Genome-wide DNA methylation and long-term ambient air pollution exposure in Korean adults

Mi Kyeong Lee, Cheng-Jian Xu, Megan U. Carnes, Cody E. Nichols, James M. Ward, The BIOS consortium, Sung Ok Kwon, Sun-Young Kim, Woo Jin Kim and Stephanie J. London

Abstract

Background: Ambient air pollution is associated with numerous adverse health outcomes, but the underlying mechanisms are not well understood; epigenetic effects including altered DNA methylation could play a role. To evaluate associations of long-term air pollution exposure with DNA methylation in blood, we conducted an epigenome-wide association study in a Korean chronic obstructive pulmonary disease cohort (N = 100 including 60 cases) using Illumina’s Infinium HumanMethylation450K Beadchip. Annual average concentrations of particulate matter ≤ 10 μm in diameter (PM10) and nitrogen dioxide (NO2) were estimated at participants’ residential addresses using exposure prediction models. We used robust linear regression to identify differentially methylated probes (DMPs) and two different approaches, DMRcate and comb-p, to identify differentially methylated regions (DMRs).

Results: After multiple testing correction (false discovery rate < 0.05), there were 12 DMPs and 27 DMRs associated with PM10 and 45 DMPs and 57 DMRs related to NO2. DMP cg06992688 (OTUB2) and several DMRs were associated with both exposures. Eleven DMPs in relation to NO2 confirmed previous findings in Europeans; the remainder were novel. Methylation levels of 39 DMPs were associated with expression levels of nearby genes in a separate dataset of 3075 individuals. Enriched networks were related to outcomes associated with air pollution including cardiovascular and respiratory diseases as well as inflammatory and immune responses.

Conclusions: This study provides evidence that long-term ambient air pollution exposure impacts DNA methylation. The differential methylation signals can serve as potential air pollution biomarkers. These results may help better understand the influences of ambient air pollution on human health.

Keywords: Air pollution, Particulate matter, Nitrogen dioxide, Epigenesis, genetic, Epigenomics

© The Author(s). 2019 Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Background
Exposure to ambient air pollution has well-documented adverse effects on health outcomes, including cardiovascular disease [1] and pulmonary function [2]. Oxidative stress and inflammation have been suggested as underlying mechanisms but specific data supporting these links are lacking. Despite mounting evidence of the negative impacts of air pollution exposure on health outcomes, the underlying mechanisms are not well understood.

DNA methylation, an epigenetic modification that can influence gene expression, has widely replicated genome-wide associations with smoking [3]. While there are fewer data, there is evidence that ambient air pollution influences methylation [4–7]. Most studies of long-term air pollution exposure and methylation have been conducted in Caucasian adult populations [5–7] and evidence for replication of differentially methylated probes (DMPs) across studies or different ethnic groups is sparse.

We performed an epigenome-wide association study (EWAS) to evaluate the relationship of long-term exposure to particulate matter ≤ 10 μm in diameter (PM10) and nitrogen dioxide (NO2) with blood DNA methylation in adults (N = 100) participating in a Korean chronic obstructive pulmonary disease (COPD) cohort. We identified differentially methylated signals in relation to air pollution exposure both at an individual C–phosphate–G (CpG) probe level and at a regional level involving several neighboring CpG probes (CpGs). We evaluated whether methylation levels of our DMPs were associated with expression levels of nearby transcripts in a large independent dataset with matched gene expression and DNA methylation in the same individuals, Biobank#based integrative omics studies (BIOS) consortium. We also replicated findings from earlier EWASes in European populations, reporting a list of DMPs showing similar associations in our Asian population.

Methods
Study population
For DNA methylation profiling, study participants (N = 100 including 60 individuals with COPD) were sampled from a Korean COPD cohort [8]. Data and biologic specimens collected at a baseline visit (between late August and early November in 2012 and 2013) were used in this study. Blood and urine samples as well as survey questionnaires were obtained for all study participants who also underwent physical examination for anthropometric measurements. A trained nurse measured height and weight using the body composition analyzer I1O 353 (Aarna Systems, Udaipur, India). Body mass index (BMI) was calculated as weight (kg) divided by height squared (m²). Information on cigarette smoking status (never, former, and current) and pack-years of smoking was obtained via questionnaires. We calculated pack-years of smoking, for current and former smokers, by multiplying the number of years smoked by the number of cigarette packs smoked per day. Current nonsmoking status was validated using urine cotinine levels (nmol/L) measured by immunoassay (Immule 2000 Xpi; Siemens Healthcare Diagnostics, Tarrytown, NY, USA). The study protocol was approved by the Institutional Review Board at Kangwon National University. We obtained informed consent from all study participants.

Air pollution exposure at residential addresses
We estimated annual average concentrations of PM10 (μg/m³) and NO2 (ppb) at each residential address obtained from the baseline survey using a national-scale exposure prediction model [9]. Using air pollution regulatory monitoring data in 2010, the prediction model estimated the annual average concentrations of the pollutants in a universal kriging framework based on geographic predictors and spatial correlation. Geographic predictors were estimated by hundreds of geographic variables that represent pollution sources including traffic, demographic characteristics, land use, physical geography, transportation facilities, emissions, vegetation, and altitude. To account for season in the prediction model, we used several inclusion criteria for monitoring sites: (1) having more than 75% (274 days) of daily data, (2) having at least one daily measurement in each of the 10 months, and (3) having no more than 45 consecutive days without daily measurements. Participants’ residential addresses at the baseline visit were geocoded using GeoCoder-Xr software (Geoservice, Seoul, South Korea).

DNA methylation profiling
DNA was extracted from blood samples collected at the baseline visit. We obtained genome-wide methylation profiles using the Infinium HumanMethylation450K BeadChip (Illumina, Inc., San Diego, CA, USA). We used a pipeline implemented in the chip analysis methylation pipeline (ChAMP) R package [10] for signal extraction and initial low-quality probe filtering, excluding probes having a detection p value > 0.01 in any sample or a bead-count < 3 in 5% or more samples. Correction for probe design bias was done using Beta Mixture Quantile dilution normalization [11]. Batch effects were corrected using Combat [12] in the sva R package [13]. To minimize false positive findings, we additionally removed non-CpG probes and probes reported to be non-specific [14, 15] or potentially influenced by nearby single-nucleotide variants [14]. We provide probe filtering steps in Additional file 2: Table S1. After excluding probes on the X and Y chromosomes, the remaining
402,508 CpGs were used for association analyses. To reduce the potential influence of extreme methylation outliers on association results, we removed methylation values more extreme than Tukey's outer fences [16] defined as more than three times the interquartile range from the 25th and 75th percentiles of methylation values at each probe, resulting in removal of 75,549 (0.19%) values across all participants. To estimate cell-type proportions including CD8+ T lymphocytes, CD4+ T lymphocytes, natural killer cells, B cells, monocytes, and granulocytes, we applied Houseman's algorithm [17] with the Reinius reference panel [18] using the minfi R package [19].

