



# Epigenetic and Mitochondrial Biomarkers Linking Air Pollution and Temperature on Human Health: The Normative Aging Study

## Citation

Peng, Cheng. 2016. Epigenetic and Mitochondrial Biomarkers Linking Air Pollution and Temperature on Human Health: The Normative Aging Study. Doctoral dissertation, Harvard T.H. Chan School of Public Health.

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:32644535>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available. Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

**Epigenetic and Mitochondrial Biomarkers Linking Air Pollution and Temperature on  
Human Health: The Normative Aging Study**

Cheng Peng

A Dissertation Submitted to the Faculty of  
The Harvard T. H. Chan School of Public Health  
in Partial Fulfillment of the Requirements  
for the Degree of *Doctor of Science*  
in the Department of Environmental Health

Harvard University

Boston Massachusetts

November 2016

**Epigenetic and mitochondrial biomarkers linking air pollution and temperature on human health: The Normative Aging Study**

**Abstract**

Fine particulate matter (particulate matter with aerodynamic diameter  $\leq 2.5$  micron or  $PM_{2.5}$ ) exposure, as well as changes in ambient meteorological conditions are associated with adverse health consequences. However, the underlying mechanisms have not been clearly delineated. Systemic inflammation and oxidative stress are two of the primary pathways proposed to account for the association of  $PM_{2.5}$  / air temperature with health related outcomes. In this dissertation work, we proposed to use two types of novel molecular biomarkers: (1) nuclear DNA (nDNA) methylation, and (2) mitochondrial DNA (mtDNA) integrity to assess inflammatory and oxidative stress pathways linking environmental insults and health.

Specifically, in Chapter I, we evaluated the mediating role of promoter region DNA methylation of inflammatory biomarkers (*IFN- $\gamma$* , *IL-6*, *ICAM-1*, and *TLR-2*) linking  $PM_{2.5}$  exposure and abnormal glucose metabolism in The Normative Aging Study. Our study showed that  $PM_{2.5}$  concentrations are associated with higher fasting blood glucose (FBG) level, and this association was in part mediated through *ICAM-1* gene methylation, particularly at the longer (28-day) moving average investigated. Our study demonstrates a novel approach of mediation analysis in epigenetic studies and highlights a mediating role of *ICAM-1* gene methylation in air-pollution associated abnormal glucose metabolism.

In Chapter II, we assessed the relative effects of  $PM_{2.5}$  mass and  $PM_{2.5}$  components on a novel oxidative stress-related marker—blood mtDNA abundance in The Normative Aging

Study. Our study showed that long-term exposure to PM<sub>2.5</sub> mass and specific PM<sub>2.5</sub> components is associated with decreased mtDNA abundance. Our findings from multi-pollutant modeling suggest that nitrate (NO<sub>3</sub><sup>-</sup>) was associated with higher mitochondrial oxidative stress independent of PM<sub>2.5</sub> mass concentration, and mass alone may not fully capture the oxidation potency of PM<sub>2.5</sub>.

In Chapter III, we explored short-term changes in daily mean and daily standard deviation (SD) (variability) of ambient air temperature with blood mtDNA lesions in The Normative Aging Study. We observed short-term increases in mean air temperature were associated with higher mtDNA lesions in elderly adults, supporting the hypothesis that changes meteorological conditions may induce pathophysiological responses among susceptible populations.

## TABLE OF CONTENTS

<b>ABSTRACT.....</b>	<b>II</b>
<b>TABLE OF CONTENTS.....</b>	<b>IV</b>
<b>FIGURES AND CAPTIONS.....</b>	<b>VIII</b>
<b>TABLES AND CAPTIONS .....</b>	<b>XII</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>XIX</b>
<b>INTRODUCTION.....</b>	<b>1</b>
<b>CHAPTER 1.....</b>	<b>5</b>
<b>PARTICULATE AIR POLLUTION AND FASTING BLOOD GLUCOSE IN NON- DIABETIC INDIVIDUALS: ASSOCIATIONS AND EPIGENETIC MEDIATION IN THE NORMATIVE AGING STUDY, 2000-2011.....</b>	<b>5</b>
ABSTRACT .....	7
INTRODUCTION .....	9
MATERIALS AND METHODS .....	10
Study Population .....	10
Air Pollution and Temperature .....	11
DNA Methylation Measurements .....	12
Fasting Blood Glucose Measurements .....	13
Statistical Analysis of the Main Association of PM <sub>2.5</sub> Levels with FBG .....	13
Statistical Analysis of DNA Methylation and Mediation Analysis .....	15
Selection of Mediators .....	15

Underlying Assumptions .....	15
Mediation Analysis .....	16
 RESULTS	
Descriptive Statistics .....	17
Main Association of PM <sub>2.5</sub> Levels with FBG .....	18
DNA Methylation and Mediation Analysis .....	18
DISCUSSION .....	20
CONCLUSIONS .....	24
ACKNOWLEDGEMENTS .....	25
FUNDINGS .....	25
CONFLICT OF INTEREST .....	25
REFERENCES .....	42
<b>CHAPTER 2.....</b>	<b>.50</b>
<b>Associations of Annual Ambient PM<sub>2.5</sub> Mass and PM<sub>2.5</sub> Components with Mitochondrial Damage in Older Adults: A Multi-pollutant Approach in The Veterans Affairs Normative Aging Study, 2000–2012.....</b>	<b>.50</b>
ABSTRACT.....	52
INTRODUCTION.....	54
MATERIALS AND METHODS.....	56
Study Participants.....	56
Estimates of PM <sub>2.5</sub> Mass and PM <sub>2.5</sub> Components.....	56
Determination of mtDNA Abundance using Quantitative Real Time PCR .....	57

Statistical Analysis .....	58
RESULTS	
Descriptive Statistics .....	61
PM <sub>2.5</sub> Mass and PM <sub>2.5</sub> Component Summary Statistics and Spatial Distribution .....	61
Main Analysis .....	62
Sensitivity Analysis .....	63
DISCUSSION .....	63
CONCLUSIONS .....	67
ACKNOWLEDGEMENTS .....	67
FUNDINGS .....	68
CONFLICT OF INTEREST .....	68
REFERENCES .....	78
<b>CHAPTER 3.....</b>	<b>.83</b>
<b>SHORT-TERM EFFECTS OF AIR TEMPERATURE AND MITOCHONDRIAL DNA</b>	
<b>LESION AMONG OLDER POPULATION.....</b>	<b>.83</b>
ABSTRACT .....	84
INTRODUCTION.....	86
MATERIALS AND METHODS	
Study Population .....	87
Ambient Temperature Measurements.....	88
Mitochondrial DNA Lesion Measurements.....	88
Statistical Analysis.....	89

Sensitivity Analysis .....	91
<b>RESULTS</b>	
Study Population.....	92
Meteorological Data.....	92
Air Temperature and Mitochondrial DNA Lesions.....	93
Sensitivity Analysis.....	94
<b>DISCUSSION.....</b>	<b>94</b>
<b>CONCLUSIONS.....</b>	<b>99</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>100</b>
<b>FUNDINGS.....</b>	<b>100</b>
<b>CONFLICT OF INTEREST.....</b>	<b>100</b>
<b>REFERENCES ....</b>	<b>118</b>
<b>SUMMARY AND CONCLUSIONS.....</b>	<b>125</b>



## FIGURES AND CAPTIONS

### **Figure 1.1: Directed acyclic graph (DAG) for mediation analysis.**

$PM_{2.5}^{i,j=J}$  represents air pollution exposure for  $i^{\text{th}}$  subject prior to  $j=J^{\text{th}}$  visit;  $M^{i,j=J}$  represents gene-specific DNA methylation for  $i^{\text{th}}$  subject at  $j^{\text{th}}$  visit;  $Y^{i,j=J}$  represents fasting blood glucose (FBG) concentrations for  $i^{\text{th}}$  subject at  $j^{\text{th}}$  visit.  $C_1^{i,j}$  represents exposure outcome confounders,  $C_2^{i,j}$  represents exposure mediator confounders;  $C_3^{i,j}$  represents mediator outcome confounders.

Note: to be simplified, correlations between repeated measures of exposures (i.e.  $PM_{2.5ij}$  and  $PM_{2.5ij+1}$ ), repeated measures of mediators (i.e.,  $M_{ij}$  and  $M_{ij+1}$ ) and repeated measures of confounders (i.e.,  $C_{ij}$  and  $C_{ij+1}$ ) are not shown in this DAG.

### **Figure 1.2: Inflammatory candidate gene methylation mediator model of the relationship between $PM_{2.5}$ concentration and fasting blood glucose level.**

*ICAM-1* mean and *TLR-2* mean DNA methylation is log-normally distributed and *IL-6* and *IFN- $\gamma$*  are normally distributed.  $\beta$  is the coefficient of the independent variable ( $PM_{2.5}$  28-day moving average) when regressing the mediator (candidate gene methylation) on the independent variable,  $\gamma$  is the coefficient of the mediator when regressing the dependent variable (FBG) on both the independent variable and the mediator. Results from regression models are adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, batch effects, percentage of lymphocytes, and percentage of neutrophils. Participants with diabetes were excluded.

**Figure S1.1: Methylation of specific CpG sites for the four candidate genes (interferon gamma (*IFN- $\gamma$* ), interleukin-6 (*IL-6*), Toll-like receptor 2 (*TLR-2*), and intracellular adhesion molecule-1 (*ICAM-1*)).**

**Figure S1.2 Timeline of the study design.**

**Figure 2.1: Graphical illustration of annual PM<sub>2.5</sub> mass and PM<sub>2.5</sub> component concentrations for the study period Normative Aging Study, 2000-2012. The green dots indicate where the fixed monitoring stations are located, while the red dots dictate participants' residential address.**

**Figure 2.2: LASSO coefficient paths: plot of PM<sub>2.5</sub> components as a function of penalty term lambda.**

Based on the Lasso shrinkage method with BIC as a selection criterion, we identified NO<sub>3</sub><sup>-</sup> independently associated with mtDNA abundance.

**Figure S2.1: Flow chart of participants' inclusion criteria (participants who came in year 2000 will be assigned to the yearly moving average of year 2000).**

**Figure S2.2: The relationship between the penalty term lambda in LASSO regression and Bayesian information criterion (BIC).**

We performed a mixed model version of Lasso regression, with subject specific intercepts, to account for correlation across multiple visits. We forced the following covariates in: age, BMI,

race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, seasonality, and PM<sub>2.5</sub> mass concentration.

**Figure 3.1: Estimated change (and 95% CI) in mtDNA lesions per IQR increase of daily mean of air temperature over 14-day moving average (by season).**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, metabolic equivalent of tasks (low:  $\leq 12$  hours/week; medium: 12-30 hours/week; high:  $\geq 30$  hours/week), smoking status (current, former, never), pack-years smoked, alcohol consumption, education level, platelet count, lymphocyte count, neutrophil count, plates, and 24-hr relative humidity.

**Figure 3.2 Estimated change (and 95% CI) in mtDNA lesions per IQR increase of daily mean of air temperature over 14-day moving average (by season). We additionally adjusted for within day temperature variability.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, metabolic equivalent of tasks (low:  $\leq 12$  hours/week; medium: 12-30 hours/week; high:  $\geq 30$  hours/week), smoking status (current, former, never), pack-years smoked, alcohol consumption, education level, platelet count, lymphocyte count, neutrophil count, plates, 24-hr relative humidity, and within day temperature variability of the corresponding moving average.

**Figure 3.3: Estimated change (and 95% CI) in mtDNA lesions per IQR increase of daily variability of air temperature over 14-day moving average.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, metabolic equivalent of tasks (low:  $\leq 12$  hours/week; medium: 12-30 hours/week; high:  $\geq 30$  hours/week), smoking status (current, former, never), pack-years smoked, alcohol consumption, education level, platelet count, lymphocyte count, neutrophil count, plates, 24-hr relative humidity and season (spring/summer/fall/winter).

\*We additionally adjusted for within day temperature mean of the corresponding moving average.

## TABLES AND CAPTIONS

**Table 1.1: Characteristics of the Normative Aging Study participants included in the analysis<sup>a</sup> (N = 551 participants), 2000-2011.**

<sup>a</sup>Cohort participants with diabetes (N=105) were excluded. Data for the first visit included in the analysis are shown.

**Table 1.2: Estimated change (and 95% CI) in fasting blood glucose (FBG) level (mg/dL) per interquartile range (IQR) increase in PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) concentration averaged over the corresponding time window before each visit.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.

**Table 1.3: Odds ratio (OR) (and 95% CI) of impaired fasting blood glucose (IFG, defined as a fasting blood glucose level greater than 100 mg/dL and less than 126 mg/dL) per interquartile range (IQR) increase in PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) concentration averaged over the corresponding time window before each visit.**

Results from GEE models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol

consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.

**Table 1.4: Mediation effect investigating whether blood *ICAM-1* methylation mediates the association between air pollution and fasting blood glucose level. Indirect effect represents the “mediated” effect through the *ICAM-1* methylation pathway. Estimates correspond to a standard deviation increase from the mean.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, batch effect, % of lymphocytes, and % of neutrophils. Participants with diabetes were excluded.

<sup>a</sup>Proportion mediated cannot be estimated in this case because  $\beta_{PM_{2.5}}$  and  $\gamma_M$  have opposite signs.

**Table S1.1: Summary statistics of  $PM_{2.5}$  (PM with aerodynamic diameter  $\leq 2.5 \mu m$ ) and temperature levels.  $PM_{2.5}$  and temperature were summarized as cumulative averaged exposures up to the previous 28-day exposure window.  $PM_{2.5}$  was estimated using spatiotemporal land-use regression models estimating levels at the participants’ residential addresses. Temperature values were obtained through the national climatic data center (NCDC); grid cells were matched to the closest weather station for meteorological variables.**

**Table S1.2: Estimated change (and 95% CI) in fasting blood glucose (FBG) level (mg/dL) per interquartile range (IQR) increase in  $PM_{2.5}$  (particulate matter with aerodynamic diameter  $\leq 2.5 \mu m$ ) concentration averaged over the corresponding time window before each visit. We**

**additionally adjusted for dietary variables (total calorie intake and glycemic index).**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, total calorie intake and glycemic index. Participants with diabetes were excluded.

**Table S1.3: Estimated change (and 95% CI) in fasting blood glucose (FBG) level (mg/dL) per interquartile range (IQR) increase in PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5 \mu\text{m}$ ) concentration averaged over the corresponding time window before each visit. We accounted for potential selection bias using inverse probability weighting.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality.

Participants with diabetes were excluded. Potential selection bias due to loss to follow-up was accounted for using inverse probability weighting. In a logistic regression, we predict the probability of coming to a subsequent visit by covariates from the previous one, which include age, BMI, regular patterns of physical activity, smoking status, pack year smoked, FEV1 and FVC ratio, medication (diuretics and beta blocker), and education level.

**Table S1.4: Estimated change (and 95% CI) in inflammatory candidate gene methylation with increase in fasting blood glucose [FBG (mg/dL)] concentrations at previous visits ( $Y_{ij} \rightarrow M_{ij+1}$ ).**

Results from linear mixed-effects regression models accounting for correlation across multiple visits, and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, batch effects, percentage of lymphocytes, and percentage of neutrophils (all from previous visits). Participants with diabetes were excluded. \**ICAM-1* and *TLR-2* DNA methylation measurements were log transformed.

**Table S1.5: Sensitivity analysis of the mediation effect of *ICAM-1* methylation on the association between PM<sub>2.5</sub> concentrations averaged over 28-day exposure window with fasting blood glucose (FBG) level (mg/dL). Natural indirect effect represents the “mediated” effect through the *ICAM-1* methylation pathway. Estimates correspond to 1 µg/m<sup>3</sup> increase in PM<sub>2.5</sub> concentration.**

Results from linear mixed-effects regression models adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, batch effect, percentage of lymphocytes, percentage of neutrophils. Participants with diabetes were excluded. <sup>a</sup> We excluded current smokers to limit potential residual confounding by smoking. <sup>b</sup> We additionally controlled for dietary intake (total calorie intake and glycemic index), to reduce potential confounding by diet. <sup>c</sup> We restricted the analysis to participants with a C-reactive protein (CRP) level less than 10 mg/L, to partially removal potential effect from acute inflammation.

**Table 2.1: Personal characteristics of participants from the Normative Aging Study, 2000-2012 (N = 646).**



**Table 2.2: Summary statistics of annual moving averages of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components, Normative Aging Study, 2000-2012.**

**Table 2.3: Spearman correlation between annual moving average concentrations of the air pollutants considered in the analysis, Normative Aging Study, 2000-2012.**

*\* $p < 0.05$ ; \*\* $p < 0.01$*

**Table 2.4: Association between annual averages of ambient air pollution and mtDNA abundance, measured as mitochondrial DNA to nuclear DNA copy number ratio (mtDNA/nDNA), Normative Aging Study, 2000-2012.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, and seasonality. \*Estimates are expressed as a SD change in mtDNA/nDNA ratio with a SD change in pollution concentration. †We additionally adjusted for PM<sub>2.5</sub> mass concentration. PM<sub>2.5</sub> mass estimate is not reported because we used this set of models to obtain estimates for PM<sub>2.5</sub> components adjusted for PM<sub>2.5</sub> mass. Each individual model provided a different estimate for PM<sub>2.5</sub> mass.

**Table 2.5: Association between annual averages of ambient air pollution and mtDNA abundance, measured as mitochondrial DNA to nuclear DNA copy number ratio (mtDNA/nDNA), Normative Aging Study, 2000-2012. Results are based on LASSO selected pollutants.**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, seasonality, PM<sub>2.5</sub> mass concentration and all other components that were selected by LASSO shrinkage method. \*Estimates are expressed as a SD change in mtDNA/nDNA ratio with a SD change in pollution concentration.

**Table S2.1: Comparison of estimated coefficients in the primary analysis and in sensitivity analysis restricted to observations with annual moving PM<sub>2.5</sub> average  $\leq 12 \mu\text{g}/\text{m}^3$  (i.e., the current NAAQS for annual PM<sub>2.5</sub> concentration).**

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, seasonality, PM<sub>2.5</sub> mass concentration and all other components that were selected by LASSO shrinkage method. \*Estimates are expressed as a SD change in mtDNA/nDNA ratio with a SD change in pollution concentration. †Sensitivity analysis where PM<sub>2.5</sub> annual moving average  $> 12 \mu\text{g}/\text{m}^3$  were excluded.

**Table 3.1: Personal characteristics of participants (N=654).**

**Table 3.2: Summary statistics and Spearman correlation of environmental variables.**

\* $p < 0.05$ ; \*\* $p < 0.01$

## ACKNOWLEDGEMENTS

I would like to express my greatest appreciation to my advisor and committee chair, Dr. Andrea A. Baccarelli, for his academic mentorship and advice in pursuing a career in public health. Without his guidance and persistent help this dissertation would not have been possible. I would also like to thank my dissertation committee members, Dr. Joel Schwartz and Dr. Brent Coull for their time, supports and constructive comments in reviewing my research work.

I have had the great pleasure of working with research teams from The Laboratory of Human Environmental Epigenetics, the Environmental Epidemiology track and The Normative Aging Study (NAS), especially with Dr. Pantel Vokonas, Dr. Marie-Abele Bind, Dr. Letizia Trevisi, Dr. Marco Sanchez-Guerra, Dr. Itai Kloog and Anna Kosheleva, who provided great support over the course of my dissertation development. I would also like to thank Dr. Diane R. Gold, who gave me strong support, encouragement during my doctoral study, and advice in life. I also extend my gratitude to all participants from The Normative Aging Study, Dr. Petros Koutrakis and his team, in particular Dr. Choong-Min Kang and Dr. Mike Wolfson for spending their time help me to understand exposure measurements.

I am grateful to spend the past five years with my best friends, Changzheng Yuan, Yuanyuan Shen and Yuan Chen, and many other friends and classmates in the U.S. Their support and help has made all the difference.

Finally, I would like to give my deepest gratitude to my husband Peng Liu, my parents Gang Peng and Ru Li, my grandparents Sujie Peng and Suliang Zhu, for believing in me, and supporting my decisions at all important stages in my life.

# INTRODUCTION

In recent years, the interplay between the environment and human genome has been extended beyond gene-environment interactions—in which genetic polymorphisms modify the effects of particular environmental exposures. Epigenetic mechanisms (refer to the heritable change in gene expression without changing the underlying DNA sequence), such as DNA methylation, histone modifications, and microRNA (miRNA) expression, have also been shown to change genome function in response to environmental challenges.<sup>1</sup> DNA methylation, through the addition of a methyl group to the 5C position of cytosine in the CpG dinucleotide sequence, serves as an important epigenetic mechanism that often silences gene expression.<sup>2</sup> In the first Chapter of this dissertation work, we evaluated the mediating role of promoter region DNA methylation of four inflammatory genes (*IFN-γ*, *IL-6*, *ICAM-1*, and *TLR-2*) linking PM<sub>2.5</sub> exposure and abnormal glucose metabolism in The Normative Aging Study. We hypothesized that higher PM<sub>2.5</sub> levels associated with increased FBG, and part of this association was mediated through methylation of inflammatory genes.

In Chapter 2 and 3, we explored another set of novel molecular biomarkers—mitochondrial DNA (mtDNA) abundance and mtDNA lesions—both of which are measurements of mtDNA integrity. Mitochondrial DNA (mtDNA) is an emerging cellular target for environmental exposures that generate oxidative damage.<sup>3</sup> Compared to the nuclear genome, the mitochondrial genome is more susceptible to oxidative damage due to lack of introns and protective histone proteins as well as limited capacity to repair.<sup>4</sup> More important, oxidation of the mitochondrial genome can lead to increased mutation rate, and, in turn, damaged mitochondria may become themselves a source of endogenous oxidative species. This makes the mitochondrial genome a central site that both reflects and intensifies oxidative damage.<sup>5</sup> Mitochondrial damage and dysfunction—as reflected in decreases in mtDNA abundance and increases in mtDNA

lesions—may provide a novel set molecular marker to assess oxidative stress that underlies PM<sub>2.5</sub> / air temperature associated pathogenesis.<sup>6,7</sup>

Specifically, in Chapter 2, we evaluated associations between annual moving averages of PM<sub>2.5</sub> mass, PM<sub>2.5</sub> components [i.e., elemental carbon (EC), organic carbon (OC), sulfate (SO<sub>4</sub><sup>2-</sup>), nitrate (NO<sub>3</sub><sup>-</sup>)] and mtDNA abundance. We measured mtDNA abundance—an indicator for oxidative damage—as the ratio between mtDNA copies to the nuclear genome. We applied a regression shrinkage and penalization method (adaptive LASSO) to identify a subset of PM<sub>2.5</sub> components most predictive of the outcome, and conducted multi-pollutant regression modeling to estimate the relative effects of individual component on mtDNA abundance. We hypothesized that some PM<sub>2.5</sub> components may have higher potential for oxidative damage to mtDNA than others, and that PM<sub>2.5</sub> mass concentration alone may not fully represent oxidative damage potential to the mitochondrial genome.

In Chapter 3, we investigated the association between short-term changes in daily mean and daily standard deviation (SD) (variability) of air temperature and blood mtDNA lesions. MtDNA lesions were expressed as the number of mtDNA lesions per 10kb and types of mtDNA lesions include single- and double-stranded breaks, abasic sites, and base damage / modification. We hypothesized that the older adults may be less adaptive to short-term changes in air temperature—especially at the extreme / sub-optimal ranges—which may lead to increased mtDNA lesions in blood.

## References

1. Bollati V, Baccarelli A. Environmental epigenetics. *Heredity (Edinb)* 2010;**105**(1):105-12.
2. Bird A. Perceptions of epigenetics. *Nature* 2007;**447**(7143):396-8.
3. Friedman JR, Nunnari J. Mitochondrial form and function. *Nature* 2014;**505**(7483):335-43.
4. Mengel-From J, Thinggaard M, Dalgard C, Kyvik KO, Christensen K, Christiansen L. Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Hum Genet* 2014;**133**(9):1149-59.
5. Shigenaga MK, Hagen TM, Ames BN. Oxidative damage and mitochondrial decay in aging. *Proc Natl Acad Sci U S A* 1994;**91**(23):10771-8.
6. Hou L, Zhang X, Dioni L, Barretta F, Dou C, Zheng Y, Hoxha M, Bertazzi PA, Schwartz J, Wu S, Wang S, Baccarelli AA. Inhalable particulate matter and mitochondrial DNA copy number in highly exposed individuals in Beijing, China: a repeated-measure study. *Part Fibre Toxicol* 2013;**10**:17.
7. Zhong J, Cayir A, Trevisi L, Sanchez-Guerra M, Lin X, Peng C, Bind MA, Prada D, Laue H, Brennan KJ, Dereix A, Sparrow D, Vokonas P, Schwartz J, Baccarelli AA. Traffic-Related Air Pollution, Blood Pressure, and Adaptive Response of Mitochondrial Abundance. *Circulation* 2016;**133**(4):378-87.

## CHAPTER 1

### **Particulate Air Pollution and Fasting Blood Glucose in non-Diabetic Individuals: Associations and Epigenetic Mediation in the Normative Aging Study, 2000-2011**

Cheng Peng<sup>1</sup>, Marie-Abele C. Bind<sup>2</sup>, Elena Colicino<sup>1</sup>, Itai Kloog<sup>3</sup>, Hyang-Min Byun<sup>4</sup>, Laura Cantone<sup>5</sup>, Letizia Trevisi<sup>1</sup>, Jia Zhong<sup>1</sup>, Kasey Brennan<sup>1</sup>, Alexandra E. Dereix<sup>1</sup>, Pantel S. Vokonas<sup>6</sup>, Brent A. Coull<sup>7</sup>, Joel D. Schwartz<sup>1,8</sup>, Andrea A. Baccarelli<sup>1</sup>

*Environ Health Perspective*; DOI:10.1289/EHP183

<sup>1</sup>Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA.

<sup>2</sup>Department of Statistics, Harvard University, Cambridge, MA, USA.

<sup>3</sup>Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Israel.

<sup>4</sup>Human Nutrition Research Center, Institute of Cellular Medicine, Newcastle University, UK.

<sup>5</sup>Molecular Epidemiology and Environmental Epigenetics Laboratory, Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy.

<sup>6</sup>VA Normative Aging Study, Veterans Affairs Boston Healthcare System and the Department of Medicine, Boston University School of Medicine, Boston, MA, USA.

<sup>7</sup>Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA.

<sup>8</sup>Channing Laboratory, Harvard Medical School, Boston, MA, USA.



Address correspondence to Cheng Peng, Department of Environmental Health, Harvard School  
of Public Health

Building 1, G-7, 665 Huntington Avenue, Boston, MA, 02115, USA. Email:  
chp950@mail.harvard.edu

Running Title: Air pollution, fasting blood glucose, and DNA methylation

## ABSTRACT

**Background:** In non-diabetic individuals, higher fasting blood glucose (FBG) independently predicts diabetes risk, cardiovascular disease and dementia. Ambient PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5 \mu\text{m}$ ) is an emerging determinant of glucose dysregulation.

PM<sub>2.5</sub> effects and mechanisms are understudied among non-diabetic individuals.

**Objectives:** Our goals were to investigate whether PM<sub>2.5</sub> is associated with increase in FBG and to explore potential mediating roles of epigenetic gene regulation.

**Methods:** In 551 non-diabetic participants in the Normative Aging Study, we measured FBG, and DNA methylation of four inflammatory genes (*IFN- $\gamma$* , *IL-6*, *ICAM-1*, and *TLR-2*), up to four times between 2000-2011 (median=2). We estimated short/medium-term (1-, 7-, and 28-day preceding each clinical visit) ambient PM<sub>2.5</sub> at each participant's address using a validated hybrid land-use regression/satellite-based model. We fitted covariate-adjusted regression models accounting for repeated measures.

**Results:** Mean FBG was 99.8 mg/dL (SD=10.7), 18% of the participants had impaired fasting glucose (IFG, i.e., 100-125 mg/dL FBG) at first visit. Interquartile increases in 1-, 7-, and 28-day PM<sub>2.5</sub> were associated with 0.57 mg/dL (95% CI: 0.02; 1.11, p=0.04), 1.02 mg/dL (95% CI: 0.41; 1.63, p=0.001), and 0.89 mg/dL (95% CI: 0.32; 1.47, p=0.003) higher FBG. The same PM<sub>2.5</sub> metrics were associated with 13% (95% CI: -3%; 33%; p=0.12), 27% (95% CI: 6%; 52%, p=0.01) and 32% (95% CI: 10%; 58%, p=0.003) higher odds of IFG, respectively. PM<sub>2.5</sub> was negatively correlated with *ICAM-1* methylation (p=0.01), but not with other genes. Mediation analysis estimated that *ICAM-1* methylation mediated 9% of the association of 28-day PM<sub>2.5</sub> with FBG.

**Conclusions:** In non-diabetics, short/medium-term PM<sub>2.5</sub> was associated with higher FBG. Mediation analysis indicated that part of this association was mediated by *ICAM-1* promoter methylation.

## INTRODUCTION

Clinically diagnosed diabetes is preceded by a long latent period of abnormal glucose metabolism<sup>8</sup>. In the asymptomatic, non-diabetic range of glycemia (<126 mg/dL)<sup>8</sup> increased fasting blood glucose (FBG) levels are already independently associated with the development of diabetes<sup>9</sup>, cardiovascular disease<sup>10-12</sup>, and dementia<sup>13</sup>. FBG variations in this range are often underappreciated and their potential determinants, especially those not directly related to lifestyle, are understudied. Ambient particulate matter (PM) pollution has recently been suggested as an emerging risk factor for metabolic disorders including impaired glucose regulation<sup>14,15</sup>. Cross-sectional and longitudinal studies have revealed that PM is associated with increased risk of diabetes<sup>16-18</sup>, and higher levels of markers of insulin resistance<sup>19</sup>. However, the relationship between PM and glycemia in the non-diabetic range has yet to be studied.

