2011. Association between markers of fatty liver disease and impaired glucose regulation in men and women from the general population: The kora-f4-study. PLoS ONE 6:e22932.

Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. 2013. A beta-mixture quantile normalization method for correcting probe design bias in illumina infinium 450 k DNA methylation data. Bioinformatics 29:189-196.

Wichmann HE, Gieger C, Illig T. 2005. Kora-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany)) 67 Suppl 1:S26-30.