**Identification of differentially methylated probes**

To evaluate associations of air pollution exposure with DNA methylation, we used robust linear regression models to decrease the influence of outlier methylation values and heteroskedasticity on association results [20]. Annual average concentrations of a pollutant (PM$_{10}$ or NO$_2$) were used as the predictor and the methylation beta values were the response variable. A methylation beta value is a ratio of methylated CpG probe intensity to total probe intensity and ranges between 0 (unmethylated) and 1 (methylated). Covariates included were age (years), sex (male, female), cigarette smoking (never, former, current), pack-years of smoking, BMI (kg/m$^2$), COPD status (cases, noncases), and estimated cell-type proportions. For genome-wide statistical significance, we set a threshold of Benjamini-Hochberg false discovery rate (FDR) adjusted $p$ value < 0.05 unless otherwise noted. We also used $p$ value < 1.2E-07 = (0.05/402,508) as a cutoff for statistically significant associations after Bonferroni correction. We used R version 3.0.2 for preprocessing methylation data from raw data (.idat files) to methylation beta values and R version 3.4.0 for association analyses and visualization of differential methylation regions.

**Identification of differentially methylated regions**

In addition to association analyses at individual CpGs, we applied two different methods to identify differential DNA methylation at the regional level in relation to air pollution exposure: DMRcate [21] and comb-p [22]. As the two methods implement different algorithms to identify differentially methylated regions (DMRs), we used both methods to find significant DMRs while reducing false positives. DMRcate uses a tunable kernel smoothing process with differential methylation association signals, whereas comb-p examines regional clustering of low $p$ values from irregularly spaced $p$ values. We used the “dmrcate” function in the DMRcate R package with input files from the epigenome-wide association results: regression coefficients, standard deviations, and uncorrected $p$ values. Comb-p, a stand-alone software, was used with input files containing uncorrected $p$ values and information on chromosomal locations (chromosome and physical position). To define significant DMRs in our study, we applied the following three criteria. First, more than one CpG should reside within a DMR. Second, regional differential methylation signals can be calculated using neighboring CpGs within 1000 base pairs (bp). Third, a region must have multiple-testing corrected $p$ value < 0.05 in both methods: Benjamini-Hochberg FDR for DMRcate and Sidak for comb-p. The use of FDR for DMRcate and Sidak for comb-p was the default setting in the two methods. As the minimum number of CpGs ($N = 2$) in a region and the minimum length of a distance ($N = 1000$ nucleotides) were the defaults in DMRcate, we used the same values for comb-p to harmonize results from the two methods. As the two methods call DMRs based on association results of neighboring probes, a significant DMR does not necessarily overlap a significant differentially methylated probe (DMP) in that region (Additional file 2: Table S2 and S3). To visualize regions of differential methylation, we used the coMET R package [23].

**Biological implications of association results**

Gene annotation for each CpG was done by using the manufacturer’s annotation file [24]; the UCSC RefGene names were obtained. For biological implications of our differential methylation signals in relation to each pollutant (PM$_{10}$ or NO$_2$), we explored curated variant annotations in the GeneticsLand software (OmicSoft, QIAGEN, NC, USA) and performed functional pathway analyses using the “Core Analysis” of ingenuity pathway analysis (IPA; Ingenuity Systems, QIAGEN, CA, USA) on genes annotated to DMPs with an uncorrected $p$ value < 1E-04 (an arbitrary cutoff for suggestive association) or significant DMRs. To assess enrichment of tissue- or cell type-specific signals, we analyzed DMPs (FDR < 0.05) and probes having the minimum $p$ value in each DMR for overlap with DNase1 hypersensitivity sites (DHSs) using the experimentally derived functional element overlap analysis of ReGions from EWAS (eFORGE, version 1.2) [25].

**Replication look-up**

To replicate our DMPs with results from previous EWASes, we looked for evidence of our DMPs (FDR < 0.05) in the two published epigenome-wide studies of PM$_{10}$ and/or NO$_2$ exposure in adults [6, 7]. Also, we examined whether DMPs reported in the two studies were replicated in our study. Across the two studies, 5001 DMPs were reported (FDR < 0.05): 9 for PM$_{10}$ and 4992 for NO$_2$. Of these, 4671 were available for the look-up analysis in our data after probe filtering: 9 for PM$_{10}$ and 4662 for NO$_2$. We set the cutoff of an uncorrected $p$ value < 0.05 for statistical significance for the look-up.
Associations of methylation levels of DMPs with gene expression levels of nearby transcripts: expression quantitative trait methylation in the BIOS data

To evaluate associations between methylation levels of DMPs and expression levels of nearby transcripts (cis-eQMTMs), we regressed the methylation M value, the log2 ratio of methylated versus unmethylated probe intensities, on gene expression, adjusting for age, sex, lymphocytes percentage, monocyte percentage, and RNA flow cell number. The inflation of models was corrected using the “bacon” method [26]. We mapped the expression quantitative trait methylation (eQMTMs) in a window of 250 kilobase pairs (kb) around the significant DMPs (FDR < 0.05).

For this analysis, we used a total of 3075 samples for which both methylation and gene expression data were available from 4 cohorts: Leiden Longevity Study, LifeLines Study, Rotterdam Study, and Netherland Twin Study. We analyzed each cohort separately and then meta-analyzed association results based on multiple-testing corrected p values (FDR from DMRcate) were visualized for regional association results (Additional file 3). No systematic inflation was observed in our results as genomic inflation factor (lambda) values were 0.83 for PM10 exposure and 1.07 for NO2 exposure.

We found numerous DMRs in relation to air pollution exposure: 22 for PM10 alone, 52 for NO2 alone, and 5 for both PM10 and NO2 (Tables 4 and 5). The five DMRs associated with NO2 exposure contain a DMP: cg06226567 (C20orf56) was positively associated with both PM10 and NO2 (FDR < 0.05). Exposure to the two pollutants was mostly positively associated with DNA methylation: 92% (N = 11/12 CpGs) for PM10 and 71% (N = 32/45 CpGs) for NO2. In Additional file 1: Figure S2, we provide Manhattan and quantile-quantile plots for visual representation of the epigenome-wide association results (Additional file 3).

Results

The average age of the study participants was 73 years (standard deviation, SD = 6) and 66% were male (Table 1). There were 39 never, 30 former, and 31 current smokers. The mean annual average concentration was 45.1 μg/m³ for PM10 and 13.1 ppb for NO2. The two air pollutants were highly correlated (Spearman correlation coefficient = 0.74, p value < 2.2E-16).

We observed numerous DMPs in relation to the two pollutants (FDR < 0.05): 11 for PM10 alone, 44 for NO2 alone, and 1 for both PM10 and NO2 (Tables 2 and 3).