The underlying mechanisms linking PM and abnormal glucose regulation are also not fully understood. Inflammation is central in both PM-associated responses and the pathogenesis of glucose dysregulation. Evidence from previous studies has linked exposures to ambient PM with lower DNA methylation in inflammatory genes<sup>20,21</sup>. DNA methylation, through the addition of a methyl group to the 5C position of cytosine in the CpG dinucleotide sequence, is a well-studied epigenetic modification that usually silences gene expression<sup>22</sup>. Conversely, lower or no methylation has been associated with upregulated gene expression. Lower global methylation content across the human epigenome has been associated with hyperglycemia and the up-regulation of inflammatory genes in peripheral leukocytes from patients with type 2 diabetes mellitus (T2DM)<sup>23,24</sup>. Yet, the role of methylation of specific genes related to inflammation in mediating the effects of PM on FBG has not been investigated.

In a repeated-measure study of older men in the Greater Boston Area, we investigated the association between ambient PM<sub>2.5</sub> (PM with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) concentrations at the participants' address—estimated over different time windows up to one month before the visit—and FBG levels among non-diabetic participants. Using recently developed repeated-measure mediation analysis, we further examined whether and to what extent PM<sub>2.5</sub> increased FBG through changes in blood leukocyte methylation in candidate inflammatory genes. We examined methylation of inflammatory cytokines (interferon gamma (*IFN- $\gamma$* ) and interleukin-6 (*IL-6*)), intercellular adhesion molecule-1 (*ICAM-1*), and Toll-like receptor 2 (*TLR-2*). We hypothesized that higher PM<sub>2.5</sub> levels associated with increased FBG, and part of this association was mediated through methylation of inflammatory genes.

## **MATERIALS AND METHODS**

### **Study Population**

The Normative Aging Study is a prospective longitudinal cohort established in 1963 by the U.S. Veterans Administration in the Greater Boston area<sup>25</sup>. Briefly, participants underwent examinations every 3-5 years. Self-administered questionnaires were collected at each visit providing information on social-demographic characteristics, medical history, medications, and lifestyle. Blood samples were collected at each clinical visit, after an overnight fast and smoking abstinence. Starting from 2000, estimated concentrations of PM<sub>2.5</sub> were obtained from a hybrid spatiotemporal prediction model, as described in the next section. A total of 656 participants had complete information on PM<sub>2.5</sub> measurements from the prediction model, FBG, and blood leukocyte DNA methylation for at least one and up to four visits between 2000 and 2011 (median=2, IQR=1). We excluded 105 participants who, at their first visit, were clinically

diabetic (FBG  $\geq$ 126 mg/dL at the visit) and/or were taking diabetes medications. Therefore, our final study population included 551 participants. Fifty-two participants were diagnosed with diabetes during subsequent visits and, for these individuals, we retained observations from the visit(s) before they became diabetics. One hundred and eighty-six participants came to just one clinical visit, 163 participants came to two clinical visits, and 202 participants came to three or more clinical visits. Participants provided signed informed consent at each visit. The study was approved by the Institutional Review Boards of the participating institutions.

### **Air Pollution and Temperature**

We estimated PM<sub>2.5</sub> concentrations at each participant's residential address using a hybrid land-use regression and satellite-based model<sup>26,27</sup>. In brief, we utilized MODIS (Moderate Resolution Imaging Spectroradiometer) satellite-derived Aerosol Optical Density (AOD) measurements to predict daily PM<sub>2.5</sub> concentration levels at a 10km spatial resolution. Daily AOD was calibrated using ground PM<sub>2.5</sub> measurements from 78 monitoring stations, land use regression and meteorological variables. To estimate PM<sub>2.5</sub> daily concentrations in each grid cell, we calibrated the AOD-PM<sub>2.5</sub> relationship using data from grid cells with both monitor and AOD values, using mixed models with random slopes for day and nested regions. In a later stage, we estimated exposures on days when AOD measures were not available (e.g., due to cloud coverage or snow). Model performance was good with high out-of-sample 10-fold cross-validated R<sup>2</sup> (mean out-of-sample R<sup>2</sup> = 0.83 and 0.81 for days with and without available AOD data, respectively)<sup>28</sup>.

Temperature values were obtained through the national climatic data center (NCDC, 2010). Only continuous operating stations with daily data running were used (26 stations). Grid cells were matched to the closest weather station for meteorological variables.

## DNA Methylation Measurements

Gene-specific DNA methylation was quantified on buffy-coat DNA using bisulfite polymerase-chain-reaction pyrosequencing<sup>29</sup>. In the Normative Aging Study, we generated pyrosequencing-based methylation data for nine genes across pathways related to oxidation, blood clotting, and inflammation<sup>20,30,31</sup>. We chose to focus our analysis on the four inflammatory genes included in these genes, including two inflammatory cytokines (*IFN- $\gamma$*  and *IL-6*), intercellular adhesion molecule 1 (*ICAM-1*), and Toll-like receptor 2 (*TLR-2*). The choice of these four inflammatory genes was aligned with previous literature on PM and inflammation, as well as subclinical inflammation and the risk of T2DM<sup>20,32-36</sup>. Specifically, *IFN- $\gamma$*  and *IL-6* encode for inflammatory cytokines that facilitate cell-to-cell communications in the inflammatory cascade, *ICAM-1* encodes for a glycoprotein that is often expressed on the cell surface of endothelial cells and leukocytes, ICAM-1 glycoproteins are important for cell surface adhesion, transmigration and homing of leukocytes from the circulation to the target tissue. *TLR-2* encodes for a surface receptor protein, which recognize conservative molecular patterns and serve as first line defense in innate immunity. DNA methylation levels were measured for each of these genes at two to five CpG sites within the promoter region except for *IL-6* methylation, which was measured within 500 base-pair downstream of the gene's promoter region where nuclear respiratory factor-1 (NRF-1) binding sites are located (see Supplementary Material, Figure S1.1). The methylation of NRF-1 region is known to suppress *IL-6* gene expression<sup>37</sup>. We calculated and used the mean level of position-specific DNA methylation for each gene because they are highly correlated and are likely to share most of the same functional complexes and traits, and we assumed that mean methylation across the promoter region reflects regional epigenetic regulation<sup>21</sup>.

## **Fasting Blood Glucose Measurement**

Blood glucose levels were measured at each visit, after an overnight fast and were analyzed using the enzymatic hexokinase method. According to American Diabetes Association criteria, FBG less than 100 mg/dL corresponds to normal levels and the 100-125mg/dL range to impaired fasting blood glucose (IFG), which in older individuals often progresses to diabetes over time. FBG larger than 125 mg/dL is defined as clinical diabetes<sup>8,38</sup>.

## **Statistical Analysis of the Main Association of PM<sub>2.5</sub> Levels with FBG**

We evaluated the association between PM<sub>2.5</sub> levels and FBG (modeled as a continuous dependent variable) using linear mixed-effects regression with subject-specific intercepts to account for the correlation among repeated FBG measurements within the same individual. Exposure variables included averages of PM<sub>2.5</sub> concentrations for 1-, 7-, and 28-day preceding each clinical visit; we considered each moving average in a separate regression model. Model estimates are expressed per interquartile range (IQR) increase in PM<sub>2.5</sub> concentration. In the models, we adjusted for the following covariates selected *a priori*, i.e., age (continuous), body mass index (BMI) (weight (kg)/height (m)<sup>2</sup>, continuous), race (white or others), regular patterns of physical activity (<12 hours/week, ≥12 and <30 hours/week, ≥30 hours/week), smoking status (never, former, or current smoker), cumulative pack-years of smoking (continuous), alcohol consumption (<2 or ≥2 drinks/day), education level (high school diploma or less, college degree, or graduate degree), statin use (nonuser, current user), temperature (continuous), and seasonality. Seasonality was modeled using Fourier series terms  $\cos(2\pi \cdot \text{doy}/365.25)$  and  $\sin(2\pi \cdot \text{doy}/365.25)$ , where doy represents day of year. We checked the linearity assumptions of the continuous covariates using cubic splines, and found no deviation from linear dose-response.



The main regression model took the general form:

$$Y_{ij} = \beta_0 + u_i + \beta_1 X_{1ij} + \dots + \beta_p X_{pij} + \beta_{PM2.5} PM_{2.5} + \varepsilon_{ij}$$

where  $i$  corresponds to each participant,  $j$  to the visit;  $\beta_0$  to the intercept for the population mean;  $u_i$  to the subject-specific random intercept.  $\beta_1 X_{1ij}$  to  $\beta_p X_{pij}$  correspond to the covariates we selected *a priori*.  $\beta_{PM2.5} PM_{2.5}$  corresponds to PM<sub>2.5</sub> levels 1-, 7-, or 28-day prior to the clinical visits, depending on the moving average used in each set of models.  $\varepsilon_{ij}$  is the within-subject error term.

In a secondary analysis, we considered a dichotomized FBG variable for impaired fasting glucose (IFG) (categorized using the 0-100mg/dL and 100-125 mg/dL ranges) as the outcome and evaluated the association between PM<sub>2.5</sub> and the odds of IFG using a logistic regression model with generalized estimating equations (GEE) and empirical variance estimates to account for repeated measurements per subject.

The logistic regression model took the general form:

$$\text{logit} [\Pr(Y_{ij}=1)] = \beta_0 + \beta_1 X_{1ij} + \dots + \beta_p X_{pij} + \beta_{PM2.5} PM_{2.5ij}$$

where  $i$  corresponds to each participant,  $j$  to the visit.  $\beta_0$  to the intercept for the population mean.  $\beta_1 X_{1ij}$  to  $\beta_p X_{pij}$  corresponds to the covariates we selected *a priori*.  $\beta_{PM2.5} PM_{2.5ij}$  corresponds to PM<sub>2.5</sub> levels 1-, 7-, and 28-day prior to the clinical visits, respectively.  $Y_{ij} = 0$  indicates subject  $i$  is not defined as IFG at visit  $j$ ;  $Y_{ij} = 1$  indicates subject  $i$  is IFG at visit  $j$ .

To account for potential selection bias due to loss of follow-up, we repeated our analyses used inverse probability weighting. Specifically, in a logistic regression, we predicted the probability of coming to a subsequent visit based on covariates from the previous one, including age, BMI, regular patterns of physical activity, smoking status, pack year smoked, FEV1 and FVC ratio, medication (diuretics and beta blocker), and education level.

## **Statistical Analysis of DNA Methylation and Mediation Analysis**

### ***Selection of mediators***

We hypothesized that associations of PM<sub>2.5</sub> with FBG could be mediated through changes in gene-specific methylation of inflammatory biomarkers. We considered the four inflammation genes (i.e., *IFN-γ*, *IL-6*, *ICAM-1* and *TLR-2*) separately in the mediation analysis. To approximate normality of the residuals, we used *IFN-γ* and *IL-6* on their original scale and log-transformed *ICAM-1* and *TLR-2*.

For DNA methylation of a specific gene to be considered as a potential mediator, we tested the following criteria (a) if there was an association between exposure and mediator; and (b) if there was an association between mediator and outcome<sup>39,40</sup>. We also examined the presence of PM-mediator interactions, and found no evidence of interactions that changed FBG levels.

### ***Underlying assumptions***

To obtain valid estimates of the natural indirect effects, we adjusted for potential exposure-outcome confounders (denoted as C<sub>1</sub>), exposure-mediator confounders (denoted as C<sub>2</sub>) and mediator-outcome confounders (denoted as C<sub>3</sub>), which included age, BMI, race, regular patterns of physical activity, smoking status, cumulative pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, batch of methylation measurement, percentages of lymphocytes, and percentage of neutrophils. We assumed no unmeasured confounding for (a) PM<sub>2.5</sub>-FBG relation, (b) methylation-FBG relation, (c) PM<sub>2.5</sub>-methylation relation, after fitting the linear mixed-effects models with subject-specific intercepts and

controlling for  $C_1$ ,  $C_2$ , and  $C_3$ . In addition, we also assumed that no methylation-FBG confounders would be affected by  $PM_{2.5}$  exposure.

Due to the longitudinal nature of the study, changes in FBG at one visit could potentially affect gene-specific methylation at the subsequent visit ( $Y_{ij} \rightarrow M_{ij+1}$ ) (dotted arrow in Figure 1.1). FBG at one visit therefore may serve as a potential mediator-outcome confounder for the subsequent visit and may introduce bias in our estimates<sup>41</sup>. Therefore, we tested the presence of an association between  $Y_{ij}$  and  $M_{ij+1}$ , to check the assumption of time-varying confounding:

$$M_{ij+1} = \alpha_0 + u_i + \alpha_1 Y_{ij} + \dots + \alpha_p X_{pij} + \alpha_{PM_{2.5}} PM_{2.5ij} + \varepsilon_{ij}$$

where  $i$  corresponds to each participant and  $j$  to the visit;  $\alpha_0$  to the intercept for the population mean;  $u_i$  to the subject-specific random intercept.  $M_{ij+1}$  corresponds to DNA methylation at subsequent visit.  $Y_{ij}$  corresponds to FBG measurement.  $\alpha_1 X_{1ij}$  to  $\alpha_p X_{pij}$  correspond to the covariates we selected a priori.  $\alpha_{PM_{2.5}} PM_{2.5}$  corresponds to  $PM_{2.5}$  levels 1-, 7-, and 28-day prior to the clinical visits, respectively.  $\varepsilon_{ij}$  is the within-subject error term.

### ***Mediation analysis***

We fitted two linear mixed-effects models with random intercepts simultaneously, one modeling the exposure-mediator association, and one modeling the mediator-outcome association<sup>42,43</sup>:

$$M_{ij} = \beta_0 + u_i + \beta_1 X_{1ij} + \dots + \beta_p X_{pij} + \beta_{PM_{2.5}} PM_{2.5ij} + \varepsilon_{ij}$$

$$Y_{ij} = \gamma_0 + g_{0i} + \gamma_1 X_{1ij} + \dots + \gamma_p X_{pij} + \gamma_{PM_{2.5}} PM_{2.5ij} + \gamma_M M_{ij} + \eta_{ij}$$

where  $i$  corresponds to each participant and  $j$  to the visit;  $\beta_0$  and  $\gamma_0$  to the intercept for the population mean;  $u_i$  and  $g_{0i}$  to the subject-specific random intercept.  $\beta_1 X_{1ij}$  to  $\beta_p X_{pij}$  and  $\gamma_p X_{pij}$  correspond to the covariates we selected a priori.  $M_{ij}$  corresponds to DNA methylation and  $Y_{ij}$  corresponds to FBG measurement.  $\beta_{PM_{2.5}} PM_{2.5}$  corresponds to  $PM_{2.5}$  levels 1-, 7-, and 28-

day prior to the clinical visits in the exposure-mediator association, and  $\gamma_M M_{ij}$  corresponds to DNA methylation in the mediator-outcome association.  $\epsilon_{ij}$  and  $\eta_{ij}$  are the within-subject error terms.

$\gamma_{PM2.5}$  corresponds to the natural direct effect, and the natural indirect effect (also called “mediated” effect) is given by the product of  $\beta_{PM2.5} * \gamma_M$ . The Delta method was used to calculate the variance of the natural indirect effect, which correspond to  $\text{Var}(\gamma_M) \beta_{PM2.5}^2 + 2\text{Cov}(\beta_{PM2.5}, \gamma_M) \beta_{PM2.5} \gamma_M + \text{Var}(\beta_{PM2.5}) \gamma_M^2$ . Proportion mediated is calculated as the percentage of natural indirect effect over the sum of natural direct and natural indirect effect (i.e.,  $[\beta_{PM2.5} * \gamma_M / (\beta_{PM2.5} * \gamma_M + \gamma_{PM2.5})]$ ).

In sensitivity analysis, we tested the robustness of the study findings to the no unmeasured-confounding assumptions: (a) we excluded participants who were current smokers to better control for residual confounding by smoking; (b) we additionally controlled for total calorie intake and glycemic index to reduce potential mediator-outcome confounding from diet; (c) we also restricted the analysis to participants with a C-reactive protein (CRP) level less than 10 mg/L, to partially remove potential effects from acute inflammation.

All analyses were conducted with SAS version 9.3 (SAS Institute Inc., Cary, NC), using PROC MIXED to fit the linear mixed effect models and PROC GENMOD to fit the GEE models.

## RESULTS

### Descriptive Statistics

A total of 1,152 FBG measurements were collected from the 551 non-diabetic participants in this study. Table 1.1 describes participant characteristics at their first visit. Mean FBG concentration

at the first visit was  $99.8 \pm 10.7$  mg/dL (mean  $\pm$  SD) and 18% of participants at the first visit had IFG (i.e., blood glucose between 100 mg/dL and 125 mg/dL). Summary statistics of PM<sub>2.5</sub> and temperature during the study period are presented in Table S1.1 (see Supplementary Material, Table S1.1).

### **Main Association of PM<sub>2.5</sub> Levels with FBG**

PM<sub>2.5</sub> levels were associated with increased FBG (Table 1.2). For an IQR increase in PM<sub>2.5</sub> concentration in the previous 1-day (IQR=5.73  $\mu\text{g}/\text{m}^3$ ), 7-day (IQR=4.25  $\mu\text{g}/\text{m}^3$ ), and 28-day (IQR=3.12  $\mu\text{g}/\text{m}^3$ ), FBG increased 0.57 mg/dL (95% CI: 0.02, 1.11, p-value=0.04), 1.02 mg/dL (95% CI: 0.41, 1.63, p-value=0.001), and 0.89 mg/dL (95% CI: 0.32, 1.47, p-value=0.003), respectively. We also found associations of PM<sub>2.5</sub> concentrations with IFG (i.e., FBG>100 mg/dl), particularly for the longer moving averages of PM<sub>2.5</sub> (Table 1.3). IQR increases in PM<sub>2.5</sub> in the previous 1-day (IQR=5.73  $\mu\text{g}/\text{m}^3$ ), 7-day (IQR=4.25  $\mu\text{g}/\text{m}^3$ ) and 28-day (IQR=3.12  $\mu\text{g}/\text{m}^3$ ) exposure windows were associated with OR equal to 1.13 (95% CI: 0.97, 1.33, p-value=0.12), 1.27 (95% CI: 1.06, 1.52, p-value=0.01) and 1.32 (95% CI: 1.10, 1.58, p-value=0.003) for IFG, respectively. We also obtained similar estimates when we used inverse probability weighting to reduce potential selection bias (see Supplementary Material, Table S1.2).

### **DNA Methylation and Mediation analysis**

PM<sub>2.5</sub> showed a negative association with *ICAM-1* methylation, which, in turn, was negatively associated with FBG (Figure 1.2). Methylation of *IFN- $\gamma$* , *IL-6* and *TLR-2* showed no association with FBG. We also found no evidence of PM<sub>2.5</sub>-mediator interactions that changed FBG levels. Since a mediator needs to be associated with both the exposure and the outcome<sup>39,40</sup>, we

conducted the analysis of mediation only for *ICAM-1* methylation. We examined the correlations between *ICAM-1* methylation and the other three genes, and found no substantial correlations, which suggests that the separate analyses are fairly appropriate.

Before conducting the mediation analysis, we examined whether FBG levels at one visit ( $Y_{ij}$ ) affected *ICAM-1* methylation at the subsequent visit ( $M_{ij+1}$ ) (i.e.,  $FBG_{ij} \rightarrow ICAM-1_{ij+1}$ ; dotted arrow in Figure 1.1), because it could potentially confound the mediator-outcome association, therefore may bias our estimates<sup>44</sup>. We found no association between  $Y_{ij}$  ( $FBG_{ij}$ ) and  $M_{ij+1}$  ( $ICAM-1_{ij+1}$ ). Point estimates were negligible (see Supplementary Material, Table S1.3). When we introduced a lag time and examined the effect of *ICAM-1* methylation on FBG levels in the subsequent visit (i.e.  $ICAM-1_{ij} \rightarrow FBG_{ij+1}$ ), we also did not find an association (point estimate = 0.14, 95% CI: -1.93, 1.65).

Table 1.4 presents the natural direct effect, the natural indirect effect, and proportion mediated for *ICAM-1* methylation over the different  $PM_{2.5}$  moving averages. We fitted the exposure-mediator and the mediator-outcome model simultaneously, and found substantial mediation effects of  $PM_{2.5}$  on FBG through a decrease in *ICAM* methylation for the 28-day exposure time window. The proportion mediated was larger (9%) for the 28-day exposure window, but small and negligible for the 1-day and 7-day moving averages.

We conducted further sensitivity analysis to examine if our results were robust to the unmeasured confounding assumptions required in the mediation analysis approach we used. Specifically, we excluded current smokers to limit residual confounding from smoking; we additionally controlled for total calorie intake and glycemic index to reduce potential confounding from diet; and we restricted the analysis to participants with a CRP level less than 10 mg/L to partially remove effect from acute inflammation. Proportion mediated for *ICAM-1*

methylation for the 28-day exposure time window was 9%, 7% and 10%, respectively (see Supplementary Material, Table S1.4).

## DISCUSSION

In the present study, we showed that PM<sub>2.5</sub> concentrations estimated at the participants' address were associated with higher FBG levels among non-diabetic individuals, as well as with higher odds of IFG. We also observed significant associations of lower blood *ICAM-1* methylation, which is expected to up-regulate the expression of ICAM-1 in blood leukocytes, with both higher PM<sub>2.5</sub> levels and higher FBG levels.

Our study is consistent with previous epidemiology studies indicating that ambient PM is associated with metabolic dysregulation<sup>14,16-18,45</sup>. Nevertheless, most previous studies either focused on T2DM or evaluated blood glucose over its entire range, including individuals with diabetes. The non-diabetic and pre-diabetic population represents an ideal target for primary prevention, which, however, has been understudied in air pollution research. Our findings of PM<sub>2.5</sub> associations with FBG in this group may help identify individuals who are particularly susceptible to changes in FBG in the non-diabetic range. It is important to notice that the small changes in FBG resulted in significant odds of IFG, owing to the fact that many people are very close to 100 mg/dL in this aged population, and a very small change in FBG may be enough to pass the threshold.

Our mediation analysis suggests that *ICAM-1* methylation in blood leukocytes served as a mediator of the association between PM<sub>2.5</sub> and FBG, and we observed significant mediated effect at 28-day exposure time window. Our finding of higher concentrations of PM<sub>2.5</sub> associated with lower methylation of the *ICAM-1* gene, which is expected to result in higher ICAM-1 expression,

is consistent with previous literature indicating that elevated concentrations of PM are associated with increase in expression of endothelial markers<sup>20,35,46,47</sup>. The ICAM-1 glycoprotein is responsible for leukocyte adhesion, homing, and transmigration during inflammatory responses<sup>48</sup>. Exposure to PM<sub>2.5</sub> may cause local inflammation in the lungs and promote circulating leukocytes in blood to transmigrate to the target tissue through the up-regulation of adhesion molecules on the endothelial cell surface. Recent observational and intervention studies have linked elevated concentrations of plasma endothelial adhesion molecules with markers of insulin resistance and increased risk of type 2 diabetes<sup>36,49,50</sup>. The ICAM-1 glycoprotein may facilitate migration of leukocytes from the blood to the adipose tissue<sup>51-53</sup>, which could result in local inflammation and subsequently insulin resistance. Alternatively, the ICAM-1 glycoprotein may also facilitate leukocyte transmigration to the pancreas<sup>54</sup>, which could affect beta-cell function and result in impaired insulin secretion. Although the observed association was relatively modest, our estimates were comparable to many other studies evaluating the association between ambient air pollution and DNA methylation with similar exposure levels<sup>55,56</sup>. DNA methylation is measured as a percentage, which indicates the proportion of cells, or more accurately of haploid genomes, which show methylation at the sequence being analyzed. The differences in DNA methylation reported in our study are related to the presence upon PM<sub>2.5</sub> exposure of higher number of circulating blood cells within no methylation at the *ICAM-1* promoter. Further research is needed to determine whether these cells correspond to a specific leukocyte population with known function and their potential roles in relation to PM<sub>2.5</sub> effects.

Conversely, methylation on the cytokine genes investigated (i.e., *IFN-γ* and *IL-6*) was not implicated as a mediator in this study. Many pro- and anti-inflammatory cytokines act in concert



to trigger the inflammatory cascade. Future research may expand the number of inflammatory cytokines investigated and examine their joint effects.

One interesting aspect of the human methylome is that it exhibits both dynamic and static patterns. For instance, methylation of imprinted genes and genes drives tissue lineage commitment and differentiation is established during embryogenesis and persists through life<sup>57</sup>; on the other hand, DNA methylation of inflammatory genes may change rapidly after environmental insults<sup>55,58,59</sup>. Our results are consistent with the dynamic nature of methylation levels in inflammatory pathways, which allows for fine tuning of inflammatory responses. Nevertheless, we are aware that differences in DNA methylation do not necessarily translate into gene expression changes. DNA methylation is only one of the regulatory machineries that control gene expression. Other regulatory mechanisms, such as transcription factor activation, histone modification, chromatin remodeling and RNA silencing may also contribute to regulation of gene expression at various stages.

This study has a number of strengths. We estimated concentrations of PM<sub>2.5</sub> at the residential address for each participant using a state-of-art hybrid model. Estimates from this hybrid model serve as better surrogates relative to the standard use of data from monitoring station for each participant's actual exposure, and limit exposure misclassifications. We conducted analyses both on FBG as a continuous variable and on IFG, a dichotomized variable constructed using a well-established cutoff for preclinical alterations in glucose metabolism. These two sets of analyses produced highly consistent results. We used repeated measures, which were accounted for using linear mixed-effects models with subject specific intercepts for FBG and generalized linear equations (GEE) model with empirical variance for IFG. We conducted mediation analysis as a novel approach for DNA methylation studies. This approach is

particularly useful for investigating effects due to environmental exposures. While Mendelian randomization—a method that relies on genotype data used as instrumental variables—is often proposed to identifying epigenetic mediation, this approach cannot be used for external risk factors, such as PM, which, due to their nature, are not associated with the participants' genetic sequences (such as single nucleotide polymorphisms) <sup>60</sup>. Finally, we measured DNA methylation in candidate genes by pyrosequencing, which yields high precision <sup>61</sup>.

Our study has a few notable limitations. We focused on correlations among DNA methylation and FBG at the same clinical visit. It is therefore difficult to disentangle the temporal relationship between DNA methylation and FBG concentrations. However, when we examined the effect of DNA methylation on FBG for the subsequent visit (i.e., *ICAM-1*<sub>ij</sub> → *FBG*<sub>ij+1</sub>), we did not find any association. We also found no association of FBG on *ICAM-1* methylation at the subsequent visit (i.e., *FBG*<sub>ij</sub> → *ICAM-1*<sub>ij+1</sub>). These analyses confirm that the effects we observed represent short-medium term responses to PM<sub>2.5</sub> and are not persistent over the 3-5 years between the medical visits. To obtain valid estimates for the natural indirect effects, we made the following assumptions: (a) no-unmeasured confounding between PM<sub>2.5</sub> concentration and FBG levels, (b) no-unmeasured confounding between PM<sub>2.5</sub> concentration and methylation, (c) no-unmeasured confounding between methylation and FBG levels, and (d) no methylation-FBG confounders affected by the exposure. We conducted a number of sensitivity analysis to test the robustness to the no-unmeasured confounding assumptions. We excluded participants who were current smokers to better control for residual confounding by smoking; we additionally controlled for total calorie intake and glycemic index, to limit potential confounding from diet; we also restricted the analysis to participants with a CRP level less than 10 mg/L, to partially removal potential effect from acute inflammation. In addition, we assessed the validity

of assumption (d) either based on subject-knowledge or empirically<sup>43</sup>. From previous subject-knowledge, we assumed that PM<sub>2.5</sub> concentration would not affect participant's age, race, smoking status, statin use, as well as batch of methylation measurements and seasonality. We also tested whether PM<sub>2.5</sub> concentration would influence participant's BMI, regular patterns of physical activity, percentage of lymphocytes and percentage of neutrophils, by regressing PM<sub>2.5</sub> concentration on each of these potential confounders. None of the above mediator-outcome confounders in current analysis were affected by the exposure (p-values were 0.10, 0.54, 0.50 and 0.78, respectively). Another limitation of our study is the potential for measurement error in both the exposure and FBG. However, we expect both measurement errors to be non-differential and therefore to attenuate – rather than to cause – the observed significant associations.

## CONCLUSION

In conclusion, we found that PM<sub>2.5</sub> concentrations are associated with higher FBG level, and this association was in part mediated through *ICAM-1* gene methylation, particularly at the longer (28-day) moving average investigated. Our study demonstrates a novel approach of mediation analysis in epigenetic studies and highlights a mediating role of *ICAM-1* gene methylation in air-pollution associated abnormal glucose metabolism. While the proportion mediated by *ICAM-1* methylation alone is relatively modest, methylation of other genes not investigated in this study, independently or in combination with *ICAM-1* methylation, may mediate larger proportions of PM<sub>2.5</sub> effects. Future epigenome-wide studies are needed to determine the extent to which DNA methylation contributes to mediate environmental effects on human metabolism.