Table 1 Descriptive characteristics of the study population

| Characteristics                        | The Korean COPD cohort (N = 100) |
|----------------------------------------|----------------------------------|
| Age, years                             | 72.8 ± 6.3                       |
| Male                                   | 66 (66%)                         |
| Body mass index, kg/m²                 | 22.9 ± 2.9                       |
| COPD, case                             | 60 (60%)                         |
| Cigarette smoking                      |                                  |
| Never                                  | 39 (39%)                         |
| Former                                 | 30 (30%)                         |
| Current                                | 31 (31%)                         |
| Pack-years                             |                                  |
| Former smoker                          | 28.9 ± 19.6                      |
| Current smoker                         | 35.7 ± 19.1                      |
| Annual average air pollution concentration at residential addresses |                           |
| PM10, μg/m³                            | 45.1 ± 2.0                       |
| NO2, ppb                               | 13.1 ± 1.4                       |

N (%) or mean ± standard deviation reported

*Chronic obstructive pulmonary disease

Of these 56 DMPs, some showed statistical significance after Bonferroni multiple testing correction: cg05454562 (WDR46), cg13999433 (AKNA), and cg11691844 (SYTL2) associated with PM10 exposure (Table 2); cg05171937 (STK38L), cg26583725 (8541 bp apart from IRS2), and cg06226567 (C20orf56) associated with NO2 exposure (Table 3). The DMP cg06992688 (OTUB2) was positively associated with both PM10 and NO2 (FDR < 0.05). Exposure to the two pollutants was mostly positively associated with DNA methylation: 92% (N = 11/12 CpGs) for PM10 and 71% (N = 32/45 CpGs) for NO2. In Additional file 1: Figure S2, we provide Manhattan and quantile-quantile plots for visual representation of the epigenome-wide association results (Additional file 3). No systematic inflation was observed in our results as genomic inflation factor (lambda) values were 0.83 for PM10 exposure and 1.07 for NO2 exposure.

We identified biological networks enriched in our association results based on genes to which either DMPs (FDR < 0.05) or CpGs having the minimum p value within the DMRs (FDR < 0.05 in DMRcate, Sidak adjusted p value < 0.05 in comb-p) were annotated: 138 for PM10 and 288 for NO2. The enriched networks included inflammatory and immune responses and cardiovascular, respiratory, and metabolic diseases (Additional file 2: Table S4 and S5). Cancer, hematological development, immunological and inflammatory diseases pathways overlap between PM10 and NO2 related differential methylation signals (Additional file 1: Figure S4, A). Of the genes associated with both PM10 and NO2 exposure, several contribute to the hematological, immunological, and inflammatory diseases pathways.

Conclusion

DMPs associated with air pollution exposure were identified in the BIOS population. A total of 56 DMPs were found, with some showing statistical significance after Bonferroni multiple testing correction. The findings suggest that DNA methylation is associated with exposure to air pollution, and that these associations are supported by biological networks enriched in inflammatory and immune responses as well as cardiovascular, respiratory, and metabolic diseases. Further studies are needed to explore the functional implications of these findings.
networks: NLRC4, RPTOR, CLUX1, SI00A12, LTA, and HLA-DMB (Additional file 1: Figure S4, B).

Using eFORGE [25], we found some enriched tissue- or cell type-specific histone marks (H3K27me3, H3K36me3, H3K4me3, H3K9me3, and H3K4me1) among the 132 probes associated with air pollution (PM10 or NO2) exposure based on either FDR < 0.05 from the DMP analyses or the minimum p value in the DMRs: 11 DMPs for PM10 exposure alone, 44 DMPs for NO2 exposure alone, 1 DMP for both PM10 and NO2 exposure, 19 probes showing the minimum p value in PM10 exposure related DMRs, 49 probes showing the minimum p value in NO2 exposure related DMRs, and 8 probes showing the minimum p value in DMRs associated with both PM10 and NO2 exposure. Enrichment of H3K4me1 in blood was observed for differential methylation related to PM10 exposure (Additional file 1: Figure S5). With respect to differential methylation related to NO2 exposure, several histone marks were enriched: H3K4me1, H3K27me3, H3K4me3, and H3K9me3 in blood; H3K4me1 and H3K27me3 in embryonic stem (ES) cell; and H3K4me1 in lung (Additional file 1: Figure S6).

Several DMPs (FDR < 0.05) in our study were reported to be associated with air pollution exposure in previous genome-wide DNA methylation studies. Of the 27 DMPs associated with NO2 (FDR < 0.05) in our study, 11 were reported to be related to NO2 exposure with the same direction of effects (Table 6) in the LifeLines cohort [7]. The 12 DMPs related to PM10 (FDR < 0.05) in our study were novel, meaning not reported to be associated with this pollutant in either of the two earlier studies [6, 7]. Notably, of the 4662 probes reported to be associated with NO2 exposure in the 2 studies and also available in our data, 26% (N = 1231) showed associations in our study of at least nominal significance (uncorrected p value < 0.05) with the same direction of effects (Additional file 2: Table S6).

From the analyses linking DNA methylation and gene expression in the BIOS data, we observed correlations of methylation levels of DMPs with gene expression levels of nearby (spanning a 250 bp window) transcripts (uncorrected p value < 0.05). Notably, of the 56 DMPs (FDR < 0.05, 70% (N = 39) were significantly related to gene expression of nearby transcripts (Additional file 2: Table S7).

**Discussion**

To our knowledge, this is the first study of genome-wide DNA methylation in relation to long-term ambient air pollution exposure, both PM10 and NO2, in an Asian population. We identified many differentially methylated signals—both individual probes and regions—related to long-term air pollution exposure in blood. We also replicated, in our Asian population, findings from earlier studies in European populations. Of our genome-wide significant findings, some provide the first replication of an earlier report from a European population [7] while others are novel. Notably, methylation levels of many DMPs were associated with gene expression levels of nearby transcripts, providing a link between ambient air

### Table 2 Differentially methylated CpGs in blood DNA in relation to PM10 exposure (FDR < 0.05), ordered by chromosomal location

| Chr | Gene (distance to gene) | Probe     | Position | Coef | SE  | P    |
|-----|-------------------------|-----------|----------|------|-----|------|
| 1   | NEGR1                   | cg07721244 | 72749275 | 0.004 | 0.001 | 1.6E-07 |
| 2   | ARID5A                  | cg04722215 | 97205147 | −0.006 | 0.001 | 1.4E-07 |
| 3   | FOXL2 (−81,364)         | cg21742790 | 138581702 | 0.005 | 0.001 | 8.6E-07 |
| 3   | XXYLT1 (−92,147)        | cg04252203 | 19469866 | 0.005 | 0.001 | 6.7E-07 |
| 6   | WDR46                   | cg05454562 | 33254447 | 0.006 | 0.001 | 4.3E-09 |
| 7   | FAM20C (−5283)          | cg16998831 | 187586  | 0.008 | 0.002 | 3.0E-07 |
| 8   | KIF13B                  | cg07023317 | 28951315 | 0.008 | 0.002 | 1.4E-06 |
| 9   | AKNA                    | cg13999433 | 117156883 | 0.007 | 0.001 | 3.9E-08 |
| 11  | SYTL2                   | cg11691844 | 85456064 | 0.006 | 0.001 | 1.1E-07 |
| 14  | OTUB2                   | cg06992688 | 94491958 | 0.008 | 0.002 | 1.0E-06 |
| 16  | MIR5093 (11,6079)       | cg26964426 | 85455911 | 0.025 | 0.005 | 8.3E-07 |
| 18  | NPC1                    | cg12709880 | 21163172 | 0.007 | 0.001 | 3.8E-07 |

*Chr* Chromosome

*Position* Physical position (base pair, National Center for Biotechnology Information human reference genome assembly Build 37.3)

*Coef* Regression coefficient from statistical model. Covariates age, sex, cigarette smoking status, pack-years of smoking, BMI, COPD status, and estimated cell-type proportions were included in the model. The coefficient can be interpreted as the difference in DNA methylation per 1 μg/m³ PM10 exposure. For example, cg07721244 showed 0.4% methylation increase per 1 μg/m³ PM10 exposure increase. Methylation values range 0–1

*SE* Standard error of regression coefficient

*p* Uncorrected p value

*p* Statistically significant after Bonferroni multiple-testing correction (1.2E-07)
### Table 3 Differentially methylated CpGs in blood DNA in relation to NO₂ exposure (FDR < 0.05), ordered by chromosomal location

| Chr | Gene (distance to gene) | Probe | Position | Coef | SE | P    |
|-----|--------------------------|-------|----------|------|----|------|
| 1   | MAN1C1 (−7282)           | cg16396978 | 25936677 | 0.008 | 0.002 | 3.9E-06 |
| 1   | ER3                      | cg13451048 | 44820073 | 0.007 | 0.001 | 8.6E-07 |
| 1   | RPL5                     | cg02769668 | 93302925 | -0.003 | 0.001 | 3.3E-07 |
| 1   | WARS2 (−29,067)          | cg06764239 | 119544777 | 0.002 | 3.5E-04 | 4.3E-06 |
| 1   | S100A12                  | cg02901136 | 153348305 | 0.012 | 0.002 | 2.7E-06 |
| 2   | STON1 (169)              | cg23256664 | 48757477 | -0.001 | 3.0E-04 | 8.9E-07 |
| 2   | NDUFB3                   | cg04865026 | 201936505 | 0.012 | 0.002 | 7.5E-07 |
| 1   | S100A12                  | cg02901136 | 153348305 | 0.012 | 0.002 | 2.7E-07 |
| 2   | PIKFYVE                  | cg19351166 | 209133632 | 0.008 | 0.002 | 5.5E-06 |
| 3   | CTDSPL                   | cg12386061 | 37906586 | 0.002 | 4.3E-04 | 5.5E-06 |
| 3   | DCBLD2 (122,596)         | cg01188562 | 98637410 | -0.004 | 0.001 | 2.0E-06 |
| 3   | AP2M1                    | cg17343451 | 183899704 | 0.009 | 0.002 | 3.3E-06 |
| 4   | CPLX1                    | cg16649791 | 816968 | -0.014 | 0.003 | 1.0E-06 |
| 4   | LINCO1097 (−3902)        | cg25913520 | 13524041 | 0.002 | 3.6E-04 | 5.5E-06 |
| 4   | LOC641518                | cg13775316 | 100993218 | 0.002 | 4.0E-04 | 6.2E-07 |
| 5   | DAP (B6217)              | cg23112301 | 1076559 | -0.005 | 0.001 | 3.8E-06 |
| 5   | ZNF366                   | cg21770462 | 71803219 | 0.008 | 0.002 | 4.7E-06 |
| 5   | ERAPI1 (−82,414)         | cg13625213 | 95915327 | -0.002 | 4.0E-04 | 3.4E-06 |
| 5   | CDHR2 (−2294)            | cg18194153 | 175967218 | 0.010 | 0.002 | 1.3E-07 |
| 6   | LT1                      | cg11586857 | 31540136 | -0.007 | 0.001 | 3.5E-06 |
| 8   | PMP2                     | cg22796481 | 82353365 | -0.019 | 0.004 | 6.7E-07 |
| 8   | OSR2                     | cg09607488 | 99963657 | 0.007 | 0.002 | 4.5E-06 |
| 9   | RORB                     | cg04130427 | 77113915 | 0.005 | 0.001 | 3.7E-06 |
| 10  | ZNF438                   | cg10575075 | 31288634 | 0.014 | 0.003 | 2.0E-06 |
| 10  | EMX2                     | cg02420850 | 119302157 | 0.002 | 4.0E-04 | 6.2E-07 |
| 11  | TMEM138                  | cg03370752 | 61136373 | 0.010 | 0.002 | 5.5E-06 |
| 11  | SORL1                    | cg17510957 | 121466629 | 0.011 | 0.002 | 5.1E-06 |
| 12  | TEAD4                    | cg12902426 | 3068889 | 0.003 | 0.001 | 3.7E-06 |
| 12  | STK38L                   | cg05171937 | 27396765 | 0.010 | 0.002 | 1.1E-08 |
| 12  | DXS55                    | cg13559144 | 124086193 | 0.002 | 4.3E-04 | 3.0E-06 |
| 13  | EDNRB                    | cg23326536 | 78491199 | -0.003 | 0.001 | 1.7E-06 |
| 13  | IR52 (−8541)             | cg26583725 | 110397643 | -0.001 | 2.3E-04 | 4.9E-08 |
| 14  | ITPK1                    | cg05284742 | 93552128 | 0.009 | 0.002 | 4.1E-06 |
| 14  | OTUB2                    | cg06992688 | 94491958 | 0.013 | 0.003 | 3.3E-06 |
| 14  | PDL4                     | cg15352829 | 10391018 | 0.010 | 0.002 | 3.0E-06 |
| 15  | LOC145663                | cg04025675 | 45671028 | 0.005 | 0.001 | 6.3E-07 |
| 16  | ZCCHC14                  | cg16727006 | 87470545 | -0.010 | 0.002 | 4.8E-06 |
| 17  | EFCAB5 (−2689)           | cg22888787 | 27950276 | 0.010 | 0.002 | 3.9E-07 |
| 17  | CD300A (−12,486)         | cg00227781 | 72450036 | 0.004 | 0.001 | 3.0E-06 |
| 19  | LOC100128675             | cg06642503 | 35597415 | -0.005 | 0.005 | 2.9E-06 |
| 19  | ZNF347                   | cg15050103 | 53642858 | -0.008 | 0.002 | 3.7E-06 |
| 19  | ZNF542 (−28,810)         | cg06109293 | 56850658 | 0.020 | 0.004 | 1.9E-07 |
pollution exposure-related differential methylation and gene expression.