## **ACKNOWLEDGEMENTS**

The authors would like to thank all Normative Aging Study participants. We are also indebted to Johanna Lepeule for sharing with us the codes we adapted to perform inverse probability weighting.

## **FUNDINGS**

This work was supported by National Institutes of Health [R01ES021733, R01ES015172, R01ES021357, and P30ES000002]; U.S. Environmental Protection Agency [RD-83479801]; and the Department of Agriculture, Agricultural Research Service [contract 53-K06-510]. The VA Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Center of the U.S. Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and Information Center, Boston, Massachusetts. The views expressed in this paper are those of the authors and do not necessarily represent the views of the US Department of Veterans Affairs.

## **DISCLOSURES**

None of the authors has any actual or potential competing financial interest

**Table 1.1** Characteristics of the Normative Aging Study participants included in the analysis, 2000-2011.

Variable	Visit 1	Visit 2	Visit 3	Visit 4
Age (years), mean $\pm$ SD	73.3 $\pm$ 6.9	75.6 $\pm$ 6.4	77.9 $\pm$ 5.9	78.5 $\pm$ 5.8
BMI (kg/m <sup>2</sup> ), mean $\pm$ SD	27.8 $\pm$ 3.7	27.4 $\pm$ 3.7	27.1 $\pm$ 3.6	27.5 $\pm$ 3.9
Smoking status, N (%)				
Never	165 (30%)	112 (31%)	71 (34%)	14 (33%)
Former	363 (66%)	244 (67%)	135 (64%)	1 (2%)
Current	23 (4%)	7 (2%)	5 (2%)	28 (65%)
Pack years, mean $\pm$ SD	19.9 $\pm$ 24.7	18.7 $\pm$ 23.1	17.1 $\pm$ 21.3	16.0 $\pm$ 11.5
Race, N (%)				
White	543 (97%)	354 (98%)	203 (96%)	42 (98%)
Other	17 (3%)	9 (2%)	8 (4%)	1 (2%)
Metabolic Equivalent of Task, N (%)				
Low ( $\leq$ 12 hours/week)	353 (64%)	222 (61%)	129 (61%)	26 (60%)
Medium (12-30 hours/week)	111 (20%)	89 (25%)	47 (22%)	10 (23%)
High ( $\geq$ 30 hours/week)	87 (16%)	52 (14%)	35 (17%)	7 (16%)
Two or more drinks per day, N (%)	102 (19%)	69 (19%)	35 (17%)	4 (9%)
Education, N (%)				
< 12 years	181 (33%)	119 (32%)	63 (30%)	11 (26%)
13-16 years	253 (46%)	163 (45%)	98 (46%)	24 (56%)
> 16 years	115 (21%)	81 (22%)	50 (24%)	8 (19%)
Statin use, N (%)	196 (36%)	180 (50%)	122 (58%)	24 (56%)
Fasting blood glucose (mg/dL)	99.8 $\pm$ 10.7	98.9 $\pm$ 10.0	99.0 $\pm$ 10.5	99.0 $\pm$ 11.5
Impaired fasting blood glucose, N (%)	100 (18%)	54 (15%)	38 (18%)	7 (16%)
Blood <i>IFN-<math>\gamma</math></i> methylation, mean $\pm$ SD	84.3 $\pm$ 5.9	85.0 $\pm$ 4.8	85.3 $\pm$ 4.7	85.3 $\pm$ 5.5
Blood <i>IL-6</i> methylation, mean $\pm$ SD	43.6 $\pm$ 10.3	43.5 $\pm$ 10.2	43.8 $\pm$ 10.3	42.8 $\pm$ 11.5
Blood <i>ICAM</i> methylation, mean $\pm$ SD	4.3 $\pm$ 1.8	3.9 $\pm$ 1.2	4.4 $\pm$ 1.2	4.8 $\pm$ 1.9
Blood <i>TLR2</i> methylation, mean $\pm$ SD	3.1 $\pm$ 1.3	3.0 $\pm$ 1.4	2.5 $\pm$ 1.4	1.9 $\pm$ 1.0

<sup>a</sup>Cohort participants with diabetes were excluded.

**Table 1.2** Estimated change (and 95% CI) in fasting blood glucose (FBG) level (mg/dL) per interquartile range (IQR) increase in PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) concentration averaged over the corresponding time window before each visit.

<b>PM<sub>2.5</sub> concentration</b>	<b>N of participants</b>	<b>N of observations</b>	<b>PM<sub>2.5</sub> Interquartile Range (IQR)</b>	<b>Estimated change (95% CI) in FBG per IQR increase in PM<sub>2.5</sub> concentrations</b>	<b>p-value</b>
1-day moving average	551	1152	5.71 $\mu\text{g}/\text{m}^3$	0.57 (0.02; 1.11)	0.04
7-day moving average	551	1152	4.28 $\mu\text{g}/\text{m}^3$	1.02 (0.41; 1.63)	0.001
28-day moving average	551	1152	3.09 $\mu\text{g}/\text{m}^3$	0.89 (0.32; 1.47)	0.003

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.

**Table 1.3** Odds ratio (OR) (and 95% CI) of impaired fasting blood glucose (IFG, defined as a fasting blood glucose level greater than 100 mg/dL and less than 126 mg/dL) per interquartile range (IQR) increase in PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) concentration averaged over the corresponding time window before each visit.

<b>PM<sub>2.5</sub> concentration</b>	<b>N of participants</b>	<b>N of observations</b>	<b>PM<sub>2.5</sub> Interquartile Range (IQR)</b>	<b>Odds ratio (95% CI) of IFG per IQR increase in PM<sub>2.5</sub> concentrations</b>	<b>p-value</b>
1-day moving average	551	1152	5.73 $\mu\text{g}/\text{m}^3$	1.13 (0.97; 1.33)	0.12
7-day moving average	551	1152	4.25 $\mu\text{g}/\text{m}^3$	1.27 (1.06; 1.52)	0.01
28-day moving average	551	1152	3.12 $\mu\text{g}/\text{m}^3$	1.32 (1.10; 1.58)	0.003

Results from GEE models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.

**Table 1.4** Mediation effect investigating whether blood *ICAM-1* methylation mediates the association between air pollution and fasting blood glucose level. Indirect effect represents the “mediated” effect through the *ICAM-1* methylation pathway. Estimates correspond to 1  $\mu\text{g}/\text{m}^3$  increase in  $\text{PM}_{2.5}$  concentration.

<b>PM<sub>2.5</sub> concentration</b>	<b>Exposure to mediator association (<math>\beta_{\text{PM}_{2.5}}</math>)</b>	<b>Mediator to outcome association (<math>\gamma_{\text{M}}</math>)</b>	<b>Mediated effect of <i>ICAM-1</i> methylation</b>	<b>Proportion mediated</b>
1-day moving average	0.004 (-0.008; 0.008)	-2.65 (-4.41; -0.89)	-0.01 (-0.02; 0.004)	- <sup>a</sup>
7-day moving average	-0.0004 (-0.007; 0.006)	-2.69 (-4.45; -0.93)	0.001 (-0.02; 0.02)	1%
28-day moving average	-0.01 (-0.02; -0.004)	-2.47 (-4.23; -0.72)	0.03 (0.0001; 0.06)	9%

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, batch effect, % of lymphocytes, and % of neutrophils. Participants with diabetes were excluded.

<sup>a</sup>Proportion mediated cannot be estimated in this case because  $\beta_{\text{PM}_{2.5}}$  and  $\gamma_{\text{M}}$  have opposite signs.

<sup>b</sup>*ICAM-1* methylation is on the logarithm scale.



## LEGEND TO FIGURES

### Figure 1.1 Directed acyclic graph (DAG) for mediation analysis.

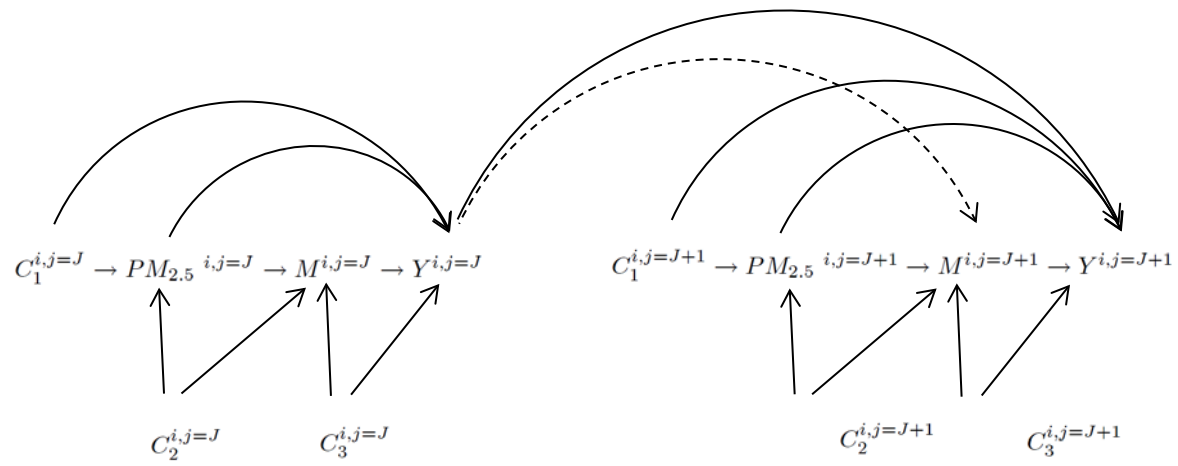
$PM_{2.5}^{i,j=J}$  represents air pollution exposure for  $i^{\text{th}}$  subject prior to  $j=J^{\text{th}}$  visit;  $M^{i,j=J}$  represents gene-specific DNA methylation for  $i^{\text{th}}$  subject at  $j^{\text{th}}$  visit;  $Y^{i,j=J}$  represents fasting blood glucose (FBG) concentrations for  $i^{\text{th}}$  subject at  $j^{\text{th}}$  visit.  $C_1^{i,j}$  represents exposure outcome confounders,  $C_2^{i,j}$  represents exposure mediator confounders;  $C_3^{i,j}$  represents mediator outcome confounders.

Note: to be simplified, correlations between repeated measures of exposures (i.e.  $PM_{2.5ij}$  and  $PM_{2.5ij+1}$ ), repeated measures of mediators (i.e.,  $M_{ij}$  and  $M_{ij+1}$ ) and repeated measures of confounders (i.e.,  $C_{ij}$  and  $C_{ij+1}$ ) are not shown in this DAG.

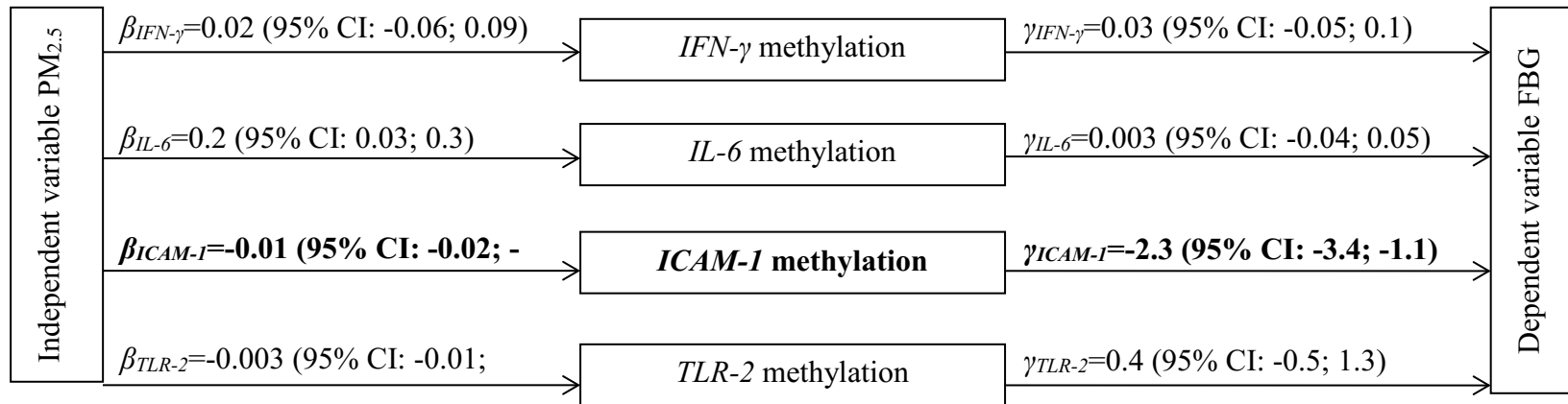
### Figure 1.2 Inflammatory candidate gene methylation mediator model of the relationship between $PM_{2.5}$ concentration and fasting blood glucose level.

*ICAM-1* mean and *TLR-2* mean DNA methylation is log-normally distributed and *IL-6* and *IFN- $\gamma$*  are normally distributed.  $\beta$  is the coefficient of the independent variable ( $PM_{2.5}$  28-day moving average) when regressing the mediator (candidate gene methylation) on the independent variable,  $\gamma$  is the coefficient of the mediator when regressing the dependent variable (FBG) on both the independent variable and the mediator. Results from regression models are adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, batch effects, percentage of lymphocytes, and percentage of neutrophils. Participants with diabetes were excluded.

Figure 1.1



**Figure 2.2**



**Table S1.1** Summary statistics of PM<sub>2.5</sub> (PM with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) and temperature levels. PM<sub>2.5</sub> and temperature were summarized as cumulative averaged exposures up to the previous 28-day exposure window during the study period. PM<sub>2.5</sub> was estimated using spatiotemporal land-use regression models estimating levels at the participants' residential addresses. Temperature values were obtained through the national climatic data center (NCDC); grid cells were matched to the closest weather station for meteorological variables.

	<b>Mean</b>	<b>SD</b>	<b>Percentiles</b>				
			<b>10th</b>	<b>25th</b>	<b>50th</b>	<b>75th</b>	<b>90th</b>
<b>PM<sub>2.5</sub> (<math>\mu\text{g}/\text{m}^3</math>)</b>							
1-day moving average	10.92	5.42	5.86	7.37	9.47	13.08	17.54
7-day moving average	10.59	3.48	6.75	8.09	10.06	12.37	14.79
28-day moving average	10.71	2.62	7.57	8.94	10.45	12.03	13.89
<b>Temperature (<math>^{\circ}\text{C}</math>)</b>							
1-day moving average	11.93	7.59	1.92	6.24	12.11	18.00	21.80
7-day moving average	11.80	7.30	1.41	6.26	12.42	18.16	20.87
28-day moving average	11.85	7.14	1.54	5.86	12.82	18.38	20.52

**Table S1.2** Estimated change (and 95% CI) in fasting blood glucose (FBG) level (mg/dL) per interquartile range (IQR) increase in PM<sub>2.5</sub> (particulate matter with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ ) concentration averaged over the corresponding time window before each visit. We accounted for potential selection bias using inverse probability weighting.

<b>PM<sub>2.5</sub> concentration</b>	<b>N of participants</b>	<b>N of observations</b>	<b>PM<sub>2.5</sub> Interquartile Range (IQR)</b>	<b>Estimated change (95% CI) in FBG per IQR increase in PM<sub>2.5</sub> concentrations</b>
1-day moving average	551	1133	5.71 $\mu\text{g}/\text{m}^3$	0.51 (-0.04; 1.05)
7-day moving average	551	1133	4.28 $\mu\text{g}/\text{m}^3$	0.91 (0.29; 1.53)
28-day moving average	551	1133	3.09 $\mu\text{g}/\text{m}^3$	0.80 (0.21; 1.39)

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.

Potential selection bias due to loss to follow-up was accounted for using inverse probability weighting. In a logistic regression, we predict the probability of coming to a subsequent visit by covariates from the previous one, which include age, BMI, regular patterns of physical activity, smoking status, pack year smoked, FEV1 and FVC ratio, medication (diuretics and beta blocker), and education level.

**Table 1.3** Estimated change (and 95% CI) in inflammatory candidate gene methylation with increase in fasting blood glucose [FBG (mg/dL)] concentrations at previous visits ( $Y_{ij} \rightarrow M_{ij+1}$ ).

Candidate gene methylation	N of subjects	Estimated change (95% CI) in $M^{i,j=J+1}$ per 1 mg/dL increase in $Y^{i,j=J}$
<i>IFN-<math>\gamma</math></i>	472	-0.02 (-0.05; 0.02)
<i>IFN-<math>\gamma</math></i> adjusted for $M^{i,j=J}$	463	-0.02 (-0.06; 0.01)
<i>IL-6</i>	472	-0.04 (-0.11; 0.04)
<i>IL-6</i> adjusted for $M^{i,j=J}$	466	-0.03 (-0.09; 0.02)
<i>ICAM-1</i> *	472	0.001 (-0.002; 0.004)
<i>ICAM-1</i> * adjusted for $M^{i,j=J}$	424	0.001 (-0.002; 0.004)
<i>TLR-2</i> *	472	0.003 (-0.002; 0.007)
<i>TLR-2</i> * adjusted for $M^{i,j=J}$	415	0.003 (-0.002; 0.008)

Results from linear mixed-effects regression models accounting for correlation across multiple visits, and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, batch effects, percentage of lymphocytes, and percentage of neutrophils (all from previous visits). Participants with diabetes were excluded. \**ICAM-1* and *TLR-2* DNA methylation measurements were log transformed.

**Table S1.4** Sensitivity analysis of the mediation effect of *ICAM-1* methylation on the association between PM<sub>2.5</sub> concentrations averaged over 28-day exposure window with fasting blood glucose (FBG) level (mg/dL). Natural indirect effect represents the “mediated” effect through the *ICAM-1* methylation pathway. Estimates correspond to 1 µg/m<sup>3</sup> increase in PM<sub>2.5</sub> concentration.

<b>PM<sub>2.5</sub> concentration averaged over 28-day moving average</b>	<b>Exposure to mediator association (<math>\beta_{PM_{2.5}}</math>) (95% CI)</b>	<b>Mediator to outcome association (<math>\gamma_M</math>) (95% CI)</b>	<b>Natural indirect effect of <i>ICAM-1</i> methylation (95% CI)</b>	<b>Proportion mediated</b>
Model 1 <sup>a</sup>	-0.01 (-0.02; -0.006)	-2.30 (-3.43; -1.17)	0.03 (-0.002; 0.06)	9%
Model 2 <sup>b</sup>	-0.01 (-0.02; -0.004)	-2.32 (-4.59; -0.05)	0.03 (-0.07; 0.07)	7%
Model 3 <sup>c</sup>	-0.01 (-0.02; -0.005)	-2.99 (-4.13; -1.85)	0.03 (0.0006; 0.07)	10%

Results from linear mixed-effects regression models adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, batch effect, percentage of lymphocytes, percentage of neutrophils. Participants with diabetes were excluded.

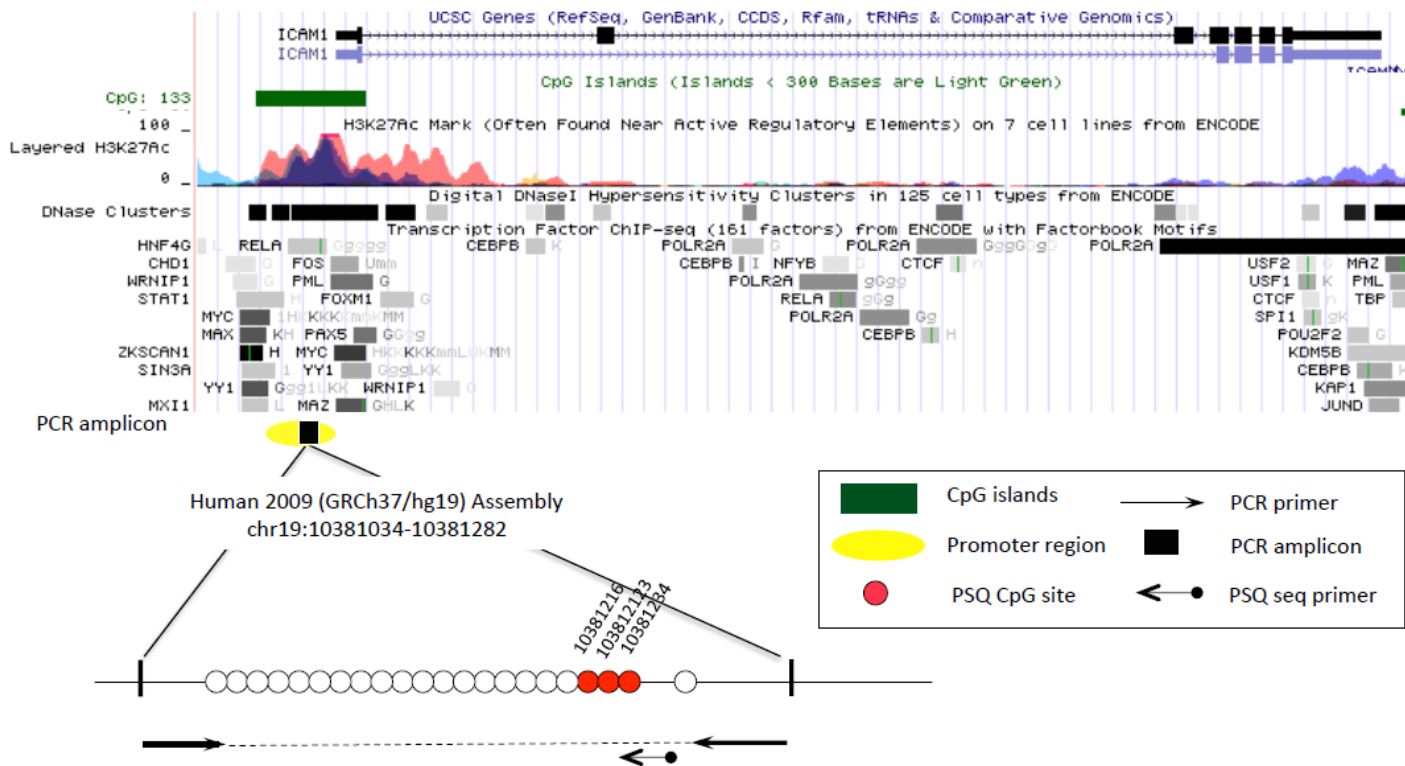
<sup>a</sup> We excluded current smokers to limit potential residual confounding by smoking.

<sup>b</sup> We additionally controlled for dietary intake (total calorie intake and glycemic index), to reduce potential confounding by diet.

<sup>c</sup> We restricted the analysis to participants with a C-reactive protein (CRP) level less than 10 mg/L, to partially removal potential effect from acute inflammation.

**Figure S1.1** Methylation of specific CpG sites for the four candidate genes (interferon gamma (*IFN- $\gamma$* ), interleukin-6 (*IL-6*), Toll-like receptor 2 (*TLR-2*), and intracellular adhesion molecule-1 (*ICAM-1*)).

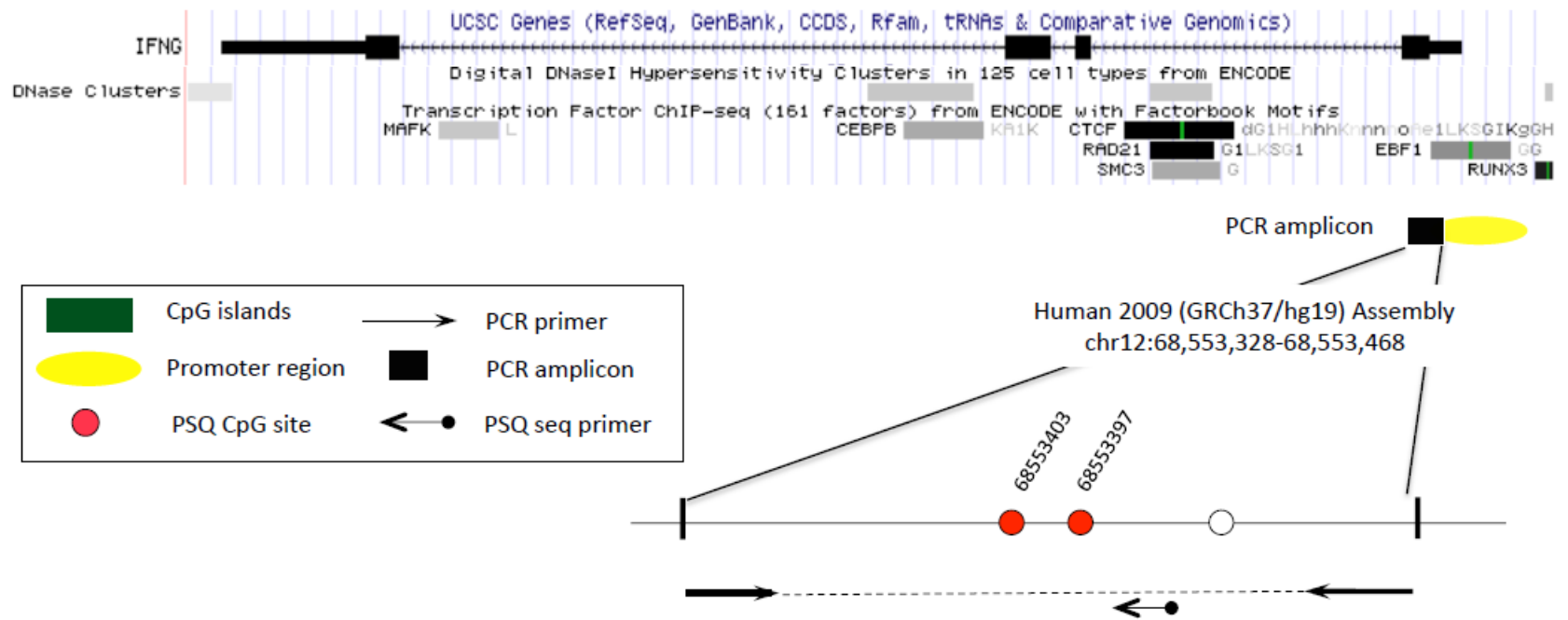
i. *ICAM-1* methylation CpG sites



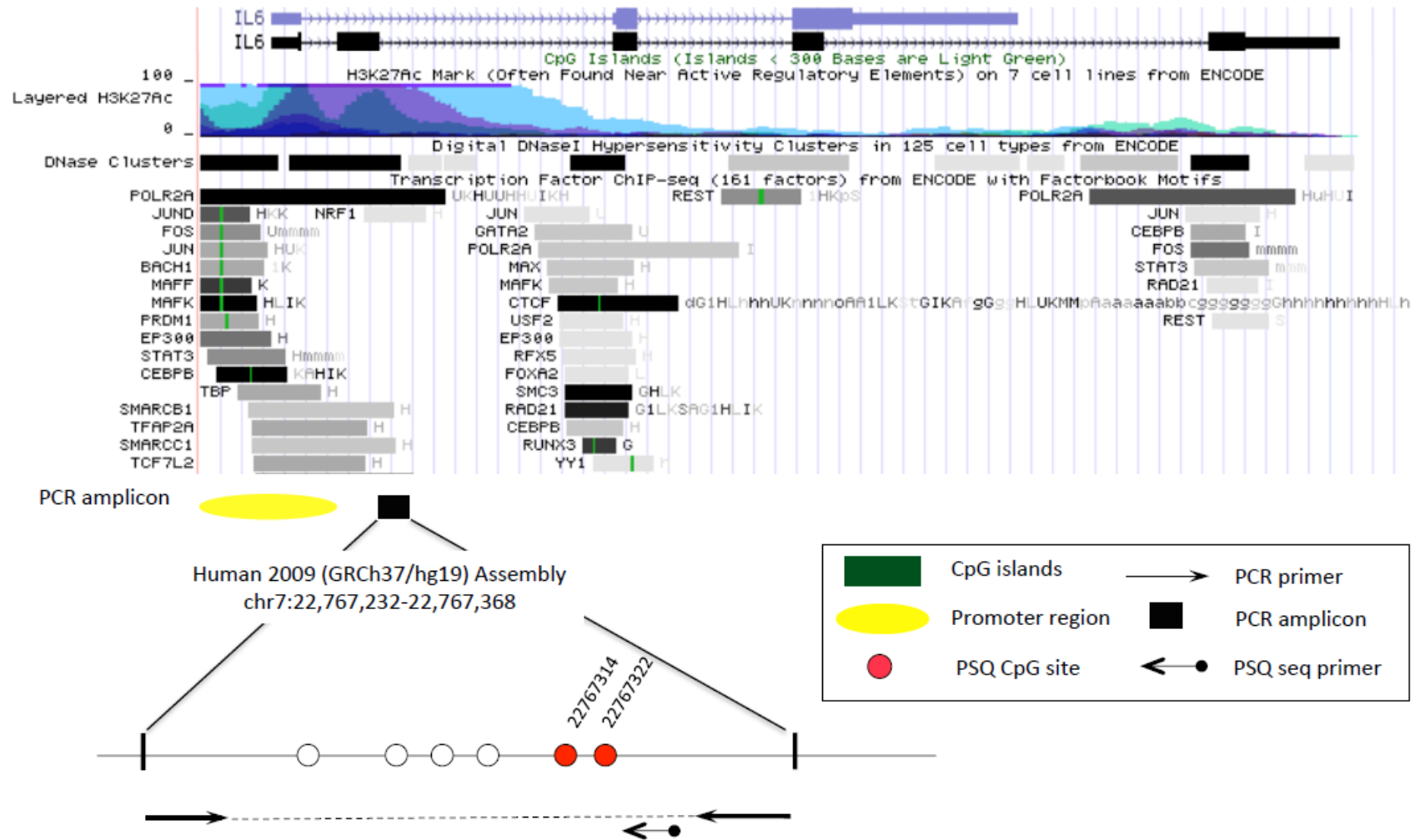
Note: PCR: polymerase chain reaction; PSQ: pyrosequencing; PSQ CpG sites are the methylation sites where pyrosequencing was performed.