Some of our DMPs annotated to genetic loci reported in published genome-wide association studies of various health outcomes that have been related to air pollution exposure. Differential methylation of cg11586857 related to both pollutants annotated to LTA in which an earlier study identified rs1799964 (p value = 3.3E-07) to be associated with blood lipid levels [28]. Cg06992688 associated with exposure to both air pollutants resides in OTUB2, a nearby gene of three genetic variants related to lung function with p values around 1.0E-04 [29]. In addition, cg05284742 related to NO2 exposure is located in ITPK1; this gene contains rs22953934 (p value = 2.3E-16) associated with myocardial infarction in Asian populations [30].

Knowledge-based pathway analyses and enrichment analyses of epigenetic elements using publicly available data provided biological implication of our study findings. Enrichment of networks, such as inflammatory and immune responses and cardiovascular, pulmonary and metabolic diseases, in our results supports previous findings of air pollution exposure and the identified disease associations. Several enriched histone marks in relevant tissue and cell types (embryonic stem cell, blood and lung) suggest additional biological relevance of our differential methylation signals.

We found five studies examining associations of DNA methylation, measured using Illumina’s Infinium 450K array, with ambient air pollution exposure in either children or adults [5–7, 31, 32]. Of the five, one reported DMPs associated with short-term exposure to particulate matter < 2.5 μm (PM2.5) [31]. Chi and colleagues [5] measured DNA methylation using the 450K array but they analyzed only a subset of probes for associations with PM2.5 and oxides of nitrogen (NOx). Gruzieva and colleagues [32] found differential methylation in children in relation to prenatal NO2 exposure. The remaining two analyzed long-term exposure to pollutants including both PM10 and NO2 for associations with genome-wide DNA methylation in adults [6, 7]. Notably, differential methylation signals in our study provide the first replication of findings from the two studies in European adults [6, 7], suggesting similar relationships between ambient air pollution exposure and DNA methylation between European and Asian populations.

In this study, we adjusted for COPD status because it may confound associations between air pollution exposure and methylation. We also explored possible effect measure modification by the disease status in a sensitivity analysis. Of the 45 CpGs related to NO2, three (cg16649791, cg13559144, and cg23326536), showed an interaction term that was nominally significant (Additional file 2: Table S8); none of the 12 PM10-related CpGs showed statistically significant interaction.

Our study has limitations and strengths. Limitations include the lack of a replication population. However, we were able to compare our findings against published lists of DMPs at genome-wide significance from two earlier studies in European populations [6, 7]. With respect to the exposure assessment, we used exposure values at residential addresses estimated from a national-scale prediction model rather than an area-specific model which could not be developed because of the limited number of monitoring sites (< 10) in the areas where our study participants resided. However, in previous US studies, estimates of PM2.5 for specific areas using national models showed association results comparable to those from area-specific models [33, 34]. Third, we used annual average concentrations estimated for 2010 and participant addresses at baseline visits in 2012 without incorporating participants’ previous exposure to air pollution. The year 2010 was used in the model because of the increased number of available monitoring sites and temporally aligned geographic data. As spatial distribution of air pollution should be relatively consistent over years in our study area with stable environments, the impact of using temporally limited exposure and address information on our methylation analysis could be small. Lastly, we have a relatively small
sample size compared to earlier genome-wide methylation studies of air pollution exposure.

The study has a number of important strengths. Participants reported residing in the same residential areas for 50 years (SD = 21) on average. This high level of residential stability improved our ability to estimate associations with long-term air pollution exposure. Further, we have included both PM$_{10}$ and NO$_2$ exposure so that we can examine whether there are common or unique differential methylation signals related to the two pollutants. In addition, we followed up our DMPs by examining relationships with gene expression and found that a majority were related to gene expression, suggesting functional importance of the associations. Further, we conducted pathway analyses and enrichment analyses of tissue- and cell-type specific histone marks to better understand the biological implication of the differentially methylated signals that we observed. Last, we identified differential methylation regions in blood DNA in relation to PM$_{10}$ exposure (adjusted $P < 0.05$ both in DMRcate and in comb-p):

| Chr  | Gene (distance to gene) | DMRcate | comb-p |
|------|-------------------------|---------|--------|
| 1    | MIB2                    | 1549615 1550031 0.020 5 (4) | 0.009 0.009 |
| 2    | NOL10 (~22,166)         | 10687583 10688726 9.4E-05 8 (5) | 10687962 10688317 2.6E-05 5 (5) 2.5E-04 |
| 2    | SNE1                    | 241975035 241976244 0.007 2 (2) | 2.011 2.5E-04 |
| 3    | IL20RB                  | 136676672 136676846 9.4E-05 7 (6) | 28874754 7.3E-04 4 (4) 0.002 |
| 3    | TRIM27                  | 28874479 28875370 9.4E-05 7 (6) | 28874754 7.3E-04 4 (4) 0.002 |
| 3    | TRIM39                  | 30297174 30297627 2.3E-08 11 (10) | 1.1E-07 8.4E-04 |
| 3    | LTA                     | 31539539 31540750 1.3E-11 19 (13) | 31540461 3.4E-06 18 (12) 4.8E-05 |
| 3    | TREM1                   | 41254471 41254997 0.018 4 (3) | 41254433 0.012 5 (3) 1.7E-04 |
| 7    | FOXL1                   | 4752951 4753002 1.3E-04 3 (3) | 7.2E-04 3.4E-04 |
| 8    | CSGALNACT1               | 19459672 19460243 0.003 7 (4) | 0.001 7.8E-04 |
| 8    | PWIL2                   | 22131675 22133356 1.2E-04 15 (6) | 22132563 0.027 13 (5) 3.8E-05 |
| 8    | KIF13B                  | 28961315 28961356 2.9E-04 3 (2) | 0.003 1.4E-06 |
| 9    | C9orf131                | 35042344 35042395 0.003 2 (2) | 0.005 5.6E-05 |
| 10   | CAMK1D                  | 12648032 12648338 3.6E-02 3 (2) | 12648526 0.011 4 (3) 0.002 |
| 10   | C10orf105               | 73498624 73498737 0.003 3 (2) | 0.032 2.7E-05 |
| 15   | PIK3R1                   | 100890907 100891257 1.1E-04 5 (4) | 100890963 0.014 4 (3) 8.8E-05 |
| 17   | TNREC6                  | 76036514 76037562 7.3E-05 7 (7) | 76037035 1.6E-05 6 (6) 0.001 |
| 19   | PRPT3                   | 846117 846354 0.010 3 (3) | 0.004 0.001 |
| 19   | HRPT3                   | 847943 848071 0.005 4 (4) | 0.003 0.001 |
| 19   | CALR                    | 13054791 13054718 0.002 5 (4) | 13054427 0.014 4 (3) 8.3E-05 |
| 19   | FBXO17                  | 39465821 39467675 0.002 8 (4) | 0.004 6.6E-04 |
| 20   | STK35                   | 2085157 2085344 0.003 2 (2) | 0.002 1.7E-05 |
| 20   | SUPI                    | 43882990 43883307 0.004 3 (3) | 43883546 8.5E-04 4 (4) 9.7E-05 |
| 20   | C20orf123               | 45179157 45179413 2.0E-04 6 (5) | 1.4E-04 2.2E-04 |
| 21   | C21orf81                | 15352848 15352983 0.013 2 (2) | 0.012 4.7E-04 |

Blanked cells in “Start,” “End,” and “#CpGs” for comb-p represent the same information compared to results in DMRcate.

- **Chr**: Chromosome
- **Gene (distance to gene)**: Minimum distance to transcription start site of the mapped gene (base pair)
- **DMRcate**: Physical position (base pair, National Center for Biotechnology Information human reference genome assembly Build 37.3)
- **comb-p**: Number of probes in the region (number of probes having uncorrected p value < 0.05)
- **FDR**: Number of probes having uncorrected p value < 0.05
- **Sidak**: P of Sidak multiple-testing correction
- **#CpGs**: Minimum p value among uncorrected p-values of CpGs in each region. When either start or end positions were different between DMRs from the two DMR approaches, we used results from DMRcate.
- **Minimum p**: Region including significant (FDR < 0.05) differentially methylated probes from our epigenome-wide association study.
Table 5  Differentially methylated regions in blood DNA in relation to NO₂ exposure (adjusted p value < 0.05 both in DMRcate and in comb-p)

| Chr | Gene (distance to gene) | DMRcate | comb-p | Minimum |
|-----|-------------------------|---------|--------|---------|
|     |                         | Start (bp) | End (bp) | FDR | #CpGs | Start (bp) | End (bp) | FDR | #CpGs |
| 1   | RUNX3                   | 25291041 | 25291905 | 0.005 | 7 (4) | 25291584 | 0.044 | 6 (3) | 0.001 |
| 1   | RPS6KA1                 | 26855423 | 26855926 | 0.006 | 4 (3) | 26855765 | 0.009 | 3 (3) | 9.1E-04 |
| 1   | TFAP2E                  | 36038468 | 36039173 | 2.1E-04 | 8 (7) | 36038701 | 3.3E-04 | 6 (6) | 0.002 |
| 1   | ARTN                    | 44398868 | 44398984 | 5.6E-04 | 10 (6) | 44399363 | 0.012 | 6 (4) | 9.8E-04 |
| 1   | S100A1Z                 | 153347819 | 153348304 | 0.005 | 3 (2) | 153389781 | 0.047 | 6 (4) | 0.001 |
| 1   | S100A14                 | 153589528 | 153590243 | 0.013 | 4 (2) | 153589781 | 0.047 | 3 (2) | 0.001 |
| 1   | S100A13                 | 153599479 | 153600156 | 0.001 | 7 (6) | 153600156 | 0.017 | 6 (6) | 0.003 |
| 1   | ATP8B2                  | 154300117 | 154300241 | 6.9E-04 | 2 (2) | 154300117 | 0.007 | 2 (2) | 3.4E-05 |
| 2   | LAX1                    | 203733971 | 203734559 | 0.004 | 6 (4) | 203734559 | 0.002 | 2 (2) | 0.002 |
| 2   | C1orf35                 | 228291118 | 228291705 | 0.002 | 6 (5) | 228291705 | 0.007 | 6 (5) | 0.009 |
| 3   | ALS2CR11                | 202483704 | 202484583 | 0.007 | 10 (5) | 202483704 | 0.008 | 7 (5) | 0.006 |
| 3   | AMT                     | 49459143 | 49460521 | 1.3E-06 | 10 (7) | 49459855 | 8.3E-05 | 9 (7) | 1.6E-04 |
| 3   | B3GALNT1                | 160475035 | 160475336 | 0.002 | 5 (5) | 160475336 | 0.003 | 5 (5) | 0.003 |
| 3   | B3GALNT1                | 160822268 | 160822911 | 0.001 | 8 (5) | 160822911 | 0.031 | 5 (4) | 0.003 |
| 4   | MGAT4                   | 179230709 | 179231109 | 0.002 | 3 (2) | 179231109 | 0.006 | 3 (2) | 0.002 |
| 5   | TRIM39                  | 228291118 | 228291705 | 0.023 | 6 (5) | 228291705 | 0.017 | 6 (5) | 0.009 |
| 5   | LTA                     | 31539539 | 31540750 | 1.9E-15 | 19 (11) | 31540461 | 4.5E-07 | 18 (11) | 3.5E-06 |
| 5   | HLA-DMA                 | 32904074 | 32905190 | 1.2E-05 | 9 (5) | 32904621 | 0.001 | 5 (3) | 8.7E-06 |
| 5   | HLA-DMB                 | 33040747 | 33040847 | 7.6E-04 | 11 (9) | 33040747 | 2.9E-05 | 11 (9) | 6.7E-04 |
| 6   | HK1                     | 71087924 | 71088038 | 0.009 | 3 (2) | 71088038 | 0.012 | 3 (2) | 0.006 |
| 6   | HTRA1                   | 1266180 | 1267228 | 8.2E-04 | 4 (4) | 1266180 | 0.001 | 3 (3) | 2.0E-04 |
| 6   | IFITM3                  | 27279101 | 27279575 | 0.009 | 3 (2) | 27279575 | 0.044 | 3 (2) | 1.8E-04 |
| 6   | STAP2                   | 89840396 | 89841345 | 1.9E-05 | 13 (5) | 89841345 | 8.9E-05 | 12 (5) | 2.1E-04 |
| 6   | CSGALNACT1              | 19459672 | 19460243 | 2.3E-05 | 7 (5) | 19460243 | 6.7E-05 | 7 (5) | 1.4E-04 |
| 6   | KIAA0146; CEBPD (−19)   | 48548112 | 48649767 | 7.6E-08 | 7 (7) | 48648112 | 3.9E-09 | 6 (6) | 7.4E-05 |
| 6   | HEY1                    | 80678770 | 80679314 | 0.002 | 4 (3) | 80679314 | 0.026 | 2 (2) | 4.4E-04 |
| 6   | NDRG1                   | 134307105 | 134307728 | 2.3E-05 | 3 (3) | 134307728 | 7.4E-04 | 2 (2) | 3.0E-05 |
| 7   | PKD1                    | 28491326 | 28492265 | 0.006 | 8 (3) | 28491326 | 0.035 | 6 (3) | 0.001 |
| 7   | PCDH20                  | 61998902 | 61999002 | 6.0E-05 | 12 (8) | 61999002 | 5.3E-04 | 8 (7) | 6.7E-04 |
| 7   | DAOA (−319,060)         | 105794690 | 105795234 | 0.023 | 3 (3) | 105795234 | 0.024 | 3 (3) | 0.003 |
| 8   | DAD1 (−58,286)          | 22974144 | 22975521 | 0.007 | 6 (5) | 22975521 | 0.029 | 5 (4) | 1.2E-04 |
| 8   | PLD4                    | 105390602 | 105391263 | 0.002 | 3 (2) | 105391263 | 0.007 | 2 (2) | 3.0E-06 |
### Table 6