ii. *IFN-γ* methylation CpG sites



iii. *IL-6* methylation CpG sites



iv. *TLR-2* methylation CpG sites

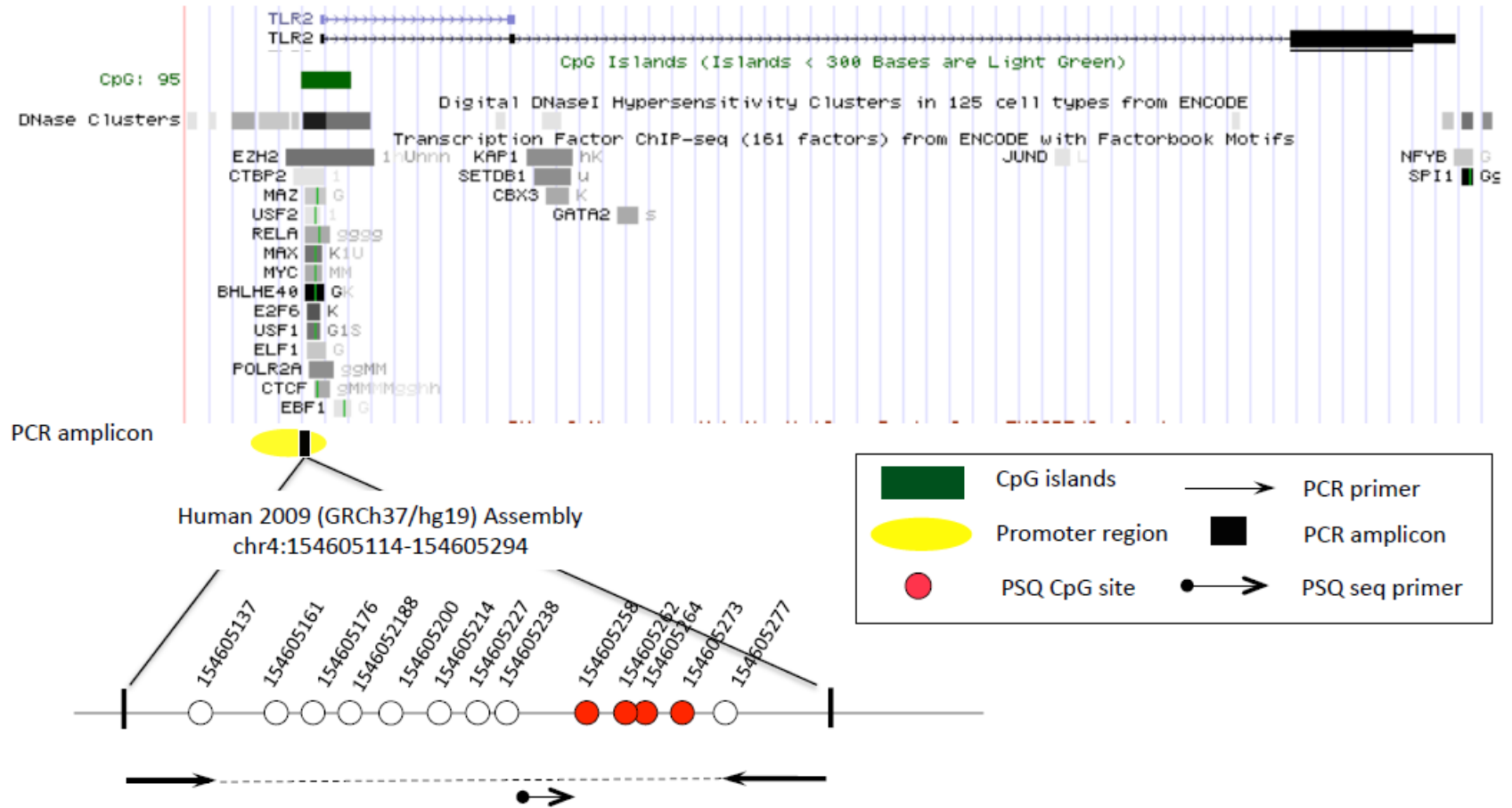
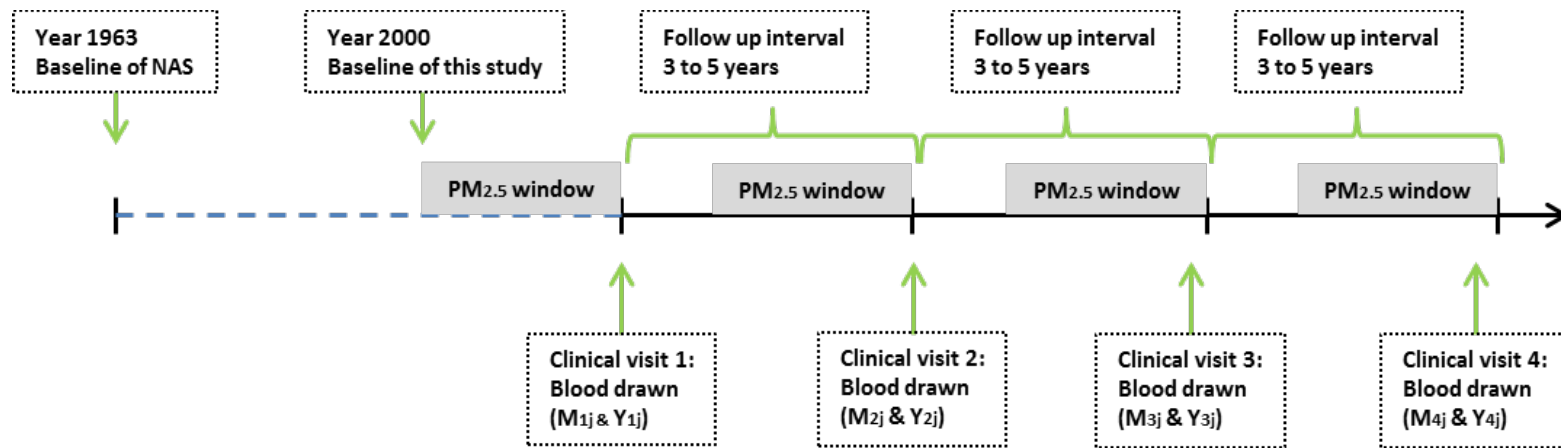


Figure S1.2 Timeline of the study design.



## REFERENCES

1. American Diabetes A. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2012;**35 Suppl 1**:S64-71.
2. Tirosh A, Shai I, Tekes-Manova D, Israeli E, Pereg D, Shochat T, Kochba I, Rudich A, Israeli Diabetes Research G. Normal fasting plasma glucose levels and type 2 diabetes in young men. *N Engl J Med* 2005;**353**(14):1454-62.
3. Duckworth WC. Hyperglycemia and cardiovascular disease. *Curr Atheroscler Rep* 2001;**3**(5):383-91.
4. Levitan EB, Song Y, Ford ES, Liu S. Is nondiabetic hyperglycemia a risk factor for cardiovascular disease? A meta-analysis of prospective studies. *Arch Intern Med* 2004;**164**(19):2147-55.
5. Sarwar N, Aspelund T, Eiriksdottir G, Gobin R, Seshasai SR, Forouhi NG, Sigurdsson G, Danesh J, Gudnason V. Markers of dysglycaemia and risk of coronary heart disease in people without diabetes: Reykjavik prospective study and systematic review. *PLoS Med* 2010;**7**(5):e1000278.
6. Crane PK, Walker R, Larson EB. Glucose levels and risk of dementia. *N Engl J Med* 2013;**369**(19):1863-4.
7. Rajagopalan S, Brook RD. Air pollution and type 2 diabetes: mechanistic insights. *Diabetes* 2012;**61**(12):3037-45.
8. Esposito K, Petrizzo M, Maiorino MI, Bellastella G, Giugliano D. Particulate matter pollutants and risk of type 2 diabetes: a time for concern? *Endocrine* 2016;**51**(1):32-7.
9. Brook RD, Jerrett M, Brook JR, Bard RL, Finkelstein MM. The relationship between diabetes mellitus and traffic-related air pollution. *J Occup Environ Med* 2008;**50**(1):32-8.

10. Andersen ZJ, Raaschou-Nielsen O, Ketznel M, Jensen SS, Hvidberg M, Loft S, Tjønneland A, Overvad K, Sørensen M. Diabetes incidence and long-term exposure to air pollution: a cohort study. *Diabetes Care* 2012;**35**(1):92-8.
11. Puett RC, Hart JE, Schwartz J, Hu FB, Liese AD, Laden F. Are particulate matter exposures associated with risk of type 2 diabetes? *Environ Health Perspect* 2011;**119**(3):384-9.
12. Kelishadi R, Mirghaffari N, Poursafa P, Gidding SS. Lifestyle and environmental factors associated with inflammation, oxidative stress and insulin resistance in children. *Atherosclerosis* 2009;**203**(1):311-9.
13. Bind MA, Baccarelli A, Zanobetti A, Tarantini L, Suh H, Vokonas P, Schwartz J. Air pollution and markers of coagulation, inflammation, and endothelial function: associations and epigene-environment interactions in an elderly cohort. *Epidemiology* 2012;**23**(2):332-40.
14. Bind MA, Lepeule J, Zanobetti A, Gasparrini A, Baccarelli A, Coull BA, Tarantini L, Vokonas PS, Koutrakis P, Schwartz J. Air pollution and gene-specific methylation in the Normative Aging Study: association, effect modification, and mediation analysis. *Epigenetics* 2014;**9**(3):448-58.
15. Newell-Price J, Clark AJ, King P. DNA methylation and silencing of gene expression. *Trends Endocrinol Metab* 2000;**11**(4):142-8.
16. Nilsson E, Jansson PA, Perfilyev A, Volkov P, Pedersen M, Svensson MK, Poulsen P, Ribel-Madsen R, Pedersen NL, Almgren P, Fadista J, Ronn T, Klarlund Pedersen B, Scheele C, Vaag A, Ling C. Altered DNA methylation and differential expression of

- genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes* 2014;**63**(9):2962-76.
17. Luttmer R, Spijkerman AM, Kok RM, Jakobs C, Blom HJ, Serne EH, Dekker JM, Smulders YM. Metabolic syndrome components are associated with DNA hypomethylation. *Obes Res Clin Pract* 2013;**7**(2):e106-e115.
  18. Power MC, Weisskopf MG, Alexeeff SE, Coull BA, Spiro A, 3rd, Schwartz J. Traffic-related air pollution and cognitive function in a cohort of older men. *Environ Health Perspect* 2011;**119**(5):682-7.
  19. Kloog I, Nordio F, Coull BA, Schwartz J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM<sub>2.5</sub> exposures in the Mid-Atlantic states. *Environ Sci Technol* 2012;**46**(21):11913-21.
  20. Madrigano J, Kloog I, Goldberg R, Coull BA, Mittleman MA, Schwartz J. Long-term exposure to PM<sub>2.5</sub> and incidence of acute myocardial infarction. *Environ Health Perspect* 2013;**121**(2):192-6.
  21. Kloog I KP, Coull BA, Lee HJ, Schwartz J. . Assessing temporally and spatially resolved PM<sub>2.5</sub> exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos Environ* 2011;**45**:6267-6275.
  22. Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res* 2004;**32**(3):e38.
  23. Lepeule J, Baccarelli A, Motta V, Cantone L, Litonjua AA, Sparrow D, Vokonas PS, Schwartz J. Gene promoter methylation is associated with lung function in the elderly: the Normative Aging Study. *Epigenetics* 2012;**7**(3):261-9.

24. Lepeule J, Bind MA, Baccarelli AA, Koutrakis P, Tarantini L, Litonjua A, Sparrow D, Vokonas P, Schwartz JD. Epigenetic influences on associations between air pollutants and lung function in elderly men: the normative aging study. *Environ Health Perspect* 2014;**122**(6):566-72.
25. Herder C, Carstensen M, Ouwens DM. Anti-inflammatory cytokines and risk of type 2 diabetes. *Diabetes Obes Metab* 2013;**15 Suppl 3**:39-50.
26. Vossoughi M, Schikowski T, Vierkotter A, Sugiri D, Hoffmann B, Teichert T, Herder C, Schulte T, Luckhaus C, Raulf-Heimsoth M, Casjens S, Bruning T, Kramer U. Air pollution and subclinical airway inflammation in the SALIA cohort study. *Immun Ageing* 2014;**11**(1):5.
27. van Eeden SF, Tan WC, Suwa T, Mukae H, Terashima T, Fujii T, Qui D, Vincent R, Hogg JC. Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (PM(10)). *Am J Respir Crit Care Med* 2001;**164**(5):826-30.
28. Madrigano J, Baccarelli A, Wright RO, Suh H, Sparrow D, Vokonas PS, Schwartz J. Air pollution, obesity, genes and cellular adhesion molecules. *Occup Environ Med* 2010;**67**(5):312-7.
29. Bluher M, Unger R, Rassoul F, Richter V, Paschke R. Relation between glycaemic control, hyperinsulinaemia and plasma concentrations of soluble adhesion molecules in patients with impaired glucose tolerance or Type II diabetes. *Diabetologia* 2002;**45**(2):210-6.



30. Choi YS, Kim S, Kyu Lee H, Lee KU, Pak YK. In vitro methylation of nuclear respiratory factor-1 binding site suppresses the promoter activity of mitochondrial transcription factor A. *Biochem Biophys Res Commun* 2004;**314**(1):118-22.
31. American Diabetes A. Standards of medical care in diabetes--2008. *Diabetes Care* 2008;**31 Suppl 1**:S12-54.
32. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**(6):1173-82.
33. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods* 2013;**18**(2):137-50.
34. Diggle P, Heagerty, P., Liang, K.Y., Zeger, S. *Analysis of Longitudinal Data*. Oxford: Oxford Statistical Science, 2013.
35. Bauer DJ, Preacher KJ, Gil KM. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. *Psychol Methods* 2006;**11**(2):142-63.
36. M-A. C. Bind TJV, B. A. Coull, J. D. Schwartz. Causal mediation analysis for longitudinal data with exogenous exposure. *Biostatistics in press*.
37. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;**15**(5):615-25.
38. Eze IC, Hemkens LG, Bucher HC, Hoffmann B, Schindler C, Kunzli N, Schikowski T, Probst-Hensch NM. Association between Ambient Air Pollution and Diabetes Mellitus in

- Europe and North America: Systematic Review and Meta-Analysis. *Environ Health Perspect* 2015.
39. Ruckerl R, Ibalid-Mulli A, Koenig W, Schneider A, Woelke G, Cyrys J, Heinrich J, Marder V, Frampton M, Wichmann HE, Peters A. Air pollution and markers of inflammation and coagulation in patients with coronary heart disease. *Am J Respir Crit Care Med* 2006;**173**(4):432-41.
  40. O'Neill MS, Veves A, Sarnat JA, Zanobetti A, Gold DR, Economides PA, Horton ES, Schwartz J. Air pollution and inflammation in type 2 diabetes: a mechanism for susceptibility. *Occup Environ Med* 2007;**64**(6):373-9.
  41. Rahman A, Fazal F. Hug tightly and say goodbye: role of endothelial ICAM-1 in leukocyte transmigration. *Antioxid Redox Signal* 2009;**11**(4):823-39.
  42. Hak AE, Pols HA, Stehouwer CD, Meijer J, Kiliaan AJ, Hofman A, Breteler MM, Witteman JC. Markers of inflammation and cellular adhesion molecules in relation to insulin resistance in nondiabetic elderly: the Rotterdam study. *J Clin Endocrinol Metab* 2001;**86**(9):4398-405.
  43. Meigs JB, Hu FB, Rifai N, Manson JE. Biomarkers of endothelial dysfunction and risk of type 2 diabetes mellitus. *JAMA* 2004;**291**(16):1978-86.
  44. Sun Q, Yue P, Deiuliis JA, Lumeng CN, Kampfrath T, Mikolaj MB, Cai Y, Ostrowski MC, Lu B, Parthasarathy S, Brook RD, Moffatt-Bruce SD, Chen LC, Rajagopalan S. Ambient air pollution exaggerates adipose inflammation and insulin resistance in a mouse model of diet-induced obesity. *Circulation* 2009;**119**(4):538-46.
  45. Rao X, Patel P, Puett R, Rajagopalan S. Air pollution as a risk factor for type 2 diabetes. *Toxicol Sci* 2015;**143**(2):231-41.

46. Mendez R, Zheng Z, Fan Z, Rajagopalan S, Sun Q, Zhang K. Exposure to fine airborne particulate matter induces macrophage infiltration, unfolded protein response, and lipid deposition in white adipose tissue. *Am J Transl Res* 2013;**5**(2):224-34.
47. Yagi N, Yokono K, Amano K, Nagata M, Tsukamoto K, Hasegawa Y, Yoneda R, Okamoto N, Moriyama H, Miki M, et al. Expression of intercellular adhesion molecule 1 on pancreatic beta-cells accelerates beta-cell destruction by cytotoxic T-cells in murine autoimmune diabetes. *Diabetes* 1995;**44**(7):744-52.
48. Guo L, Byun HM, Zhong J, Motta V, Barupal J, Zheng Y, Dou C, Zhang F, McCracken JP, Diaz A, Marco SG, Colicino S, Schwartz J, Wang S, Hou L, Baccarelli AA. Effects of short-term exposure to inhalable particulate matter on DNA methylation of tandem repeats. *Environ Mol Mutagen* 2014;**55**(4):322-35.
49. Madrigano J, Baccarelli A, Mittleman MA, Wright RO, Sparrow D, Vokonas PS, Tarantini L, Schwartz J. Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. *Environ Health Perspect* 2011;**119**(7):977-82.
50. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. *Nature* 1993;**366**(6453):362-5.
51. Tarantini L, Bonzini M, Apostoli P, Pegoraro V, Bollati V, Marinelli B, Cantone L, Rizzo G, Hou L, Schwartz J, Bertazzi PA, Baccarelli A. Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. *Environ Health Perspect* 2009;**117**(2):217-22.
52. Hou L, Zhang X, Zheng Y, Wang S, Dou C, Guo L, Byun HM, Motta V, McCracken J, Diaz A, Kang CM, Koutrakis P, Bertazzi PA, Li J, Schwartz J, Baccarelli AA. Altered

- methylation in tandem repeat element and elemental component levels in inhalable air particles. *Environ Mol Mutagen* 2014;**55**(3):256-65.
53. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? *Int J Epidemiol* 2012;**41**(1):5-9.
54. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc* 2007;**2**(9):2265-75.

## CHAPTER 2

### **Associations of Annual Ambient PM<sub>2.5</sub> Mass and PM<sub>2.5</sub> Components with Mitochondrial Damage in Older Adults: A Multi-pollutant Approach in The Veterans Affairs Normative Aging Study, 2000–2012**

Cheng Peng <sup>a</sup>, Akin Cayir <sup>b</sup>, Marco Sanchez-Guerra <sup>a</sup>, Qian Di <sup>a</sup>, Ander Wilson <sup>c</sup>, Jia Zhong <sup>a</sup>, Anna Kosheleva <sup>a</sup>, Letizia Trevisi <sup>a</sup>, Elena Colicino <sup>a</sup>, Kasey Brennan <sup>a</sup>, Alexandra E. Dereix <sup>a</sup>, Lingzhen Dai <sup>a</sup>, Jia Chen <sup>d</sup>, Brent A. Coull <sup>b</sup>, Pantel Vokonas <sup>e</sup>, Joel Schwartz <sup>a,f</sup>, Andrea A. Baccarelli <sup>a</sup>

<sup>a</sup> Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA

<sup>b</sup> Vocational Health College, Canakkale Onsekiz Mart University, Çanakkale, Turkey

<sup>c</sup> Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA

<sup>d</sup> Department of Preventive Medicine, Mount Sinai Hospital, New York, NY, USA

<sup>e</sup> Normative Aging Study, Veterans Affairs Boston Healthcare System and the Department of Medicine, Boston University School of Medicine, Boston, MA, USA

<sup>f</sup> Channing Laboratory, Harvard Medical School, Boston, MA, USA

Correspondence: Cheng Peng, Department of Environmental Health, Harvard School of Public Health, Building 1, G-7, 665 Huntington Avenue, Boston, MA 02115, USA. Email: [chp950@mail.harvard.edu](mailto:chp950@mail.harvard.edu)

Running title: Multi-pollutant; Mitochondrial DNA abundance; LASSO

## ABSTRACT

**Background:** Fine particulate matter (PM<sub>2.5</sub>) represents a mixture of components with different toxicities. However, little is known about the relative effects of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components on oxidative stress-related markers, such as mitochondrial DNA (mtDNA) abundance, that may lie on the pathway of PM<sub>2.5</sub>-associated disease.

**Methods:** We studied 646 elderly male participants in the Normative Aging Study (2000-2012) to investigate associations of long-term exposure to PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components with mtDNA abundance. Annual concentrations of pollutants were estimated at each participant's residential address using spatial- and temporal-resolved chemical transport models. We measured blood mtDNA abundance using real-time PCR. We applied a shrinkage and selection method (adaptive LASSO) to identify pollutants most predictive of mtDNA abundance, and fit multi-pollutant linear mixed-effects models with subject-specific intercept to estimate the relative effects of individual PM component.

**Results:** MtDNA abundance was negatively associated with PM<sub>2.5</sub> mass concentration in the previous 1-year moving average exposure window and—after adjusting for PM<sub>2.5</sub> mass—several PM<sub>2.5</sub> components, including organic carbon, sulfate (marginally) and nitrate. In multi-pollutant models including as independent variables PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components selected by LASSO, nitrate components of PM<sub>2.5</sub> were associated with mtDNA abundance. A standard deviation (SD) increase in annual PM<sub>2.5</sub>-associated nitrate concentration was associated with a 0.12 SD (95% CI: -0.18; -0.07) decrease in mtDNA abundance. Analyses restricted to PM<sub>2.5</sub> annual moving average below the current 1-year EPA standard ( $\leq 12 \mu\text{g}/\text{m}^3$ ) produced similar results.

**Conclusions:** Long-term exposures to PM<sub>2.5</sub>-associated nitrate were related with decreased mtDNA abundance independent of PM<sub>2.5</sub> mass. Mass alone may not fully capture the oxidation potency of PM<sub>2.5</sub> to the mitochondrial genome.



## INTRODUCTION

Epidemiological studies have demonstrated consistent associations between fine particulate matter (PM with aerodynamic diameter  $\leq 2.5$   $\mu\text{m}$ , or  $\text{PM}_{2.5}$ ) exposures and increased cardiopulmonary risks.<sup>62-64</sup> However,  $\text{PM}_{2.5}$  is a complex mixture, so mass concentration may not capture how individual PM components vary in physicochemical and toxicological properties.<sup>65</sup> Further, because  $\text{PM}_{2.5}$  chemical composition often reflects source of emission, identifying sources that contribute to greater risk could help inform more effective emission control strategies.

To foster understanding of differential toxicities of  $\text{PM}_{2.5}$  components, recent epidemiological studies have studied the adverse health effects of  $\text{PM}_{2.5}$  by its constituent components. A growing body of evidence suggests that PM from vehicle exhausts, diesel, residual oil burning, and road dust are associated with higher risks of respiratory and cardiovascular hospital admissions and mortality.<sup>66-68</sup> Several studies also demonstrated that specific  $\text{PM}_{2.5}$  components, such as elemental carbon (EC), silicon (Si), sulfate ( $\text{SO}_4^{2-}$ ) and nitrate ( $\text{NO}_3^-$ ), modify the associations between  $\text{PM}_{2.5}$  and cardiovascular mortality.<sup>66,69</sup>  $\text{PM}_{2.5}$  containing higher proportions of metal components has also been shown to have stronger association with mortality.<sup>70</sup> Oxidative stress is one of the primary pathways proposed to account for the association of  $\text{PM}_{2.5}$  with health related outcomes. There is little information, however, about the relative effects of  $\text{PM}_{2.5}$  components on early oxidation markers, especially for long-term exposures.

Mitochondrial DNA (mtDNA) is an emerging cellular target for environmental exposures that generate oxidative damage.<sup>3</sup> Compared to the nuclear genome, the mitochondrial genome is more susceptible to oxidative damage due to lack of introns and protective histone proteins as well as limited capacity to repair.<sup>4</sup> More important, oxidation of the mitochondrial genome can lead to increased mutation rate, and, in turn, damaged mitochondria may become themselves a source of endogenous oxidative species. This makes the mitochondrial genome a central site that both reflects and intensifies oxidative damage.<sup>5</sup> Further, mitochondrial damage and dysfunction is associated with a wide range of cardiac outcomes, including atherosclerosis,<sup>71</sup> hypercholesterolemia,<sup>72</sup> ischemic heart disease,<sup>73</sup> and hypertension.<sup>74</sup> Mitochondrial damage and dysfunction—as reflected in decreases in mtDNA abundance relative to the nuclear genome—may provide a novel molecular marker to assess oxidative stress that underlies PM associated pathogenesis.<sup>6,7</sup>

In this study, we examined associations between annual moving averages of PM<sub>2.5</sub> mass, PM<sub>2.5</sub> components [i.e., elemental carbon (EC), organic carbon (OC), sulfate (SO<sub>4</sub><sup>2-</sup>), nitrate (NO<sub>3</sub><sup>-</sup>)] and mtDNA abundance in an elderly population. We estimated concentrations of PM<sub>2.5</sub> mass and components at the residential address for each participant using spatial- and temporal-resolved models, which incorporated both land-use regression and chemical transport modeling. We measured mtDNA abundance—an indicator for oxidative damage—as the ratio between mtDNA copies to the nuclear genome. We applied a regression shrinkage and penalization method to identify a subset of PM<sub>2.5</sub> components most predictive of the outcome, and conducted multi-pollutant regression modeling to estimate the relative effects of individual component on mtDNA abundance. We hypothesized that some PM<sub>2.5</sub> components may have higher potential for

oxidative damage to mtDNA than others, and that PM<sub>2.5</sub> mass concentration alone may not fully represent oxidative damage potential to the mitochondrial genome.

## **METHODS**

### **Study participants**

We based this analysis on the Normative Aging Study, a prospective longitudinal cohort established in 1963 by the U.S. Veterans Administration in the greater Boston, Massachusetts area. A detailed study description was published recently.<sup>25</sup> In brief, study participants underwent physical examinations every 3–5 years. Self-administered questionnaires—which included social-demographic characteristics, medical history, medications, and lifestyle—were completed during each visit. Blood samples were drawn after an overnight fast and smoking abstinence. Beginning in 2000, we estimated air pollution concentrations at the participant’s home address, including PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components (OC, EC, SO<sub>4</sub><sup>2-</sup>, NO<sub>3</sub><sup>-</sup>), using hybrid spatiotemporal prediction models. In total, we collected complete information for 646 male participants, including air pollution measurements estimated from prediction models and mtDNA measurements for at least one visit between 2000 and 2012 (total mtDNA measurements = 1,178; number of visits per participant = 1–4) (Supplementary Figure 2.1). Of the 646 participants, 273 participants completed only one visit, 236 participants completed two visits, and 137 participants completed three or more visits. Participants signed informed consents at each clinical visit. Institutional Review Boards of the participating institutions approved this study.

### **Estimates of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components**

We constructed separate hybrid models to estimate PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components concentrations at each participant's residential address.<sup>75</sup> Each hybrid model integrates simulation outputs from a chemical transport model, land-use terms, meteorological variables, and other correction terms. In brief, we used Goddard Earth Observing System Chemistry (GEOS-Chem) model, a chemical transport model that simulates formation and transportation of atmospheric components, to obtain raw estimation of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components. Raw outputs were then calibrated by land-use terms, meteorological variables, and weighted sum of nearby monitoring station data using a backward propagation neural network. Neural networks were used to model possible complex and nonlinear relationships between variables. Land-use terms included road density, normalized difference vegetation index (a measurement of greenness in the surrounding space), elevation, population density, percentage of urban space, and EPA emission inventory. Meteorological variables included air temperature, accumulated total precipitation, downward shortwave radiation flux, accumulated total evaporation, planetary boundary layer height, low cloud area fraction, precipitable water for the entire atmosphere, barometric pressure, specific humidity at 2 m, visibility, and wind speed. Wind speed was computed as the vector sum of u-wind (east–west component of the wind) and v-wind (north–south component of the wind), both at 10 m. Ten-fold cross-validation showed that the calibration model performed well for PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components, with total R<sup>2</sup> between 0.70 and 0.80 for PM<sub>2.5</sub> mass and each component. Calibration models also revealed little bias—the slope of predicted versus measured was ~1.01. Annual moving average values were calculated when at least 75% of daily estimates were available.

### **Determination of mtDNA abundance using quantitative real time PCR**

Relative mtDNA abundance was defined as total mtDNA content divided by total nDNA. A detailed protocol has been published.<sup>76</sup> In brief, we used multiplex real-time PCR (RT-PCR), with minor modifications, to measure mtDNA content.<sup>77</sup> We used a reference DNA pool consisting of a subset of 300 test samples from this same study to plot a titration curve for standardization ( $R^2 \geq 0.99$ ). We calculated a ratio between mtDNA and nDNA, which we standardized using the reference DNA pool. The standardized ratio was used in subsequent statistical analysis. A ratio of 1 indicated that mtDNA/nDNA ratio for the test sample was equivalent to mtDNA/nDNA ratio in the reference DNA pool. All samples were run in triplicate.