Look-up analysis of CpGs associated with NO₂ exposure in the Korean COPD Cohort (FDR < 0.05) in a previous publication from the LifeLines Cohort from the Netherlands

| Chr | Gene (distance to gene) | Probe | The Korean COPD study | The LifeLines cohort study [7] |
|-----|-------------------------|-------|------------------------|-------------------------------|
|     |                         |       | Coef (per 1 ppb NO₂) ± SE | Coef (per 10 μg/m³ NO₂) ± SE |
|     |                         |       | P                      | P                             |
| 1   | MAN1C1 (~7282)          | cg16396978 | 0.008 ± 0.002 | 0.013 ± 0.004 | 5.4E-04 |
| 1   | SI00412                 | cg02901136 | 0.012 ± 0.002 | 0.027 ± 0.006 | 3.1E-05 |
| 2   | PLEKH3                  | cg09950920 | 0.013 ± 0.003 | 0.024 ± 0.007 | 3.4E-04 |
| 3   | AP2M1                   | cg17343451 | 0.009 ± 0.002 | 0.020 ± 0.005 | 1.4E-05 |
| 5   | ZNF366                  | cg21770462 | 0.008 ± 0.002 | 0.015 ± 0.004 | 4.1E-05 |
| 10  | ZNF438                  | cg10575075 | 0.014 ± 0.003 | 0.026 ± 0.007 | 2.7E-04 |
| 11  | TMEM138                 | cg03370752 | 0.010 ± 0.002 | 0.028 ± 0.008 | 2.4E-04 |
| 11  | SORL1                   | cg17510957 | 0.011 ± 0.002 | 0.023 ± 0.007 | 4.7E-04 |
| 12  | STK38L                  | cg05171937 | 0.010 ± 0.002 | 0.036 ± 0.009 | 4.0E-05 |
| 14  | OTUB2                   | cg06992688 | 0.013 ± 0.003 | 0.026 ± 0.007 | 1.4E-04 |
| 21  | MOR3                    | cg01261013 | 0.010 ± 0.002 | 0.023 ± 0.006 | 3.3E-04 |

*aChromosome  
bDistance to transcription start site of the mapped gene (base pair)  
cRegression coefficient from statistical model  
dStandard error of regression coefficient  
eStatistical significance from statistical model
DMRs by combining association signals at neighboring CpGs using two different methods in addition to identifying DMPs.

Conclusions
We identified differential DNA methylation signals in blood associated with long-term ambient air pollution exposure and linked differential methylation to differential gene expression. Replication of many of our results from an Asian population, in a European population, suggests similar influences of air pollution exposure across ancestry. Our CpGs and regions showing differential methylation are potential biomarkers for long-term ambient air pollution exposure. These findings may better inform mechanisms linking air pollution exposure to adverse health outcomes.

Additional files

Additional file 1: Figure S1. Workflow of the epigenome-wide association study of long-term ambient air pollution exposure.

Figure S2. Manhattan and quantile-quantile plots. Figure S3. Regional visualization of the association of air pollution exposure (PM_{10} and NO_{2}) with blood DNA methylation. Figure S4. Visualization of pathway analysis results. Figure S5. Tissue- and cell-type specific enrichment pattern in CpGs significantly associated (FDR < 0.05) with PM_{10} exposure. Figure S6. Tissue- and cell-type specific enrichment pattern in CpGs significantly associated (FDR < 0.05) with NO_{2} exposure (DOCX 6165 kb).

Additional file 2: Table S1. CpG probe filtering criteria in the 450K array. Table S2. CpGs included in the top five differentially methylated regions in relation to PM_{10} from each analysis: DMRCrate and comb-p (ordered by software and chromosomal location). Table S3. CpGs included in the top five differentially methylated regions in relation to NO_{2} from each analysis: DMRCrate and comb-p (ordered by software and chromosomal location). Table S4. Enriched networks in genes related to PM_{10} exposure. Table S5. Enriched networks in genes related to NO_{2} exposure. Table S6. Look-up analysis of CpGs associated with NO_{2} exposure (FDR < 0.05 in earlier epigenome-wide association studies) in the Korean COPD cohort, sorted by uncorrected P in the Korean COPD Cohort. Table S7. Associations between methylation levels at air pollution associated CpGs (FDR < 0.05) and the expression levels of nearby genes: cis-eQTM. Table S8. Differential methylation of an interaction between NO_{2} exposure and COPD status (XLSX 115 kb).

Additional file 3: Table S9. Differential methylation in relation to NO_{2} exposure (XLSX 46770 kb).

Additional file 4: Table S10. Differential methylation in relation to PM_{10} exposure. Table S11. Differential methylation in relation to COPD status (XLSX 46770 kb).

Abbreviations
BIOS: Biobank-based integrative omics studies; BMI: Body mass index; ChAMP: Chip analysis methylation pipeline; COPD: Chronic obstructive pulmonary disease; CpGs: C–phosphate–G probes; DMPs: Differentially methylated probes; DMRs: Differentially methylated regions; eFORGE: Experimentally-derived functional element overlap analysis of regions from EWAS; eQTM: Expression quantitative trait methylation; EWAS: Epigenome-wide association study; FDR: False discovery rate; IPA: Ingenuity pathway analysis; NO_{2}: Nitrogen dioxide; PM_{10}: Particulate matter ≤ 10 μm in diameter; SD: Standard deviation

Acknowledgements
We appreciate all the study participants for their contribution to this research. We thank Dr. Frank Day of NIEHS for expert computational assistance. The Biobank-Based Integrative Omics Studies (BIOS) Consortium is funded by BBMRI-NL, a research infrastructure financed by the Dutch government (NWO 184.021.007).

References
1. Patel V, Kantipudi N, Jones G, Upton A, Kamath MV. Air pollution and cardiovascular disease: a review. Crit Rev Biomed Eng. 2016;44:327–46.
2. Adam M, Schikowski T, Cansin AE, Cai Y, Jacquemin B, Sanchez M, Vierkott A, Marcon A, Kielidi D, Sugiri D, et al. Adult lung function and long-term air pollution exposure. ESCAPE: a multicentre cohort study and meta-analysis. Eur Respir J. 2015;45:38–50.
3. Johanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, Guan W, Xu T, Elks CE, Aslibekyan S, et al. Epigenetic signatures of cigarette smoking. Circ Cardiovasc Genet. 2016;9:436–47.

Funding
This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2018R1A1A1057961, 2017R1A2B4003790, 2018R1A2B604060), the Ministry of Education (2013R1A6A4A04055017), National Cancer Center (NCC-1810200-01), and grants from the Environmental Health Center funded by the Ministry of Environment, Republic of Korea. Supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (Z01 ES04012). All authors have no relationships or financial involvement with an organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript.