### **Statistical analysis**

We evaluated associations of annual PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components with mtDNA abundance using linear mixed-effects regressions with subject-specific intercepts to account for repeated measurements of mtDNA abundance from the same individual. First, we used a single-pollutant linear mixed-effects model to estimate associations between each pollutant (PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components) and mtDNA abundance in separate regression models. Since PM<sub>2.5</sub> total mass concentration often correlates with PM<sub>2.5</sub> components and with mtDNA abundance, single-pollutant models do not take into account potential confounding by PM<sub>2.5</sub> mass or by other PM<sub>2.5</sub> components. We consider two approaches to adjust for confounding by PM<sub>2.5</sub> mass and other PM<sub>2.5</sub> components. First, to obtain estimates for PM<sub>2.5</sub> components adjusted by total PM<sub>2.5</sub> mass, we added PM<sub>2.5</sub> mass as a covariate in each single-component model. This yielded a second set of model estimates of the association between each PM<sub>2.5</sub> component and mtDNA abundance, adjusted for confounding by PM<sub>2.5</sub> mass.

However, this method still does not take into account confounding by other chemical components of PM<sub>2.5</sub>. Therefore, we also fit multi-pollutant models that adjust for PM<sub>2.5</sub> total mass concentration and other PM<sub>2.5</sub> components. To obtain more parsimonious models that only include exposures predictive of mtDNA abundance, we applied the adaptive least absolute shrinkage and selection operator (LASSO). LASSO is a regression analysis method with penalization on the absolute values of regression coefficient. The resulting estimates has regression coefficient for components that are not predictive of outcome set equal to zero. Hence, it provides a sparse selection of the PM<sub>2.5</sub> components that are associated with mtDNA abundance. We created a set of component-specific weights for the LASSO as the inverse of coefficients estimated with the ordinary linear mixed-effects models. LASSO requires specifying a tuning parameter,  $\lambda$ . We ran models across a range of  $\lambda$  values and chose the model that gave the smallest Bayesian information criterion (BIC). We only penalized PM<sub>2.5</sub> components and not PM<sub>2.5</sub> mass or other covariates.

We adjusted all models for the following covariates selected *a priori*: age (continuous), body mass index (BMI) [weight (kg)/height (m)<sup>2</sup>, continuous], race (white, black, or other), regular patterns of physical activity (<12 kcal/kg\*hours/week, 12–30 kcal/kg\*hours/week, or >30 kcal/kg\*hours/week), smoking status (never, former, or current smoker), cumulative pack-years of smoking (continuous), education level (high school diploma or less, college degree, or graduate degree), seasonality, and blood counts of lymphocytes, neutrophils, and platelets,. Seasonality was modeled as Fourier series terms  $\cos(2\pi*\text{doy}/365.25)$  and  $\sin(2\pi*\text{doy}/365.25)$ , where *doy* represents day of year. We standardized regression coefficients, and results were

expressed as a standard deviation (SD) change in mtDNA/nDNA ratio per a SD change in pollution concentration.

The main regression model took the general form:

$$mtDNA_{ij} = \beta_0 + u_i + \alpha X_{ij} + \beta Z_{ij} + \varepsilon_{ij}$$

where  $i$  corresponds to  $i^{\text{th}}$  subject;  $j$  corresponds to  $j^{\text{th}}$  visit;  $\beta_0$  is the intercept for the population average;  $u_i$  is the subject-specific random intercept;  $X_{ij} = (X_{ij1}, \dots, X_{ijN})^T$  is a vector of PM<sub>2.5</sub> major chemical component concentrations 365 days prior to each clinical visit;  $Z_{ij} = (Z_{ij1}, \dots, Z_{ijP})^T$  is a vector of PM<sub>2.5</sub> total mass concentration 365 days prior to each clinical visit and other covariates we adjusted for in the regression model. Therefore,  $\alpha$  is the fixed effect of PM<sub>2.5</sub> mass and other covariates ( $X_{ij}$ ), and  $\beta$  represents the penalized effects of PM<sub>2.5</sub> components ( $Z_{ij}$ ) selected by LASSO. The main regression analysis with adaptive LASSO accounts for confounding by PM<sub>2.5</sub> total mass concentration as well as other PM<sub>2.5</sub> components, reducing potential multi-collinearity issues via penalization.

In sensitivity analysis, potential selection bias due to loss of participant follow-up was accounted for using inverse probability of weighting. Using logistic regression, we predicted the probability of participants coming to a subsequent visit using covariates from previous visits that predicted if participants would return. The set of covariates in the logistic regression are: age, BMI, regular patterns of physical activity, smoking status, pack-year smoked, forced expiratory volume in one second (FEV1) versus forced vital capacity (FVC) ratio (a measurement of airway obstruction),

clinically diagnosed hypertension, clinically diagnosed diabetes, medications (diuretics and beta blockers), and education level. We also restricted our analysis to annual PM<sub>2.5</sub> total mass concentration of  $\leq 12 \mu\text{g}/\text{m}^3$ , which is the current National Ambient Air Quality Standard for annual PM<sub>2.5</sub>.

All analyses were conducted with SAS version 9.3 (SAS Institute Inc., Cary, NC) or, with R 3.0.1 (<http://www.r-project.org/>).

## **RESULTS**

### **Descriptive Statistics**

Table 2.1 presents characteristics of the 646 Normative Aging Study participants at each visit. Mean age at first visit was  $73.6 \pm 6.8$  years. The study population was predominantly white (97%). Four percent of participants were current smokers, and 66% were former smokers at first visit. Of the 646 participants, 373 participants came to more than one visit (58%).

### **PM<sub>2.5</sub> Mass and PM<sub>2.5</sub> Component Summary Statistics and Spatial Distribution**

Tables 2.2 and 2.3 present summary statistics and Spearman correlations for annual moving averages of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components. Annual moving averages of EC, OC, SO<sub>4</sub><sup>2-</sup>, and NO<sub>3</sub><sup>-</sup> were positively correlated with PM<sub>2.5</sub> mass ( $\rho = 0.58, 0.68, 0.43, \text{ and } 0.41$ , respectively). Notably, OC, SO<sub>4</sub><sup>2-</sup>, and NO<sub>3</sub><sup>-</sup> contributed to a large proportion of PM<sub>2.5</sub> mass (Table 2.2).

Figure 21 shows the spatial and temporal distribution of annual concentrations of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components in the study area, including participants' residential addresses and locations of



fixed monitoring stations. EC serves as a tracer for local traffic emissions. We observed higher concentrations of EC near major roadways. OC is either emitted from biomass or formed as secondary organic aerosols. OC concentrations are closely related to anthropogenic activities in the study region.  $\text{SO}_4^{2-}$  particles mainly come from coal-burning power plants, they are regional air pollutants and show more dispersed concentrations.  $\text{NO}_3^-$  originates from both power plants and traffic emissions. In the northeastern region, we observed higher  $\text{NO}_3^-$  concentrations near major roadways, which suggests that traffic contribute to a significant part of  $\text{NO}_3^-$  emission in this area.

### **Main Analysis**

Table 2.4 reports the associations of  $\text{PM}_{2.5}$  mass and  $\text{PM}_{2.5}$  components with mtDNA abundance. MtDNA abundance was negatively associated with  $\text{PM}_{2.5}$  mass concentration in the previous annual moving average exposure window. One SD increase in  $\text{PM}_{2.5}$  mass concentration was associated with a 0.07 SD decrease in mtDNA abundance (95% CI:  $-0.13, -0.02$ ). Several  $\text{PM}_{2.5}$  components (OC,  $\text{SO}_4^{2-}$ , and  $\text{NO}_3^-$ ) were negatively associated with mtDNA abundance in single-pollutant models. Further, OC and  $\text{NO}_3^-$  remained statistically significant in two-pollutant models adjusted for  $\text{PM}_{2.5}$  mass. Sulfate levels were also marginally negatively associated with mtDNA abundance in models adjusted for  $\text{PM}_{2.5}$  mass.

Finally, we used multi-pollutant models, which included  $\text{PM}_{2.5}$  mass and  $\text{PM}_{2.5}$  components as independent variables in the same model, and applied LASSO to obtain a more parsimonious model that only included exposures predictive of mtDNA abundance. Figure 2.2 shows LASSO selection path as a function of the penalty term  $\lambda$ . We ran models across a range of  $\lambda$  values and

chose the model that gave the smallest Bayesian Information Criterion (BIC), which was  $\lambda = 1.8$  (Supplementary Figure 2.2). At  $\lambda = 1.8$ , LASSO selected  $\text{NO}_3^-$ , dropping EC, OC, and  $\text{SO}_4^{2-}$  in the regression. Models including  $\text{PM}_{2.5}$  mass and pollutants selected by LASSO showed negative associations of  $\text{NO}_3^-$  with mtDNA abundance (Table 2.5). Specifically, one SD increases in annual nitrate concentration was associated with a 0.12 SD (95% CI: -0.18; -0.07) decrease in mtDNA abundance.

### **Sensitivity Analysis**

We conducted a sensitivity analysis restricted to observations with  $\text{PM}_{2.5}$  mass concentration of  $\leq 12 \mu\text{g}/\text{m}^3$ , the current National Ambient Air Quality Standards (NAAQS) for annual  $\text{PM}_{2.5}$  exposure. In this analysis, we also found a negative association between  $\text{NO}_3^-$  and mtDNA abundance (Supplementary Table 2.2). Further, to account for potential selection bias, we repeated all analyses using inverse probability weighting and obtained result estimates similar to those from primary analyses (results not shown).

### **DISCUSSION**

In this study of a cohort of older individuals, we showed that mtDNA abundance was negatively associated with long-term exposures to  $\text{PM}_{2.5}$  mass and—after adjusting for  $\text{PM}_{2.5}$  mass—certain  $\text{PM}_{2.5}$  components, including OC,  $\text{SO}_4^{2-}$  (marginally) and  $\text{NO}_3^-$ . In multipollutant models that included  $\text{PM}_{2.5}$  mass and  $\text{PM}_{2.5}$  components selected by LASSO,  $\text{NO}_3^-$  remained negatively associated with mtDNA abundance, suggesting that  $\text{NO}_3^-$  was associated with mitochondrial oxidative stress independent of  $\text{PM}_{2.5}$  mass concentration. To our knowledge, this is the first

study describing long-term exposures to PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components with changes in oxidative stress biomarker of mitochondrial damage.

Among the PM<sub>2.5</sub> components we examined, EC is a surrogate for traffic pollution, EC is highly poly-aromatic and exhibits redox cycling ability.<sup>78</sup> OC is the organic fraction of PM<sub>2.5</sub>, and it comprises a variety of redox-active chemicals, including organic acids, alkane, polycyclic aromatic hydrocarbons, and aldehydes. OC is either emitted directly from biomass and fossil fuel combustions or formed as secondary organic aerosols.<sup>79</sup> SO<sub>4</sub><sup>2-</sup> and NO<sub>3</sub><sup>-</sup> are secondary aerosols derived from photochemical reactions of sulfur dioxide (SO<sub>2</sub>) and nitrogen oxides (NO<sub>x</sub>). Power plants are the major source of SO<sub>4</sub><sup>2-</sup>, while NO<sub>3</sub><sup>-</sup> originates from both power plants and traffic emissions.<sup>80</sup>

In this study, we first fit single- and two-pollutant models to estimate the associations between PM<sub>2.5</sub> components with mtDNA abundance. We included in the two-pollutant models PM<sub>2.5</sub> to avoid potential confounding by mass concentration. Further, we applied a multi-pollutant approach and used dimension reduction method *a priori* to reduce potential multi-collinearity issues while adjusting for exposure to other components. We found that NO<sub>3</sub><sup>-</sup> concentration in PM<sub>2.5</sub> negatively associated with mtDNA abundance in multi-pollutant modeling, suggesting that the oxidative stress potential of PM<sub>2.5</sub> may not be fully captured by mass concentration.

The biological mechanisms by which PM<sub>2.5</sub> may impact health are still not fully understood, but oxidative stress is thought to play a key role. In this study, we evaluated mtDNA abundance – a novel biomarker indicative of mtDNA damage. Distinct from nuclear DNA, mitochondria have

circular genomes due to their proteo-bacterial origin. Each cell contains hundreds to thousands of mitochondria, with each mitochondrion carrying 2-10 copies of circular DNA.<sup>3</sup> Maintaining mitochondrial integrity is critical for short-term cellular homeostasis and long-term organism survival. Mitochondrial homeostasis is usually maintained through a dynamic process of fission/fusion, coupled with biogenesis/mitophagy. If damage to the mitochondrial genome accumulates and is beyond repair, mtDNA copies are cleared, resulting in lower mtDNA abundance.<sup>5</sup> Mitochondrial damage may directly impact cellular processes that take place in the mitochondria, such as oxidative phosphorylation and hence ATP production. Mitochondria are also central sites for fatty acid beta oxidation, where fatty acids are catabolized into acetyl-CoA, which in turn feeds into the citric acid cycle.<sup>81</sup> Mitochondrial damage may disturb fatty acid breakdown and dysregulate triglyceride metabolic pathways.<sup>81</sup> In addition, mitochondria play important roles in calcium homeostasis via transient calcium ion uptake and storage for later release. Damage to the mitochondria therefore could disrupt calcium homeostasis and calcium signaling pathways.<sup>82</sup> Further, increased production of oxidative species may cause damage elsewhere—e.g., to macromolecules, such as nuclear DNA, proteins, and lipids—ultimately disrupting normal cell structure and function.

Previous epidemiological studies have reported associations between ambient air pollution exposure and mtDNA abundance. However, most of these studies focus on occupational settings and report somewhat inconsistent results, possibly owing to variations in exposure concentration, exposure duration, type of exposure, and the dynamic balance between oxidative damage and anti-oxidative defense pathways.<sup>6,83,84</sup> For instance, exposure to low-level benzene was associated with increased mtDNA abundance, probably due to an adaptive biogenesis response

to compensate moderate oxidative damage.<sup>83</sup> On the other hand, exposure to high levels of EC and PM<sub>10</sub> (particulate matter with aerodynamic diameter  $\leq 10 \mu\text{m}$ ) was associated with lower levels of mtDNA abundance.<sup>84</sup> These studies focused on short-term exposures, which may not capture aggregated risk due to prolonged environmental exposure.

Our study has a number of strengths. Our study population consists primarily of the older individuals (mean age = 73.6 years at first visit), which represent a susceptible subgroup for air pollution exposures due to lower physiological functions and high prevalence of preexisting diseases. The National Research Council sets high research priority to understand population susceptibility to air pollution exposure, to inform more effective policy that yields maximal health benefits for the entire population. We estimated concentrations of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components at each participant's residential address using state-of-the-art hybrid spatiotemporal prediction models, which incorporate chemical transport models, land-use terms, and meteorological variables. This method allowed us to estimate pollutant concentrations with a high degree of spatial and temporal resolution and introduced sufficient exposure variations among participants to study long-term exposure of ambient air pollutants. Estimated pollution concentrations from these models may serve as better surrogates for personal exposure and hence may reduce potential exposure misclassifications. Our assay for analyzing mtDNA yields high precision; within- and between-plate coefficients of variation were 3.35% and 3.26%, respectively. We used a multi-pollutant approach as our primary analysis, and applied LASSO shrinkage and selection *a priori* to obtain a more parsimonious model. LASSO removes predictors that do not have explanatory power, which helps to address multi-collinearity issues and yielded more precise estimates. Further, we accounted for potential selection bias due to loss

to follow-up using inverse probability of weighting (IPW). Point estimates were similar for models using or not using IPW, indicating that little selection bias was introduced due to loss to follow-up.

Our study also suffers from several limitations. In the current analysis, we modeled four PM<sub>2.5</sub> components that contribute greatly to PM<sub>2.5</sub> mass; however, other components not included in our analysis—such as metals and specific organic species—may also adversely affect health. However, due to their minimal contribution to PM<sub>2.5</sub> mass, these components often fall below method detection limits, and are difficult to estimate with good accuracy. Additionally, exposure estimates from our cohort focused on the greater Boston area and may not be generalizable to other regions of the U.S. due to the variations in PM chemical compositions in different regions.

## **CONCLUSION**

Long-term exposure to PM<sub>2.5</sub> mass and specific PM<sub>2.5</sub> components is associated with decreased mtDNA abundance in this elderly study population. Our findings from multi-pollutant modeling suggest that NO<sub>3</sub><sup>-</sup>—as a chemical component of PM<sub>2.5</sub> was associated with higher mitochondrial oxidative stress independent of PM<sub>2.5</sub> mass concentration, and mass alone may not fully capture the oxidation potency of PM<sub>2.5</sub>.

## **ACKNOWLEDGEMENTS**

The authors would like to thank all Normative Aging Study participants. The authors also thank Johanna Lepeule for sharing the codes adapted to perform inverse probability weighting.

## **FUNDINGS**

This work was supported by National Institutes of Health grants (R01ES015172, R21ES021895, R01ES021733, R01ES021357, and P30ES000002) and by U.S. Environmental Protection Agency funding (RD-83479801). The Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Center of the U.S. Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and Information Center, Boston, Massachusetts. Views expressed in this paper are those of the authors and do not necessarily represent views of the funding or supporting entities.

## **CONFLICTS OF INTEREST**

None declared.

**Table 2.1** Personal characteristics of participants from the Normative Aging Study, 2000-2012 (N = 646).

<b>Variable</b>	<b>Visit 1 (N = 646)</b>	<b>Visit 2 (N = 373)</b>	<b>Visit 3 and 4 (N = 159)</b>
Age [years (mean ± SD)]	73.6 ± 6.8	76.2 ± 6.7	80.0 ± 6.2
BMI [kg/m <sup>2</sup> (mean ± SD)]	28.1 ± 4.0	27.8 ± 4.1	27.8 ± 4.3
Race [n (%)]			
White	627 (97%)	362 (97%)	155 (97%)
Black	13 (2%)	7 (2%)	3 (2%)
Others	6 (1%)	4 (1%)	1 (1%)
Smoking [n (%)]			
Never	191 (30%)	109 (29%)	46 (29%)
Former	429 (66%)	253 (68%)	110 (69%)
Current	26 (4%)	11 (3%)	3 (2%)
Pack year smoked [year (mean ± SD)]	19.9 ± 24.2	20.9 ± 24.5	19.2 ± 22.9
Regular patterns of physical activity [n (%)]			
Low (≤ 12 kcal/kg*hours/week)	427 (66%)	247 (66%)	106 (67%)
Median (12-30 kcal/kg*hours/week)	119 (18%)	87 (23%)	26 (16%)
High (≥30 kcal/kg*hours/week)	100 (16%)	39 (11%)	27 (17%)
Education level [n (%)]			
< 12 years	226 (35%)	125 (33%)	50 (31%)
13-16 years	290 (45%)	171 (46%)	79 (50%)
> 16 years	130 (20%)	77 (21%)	30 (19%)
Lymphocyte count [cells/cm <sup>3</sup> (mean ± SD)]	25.8 ± 9.2	25.1 ± 8.2	23.9 ± 9.1
Neutrophil count [cells/cm <sup>3</sup> (mean ± SD)]	61.7 ± 9.5	62.4 ± 8.8	63.3 ± 9.4
Platelet count [1000/mm <sup>3</sup> (mean ± SD)]	222.5 ± 56.0	225.6 ± 60.8	201.8 ± 45.4



**Table 2.2** Summary statistics of annual moving averages of PM<sub>2.5</sub> mass and PM<sub>2.5</sub> components, Normative Aging Study, 2000-2012.

Pollutant	Mean	SD	Percentiles		
			25th	50th	75th
PM <sub>2.5</sub> mass (μg/m <sup>3</sup> )	10.24	1.62	9.16	10.31	11.35
<i>PM<sub>2.5</sub> chemical components</i>					
EC (μg/m <sup>3</sup> )	0.56	0.17	0.44	0.56	0.67
OC (μg/m <sup>3</sup> )	2.93	0.96	2.21	3.05	3.60
SO <sub>4</sub> <sup>2-</sup> (μg/m <sup>3</sup> )	3.38	1.40	2.90	3.39	3.84
NO <sub>3</sub> <sup>-</sup> (μg/m <sup>3</sup> )	1.13	0.34	1.17	0.91	1.38

**Table 2.3** Spearman correlation between annual moving average concentrations of the air pollutants considered in the analysis, Normative Aging Study, 2000-2012.

	PM <sub>2.5</sub>	EC	OC	SO <sub>4</sub> <sup>2-</sup>	NO <sub>3</sub> <sup>-</sup>
PM <sub>2.5</sub> (µg/m <sup>3</sup> )	1	0.58**	0.68**	0.43**	0.41**
EC (µg/m <sup>3</sup> )	-	1	0.64**	0.49**	0.40**
OC (µg/m <sup>3</sup> )	-	-	1	0.46**	0.56**
SO <sub>4</sub> <sup>2-</sup> (µg/m <sup>3</sup> )	-	-	-	1	0.32**
NO <sub>3</sub> <sup>-</sup> (µg/m <sup>3</sup> )	-	-	-	-	1

\* $p < 0.05$ ; \*\* $p < 0.01$

**Table 2.4** Association between annual averages of ambient air pollution and mtDNA abundance, measured as mitochondrial DNA to nuclear DNA copy number ratio (mtDNA/nDNA), Normative Aging Study, 2000-2012.

Pollutant	Estimates* (95% CI)	p-value	Estimates* (95% CI)	p-value
	Single pollutant model		Adjusted for PM <sub>2.5</sub> <sup>†</sup>	
PM <sub>2.5</sub> mass	-0.07 (-0.13; -0.02)	0.005	- <sup>†</sup>	- <sup>†</sup>
<i>PM<sub>2.5</sub> chemical components</i>				
EC	-0.04 (-0.10; 0.009)	0.11	-0.008 (-0.07; 0.05)	0.79
OC	-0.10 (-0.15; -0.05)	0.0001	-0.09 (-0.15; -0.02)	0.007
SO <sub>4</sub> <sup>2-</sup>	-0.07 (-0.12; -0.02)	0.01	-0.05 (-0.10; 0.004)	0.07
NO <sub>3</sub> <sup>-</sup>	-0.13 (-0.18; -0.08)	<0.0001	-0.12 (-0.18; -0.07)	<0.0001

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, and seasonality.

\* Estimates are expressed as a SD change in mtDNA/nDNA ratio with a SD change in pollution concentration.

<sup>†</sup> We additionally adjusted for PM<sub>2.5</sub> mass concentration. PM<sub>2.5</sub> mass estimate is not reported because we used this set of models to obtain estimates for PM<sub>2.5</sub> components adjusted for PM<sub>2.5</sub> mass. Each individual model provided a different estimate for PM<sub>2.5</sub> mass.

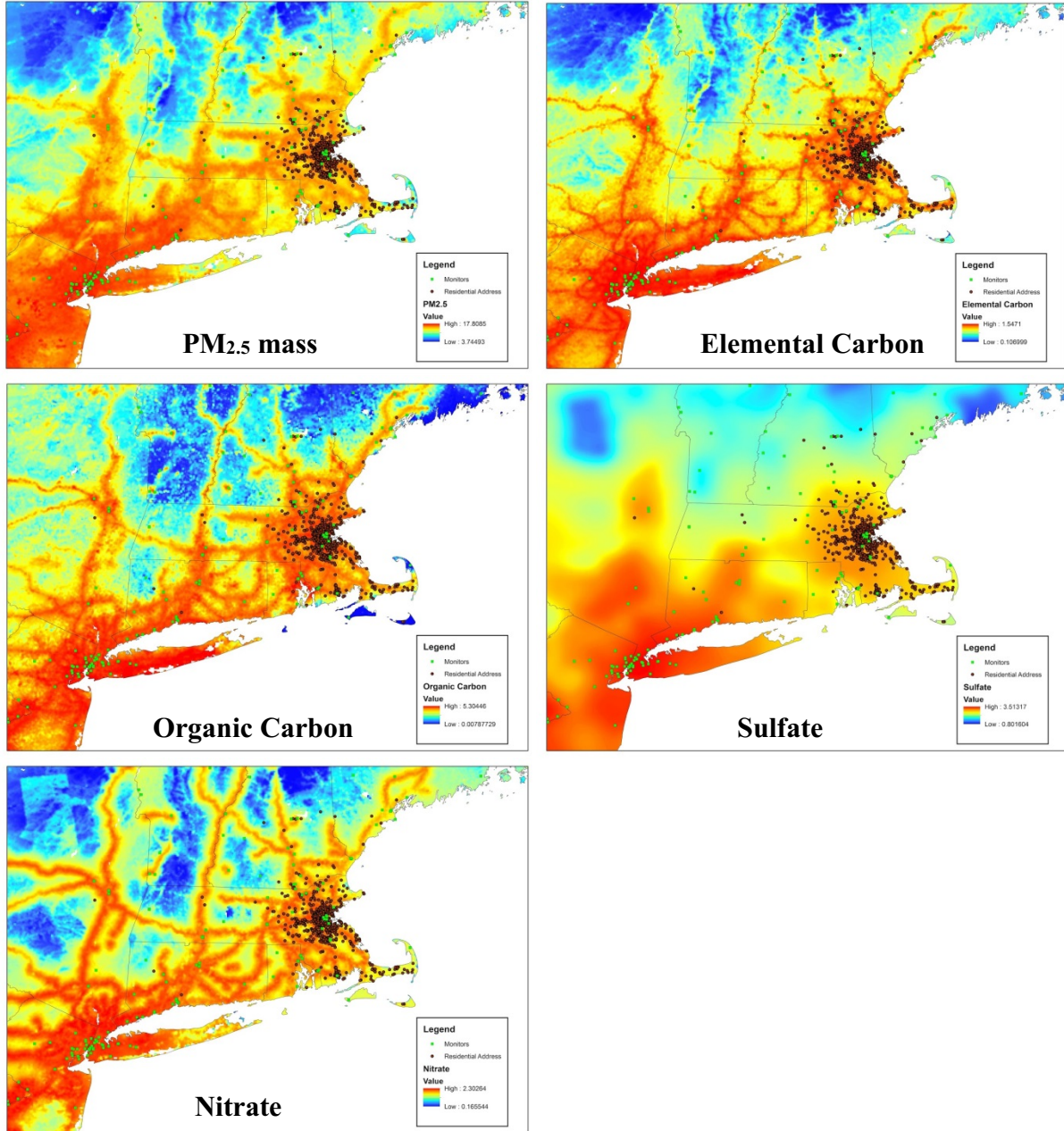
**Table 2.5** Association between annual averages of ambient air pollution and mtDNA abundance, measured as mitochondrial DNA to nuclear DNA copy number ratio (mtDNA/nDNA), Normative Aging Study, 2000-2012. Results are based on LASSO selected pollutants.

<b>Pollutant</b>	<b>Estimates* (95% CI)</b>	<b>p-value</b>
<i>PM<sub>2.5</sub> chemical components</i>		
EC	Not selected by LASSO	-
OC	Not selected by LASSO	-
SO <sub>4</sub> <sup>2-</sup>	Not selected by LASSO	-
NO <sub>3</sub> <sup>-</sup>	-0.12 (-0.18; -0.07)	<0.0001

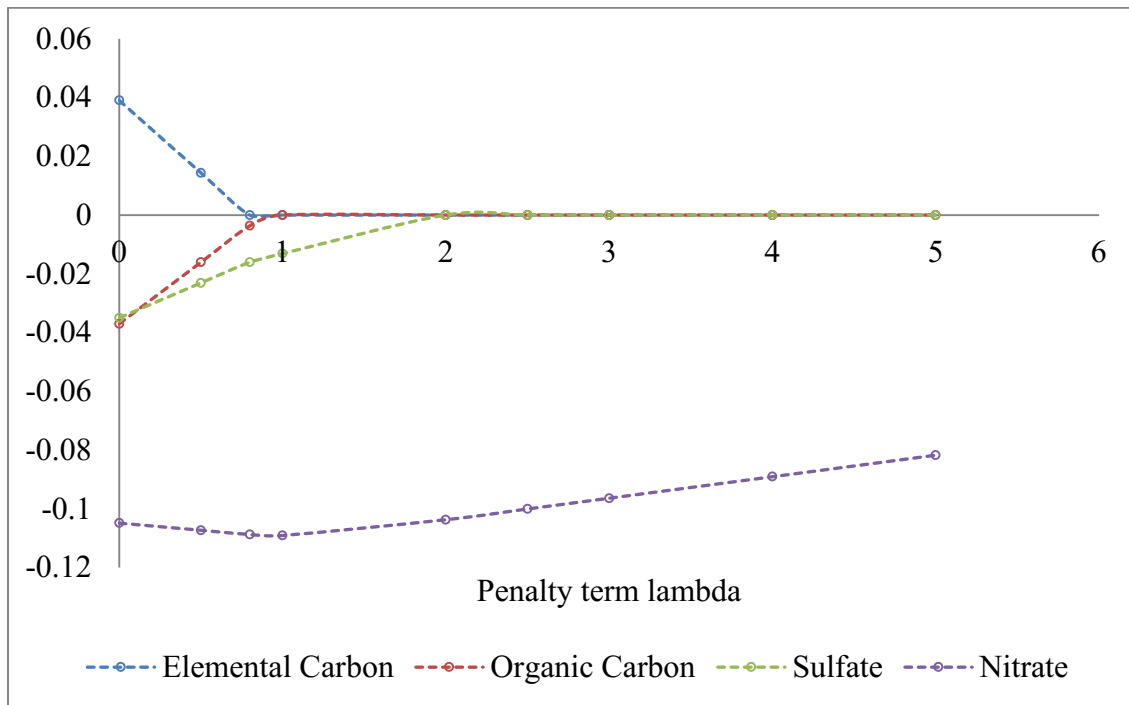
Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, seasonality, PM<sub>2.5</sub> mass concentration and all other components that were selected by LASSO shrinkage method.