Availability of data and materials
The results of epigenome-wide association study of PM_{10} and NO_{2} exposure using Infinium HumanMethylation450 BeadChip are provided in Additional file 3: Table S9 and S10 of this manuscript.

Authors’ contributions
MKL, CJX, MUC, CN, JW, and SOK contributed to manuscript preparation, data analysis or data interpretation. SJL, WJK, and SYK advised on analytic approaches and interpretation of results and contributed to the drafting of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate
The Institute Review Board of the Kangwon National University Hospital approved analyses of the data (Institutional Review Board of Kangwon National University Hospital 2012-06-007-001 and KNUH-2016-05-003-001). Informed written consent was obtained from all participants. The study adhered to the tenets of the Helsinki Declaration of 1975, as revised in 2008.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1. Epidemiology Branch, Division of Intramural Research, Department of Health and Human Services, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA.
2. Department of Pediatric Pulmonology and Pediatric Allergy, Beatrix Children’s Hospital, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 3. Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 4. GRIAC Research Institute, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. 5. OmicSoft, QiAGEN Bioinformatics, Cary, NC, 27513, USA. 6. Department of Internal Medicine and Environmental Health Center, Kangwon National University Hospital, School of Medicine, Kangwon National University, Chuncheon 24289, South Korea. 7. Department of Cancer Control and Population Health, Graduate School of Cancer Science and Policy, National Cancer Center, Goyang 10408, South Korea.

Received: 4 December 2018 Accepted: 18 February 2019
Published online: 28 February 2019
4. Hou L, Zhang X, Wang D, Baccarelli A. Environmental chemical exposures and human epigenetics. Int J Epidemiol. 2012;41:79–105.

5. Chi GC, Liu Y, MacDonald JW, Barr RG, Donohue KM, Hensley MD, Hou L, McCauley CE, Reynolds LM, Siscovick DS, Kaufman JD. Long-term outdoor air pollution and DNA methylation in circulating monocytes: results from the multi-ethnic study of atherosclerosis (MESA). Environ Health. 2016;15:119.

6. Plusquin M, Guida F, Poldoro S, Vermeulen R, Raaschou-Nielsen O, Campiello G, Hoek G, Kyrtopoulos SA, Georgiadis P, Naccarato A, et al. DNA methylation and exposure to ambient air pollution in two prospective cohorts. Environ Int. 2017;108:127–36.

7. de FC, Lichtenfels AJ, van der Plaat DA, de Jong K, van Diemen CC, Postma DS, Nededjov K, van Duijn CM, Armin N, la Bastide-van Gernert, de Vries M, et al. Long-term air pollution exposure, genome-wide DNA methylation and lung function in the LifeLines cohort study. Environ Health Perspect. 2018;126:027004.

8. Hong Y, Kwon J, Lee S, Han YJ, Moon JY, Kim HY, Han S, Lee S, Kim WJ. Methodology of an observational cohort study for subjects with chronic obstructive pulmonary disease in dusty areas near cement plants. J Pulm Respir Med. 2014;04:169–74.

9. Kim SY, Song I. National-scale exposure prediction for long-term concentrations of particulate matter and nitrogen dioxide. Environ Pollut. 2017;226:21–9.

10. Morris TJ, Butterle LM, Feber A, Teschendorff AE, Chakravarthy AR, Bond B, Beek P, Chamon XM, Chavatte-Palmer P, Chen S, et al. A genome-wide association study of DNA methylation and pest resistance in rice. Genome Biol. 2014;15:R37.

11. Teschendorff AE, Marabito F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. Meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. Environ Health Perspect. 2017;125:104–10.

12. Panni T, Mehta AJ, Schwartz JD, Baccarelli AA, Just AC, Wolf K, Wahl S, Cypys J, Kunze S, Strauch K, et al. Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. Environ Health Perspect. 2016;124:983–90.

13. Gruzieva O, Xu CJ, Breton CV, Arnes-Maeso L, Anto JM, Auffray C, Ballereau S, Bellanger T, Bousquet J, Bustamante M, et al. Epigenome-wide meta-analysis of methylation in pediatric acute lymphoblastic leukemia. Genome Bio. 2013;14:R105.

14. Price ME, Cotton AM, Lam LL, Farme P, Emberley E, Brown CJ, Robinson WP, Kobor MS. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics. 2013;6:4.

15. Tukey JW. Exploratory data analysis. MA: Addison-Wesley Reading; 1977.

16. Houseman EA, Accomando WP, Forestier E, Frost BM, Harila-Saari A, et al. Genome-wide DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. Environ Health Perspect. 2013;121:1017–25.

17. Kim SY, Sheppard L, Sampson PD, Kim SY, Richards M, Veldal S, Kaufman JD, Szpiro AA. A national prediction model for PM2.5 component exposures and measurement error-corrected health effect inference. Environ Health Perspect. 2013;121:1017–25.

18. Hancock DB, Eijgelsheim M, Vilk JB, Gharib SA, Loefle LR, Marcianiet RD, Franceschini N, van Dume YM, Chen TH, Barr RG, et al. Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet. 2010;42:45–52.

19. Yamada Y, Nishida T, Ichihara S, Sawabe M, Fuku N, Nishigaki Y, Ayagi Y, Tanaka M, Fujiwara Y, Yoshida H, et al. Association of a polymorphism of BTN2A1 with myocardial infarction in East Asian populations. Atherosclerosis. 2011;215:145–52.

20. Kim SY, Sheppard L, Kaufman JD, Baccarelli AA, Just AC, Wolf K, Wahl S, Cypys J, Kunze S, Strauch K, et al. Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. Environ Health Perspect. 2016;124:983–90.

21. Gruzieva O, Xu CJ, Breton CV, Arnes-Maeso L, Anto JM, Auffray C, Ballereau S, Bellanger T, Bousquet J, Bustamante M, et al. Epigenome-wide meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. Environ Health Perspect. 2017;125:104–10.

22. Teschendorff AE, Marabito F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. Meta-analysis of methylation in children related to prenatal NO2 air pollution exposure. Environ Health Perspect. 2017;125:104–10.

23. Bergen S, Sheppard L, Sampson PD, Kim SY, Richards M, Veldal S, Kaufman JD, Szpiro AA. A national prediction model for PM2.5 component exposures and measurement error-corrected health effect inference. Environ Health Perspect. 2013;121:1017–25.

24. Kim SY, Sheppard L, Kaufman JD, Bergen S, Szpiro AA, Larson TV, Adar SD, Diez Roux AV, Polak JF, Veldal S, Individual-level concentrations of fine particulate matter chemical substances and subclinical atherosclerosis: a cross-sectional analysis based on 2 advanced exposure prediction models in the multi-ethnic study of atherosclerosis. Am J Epidemiol. 2014;180:718–28.