\* Estimates are expressed as a SD change in mtDNA/nDNA ratio with a SD change in pollution concentration.

**Figure 2.1** Graphical illustration of annual PM<sub>2.5</sub> mass and PM<sub>2.5</sub> component concentrations for the study period Normative Aging Study, 2000-2012. The green dots indicate where the fixed monitoring stations are located, while the red dots dictate participants' residential address.



**Figure 2.2** LASSO coefficient paths: plot of PM<sub>2.5</sub> components as a function of penalty term lambda.



Based on the Lasso shrinkage method with BIC as a selection criterion, we identified NO<sub>3</sub><sup>-</sup> independently associated with mtDNA abundance.

**Table S2.1** Comparison of estimated coefficients in the primary analysis and in sensitivity analysis restricted to observations with annual moving PM<sub>2.5</sub> average  $\leq 12 \mu\text{g}/\text{m}^3$  (i.e., the current NAAQS for annual PM<sub>2.5</sub> concentration).

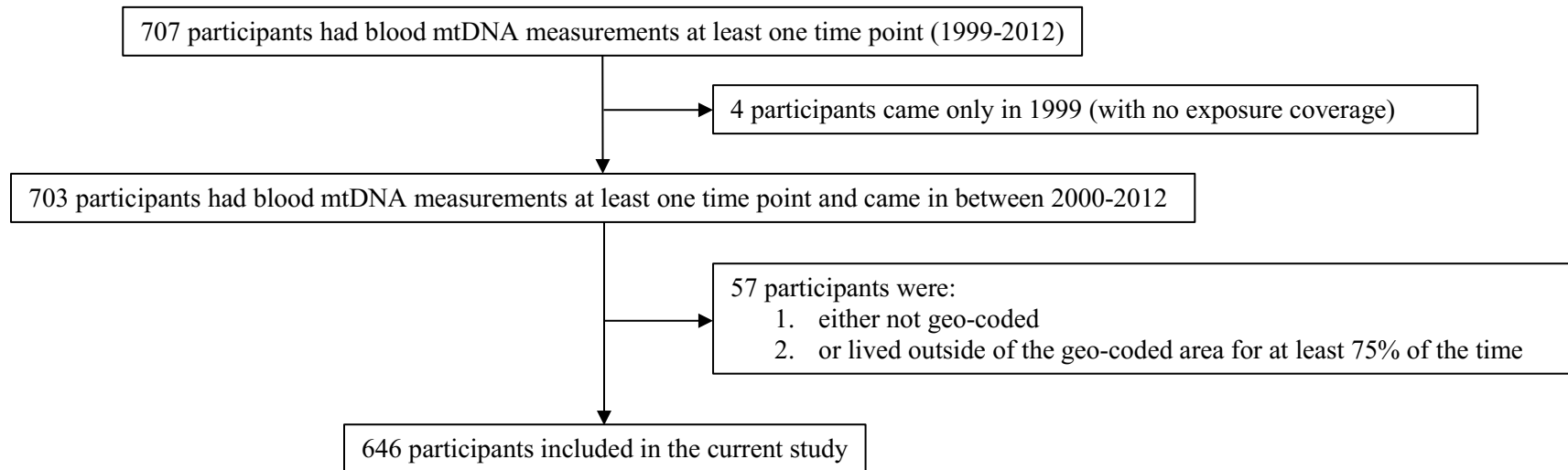
Pollutant	Estimates* (95% CI)	p-value	Estimates* (95% CI)	p-value
	Primary analysis (N = 1178)		Sensitivity analysis† (N=1052)	
NO <sub>3</sub> <sup>-</sup>	-0.12 (-0.18; -0.07)	<0.0001	-0.11 (-0.16; -0.05)	0.0003

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, seasonality, PM<sub>2.5</sub> mass concentration and all other components that were selected by LASSO shrinkage method.

\* Estimates are expressed as a SD change in mtDNA/nDNA ratio with a SD change in pollution concentration.

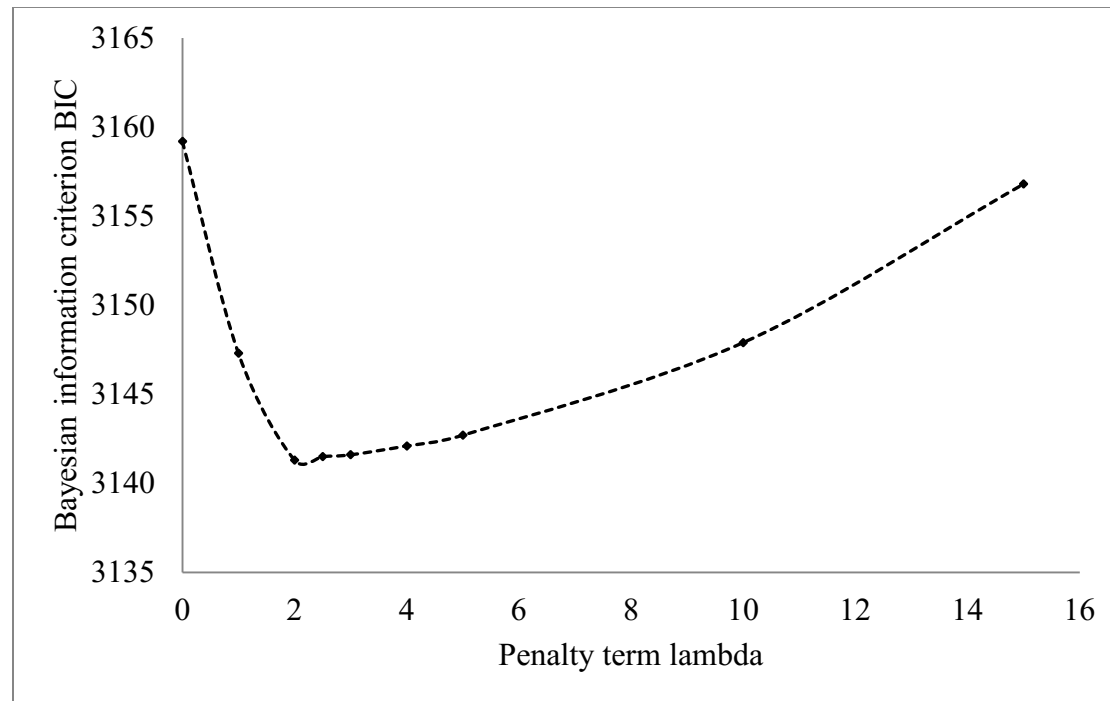
† Sensitivity analysis where PM<sub>2.5</sub> annual moving average  $> 12 \mu\text{g}/\text{m}^3$  were excluded.

**Figure S2.1** Flow chart of participants' inclusion criteria (participants who came in year 2000 will be assigned to the yearly moving average of year 2000).





**Figure S2.2** The relationship between the penalty term lambda in LASSO regression and Bayesian information criterion (BIC).



We performed a mixed model version of Lasso regression, with subject specific intercepts, to account for correlation across multiple visits. We forced the following covariates in: age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, education level, lymphocyte count, neutrophil count, platelet count, seasonality, and PM<sub>2.5</sub> mass concentration

## REFERENCE

1. Bollati V, Baccarelli A. Environmental epigenetics. *Heredity (Edinb)* 2010;**105**(1):105-12.
2. Bird A. Perceptions of epigenetics. *Nature* 2007;**447**(7143):396-8.
3. Friedman JR, Nunnari J. Mitochondrial form and function. *Nature* 2014;**505**(7483):335-43.
4. Mengel-From J, Thinggaard M, Dalgard C, Kyvik KO, Christensen K, Christiansen L. Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Hum Genet* 2014;**133**(9):1149-59.
5. Shigenaga MK, Hagen TM, Ames BN. Oxidative damage and mitochondrial decay in aging. *Proc Natl Acad Sci U S A* 1994;**91**(23):10771-8.
6. Hou L, Zhang X, Dioni L, Barretta F, Dou C, Zheng Y, Hoxha M, Bertazzi PA, Schwartz J, Wu S, Wang S, Baccarelli AA. Inhalable particulate matter and mitochondrial DNA copy number in highly exposed individuals in Beijing, China: a repeated-measure study. *Part Fibre Toxicol* 2013;**10**:17.
7. Zhong J, Cayir A, Trevisi L, Sanchez-Guerra M, Lin X, Peng C, Bind MA, Prada D, Laue H, Brennan KJ, Dereix A, Sparrow D, Vokonas P, Schwartz J, Baccarelli AA. Traffic-Related Air Pollution, Blood Pressure, and Adaptive Response of Mitochondrial Abundance. *Circulation* 2016;**133**(4):378-87.
8. American Diabetes A. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2012;**35 Suppl 1**:S64-71.

9. Tirosh A, Shai I, Tekes-Manova D, Israeli E, Pereg D, Shochat T, Kochba I, Rudich A, Israeli Diabetes Research G. Normal fasting plasma glucose levels and type 2 diabetes in young men. *N Engl J Med* 2005;**353**(14):1454-62.
10. Duckworth WC. Hyperglycemia and cardiovascular disease. *Curr Atheroscler Rep* 2001;**3**(5):383-91.
11. Levitan EB, Song Y, Ford ES, Liu S. Is nondiabetic hyperglycemia a risk factor for cardiovascular disease? A meta-analysis of prospective studies. *Arch Intern Med* 2004;**164**(19):2147-55.
12. Sarwar N, Aspelund T, Eiriksdottir G, Gobin R, Seshasai SR, Forouhi NG, Sigurdsson G, Danesh J, Gudnason V. Markers of dysglycaemia and risk of coronary heart disease in people without diabetes: Reykjavik prospective study and systematic review. *PLoS Med* 2010;**7**(5):e1000278.
13. Crane PK, Walker R, Larson EB. Glucose levels and risk of dementia. *N Engl J Med* 2013;**369**(19):1863-4.
14. Rajagopalan S, Brook RD. Air pollution and type 2 diabetes: mechanistic insights. *Diabetes* 2012;**61**(12):3037-45.
15. Esposito K, Petrizzo M, Maiorino MI, Bellastella G, Giugliano D. Particulate matter pollutants and risk of type 2 diabetes: a time for concern? *Endocrine* 2016;**51**(1):32-7.
16. Brook RD, Jerrett M, Brook JR, Bard RL, Finkelstein MM. The relationship between diabetes mellitus and traffic-related air pollution. *J Occup Environ Med* 2008;**50**(1):32-8.
17. Andersen ZJ, Raaschou-Nielsen O, Ketznel M, Jensen SS, Hvidberg M, Loft S, Tjønneland A, Overvad K, Sorensen M. Diabetes incidence and long-term exposure to air pollution: a cohort study. *Diabetes Care* 2012;**35**(1):92-8.

18. Puett RC, Hart JE, Schwartz J, Hu FB, Liese AD, Laden F. Are particulate matter exposures associated with risk of type 2 diabetes? *Environ Health Perspect* 2011;**119**(3):384-9.
19. Kelishadi R, Mirghaffari N, Poursafa P, Gidding SS. Lifestyle and environmental factors associated with inflammation, oxidative stress and insulin resistance in children. *Atherosclerosis* 2009;**203**(1):311-9.
20. Bind MA, Baccarelli A, Zanobetti A, Tarantini L, Suh H, Vokonas P, Schwartz J. Air pollution and markers of coagulation, inflammation, and endothelial function: associations and epigene-environment interactions in an elderly cohort. *Epidemiology* 2012;**23**(2):332-40.
21. Bind MA, Lepeule J, Zanobetti A, Gasparrini A, Baccarelli A, Coull BA, Tarantini L, Vokonas PS, Koutrakis P, Schwartz J. Air pollution and gene-specific methylation in the Normative Aging Study: association, effect modification, and mediation analysis. *Epigenetics* 2014;**9**(3):448-58.
22. Newell-Price J, Clark AJ, King P. DNA methylation and silencing of gene expression. *Trends Endocrinol Metab* 2000;**11**(4):142-8.
23. Nilsson E, Jansson PA, Perfilyev A, Volkov P, Pedersen M, Svensson MK, Poulsen P, Ribel-Madsen R, Pedersen NL, Almgren P, Fadista J, Ronn T, Klarlund Pedersen B, Scheele C, Vaag A, Ling C. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes* 2014;**63**(9):2962-76.

24. Luttmer R, Spijkerman AM, Kok RM, Jakobs C, Blom HJ, Serne EH, Dekker JM, Smulders YM. Metabolic syndrome components are associated with DNA hypomethylation. *Obes Res Clin Pract* 2013;**7**(2):e106-e115.
25. Power MC, Weisskopf MG, Alexeeff SE, Coull BA, Spiro A, 3rd, Schwartz J. Traffic-related air pollution and cognitive function in a cohort of older men. *Environ Health Perspect* 2011;**119**(5):682-7.
26. Kloog I, Nordio F, Coull BA, Schwartz J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM<sub>2.5</sub> exposures in the Mid-Atlantic states. *Environ Sci Technol* 2012;**46**(21):11913-21.
27. Madrigano J, Kloog I, Goldberg R, Coull BA, Mittleman MA, Schwartz J. Long-term exposure to PM<sub>2.5</sub> and incidence of acute myocardial infarction. *Environ Health Perspect* 2013;**121**(2):192-6.
28. Kloog I KP, Coull BA, Lee HJ, Schwartz J. . Assessing temporally and spatially resolved PM<sub>2.5</sub> exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos Environ* 2011;**45**:6267-6275.
29. Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res* 2004;**32**(3):e38.
30. Lepeule J, Baccarelli A, Motta V, Cantone L, Litonjua AA, Sparrow D, Vokonas PS, Schwartz J. Gene promoter methylation is associated with lung function in the elderly: the Normative Aging Study. *Epigenetics* 2012;**7**(3):261-9.
31. Lepeule J, Bind MA, Baccarelli AA, Koutrakis P, Tarantini L, Litonjua A, Sparrow D, Vokonas P, Schwartz JD. Epigenetic influences on associations between air pollutants

- and lung function in elderly men: the normative aging study. *Environ Health Perspect* 2014;**122**(6):566-72.
32. Herder C, Carstensen M, Ouwens DM. Anti-inflammatory cytokines and risk of type 2 diabetes. *Diabetes Obes Metab* 2013;**15 Suppl 3**:39-50.
33. Vossoughi M, Schikowski T, Vierkotter A, Sugiri D, Hoffmann B, Teichert T, Herder C, Schulte T, Luckhaus C, Raulf-Heimsoth M, Casjens S, Bruning T, Kramer U. Air pollution and subclinical airway inflammation in the SALIA cohort study. *Immun Ageing* 2014;**11**(1):5.
34. van Eeden SF, Tan WC, Suwa T, Mukae H, Terashima T, Fujii T, Qui D, Vincent R, Hogg JC. Cytokines involved in the systemic inflammatory response induced by exposure to particulate matter air pollutants (PM(10)). *Am J Respir Crit Care Med* 2001;**164**(5):826-30.
35. Madrigano J, Baccarelli A, Wright RO, Suh H, Sparrow D, Vokonas PS, Schwartz J. Air pollution, obesity, genes and cellular adhesion molecules. *Occup Environ Med* 2010;**67**(5):312-7.
36. Bluher M, Unger R, Rassoul F, Richter V, Paschke R. Relation between glycaemic control, hyperinsulinaemia and plasma concentrations of soluble adhesion molecules in patients with impaired glucose tolerance or Type II diabetes. *Diabetologia* 2002;**45**(2):210-6.
37. Choi YS, Kim S, Kyu Lee H, Lee KU, Pak YK. In vitro methylation of nuclear respiratory factor-1 binding site suppresses the promoter activity of mitochondrial transcription factor A. *Biochem Biophys Res Commun* 2004;**314**(1):118-22.

38. American Diabetes A. Standards of medical care in diabetes--2008. *Diabetes Care* 2008;**31 Suppl 1**:S12-54.
39. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;**51**(6):1173-82.
40. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods* 2013;**18**(2):137-50.
41. Diggle P, Heagerty, P., Liang, K. Y., Zeger, S. *Analysis of Longitudinal Data*. Oxford: Oxford Statistical Science, 2013.
42. Bauer DJ, Preacher KJ, Gil KM. Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. *Psychol Methods* 2006;**11**(2):142-63.
43. M-A. C. Bind TJV, B. A. Coull, J. D. Schwartz. Causal mediation analysis for longitudinal data with exogenous exposure. *Biostatistics in press*.
44. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;**15**(5):615-25.
45. Eze IC, Hemkens LG, Bucher HC, Hoffmann B, Schindler C, Kunzli N, Schikowski T, Probst-Hensch NM. Association between Ambient Air Pollution and Diabetes Mellitus in Europe and North America: Systematic Review and Meta-Analysis. *Environ Health Perspect* 2015.
46. Ruckerl R, Ibald-Mulli A, Koenig W, Schneider A, Woelke G, Cyrys J, Heinrich J, Marder V, Frampton M, Wichmann HE, Peters A. Air pollution and markers of

- inflammation and coagulation in patients with coronary heart disease. *Am J Respir Crit Care Med* 2006;**173**(4):432-41.
47. O'Neill MS, Veves A, Sarnat JA, Zanobetti A, Gold DR, Economides PA, Horton ES, Schwartz J. Air pollution and inflammation in type 2 diabetes: a mechanism for susceptibility. *Occup Environ Med* 2007;**64**(6):373-9.
48. Rahman A, Fazal F. Hug tightly and say goodbye: role of endothelial ICAM-1 in leukocyte transmigration. *Antioxid Redox Signal* 2009;**11**(4):823-39.
49. Hak AE, Pols HA, Stehouwer CD, Meijer J, Kiliaan AJ, Hofman A, Breteler MM, Witteman JC. Markers of inflammation and cellular adhesion molecules in relation to insulin resistance in nondiabetic elderly: the Rotterdam study. *J Clin Endocrinol Metab* 2001;**86**(9):4398-405.
50. Meigs JB, Hu FB, Rifai N, Manson JE. Biomarkers of endothelial dysfunction and risk of type 2 diabetes mellitus. *JAMA* 2004;**291**(16):1978-86.
51. Sun Q, Yue P, DeLuca JA, Lumeng CN, Kampfrath T, Mikolaj MB, Cai Y, Ostrowski MC, Lu B, Parthasarathy S, Brook RD, Moffatt-Bruce SD, Chen LC, Rajagopalan S. Ambient air pollution exaggerates adipose inflammation and insulin resistance in a mouse model of diet-induced obesity. *Circulation* 2009;**119**(4):538-46.
52. Rao X, Patel P, Puett R, Rajagopalan S. Air pollution as a risk factor for type 2 diabetes. *Toxicol Sci* 2015;**143**(2):231-41.
53. Mendez R, Zheng Z, Fan Z, Rajagopalan S, Sun Q, Zhang K. Exposure to fine airborne particulate matter induces macrophage infiltration, unfolded protein response, and lipid deposition in white adipose tissue. *Am J Transl Res* 2013;**5**(2):224-34.



54. Yagi N, Yokono K, Amano K, Nagata M, Tsukamoto K, Hasegawa Y, Yoneda R, Okamoto N, Moriyama H, Miki M, et al. Expression of intercellular adhesion molecule 1 on pancreatic beta-cells accelerates beta-cell destruction by cytotoxic T-cells in murine autoimmune diabetes. *Diabetes* 1995;**44**(7):744-52.
55. Guo L, Byun HM, Zhong J, Motta V, Barupal J, Zheng Y, Dou C, Zhang F, McCracken JP, Diaz A, Marco SG, Colicino S, Schwartz J, Wang S, Hou L, Baccarelli AA. Effects of short-term exposure to inhalable particulate matter on DNA methylation of tandem repeats. *Environ Mol Mutagen* 2014;**55**(4):322-35.
56. Madrigano J, Baccarelli A, Mittleman MA, Wright RO, Sparrow D, Vokonas PS, Tarantini L, Schwartz J. Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. *Environ Health Perspect* 2011;**119**(7):977-82.
57. Li E, Beard C, Jaenisch R. Role for DNA methylation in genomic imprinting. *Nature* 1993;**366**(6453):362-5.
58. Tarantini L, Bonzini M, Apostoli P, Pegoraro V, Bollati V, Marinelli B, Cantone L, Rizzo G, Hou L, Schwartz J, Bertazzi PA, Baccarelli A. Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. *Environ Health Perspect* 2009;**117**(2):217-22.
59. Hou L, Zhang X, Zheng Y, Wang S, Dou C, Guo L, Byun HM, Motta V, McCracken J, Diaz A, Kang CM, Koutrakis P, Bertazzi PA, Li J, Schwartz J, Baccarelli AA. Altered methylation in tandem repeat element and elemental component levels in inhalable air particles. *Environ Mol Mutagen* 2014;**55**(3):256-65.

60. Relton CL, Davey Smith G. Is epidemiology ready for epigenetics? *Int J Epidemiol* 2012;**41**(1):5-9.
61. Tost J, Gut IG. DNA methylation analysis by pyrosequencing. *Nat Protoc* 2007;**2**(9):2265-75.
62. Dockery DW, Pope CA, 3rd, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG, Jr., Speizer FE. An association between air pollution and mortality in six U.S. cities. *N Engl J Med* 1993;**329**(24):1753-9.
63. Pope CA, 3rd, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW, Jr. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am J Respir Crit Care Med* 1995;**151**(3 Pt 1):669-74.
64. Pope CA, 3rd, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K, Thurston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002;**287**(9):1132-41.
65. Council NR. Research Priorities for Airborne Particulate Matter IV - Continuing Research Progress. Committee on Research Priorities for Airbone Particulate Matter. In: Council) NNR, ed. Washington, DC: National Academy Press, 2004.
66. Dai L, Zanobetti A, Koutrakis P, Schwartz JD. Associations of fine particulate matter species with mortality in the United States: a multicity time-series analysis. *Environ Health Perspect* 2014;**122**(8):837-42.
67. Peng RD, Bell ML, Geyh AS, McDermott A, Zeger SL, Samet JM, Dominici F. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. *Environ Health Perspect* 2009;**117**(6):957-63.

68. Ito K, Mathes R, Ross Z, Nadas A, Thurston G, Matte T. Fine particulate matter constituents associated with cardiovascular hospitalizations and mortality in New York City. *Environ Health Perspect* 2011;**119**(4):467-73.
69. Kioumourtzoglou MA, Austin E, Koutrakis P, Dominici F, Schwartz J, Zanobetti A. PM2.5 and survival among older adults: effect modification by particulate composition. *Epidemiology* 2015;**26**(3):321-7.
70. Franklin M, Koutrakis P, Schwartz P. The role of particle composition on the association between PM2.5 and mortality. *Epidemiology* 2008;**19**(5):680-9.
71. Madamanchi NR, Runge MS. Mitochondrial dysfunction in atherosclerosis. *Circ Res* 2007;**100**(4):460-73.
72. Ballinger SW. Mitochondrial dysfunction in cardiovascular disease. *Free Radic Biol Med* 2005;**38**(10):1278-95.
73. Ferrari R. The role of mitochondria in ischemic heart disease. *J Cardiovasc Pharmacol* 1996;**28 Suppl 1**:S1-10.
74. Dikalov SI, Ungvari Z. Role of mitochondrial oxidative stress in hypertension. *Am J Physiol Heart Circ Physiol* 2013;**305**(10):H1417-27.
75. Qian Di PK, Joel Schwartz. A Hybrid Prediction Model for PM2.5 Mass and Components Using a Chemical Transport Model and Land Use Regression. *Atmospheric Environment* 2016;**131**:390-399.
76. Lin C-S, Wang L-S, Chou T-Y, Hsu W-H, Lin H-C, Lee S-Y, Lee M-H, Chang S-C, Wei Y-H. Cigarette smoking and hOGG1 Ser326Cys polymorphism are associated with 8-OHdG accumulation on mitochondrial DNA in thoracic esophageal squamous cell carcinoma. *Annals of surgical oncology* 2013;**20**(3):379-388.

77. Andreu AL, Martinez R, Marti R, Garcia-Arumi E. Quantification of mitochondrial DNA copy number: pre-analytical factors. *Mitochondrion* 2009;**9**(4):242-6.
78. Antinolo M, Willis MD, Zhou S, Abbatt JP. Connecting the oxidation of soot to its redox cycling abilities. *Nat Commun* 2015;**6**:6812.
79. Seinfeld JHaP, S. N. Atmospheric chemistry and physics: From air pollution to climate change. *John Wiley, New York, USA* 1998:700-765.
80. Bell ML, Dominici F, Ebisu K, Zeger SL, Samet JM. Spatial and temporal variation in PM(2.5) chemical composition in the United States for health effects studies. *Environ Health Perspect* 2007;**115**(7):989-95.
81. Bartlett K, Eaton S. Mitochondrial beta-oxidation. *Eur J Biochem* 2004;**271**(3):462-9.
82. Vandecasteele G, Szabadkai G, Rizzuto R. Mitochondrial calcium homeostasis: mechanisms and molecules. *IUBMB Life* 2001;**52**(3-5):213-9.
83. Carugno M, Pesatori AC, Dioni L, Hoxha M, Bollati V, Albetti B, Byun HM, Bonzini M, Fustinoni S, Cocco P, Satta G, Zucca M, Merlo DF, Cipolla M, Bertazzi PA, Baccarelli A. Increased mitochondrial DNA copy number in occupations associated with low-dose benzene exposure. *Environ Health Perspect* 2012;**120**(2):210-5.
84. Pavanello S, Dioni L, Hoxha M, Fedeli U, Mielzynska-Svach D, Baccarelli AA. Mitochondrial DNA copy number and exposure to polycyclic aromatic hydrocarbons. *Cancer Epidemiol Biomarkers Prev* 2013;**22**(10):1722-9.
85. Braga AL, Zanobetti A, Schwartz J. The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. *Environ Health Perspect* 2002;**110**(9):859-63.
86. Wolf K, Schneider A, Breitner S, von Klot S, Meisinger C, Cyrus J, Hymer H, Wichmann HE, Peters A, Cooperative Health Research in the Region of Augsburg Study

- G. Air temperature and the occurrence of myocardial infarction in Augsburg, Germany. *Circulation* 2009;**120**(9):735-42.
87. Schwartz J, Samet JM, Patz JA. Hospital admissions for heart disease: the effects of temperature and humidity. *Epidemiology* 2004;**15**(6):755-61.
88. Koken PJ, Piver WT, Ye F, Elixhauser A, Olsen LM, Portier CJ. Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver. *Environ Health Perspect* 2003;**111**(10):1312-7.
89. Schneider A, Panagiotakos D, Picciotto S, Katsouyanni K, Lowel H, Jacquemin B, Lanki T, Stafoggia M, Bellander T, Koenig W, Peters A, Group AS. Air temperature and inflammatory responses in myocardial infarction survivors. *Epidemiology* 2008;**19**(3):391-400.
90. Kim Y, Joh S. A vulnerability study of the low-income elderly in the context of high temperature and mortality in Seoul, Korea. *Sci Total Environ* 2006;**371**(1-3):82-8.
91. Barnett AG, Sans S, Salomaa V, Kuulasmaa K, Dobson AJ, Project WM. The effect of temperature on systolic blood pressure. *Blood Press Monit* 2007;**12**(3):195-203.
92. Alperovitch A, Lacombe JM, Hanon O, Dartigues JF, Ritchie K, Ducimetiere P, Tzourio C. Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: the Three-City study. *Arch Intern Med* 2009;**169**(1):75-80.
93. Ren C, O'Neill MS, Park SK, Sparrow D, Vokonas P, Schwartz J. Ambient temperature, air pollution, and heart rate variability in an aging population. *Am J Epidemiol* 2011;**173**(9):1013-21.
94. Halonen JJ, Zanobetti A, Sparrow D, Vokonas PS, Schwartz J. Relationship between outdoor temperature and blood pressure. *Occup Environ Med* 2011;**68**(4):296-301.

95. Wilker EH, Yeh G, Wellenius GA, Davis RB, Phillips RS, Mittleman MA. Ambient temperature and biomarkers of heart failure: a repeated measures analysis. *Environ Health Perspect* 2012;**120**(8):1083-7.
96. Schauble CL, Hampel R, Breitner S, Ruckerl R, Phipps R, Diaz-Sanchez D, Devlin RB, Carter JD, Soukup J, Silbajoris R, Dailey L, Koenig W, Cyrus J, Geruschkat U, Belcredi P, Kraus U, Peters A, Schneider AE. Short-term effects of air temperature on blood markers of coagulation and inflammation in potentially susceptible individuals. *Occup Environ Med* 2012;**69**(9):670-8.
97. Keatinge WR, Coleshaw SR, Cotter F, Mattock M, Murphy M, Chelliah R. Increases in platelet and red cell counts, blood viscosity, and arterial pressure during mild surface cooling: factors in mortality from coronary and cerebral thrombosis in winter. *Br Med J (Clin Res Ed)* 1984;**289**(6456):1405-8.
98. Hampel R, Breitner S, Ruckerl R, Frampton MW, Koenig W, Phipps RP, Wichmann HE, Peters A, Schneider A. Air temperature and inflammatory and coagulation responses in men with coronary or pulmonary disease during the winter season. *Occup Environ Med* 2010;**67**(6):408-16.
99. Bind MA, Zanobetti A, Gasparini A, Peters A, Coull B, Baccarelli A, Tarantini L, Koutrakis P, Vokonas P, Schwartz J. Effects of temperature and relative humidity on DNA methylation. *Epidemiology* 2014;**25**(4):561-9.
100. Dhalla NS, Temsah RM, Netticadan T. Role of oxidative stress in cardiovascular diseases. *J Hypertens* 2000;**18**(6):655-73.

101. Yakes FM, Van Houten B. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc Natl Acad Sci U S A* 1997;**94**(2):514-9.
102. Zastawny TH, Dabrowska M, Jaskolski T, Klimarczyk M, Kulinski L, Koszela A, Szczesniewicz M, Sliwinska M, Witkowski P, Olinski R. Comparison of oxidative base damage in mitochondrial and nuclear DNA. *Free Radic Biol Med* 1998;**24**(5):722-5.
103. Demple B, Harrison L. Repair of oxidative damage to DNA: enzymology and biology. *Annu Rev Biochem* 1994;**63**:915-48.
104. Corral-Debrinski M, Stepien G, Shoffner JM, Lott MT, Kanter K, Wallace DC. Hypoxemia is associated with mitochondrial DNA damage and gene induction. Implications for cardiac disease. *JAMA* 1991;**266**(13):1812-6.
105. Lesnefsky EJ, Moghaddas S, Tandler B, Kerner J, Hoppel CL. Mitochondrial dysfunction in cardiac disease: ischemia--reperfusion, aging, and heart failure. *J Mol Cell Cardiol* 2001;**33**(6):1065-89.
106. Corral-Debrinski M, Shoffner JM, Lott MT, Wallace DC. Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease. *Mutat Res* 1992;**275**(3-6):169-80.
107. Mehta AJ, Kloog I, Zanobetti A, Coull BA, Sparrow D, Vokonas P, Schwartz J. Associations between changes in city and address specific temperature and QT interval--the VA Normative Aging Study. *PLoS One* 2014;**9**(9):e106258.
108. Kloog I, Chudnovsky A, Koutrakis P, Schwartz J. Temporal and spatial assessments of minimum air temperature using satellite surface temperature measurements in Massachusetts, USA. *Sci Total Environ* 2012;**432**:85-92.

109. Kloog I, Nordio, F., Coull, B. A., Schwartz, J. . Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. *Remote Sens. Environ.* 2014(150):132–139.
110. Furda A, Santos JH, Meyer JN, Van Houten B. Quantitative PCR-based measurement of nuclear and mitochondrial DNA damage and repair in mammalian cells. *Methods Mol Biol* 2014;**1105**:419-37.
111. Ponti M, Forrow SM, Souhami RL, D'Incalci M, Hartley JA. Measurement of the sequence specificity of covalent DNA modification by antineoplastic agents using Taq DNA polymerase. *Nucleic Acids Res* 1991;**19**(11):2929-33.
112. Gasparrini A, Guo Y, Hashizume M, Lavigne E, Zanobetti A, Schwartz J, Tobias A, Tong S, Rocklov J, Forsberg B, Leone M, De Sario M, Bell ML, Guo YL, Wu CF, Kan H, Yi SM, de Sousa Zanotti Stagliorio Coelho M, Saldiva PH, Honda Y, Kim H, Armstrong B. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet* 2015;**386**(9991):369-75.
113. Yang J, Yin P, Zhou M, Ou CQ, Li M, Li J, Liu X, Gao J, Liu Y, Qin R, Xu L, Huang C, Liu Q. The burden of stroke mortality attributable to cold and hot ambient temperatures: Epidemiological evidence from China. *Environ Int* 2016;**92-93**:232-238.
114. O'Neill MS, Ebi KL. Temperature extremes and health: impacts of climate variability and change in the United States. *J Occup Environ Med* 2009;**51**(1):13-25.
115. Leon LR, Helwig BG. Heat stroke: role of the systemic inflammatory response. *J Appl Physiol (1985)* 2010;**109**(6):1980-8.



116. Hall DM, Baumgardner KR, Oberley TD, Gisolfi CV. Splanchnic tissues undergo hypoxic stress during whole body hyperthermia. *Am J Physiol* 1999;**276**(5 Pt 1):G1195-203.
117. Hall DM, Buettner GR, Oberley LW, Xu L, Matthes RD, Gisolfi CV. Mechanisms of circulatory and intestinal barrier dysfunction during whole body hyperthermia. *Am J Physiol Heart Circ Physiol* 2001;**280**(2):H509-21.
118. Peng C, Bind MC, Colicino E, Kloog I, Byun HM, Cantone L, Trevisi L, Zhong J, Brennan K, Dereix AE, Vokonas PS, Coull BA, Schwartz JD, Baccarelli AA. Particulate Air Pollution and Fasting Blood Glucose in Non-Diabetic Individuals: Associations and Epigenetic Mediation in the Normative Aging Study, 2000-2011. *Environ Health Perspect* 2016.
119. Murphy MP. How mitochondria produce reactive oxygen species. *Biochem J* 2009;**417**(1):1-13.
120. Crecelius AR, Kirby BS, Dinunno FA. Intravascular ATP and the regulation of blood flow and oxygen delivery in humans. *Exerc Sport Sci Rev* 2015;**43**(1):5-13.
121. Bouchama A, Knochel JP. Heat stroke. *N Engl J Med* 2002;**346**(25):1978-88.
122. Stauss HM, Morgan DA, Anderson KE, Massett MP, Kregel KC. Modulation of baroreflex sensitivity and spectral power of blood pressure by heat stress and aging. *Am J Physiol* 1997;**272**(2 Pt 2):H776-84.
123. Inoue Y, Shibasaki M, Ueda H, Ishizashi H. Mechanisms underlying the age-related decrement in the human sweating response. *Eur J Appl Physiol Occup Physiol* 1999;**79**(2):121-6.

124. Kenney MJ, Fels RJ. Sympathetic nerve regulation to heating is altered in senescent rats.  
*Am J Physiol Regul Integr Comp Physiol* 2002;**283**(2):R513-20.

## CHAPTER 3

### Short-term effects of air temperature and mitochondrial DNA lesions among older population

Cheng Peng <sup>a</sup>, Marco Sanchez-Guerra <sup>a</sup>, Ander Wilson <sup>b</sup>, Amar J. Mehta <sup>a</sup>, Jia Zhong <sup>c</sup>, Antonella Zanobetti <sup>a</sup>, Kasey Brennan <sup>c</sup>, Alexandra E. Dereix <sup>c</sup>, Brent A. Coull <sup>b</sup>, Pantel Vokonas <sup>d</sup>, Joel Schwartz <sup>a, e</sup>, Andrea Baccarelli <sup>c</sup>

<sup>a</sup> Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>b</sup> Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA

<sup>c</sup> Department of Environmental Health, Mailman School of Public Health, Columbia University, New York, NY, USA

<sup>d</sup> Normative Aging Study, Veterans Affairs Boston Healthcare System and the Department of Medicine, Boston University School of Medicine, Boston, MA, USA

<sup>e</sup> Channing Laboratory, Harvard Medical School, Boston, MA, USA

Correspondence: Cheng Peng, Department of Environmental Health, Harvard T.H. Chan School of Public Health, Building 1, G-7, 665 Huntington Avenue, Boston, MA 02115, USA. Email: [chp950@mail.harvard.edu](mailto:chp950@mail.harvard.edu)

Running Title: Temperature, mitochondrial DNA lesion, spline

## **Abstract**

**Background:** Previous studies have linked both extreme and sub-optimal air temperature to cardiopulmonary morbidity and mortality, especially in older individuals. However, the underlying mechanisms are yet to be determined.

**Objectives:** We hypothesized that short-term increases in air temperature may induce blood mitochondrial DNA (mtDNA) lesions, which could contribute to the path of temperature related pathogenesis.

**Methods:** We repeatedly measured mtDNA lesions in blood sample from 654 participants in the Normative Aging Study from 1999 to 2013 (1142 observations). Hourly temperature data were obtained from the Boston Logan Airport weather station (located approximately 12 km from the clinical site). We calculated the 2-, 7-, and 14-day moving averages of 24-hour mean and 24-hour variability of temperature. We measured mtDNA lesions using real-time polymerase chain reaction, and mtDNA lesion was expressed as the number of lesions per 10kb. We fit covariate-adjusted linear-mixed models accounting for repeated measures to evaluate the association between short-term increases in mean and variability of temperature with mtDNA lesions.

**Results:** Interquartile increases in the 7-, and 14-day moving average of 24-hour mean temperature in summer were associated with a 0.15 (95% CI: [0.05; 0.25],  $p=0.0033$ ), and 0.19 (95% CI: [0.07; 0.30],  $p=0.0011$ ) increase in the number of mtDNA lesion per 10kb, respectively. We also observed positive associations between increases in mean air temperature and mtDNA lesions in winter. For example, an IQR increase in 7-day (5.2 °C) and 14-day (4.2 °C) moving average of 24-hour mean temperature was associated with a 0.26 (95% CI: [0.08; 0.43],  $p=0.0037$ ), and 0.23 (95% CI: [0.08; 0.37],  $p=0.0019$ ) increase in number of mtDNA lesions per 10kb, respectively. Results were similar when we further adjusted for temperature

variability. We also observed significant associations between increases in temperature variability and mtDNA lesions independent of mean air temperature.

**Conclusions:** Short-term exposures to higher mean and variability of air temperature were associated with increased mtDNA lesions in older adults, supporting the hypothesis that changes meteorological conditions may induce pathophysiological responses among susceptible populations.

## Introduction

A growing body of evidence has shown that changes in air temperature are associated with increased cardiopulmonary morbidity and mortality,<sup>85-89</sup> especially in the older population.<sup>90</sup> However, the underlying mechanisms have not been clearly delineated. Previous epidemiological studies have linked changes in air temperature with abnormal cardiac output,<sup>91-94</sup> myocyte dysfunction,<sup>95</sup> increased blood pro-coagulatory and pro-inflammatory markers,<sup>96-98</sup> as well as epigenetic modifications.<sup>99</sup> Yet, the role of mitochondrial-related oxidative stress—an emerging contributor to a range of cardiovascular disorders, such as atherosclerosis, ischemic heart disease, and hypertension<sup>100</sup>—has not been studied.

Mitochondrial DNA (mtDNA) is a critical cellular target for reactive oxygen species (ROS).<sup>101,102</sup> Distinct from other cellular organelles, mitochondria contain their own circular double-stranded DNA molecules.<sup>4</sup> Mitochondrial genome appears to be more vulnerable to oxidative attack than nuclear DNA, due to the lack of protective barriers, such as histone proteins and chromatin organizations, as well as the relative limited activity to repair DNA damage.<sup>3,4</sup> . Moreover, since 1%-2% of the molecular oxygen consumed by aerobic respiration is released as ROS via the electron transport chain (ETC) reaction in the mitochondria, disturbance of normal ETC activity may further exacerbate mitochondrial-ROS formation.<sup>103</sup> As a result, damaged mitochondria may themselves become an important source of endogenous ROS generation and lead to further damage. Oxidative damage to the mitochondria may induce mtDNA lesions, and types of lesions include single- and double-strand breaks, abasic sites, and base damage / modification.<sup>101</sup> Several lines of evidence indicate associations between mtDNA lesions and cardiovascular diseases,<sup>104-106</sup> which suggest that mitochondrial genome damage, as reflected by

the increase in the number of mtDNA lesions, may represent a plausible biological mechanism of air temperature related cardiovascular events.

In a repeated-measure study of older men in the Greater Boston area, we investigated the association between short-term increases in 24-hour mean and variability standard deviation (SD) of air temperature and blood mtDNA lesions. MtDNA lesions were expressed as the number of lesions per 10 kb. We hypothesized that older adults may be less adaptive to short-term increases in air temperature—especially at the extreme / sub-optimal ranges—which may lead to increased mtDNA lesions in blood.

## **Materials and Methods**

### **Study Population**

Our study population came from the Normative Aging Study, a prospective longitudinal cohort established in 1963 by the U.S. Veterans Administration from the Great Boston area. A detailed protocol has been published previously.<sup>25</sup> In brief, participants underwent clinical examinations every 3 to 5 years. At each visit, self-administered questionnaires were collected – providing information on social-demographic characteristics, medical history, medication and lifestyle. Blood samples were drawn after an overnight fast and smoking abstinence. A total of 654 participants had complete information on temperature measurements and blood mitochondrial markers for at least one and up to four visits between 1999-2013. Specifically, 173 participants came to one clinical visit, 387 participants came to two clinical visits, and 94 participants came to three or more clinical visits. This study was approved by the Institutional Review Boards of the participating institutions.

### **Ambient Temperature Measurements**

We obtained hourly temperature and humidity data from the National Weather Service Station at Boston Logan airport – located approximately 12km from the examination site. For each calendar day, we calculated 24-hour mean and 24-hour standard deviation of temperature. We then computed moving average of the 24-hour mean and 24-hour SD of temperature (i.e., 2-, 7- and 14-day prior to clinical visit).<sup>107</sup>

As a secondary analysis, we estimated air temperature at each participant’s residential address for each calendar day using a hybrid land-use regression and satellite-based model.<sup>108</sup>

Specifically, we utilized MODIS (Moderate Resolution Imaging Spectroradiometer) satellite-derived land surface temperature measurements to predict daily air temperature at a 1km by 1km spatial resolution. The model showed excellent prediction performance with cross validation  $R^2 = 0.94$  and a root mean square predicted error (RMSPE) of 1.48. For more in depth details please refer to Kloog 2014.<sup>109</sup> We then computed moving average of 24-hour mean of temperature for 2-, 7- and 14-days prior to each clinical visit.

### **Mitochondrial DNA Lesion Measurements**

MtDNA lesions were defined as the number of mtDNA lesions per 10kb. A detailed protocol has been published previously.<sup>110</sup> In brief, we adapted a quantitative polymerase chain reaction assay (QPCR) to amplify a long fragment (8.9 kb) and a short fragment (222bp) of the mitochondrial genome.<sup>111</sup> Because each mitochondrion contains multiple copies of circular DNA, the short fragment was amplified for normalization purposes, and this is based on the premise that the



short fragment would have no lesion due to its short in length. Oxidative damage associated mtDNA lesions, such as single- and double-stranded breaks, abasic sites, and base damage / modification would halt the progression of the DNA polymerase, which in turn lead to decreased amplification of the desire sequence. Therefore, only those mtDNA fragments with no lesions would be amplified as a result. We measured the intensity of the fluorescence using PicoGreen dsDNA (Thermo Fisher Scientific, Life Technologies, Carlsbad, California). The amount of fluorescence released is inversely proportional to the number of mtDNA lesions. All samples were run in duplicate. The between batch coefficient of variation for the current assay was 10%.

### **Statistical Analysis**

In this study, we evaluated the associations of short-term increases in mean and variability of temperature with number of mtDNA lesions per 10kb in blood. Since previous literatures often indicated J- or U-shape relationships when modeling temperature with health-related outcomes,<sup>112-114</sup> we examined the dose-response relationships between air temperature and mtDNA lesions using generalized additive mixed-effects models with penalized cubic splines. We observed deviations from linear dose-response relationships between increases in mean air temperature and mtDNA lesions, but not for increases in temperature variability and mtDNA lesions. We further tested for effect modification by the four categories of season (spring/summer/fall/winter) and observed significant effect modification by season when we modeled the relationship between mean temperature and mtDNA lesions. Within each seasonal category, however, we observed no deviations from linear dose-response. We therefore stratified

our analysis by season and modeled mean air temperature as a linear term in the primary analysis.

We fitted linear mixed-effects models with subject-specific intercepts to account for repeated measurements of mtDNA lesion from the same individual and examined the association of increases in mean and variability of air temperature and number of mtDNA lesions per 10kb. We modeled cumulative effects of air temperature as the 2-, 7- and 14-day moving average of 24-hour mean and 24-hour variability of temperature. We considered each moving average separately in the main analysis. Model estimates are expressed per interquartile range (IQR) increases in mean and variability of air temperature for the specific moving average.

In each model, we controlled for potential confounders and risk factors of mtDNA lesion: season (spring/summer/fall/winter), relative humidity (continuous), age (continuous), body mass index (BMI) (weight (kg)/height (m)<sup>2</sup>, continuous), race (white or others), regular patterns of physical activity (<12 hours/week, ≥12 and <30 hours/week, ≥30 hours/week), smoking status (never, former, or current smoker), cumulative pack-years of smoking (continuous), education level (high school diploma or less, college degree, or graduate degree), platelet, lymphocyte and neutrophil counts in the blood, and batch effects.

The main regression model took the general form:

$$Y_{ij} = \beta_0 + u_i + \beta_1 X_{1ij} + \dots + \beta_p X_{pij} + \beta_{\text{temperature}} \text{Temperature}_{ij} + \varepsilon_{ij}$$

where  $i$  indexes participants,  $j$  indexes visit number;  $\beta_0$  is the intercept for the population mean;  $u_i$  is the subject-specific random intercept. We assumed a linear effect of the moving

average of mean or variability temperature represented by  $\beta_{\text{temperature}} \text{Temperature}_{ij}$ . Finally,  $\beta_1 X_{1ij}$  to  $\beta_p X_{pij}$  correspond to the covariates we selected *a priori*.  $\epsilon_{ij}$  is the within-subject error term. We consider separate models for moving averages of mean temperature and variability temperature as well as a joint model that includes both a moving average of mean and variability temperature in the same model. Model estimates are expressed per interquartile range (IQR) increase in mean and variability of air temperature for the specific moving averages.

### **Sensitivity Analysis**

Among the 654 participants included in our analysis, 27% of participants did not come for a subsequent clinical visit. We therefore accounted for potential selection bias due to loss to follow-up using inverse probability of weighting. Specifically, we predict the probability of coming to a subsequent clinical visit using covariates from the previous visits in a logistic regression: age, BMI, regular patterns of physical activity, smoking status, pack year smoked, forced expiratory volume in one second (FEV1) versus forced vital capacity (FVC) ratio, clinically diagnosed hypertension, clinically diagnosed diabetes and medication (diuretics, and beta blocker).

In the main analysis, we obtained temperature measurements from the stationary monitoring station, which is located approximately 12km from the clinical site. This may introduce potential exposure misclassifications, as two participants came on the same day would be assigned to the same value. As a sensitivity analysis, we obtained temperature measurements from a spatial and temporal resolved model, and examined the association between increases in mean air temperature from the modeling data and number of mtDNA lesions per 10kb.

All analyses were conducted with SAS version 9.3 (SAS Institute Inc., Cary, NC) or, with R 3.0.1 (<http://www.r-project.org/>).

## **Results**

### **Study Population**

Table 3.1 shows participant characteristics at each visit. Of the 654 participants who had blood mtDNA lesion measurements, 481 participants came to more than one clinical visit (74%). Mean age at the first visit was  $73.1 \pm 6.8$  years, and mean BMI at the first visit was  $28.1 \pm 4.0$  kg/m<sup>2</sup> (mean  $\pm$  SD). Four percent of the participants were current smokers, and 66% were former smokers at the first visit. The study population was predominately white (97%).

### **Meteorological Data**

Table 3.2 describes the summary statistics and Spearman correlations of the meteorological variables. Twenty-four hour mean air temperature was  $13.0 \pm 8.6$  °C across the four different seasons, and was  $21.8 \pm 4.5$  °C and  $1.2 \pm 5.0$  °C (mean  $\pm$  SD), respectively, in summer and winter. Minimum 24-hour mean temperature in the winter was  $-13.9$  °C; while maximum 24-hour mean temperature in the summer was  $31.2$  °C. Twenty-four hour mean and 24-hour SD of temperature showed moderate correlation in spring and summer ( $\rho=0.33$  and  $0.36$ , respectively).

### **Air Temperature and Mitochondrial DNA Lesions**

Short-term increases in mean air temperature were associated with increased mtDNA lesions in blood in summer. For example, an IQR increase in 7-day ( $3.6$  °C) and 14-day ( $3.8$  °C) moving

average of 24-hour mean air temperature was associated with a 0.15 (95% CI: [0.05; 0.25],  $p=0.0033$ ), and 0.19 (95% CI: [0.07; 0.30],  $p=0.0011$ ) increase in number of mtDNA lesions per 10kb, respectively. Further, we also observed positive associations between increases in mean air temperature and mtDNA lesions in winter (Figure 3.1). For example, an IQR increase in 7-day (5.2 °C) and 14-day (4.2 °C) moving average of 24-hour mean temperature was associated with a 0.26 (95% CI: [0.08; 0.43],  $p=0.0037$ ), and 0.23 (95% CI: [0.08; 0.37],  $p=0.0019$ ) increase in number of mtDNA lesions per 10kb, respectively (Figure 3.1). Similar associations were also observed when we further adjusted for temperature variability (Figure 3.2). Short-term increases in temperature variability were also associated with higher number of mtDNA lesion independent of mean air temperature (Figure 3.3).

### **Sensitivity Analysis**

In the sensitivity analysis, we used inverse probability weighing to account for potential selection bias due to loss to follow up. Result estimates were similar compared with those from primary analysis (data not shown). In addition, we obtained temperature measurements from a spatial and temporal resolved model, and compared results from the modeling data to monitoring stations. Results were fairly comparable. For example, a 5°C increase in 7-day moving average of 24-hour mean air temperature was associated with a 0.21 (95% CI: [0.07; 0.35]) increase in number of mtDNA lesions per 10kb based on stationary monitoring station data; and a 5°C increase in 7-day moving average of 24-hour mean air temperature was associated with a 0.31 (95% CI: [0.50; 0.57]) increase in number of mtDNA lesions per 10kb from the spatiotemporal modeling approach.

## Discussion

In this analysis of the association between increases in air temperature and blood mitochondrial DNA lesions in a cohort of older individuals, we showed that higher mean temperature in the short-term was associated with increased blood mitochondrial oxidative damage—measured as the relative number of mtDNA lesions per 10kb. For example, an IQR increase in 7-day and 14-day moving average of 24-hour mean temperature was associated with a 0.15 (95% CI: [0.05; 0.25],  $p=0.0033$ ), and 0.19 (95% CI: [0.07; 0.30],  $p=0.0011$ ) increase in mtDNA lesions, respectively, in summer. We also observed positive associations between increases in mean air temperature and mtDNA lesions in winter. Results were similar when we further adjusted for within day temperature variability. Further, we also found significant associations between increased within day temperature variability and mtDNA lesions independent of mean air temperature. To our knowledge, this is the first study describing increases in air temperature and mitochondrial DNA damage.

The observed increase in mtDNA lesions in the summer ( $\sim 20$  °C and above) may be due to heat stress. In human, core body temperature is usually tightly regulated and maintained within a narrow range ( $\sim 37$  °C). When external temperature increases, thermoreceptors under the skin would sense such change in temperature and send information to the integrating center in brain via the afferent neural pathway. The integrating center, located in the hypothalamus would sense such temperature discrepancy between external environment and the set point of core body temperature, and send signals via the efferent neural pathway to dissipate heat.<sup>115</sup> There are both physiological and behavioral responses for heat dissipation. For example, increases in air temperature would impose large demand on the heart and the vasculature system, resulting in

blood jeopardizing from core organs to the body's surface for heat dissipation.<sup>115</sup> Animal studies have shown that heat stress increase metabolic demand, and reduce splanchnic blood flow, which in turn reduce visceral perfusion; the prolonged ischemic environment may results in the generation of reactive oxygen and nitrogen species, which may lead to visceral – and in turn – systemic oxidative stress.<sup>116,117</sup> Therefore, the observed increased mtDNA lesions in response to higher mean temperature during hot days may be reflecting systemic oxidative stress due to heat stress. Further, higher temperature in hot days may increase basal metabolic rate, which in turn lead to increased mitochondrial activity. Since mitochondria are the primary site of endogenous ROS generation, increased mitochondrial activity in hot days may also contribute to mtDNA lesions.

Interestingly, we also observed positive associations between increase in mean air temperature and mtDNA lesions at low temperature range (~ -10 to 0 °C). Previous literatures indicate J- or U-shape relationships between changes in mean air temperature and mortality.<sup>112</sup>

Epidemiological studies however reported inconsistent associations between changes in mean air temperature and blood biomarkers. For instance, Hampel *et al.* and Schauble *et al.* reported increases in blood inflammatory and coagulation biomarkers with decreased air temperature in patients with cardiopulmonary and metabolic diseases;<sup>96,98</sup> whereas in individuals with stable heart failure, Wilker *et al.* showed increased blood B-type natriuretic peptide and C-reactive protein levels with increased mean air temperature.<sup>95</sup> Nevertheless, these studies focused on diseased populations, and it is likely that individuals with chronic diseases, such as coronary heart disease (CHD) or diabetes spend less time outdoor at extreme or sub-optimal temperatures, which may result in potential exposure misclassifications. As sensitivity analysis, we excluded

participants with chronic disease status (CHD and diabetes) and we observed slightly stronger associations between increases in mean air temperature and mtDNA lesions. For instance, an IQR increase in 7-day and 14-day moving average of 24-hour mean temperature was associated with a 0.35 (95% CI: 0.09; 0.62), and 0.40 (95% CI: 0.18; 0.62) increase in mtDNA lesions, respectively. From the same cohort, Bind *et al.* found increased DNA methylation in several inflammatory and oxidative stress related genes with decreased air temperature.<sup>99</sup> DNA methylation in inflammatory genes have been shown to change rapidly after environmental insults for fine-tuning of inflammatory responses,<sup>58,59,118</sup> and it is well-known that increased methylation often silences gene expression.

The observed positive associations between increases in mean air temperature and mtDNA lesions in cold weather may have resulted from decreased metabolic rate, which in turn reduces mitochondrial activity. Mitochondria are the primary sites for cellular energy production, as well as endogenous ROS generation.<sup>119</sup> Since ROS are the byproducts of ETC where glucose are metabolized to cellular energy of adenosine triphosphate (ATP),<sup>119</sup> a decrease in ROS generation due to reduced mitochondrial activity may be coupled with reduced ATP production. Hence, although we observed decreased blood mtDNA lesions as mean air temperature decreases in cold weather, other cellular processes that are ATP-dependent—such as vascular control—may also be depressed.<sup>120</sup> Impaired vascular control may be particularly of a problem to the elderly population.

Our study participants consist mainly of old adults (mean age = 73.1 years at first visit). In healthy young adults, increase in temperature may result in climate adaptation, which may



involve enhanced cardiovascular performance, activation of the renin-angiotensin axis and increased capacity to secrete sweats, and expansion of plasma volume.<sup>115,121</sup> However, older individuals are less adaptive to changes in air temperature, due to their impaired homeostatic control mechanisms (i.e., impaired baroreceptor reflex mechanism,<sup>122</sup> lower sweating rate,<sup>123</sup> and diminished renal sympathetic nerve activity<sup>124</sup>). As a result, increase in air temperature would disproportionately affect elderly individuals. Our study highlights the relevance of investigating temperature influences in such susceptible population.

Our study has a number of strengths. Our assay for analyzing mtDNA lesion yields high precision: the between batch coefficient of variation was 10%. Further, we accounted for potential selection bias due to loss to follow-up using inverse probability of weighting (IPW). Point estimates were similar for models using or not using IPW, indicating that little selection bias was introduced due to loss to follow-up.

Our study has several limitations. In the primary analysis, we obtained air temperature data from fixed monitoring station sites located approximately 12km away from the clinical site. This may introduce potential exposure misclassifications. In the sensitivity analysis, we estimated air temperature at each participant's residential address for each calendar day using a hybrid land-use regression and satellite-based model, and examined the association between increases in mean air temperature and mtDNA lesions. Temperature estimates from our modeling approach would introduce sufficient spatial and temporal variability, and minimize exposure misclassifications. Results were similar when we compared estimates from stationary monitoring stations and modeling approach. Our exposure matrix is ambient air temperature, however, we

have no indoor measures and no information about how much time each participant spent outdoor. Therefore, our study may suffer from potential exposure misclassifications, as elderly tend to stay indoor during hot and cold days, where air conditioning is available. In addition, since our study is based in Boston Massachusetts, the ambient air temperature range in our study may not be generalizable to other climate zones.

### **Conclusion**

In conclusion, we observed short-term exposures to higher mean and variability of air temperature with increased mtDNA lesions in elderly adults, supporting the hypothesis that changes meteorological conditions may induce molecular mitochondrial alterations in older individuals. The potential health consequences associated with increased mtDNA lesions must be further elucidated.

### **ACKNOWLEDGEMENTS**

The authors would like to thank all Normative Aging Study participants.

### **FUNDINGS**

This work was supported by National Institutes of Health [R01ES021733, R01ES015172, RO1ES025225, U34AG051918, P30ES009089 and T32ES007142]; U.S. Environmental Protection Agency [RD-83479801]; and the Department of Agriculture, Agricultural Research Service [contract 53-K06-510]. The VA Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Center of the U.S. Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and

Information Center, Boston, Massachusetts. The views expressed in this paper are those of the authors and do not necessarily represent the views of the US Department of Veterans Affairs.

## **DISCLOSURES**

None of the authors has any actual or potential competing financial interests.

**Table 3.1.** Personal characteristics of participants (N=654).

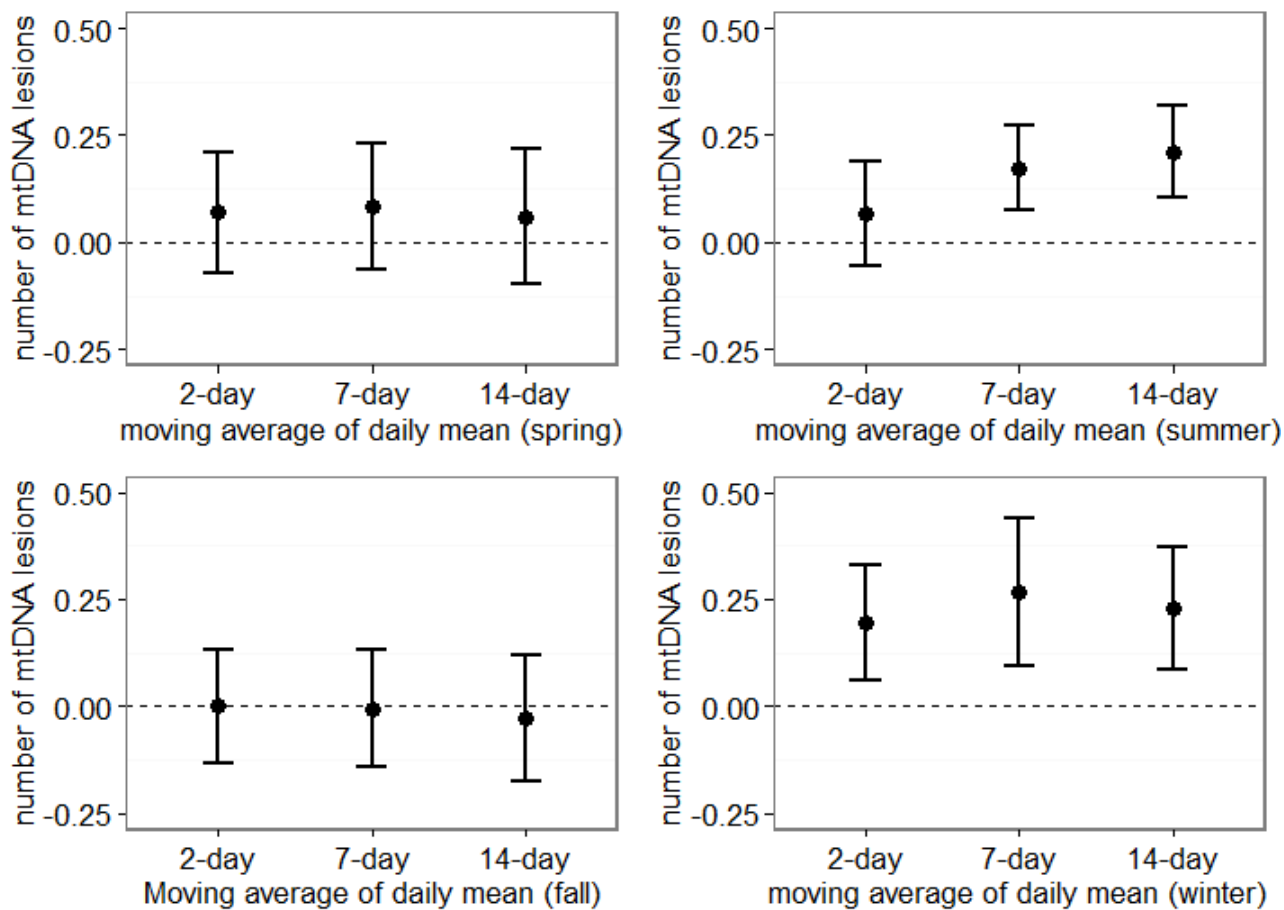
Variable	Visit 1	Visit 2	Visit 3 and more
	(n = 654)	(n = 387)	(n = 94)
Age (years), mean $\pm$ SD	73.1 $\pm$ 6.8	76.1 $\pm$ 6.6	79.8 $\pm$ 6.2
BMI (kg/m <sup>2</sup> ), mean $\pm$ SD	28.1 $\pm$ 4.0	27.9 $\pm$ 4.2	28.1 $\pm$ 4.2
Smoking status, N (%)			
Never	196 (30%)	121 (32%)	31 (33%)
Former	432 (66%)	249 (65%)	60 (65%)
Current	26 (4%)	14 (4%)	2 (2%)
Pack years, mean $\pm$ SD	19.9 $\pm$ 24.2	20.3 $\pm$ 24.7	21.0 $\pm$ 25.0
Race, N (%)			
White	631 (97%)	372 (97%)	91 (97%)
Other	17 (3%)	11 (3%)	3 (3%)
Metabolic Equivalent of Task, N (%)			
Low ( $\leq$ 12 hours/week)	412 (63%)	247 (66%)	66 (70%)
Medium (12-30 hours/week)	145 (22%)	82 (21%)	13 (14%)
High ( $\geq$ 30 hours/week)	95 (15%)	48 (12%)	15 (16%)
Two or more drinks per day, N (%)	133 (20%)	61 (16%)	14 (15%)
Education, N (%)			
< 12 years	216 (33%)	128 (33%)	33 (35%)
13-16 years	306 (47%)	180 (47%)	49 (52%)
> 16 years	132 (20%)	79 (20%)	12 (13%)
Season, N (%)			
Spring (March-May)	156 (24%)	95 (25%)	27 (29%)
Summer (June-August)	187 (29%)	104 (27%)	26 (28%)
Fall (September-November)	214 (33%)	124 (32%)	24 (26%)
Winter (December-February)	97 (15%)	63 (16%)	17 (18%)
Lymphocytes count [cells/cm <sup>3</sup> (mean $\pm$ SD)]	25.9 $\pm$ 8.7	24.7 $\pm$ 7.8	23.6 $\pm$ 7.5
Neutrophils count [cells/cm <sup>3</sup> (mean $\pm$ SD)]	61.7 $\pm$ 9.0	62.9 $\pm$ 8.5	64.1 $\pm$ 8.2
Platelets [1000/mm <sup>3</sup> (mean $\pm$ SD)]	220.6 $\pm$ 55.7	220.8 $\pm$ 54.9	207.0 $\pm$ 57.5
mtDNA lesions, mean $\pm$ SD	1.9 $\pm$ 0.9	1.4 $\pm$ 0.9	1.8 $\pm$ 0.9

**Table 3.2.** Summary statistics and Spearman correlation of environmental variables.

	Summary statistics		Spearman correlation	
	Mean $\pm$ SD	Median (p25, p75)	24-hr T <sub>MEAN</sub>	24-hr T <sub>SD</sub>
Overall (n = 1141)				
24-hr T <sub>MEAN</sub> (°C)	13.0 $\pm$ 8.6	13.3 (6.7, 20.0)	1.00	0.33**
24-hr T <sub>SD</sub> (°C)	2.4 $\pm$ 1.2	2.2 (1.5, 3.1)	-	1.00
Spring (n = 281)				
24-hr T <sub>MEAN</sub> (°C)	9.8 $\pm$ 6.1	10.2 (5.8, 13.1)	1.00	0.36**
24-hr T <sub>SD</sub> (°C)	2.4 $\pm$ 1.3	2.1 (1.5, 2.9)	-	1.00
Summer (n = 318)				
24-hr T <sub>MEAN</sub> (°C)	21.8 $\pm$ 4.5	22.1 (19.0, 24.8)	1.00	0.44**
24-hr T <sub>SD</sub> (°C)	2.6 $\pm$ 1.2	2.6 (1.7, 3.3)	-	1.00
Fall (n = 365)				
24-hr T <sub>MEAN</sub> (°C)	13.5 $\pm$ 5.6	13.7 (9.2, 17.7)	1.00	0.27**
24-hr T <sub>SD</sub> (°C)	2.3 $\pm$ 1.0	2.1 (1.4, 3.0)	-	1.00
Winter (n = 177)				
24-hr T <sub>MEAN</sub> (°C)	1.2 $\pm$ 5.0	1.6 (-2.1, 4.0)	1.00	0.02
24-hr T <sub>SD</sub> (°C)	2.2 $\pm$ 1.3	1.9 (1.4, 2.6)	-	1.00

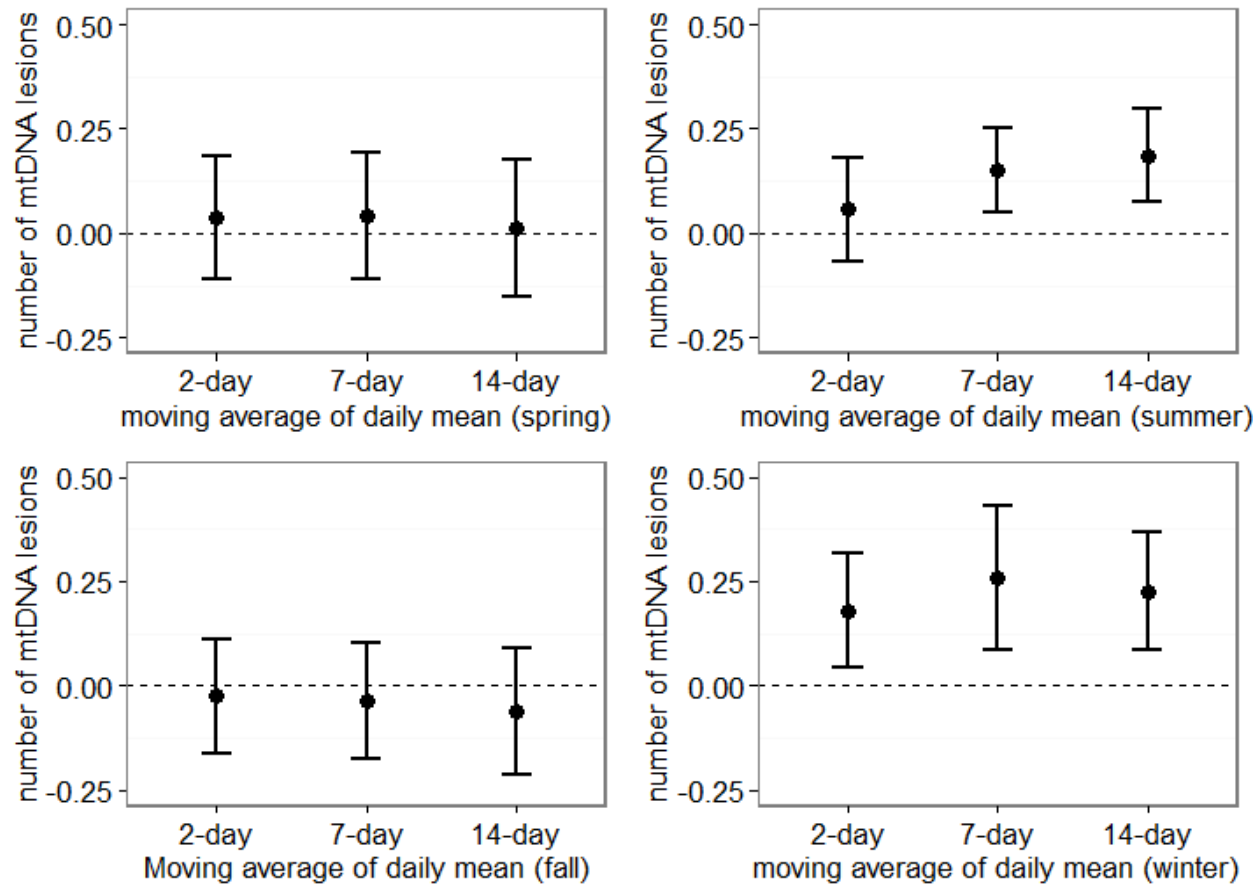
\*  $p < 0.05$ ; \*\*  $p < 0.01$

**Figure 3.1** Estimated change (and 95% CI) in mtDNA lesions per IQR increase of daily mean of air temperature over 14-day moving average (by season).



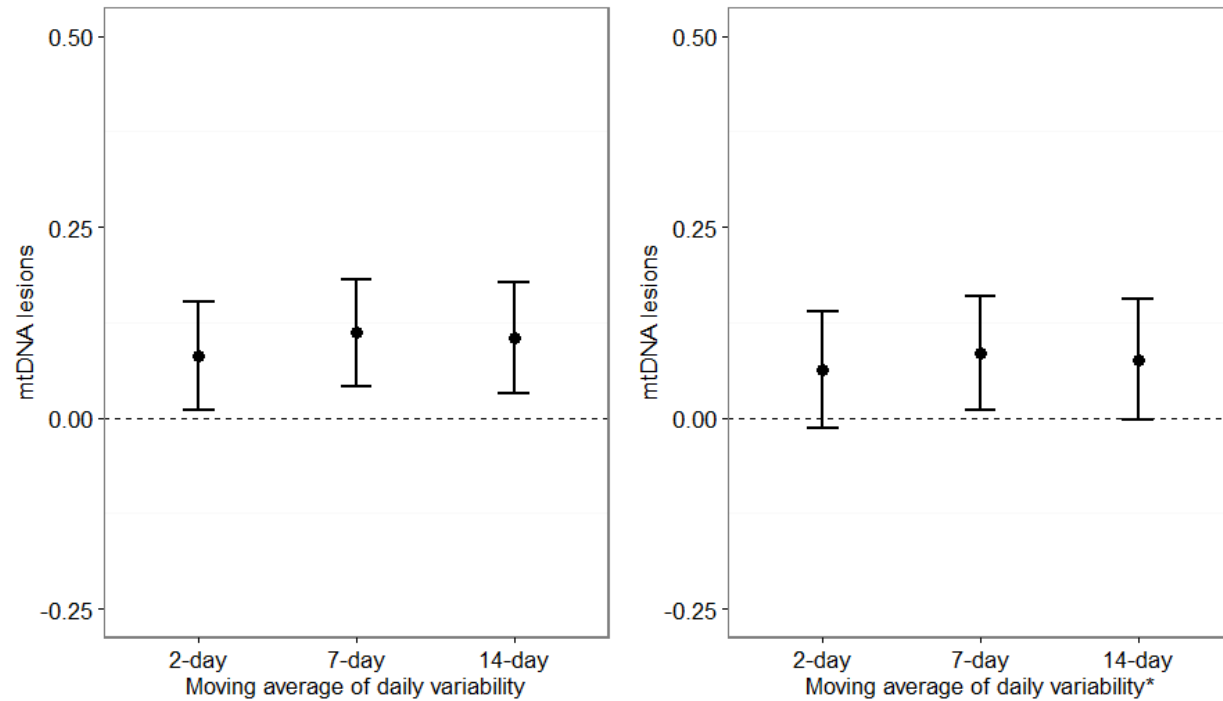
Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, metabolic equivalent of tasks (low:  $\leq 12$  hours/week; medium: 12-30 hours/week; high:  $\geq 30$  hours/week), smoking status (current, former, never), pack-years smoked, alcohol consumption, education level, platelet count, lymphocyte count, neutrophil count, plates, and 24-hr relative humidity.

**Figure 3.2** Estimated change (and 95% CI) in mtDNA lesions per IQR increase of daily mean of air temperature over 14-day moving average (by season). We additionally adjusted for within day temperature variability.



Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, metabolic equivalent of tasks (low:  $\leq 12$  hours/week; medium: 12-30 hours/week; high:  $\geq 30$  hours/week), smoking status (current, former, never), pack-years smoked, alcohol consumption, education level, platelet count, lymphocyte count, neutrophil count, plates, 24-hr relative humidity, and within day temperature variability of the corresponding moving average.

**Figure 3.3** Estimated change (and 95% CI) in mtDNA lesions per IQR increase of daily variability of air temperature over 14-day moving average.



Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, metabolic equivalent of tasks (low:  $\leq 12$  hours/week; medium: 12-30 hours/week; high:  $\geq 30$  hours/week), smoking status (current, former, never), pack-years smoked, alcohol consumption, education level, platelet count, lymphocyte count, neutrophil count, plates, 24-hr relative humidity and season (spring/summer/fall/winter).

\*We additionally adjusted for within day temperature mean of the corresponding moving average.



## REFERENCES

1. Braga AL, Zanobetti A, Schwartz J. The effect of weather on respiratory and cardiovascular deaths in 12 U.S. cities. *Environ Health Perspect* 2002;**110**(9):859-63.
2. Wolf K, Schneider A, Breitner S, von Klot S, Meisinger C, Cyrus J, Hymer H, Wichmann HE, Peters A, Cooperative Health Research in the Region of Augsburg Study G. Air temperature and the occurrence of myocardial infarction in Augsburg, Germany. *Circulation* 2009;**120**(9):735-42.
3. Schwartz J, Samet JM, Patz JA. Hospital admissions for heart disease: the effects of temperature and humidity. *Epidemiology* 2004;**15**(6):755-61.
4. Koken PJ, Piver WT, Ye F, Elixhauser A, Olsen LM, Portier CJ. Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver. *Environ Health Perspect* 2003;**111**(10):1312-7.
5. Schneider A, Panagiotakos D, Picciotto S, Katsouyanni K, Lowel H, Jacquemin B, Lanki T, Stafoggia M, Bellander T, Koenig W, Peters A, Group AS. Air temperature and inflammatory responses in myocardial infarction survivors. *Epidemiology* 2008;**19**(3):391-400.
6. Kim Y, Joh S. A vulnerability study of the low-income elderly in the context of high temperature and mortality in Seoul, Korea. *Sci Total Environ* 2006;**371**(1-3):82-8.
7. Barnett AG, Sans S, Salomaa V, Kuulasmaa K, Dobson AJ, Project WM. The effect of temperature on systolic blood pressure. *Blood Press Monit* 2007;**12**(3):195-203.
8. Alperovitch A, Lacombe JM, Hanon O, Dartigues JF, Ritchie K, Ducimetiere P, Tzourio C. Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: the Three-City study. *Arch Intern Med* 2009;**169**(1):75-80.

9. Ren C, O'Neill MS, Park SK, Sparrow D, Vokonas P, Schwartz J. Ambient temperature, air pollution, and heart rate variability in an aging population. *Am J Epidemiol* 2011;**173**(9):1013-21.
10. Halonen JJ, Zanobetti A, Sparrow D, Vokonas PS, Schwartz J. Relationship between outdoor temperature and blood pressure. *Occup Environ Med* 2011;**68**(4):296-301.
11. Wilker EH, Yeh G, Wellenius GA, Davis RB, Phillips RS, Mittleman MA. Ambient temperature and biomarkers of heart failure: a repeated measures analysis. *Environ Health Perspect* 2012;**120**(8):1083-7.
12. Schauble CL, Hampel R, Breitner S, Ruckerl R, Phipps R, Diaz-Sanchez D, Devlin RB, Carter JD, Soukup J, Silbajoris R, Dailey L, Koenig W, Cyrus J, Gerschkat U, Belcredi P, Kraus U, Peters A, Schneider AE. Short-term effects of air temperature on blood markers of coagulation and inflammation in potentially susceptible individuals. *Occup Environ Med* 2012;**69**(9):670-8.
13. Keatinge WR, Coleshaw SR, Cotter F, Mattock M, Murphy M, Chelliah R. Increases in platelet and red cell counts, blood viscosity, and arterial pressure during mild surface cooling: factors in mortality from coronary and cerebral thrombosis in winter. *Br Med J (Clin Res Ed)* 1984;**289**(6456):1405-8.
14. Hampel R, Breitner S, Ruckerl R, Frampton MW, Koenig W, Phipps RP, Wichmann HE, Peters A, Schneider A. Air temperature and inflammatory and coagulation responses in men with coronary or pulmonary disease during the winter season. *Occup Environ Med* 2010;**67**(6):408-16.

15. Bind MA, Zanobetti A, Gasparri A, Peters A, Coull B, Baccarelli A, Tarantini L, Koutrakis P, Vokonas P, Schwartz J. Effects of temperature and relative humidity on DNA methylation. *Epidemiology* 2014;**25**(4):561-9.
16. Dhalla NS, Temsah RM, Netticadan T. Role of oxidative stress in cardiovascular diseases. *J Hypertens* 2000;**18**(6):655-73.
17. Yakes FM, Van Houten B. Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc Natl Acad Sci U S A* 1997;**94**(2):514-9.
18. Zastawny TH, Dabrowska M, Jaskolski T, Klimarczyk M, Kulinski L, Koszela A, Szczesniewicz M, Sliwinska M, Witkowski P, Olinski R. Comparison of oxidative base damage in mitochondrial and nuclear DNA. *Free Radic Biol Med* 1998;**24**(5):722-5.
19. Mengel-From J, Thinggaard M, Dalgard C, Kyvik KO, Christensen K, Christiansen L. Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Hum Genet* 2014;**133**(9):1149-59.
20. Friedman JR, Nunnari J. Mitochondrial form and function. *Nature* 2014;**505**(7483):335-43.
21. Demple B, Harrison L. Repair of oxidative damage to DNA: enzymology and biology. *Annu Rev Biochem* 1994;**63**:915-48.
22. Corral-Debrinski M, Stepien G, Shoffner JM, Lott MT, Kanter K, Wallace DC. Hypoxemia is associated with mitochondrial DNA damage and gene induction. Implications for cardiac disease. *JAMA* 1991;**266**(13):1812-6.

23. Lesnefsky EJ, Moghaddas S, Tandler B, Kerner J, Hoppel CL. Mitochondrial dysfunction in cardiac disease: ischemia--reperfusion, aging, and heart failure. *J Mol Cell Cardiol* 2001;**33**(6):1065-89.
24. Corral-Debrinski M, Shoffner JM, Lott MT, Wallace DC. Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease. *Mutat Res* 1992;**275**(3-6):169-80.
25. Power MC, Weisskopf MG, Alexeeff SE, Coull BA, Spiro A, 3rd, Schwartz J. Traffic-related air pollution and cognitive function in a cohort of older men. *Environ Health Perspect* 2011;**119**(5):682-7.
26. Mehta AJ, Kloog I, Zanobetti A, Coull BA, Sparrow D, Vokonas P, Schwartz J. Associations between changes in city and address specific temperature and QT interval--the VA Normative Aging Study. *PLoS One* 2014;**9**(9):e106258.
27. Kloog I, Chudnovsky A, Koutrakis P, Schwartz J. Temporal and spatial assessments of minimum air temperature using satellite surface temperature measurements in Massachusetts, USA. *Sci Total Environ* 2012;**432**:85-92.
28. Kloog I, Nordio, F., Coull, B. A., Schwartz, J. . Predicting spatiotemporal mean air temperature using MODIS satellite surface temperature measurements across the Northeastern USA. *Remote Sens. Environ.* 2014(150):132–139.
29. Furda A, Santos JH, Meyer JN, Van Houten B. Quantitative PCR-based measurement of nuclear and mitochondrial DNA damage and repair in mammalian cells. *Methods Mol Biol* 2014;**1105**:419-37.

30. Ponti M, Forrow SM, Souhami RL, D'Incalci M, Hartley JA. Measurement of the sequence specificity of covalent DNA modification by antineoplastic agents using Taq DNA polymerase. *Nucleic Acids Res* 1991;**19**(11):2929-33.
31. Gasparrini A, Guo Y, Hashizume M, Lavigne E, Zanobetti A, Schwartz J, Tobias A, Tong S, Rocklov J, Forsberg B, Leone M, De Sario M, Bell ML, Guo YL, Wu CF, Kan H, Yi SM, de Sousa Zanotti Stagliorio Coelho M, Saldiva PH, Honda Y, Kim H, Armstrong B. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet* 2015;**386**(9991):369-75.
32. Yang J, Yin P, Zhou M, Ou CQ, Li M, Li J, Liu X, Gao J, Liu Y, Qin R, Xu L, Huang C, Liu Q. The burden of stroke mortality attributable to cold and hot ambient temperatures: Epidemiological evidence from China. *Environ Int* 2016;**92-93**:232-238.
33. O'Neill MS, Ebi KL. Temperature extremes and health: impacts of climate variability and change in the United States. *J Occup Environ Med* 2009;**51**(1):13-25.
34. Leon LR, Helwig BG. Heat stroke: role of the systemic inflammatory response. *J Appl Physiol (1985)* 2010;**109**(6):1980-8.
35. Hall DM, Baumgardner KR, Oberley TD, Gisolfi CV. Splanchnic tissues undergo hypoxic stress during whole body hyperthermia. *Am J Physiol* 1999;**276**(5 Pt 1):G1195-203.
36. Hall DM, Buettner GR, Oberley LW, Xu L, Matthes RD, Gisolfi CV. Mechanisms of circulatory and intestinal barrier dysfunction during whole body hyperthermia. *Am J Physiol Heart Circ Physiol* 2001;**280**(2):H509-21.
37. Peng C, Bind MC, Colicino E, Kloog I, Byun HM, Cantone L, Trevisi L, Zhong J, Brennan K, Dereix AE, Vokonas PS, Coull BA, Schwartz JD, Baccarelli AA. Particulate

- Air Pollution and Fasting Blood Glucose in Non-Diabetic Individuals: Associations and Epigenetic Mediation in the Normative Aging Study, 2000-2011. *Environ Health Perspect* 2016.
38. Hou L, Zhang X, Zheng Y, Wang S, Dou C, Guo L, Byun HM, Motta V, McCracken J, Diaz A, Kang CM, Koutrakis P, Bertazzi PA, Li J, Schwartz J, Baccarelli AA. Altered methylation in tandem repeat element and elemental component levels in inhalable air particles. *Environ Mol Mutagen* 2014;**55**(3):256-65.
39. Tarantini L, Bonzini M, Apostoli P, Pegoraro V, Bollati V, Marinelli B, Cantone L, Rizzo G, Hou L, Schwartz J, Bertazzi PA, Baccarelli A. Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. *Environ Health Perspect* 2009;**117**(2):217-22.
40. Murphy MP. How mitochondria produce reactive oxygen species. *Biochem J* 2009;**417**(1):1-13.
41. Crecelius AR, Kirby BS, Dinunno FA. Intravascular ATP and the regulation of blood flow and oxygen delivery in humans. *Exerc Sport Sci Rev* 2015;**43**(1):5-13.
42. Bouchama A, Knochel JP. Heat stroke. *N Engl J Med* 2002;**346**(25):1978-88.
43. Stauss HM, Morgan DA, Anderson KE, Massett MP, Kregel KC. Modulation of baroreflex sensitivity and spectral power of blood pressure by heat stress and aging. *Am J Physiol* 1997;**272**(2 Pt 2):H776-84.
44. Inoue Y, Shibasaki M, Ueda H, Ishizashi H. Mechanisms underlying the age-related decrement in the human sweating response. *Eur J Appl Physiol Occup Physiol* 1999;**79**(2):121-6.

45. Kenney MJ, Fels RJ. Sympathetic nerve regulation to heating is altered in senescent rats.  
*Am J Physiol Regul Integr Comp Physiol* 2002;**283**(2):R513-20.

## SUMMARY AND CONCLUSIONS

The goal of this dissertation was to evaluate the use of epigenetic and mitochondrial biomarkers linking environmental stressors (i.e. PM<sub>2.5</sub> exposure and changes in ambient air temperature) and health in the field of Environmental Epidemiology. Our study leveraged an existing prospective longitudinal cohort – The Normative Aging Study. Through applying advanced statistical methods, such as causal mediation analysis and dimension reduction methods (adaptive LASSO), we sought to address three public health questions: (1) the potential mediating role of DNA methylation of inflammatory cytokines linking PM<sub>2.5</sub> exposure and health outcomes; (2) the relative oxidation potency of PM<sub>2.5</sub> components independent of PM<sub>2.5</sub> mass concentration; (3) short-term changes in mean ambient air temperature and oxidative damage.

The results from Chapter 1 showed that PM<sub>2.5</sub> concentrations were associated with higher FBG level, and this association was in part mediated through *ICAM-1* gene methylation, particularly at the longer (28-day) moving average investigated. Our study demonstrates a novel approach of mediation analysis in epigenetic studies and highlights a mediating role of *ICAM-1* gene methylation in air-pollution associated abnormal glucose metabolism. While the proportion mediated by *ICAM-1* methylation alone is relatively modest, methylation of other genes not investigated in this study, independently or in combination with *ICAM-1* methylation, may mediate larger proportions of PM<sub>2.5</sub> effects. Future epigenome-wide studies are needed to determine the extent to which DNA methylation contributes to mediate environmental effects on human metabolism.

In Chapter 2, our study showed that long-term exposure to PM<sub>2.5</sub> mass and specific PM<sub>2.5</sub> components were associated with decreased mtDNA abundance in this elderly study population.



We first fit single- and two-pollutant models to estimate the associations between PM<sub>2.5</sub> components with mtDNA abundance. We included in the two-pollutant models PM<sub>2.5</sub> to avoid potential confounding by mass concentration. Further, we applied a multi-pollutant approach and used dimension reduction method *a priori* to reduce potential multi-collinearity issues while adjusting for exposure to other components. We found that NO<sub>3</sub><sup>-</sup> concentration in PM<sub>2.5</sub> negatively associated with mtDNA abundance in multi-pollutant modeling, suggesting that the oxidative stress potential of PM<sub>2.5</sub> may not be fully captured by mass concentration.

In Chapter 3, we observed short-term exposures to higher mean air temperature with increased mtDNA lesions in elderly adults, supporting the hypothesis that changes meteorological conditions may induce pathophysiological responses among susceptible subgroups—such as the elderly population—which represents a group of unacclimatized and immunocompromised individuals. Nevertheless, the potential health consequences of increased mtDNA lesions must be further elucidated.

Taken together, our study highlight the use of two sets of novel molecular biomarkers (DNA methylation at the promoter region of inflammatory cytokines and oxidative stress-related mtDNA integrity) in environmental epidemiology studies. Since inflammation and oxidative stress are two of the plausible pathways linking environmental challenges and health, this dissertation work may span the scope of commonly used biomarkers today, and help to advance our current understanding of body's response to external environmental stressors.