Short-term airborne particulate matter exposure alters the epigenetic landscape of human genes associated with the mitogen-activated protein kinase network: a cross-sectional study

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters
Short-term airborne particulate matter exposure alters the epigenetic landscape of human genes associated with the mitogen-activated protein kinase network: a cross-sectional study

Juan Jose Carmona1,2,3, Tamar Sofer3, John Hutchinson4, Laura Cantone5, Brent Coull3, Arnab Maity6, Pantel Vokonas7, Xihong Lin3, Joel Schwartz1,2 and Andrea A Baccarelli1,2,3*

Abstract

Background: Exposure to air particulate matter is known to elevate blood biomarkers of inflammation and to increase cardiopulmonary morbidity and mortality. Major components of airborne particulate matter typically include black carbon from traffic and sulfates from coal-burning power plants. DNA methylation is thought to be sensitive to these environmental toxins and possibly mediate environmental effects on clinical outcomes via regulation of gene networks. The underlying mechanisms may include epigenetic modulation of major inflammatory pathways, yet the details remain unclear.

Methods: We sought to elucidate how short-term exposure to air pollution components, singly and/or in combination, alter blood DNA methylation in certain inflammation-associated gene networks, MAPK and NF-κB, which may transmit the environmental signal(s) and influence the inflammatory pathway in vivo. To this end, we utilized a custom-integrated workflow—molecular processing, pollution surveillance, biostatistical analysis, and bioinformatic visualization—to map novel human (epi)gene pathway-environment interactions.

Results: Specifically, out of 84 MAPK pathway genes considered, we identified 11 whose DNA methylation status was highly associated with black carbon exposure, after adjusting for potential confounders—age, sulfate exposure, smoking, blood cell composition, and blood pressure. Moreover, after adjusting for these confounders, multi-pollutant analysis of synergistic DNA methylations significantly associated with sulfate and BC exposures yielded 14 MAPK genes. No associations were found with the NF-κB pathway.

Conclusion: Exposure to short-term air pollution components thus resulted in quantifiable epigenetic changes in the promoter areas of MAPK pathway genes. Bioinformatic mapping of single- vs. multi-exposure-associated epigenetic changes suggests that these alterations might affect biological pathways in nuanced ways that are not simply additive or fully predictable via individual-level exposure assessments.

* Correspondence: abaccare@hsph.harvard.edu
1Laboratory of Human Environmental Epigenetics, Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA
2Exposure, Epidemiology, and Risk Program, Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA
Full list of author information is available at the end of the article

© 2014 Carmona et al; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. 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 air particulate matter (PM) is well known to augment oxidative stress in exposed individuals, and it has been consistently linked to reduced lung function as well as hospitalization and mortality for various cardiopulmonary diseases [1]. Indeed, ambient air pollution is a critical public health concern in US cities and worldwide, accounting for a staggering ~3.7 million premature deaths globally in 2012, according to a World Health Organization estimate [2,3]. In highly PM-exposed individuals, blood leukocyte gene expression profiling has revealed responses related to worsened systemic oxidative stress and inflammation, which exacerbate aggravation of the airways and disease outcomes [4], yet the underlying mechanisms remain unclear.

Recent interest has focused on examining different components of air pollution particles to assess their relative toxicity in vivo. For example, black carbon (BC), a component of PM commonly used as a marker of traffic particles, has been reported to be more strongly associated with increased blood pressure in the elderly compared to other pollutants [5]. But sulfates—i.e., PM components from industrial emissions—had comparable effect as BC on dilation of the brachial artery [6]. Similarly, traffic pollutants are often more commonly cited as associated with asthma [7]. However, PM with an aerodynamic diameter <2.5 µm (PM2.5) had stronger associations than BC in some studies [7,8], suggesting that PM from sources other than traffic has stronger effects on asthma. Interestingly, sulfates have also been associated with asthma phenotypes [9]. While there are many components of particles in PM, BC as a general index of traffic-derived particles and sulfate as a marker of secondary particles, such as coal-burning power plants, are reasonable surrogates for two of its more important sources. At present, however, questions remain about how best to model these two pollutants, individually and/or in combination, in order to understand their effects on human populations [10].

DNA methylation is an emerging field that can provide unique opportunities to study source- and component-specific biological effects from PM-related exposures in vivo [11-14]. Many nuclear-encoded genes have been shown to undergo exposure-related changes in their DNA methylation status, an epigenetic mechanism highly sensitive to chemical components found in PM [13-17]. DNA methylation, the best understood of the epigenetic mechanisms, is the covalent addition of methyl groups to cytosine to form 5-methyl-cytosine (5mC). Methylation of promoter regions and other regulatory sequences usually tends to repress gene expression, although examples leading to gene activation exist; consequently, DNA methylation is now recognized as an important regulator of transcription [11,18,19]. Since DNA methylation is responsive to environment signals but changes less rapidly than mRNA or protein/metabolite levels, it may represent a more stable bio-archive of environmental exposures [20-22]. Importantly, these genome-wide chemical “annotations” hold considerable promise as molecular indicators of toxic exposures and possible predictors of disease risk [16,17].

Our work aims to apply recently validated biostatistical and biotechnological platforms to elucidate further the connection between epigenetic changes in genes with ambient air pollution exposures in vivo. To this end, we evaluated methylation microarray data from a subset of 141 male participants in the Normative Aging Study (NAS), together with air pollution data from their corresponding geographical regions. We focused on a key molecular pathway and its downstream target—MAPK (the mitogen-activated protein kinase) and NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells), respectively—which have been shown to be activated by PM exposure and may operate as biological mediators of pathophysiological responses to PM [23-28].

The MAPK cascade transduces a broad range of extracellular stress and physiological signals, and it mediates cellular responses to diverse processes from cell proliferation and differentiation to inflammation [29-31]. The various signaling branches of this pathway share a multi-tiered control system: MAPK proteins are activated by dual phosphorylation on tyrosine and serine/threonine residues via an upstream layer of dual-specificity kinases, MAPK kinases (MAPKK), which are themselves phosphorylated by a third tier of kinases, MAPKKKs. Previous experimental work has already identified activation of the MAPK cascade as a possible mechanistic link between air pollution exposures and respiratory and cardiovascular health outcomes [27,28]. MAPK signaling typically originates from physiological stimuli, e.g., via cell-surface receptors—and in some instances these receptors are coupled to small GTPases [29].

Although a comprehensive MAPK signaling-interaction map, or “interactome,” has recently been created [32], drawn using experimental data from molecular studies and predictive bioinformatic models, very little is known about the epigenetic regulation of these MAPK pathway genes, individually or collectively, in humans. During signaling, communication is known to exist between the MAPK network (upstream) and associated players (downstream) like NF-κB [24]. Despite some cell-based studies which have found that PM causes expression of NF-κB-related genes and oxidant-dependent activation of NF-κB in vitro [25], and that air pollution particles activate NF-κB on contact with airway epithelial cell surfaces in a rodent tracheal explant model [26], in vivo data are lacking.

Our results demonstrate, for the first time, that methylation signals in certain (epi)gene clusters of the MAPK pathway are significantly associated with ambient air pollution.
exposure in vivo. Within the MAPK pathway, DNA methylation showed responses in genes that were specific to BC and BC with sulfates. Strikingly, the combination of BC and sulfates induced DNA methylation responses that were wholly different from those detected in BC or sulfates alone. These new data, therefore, provide evidence that the biological effects of airborne particles on the human epigenome may vary depending on emission source as well as on the combination of PM components [10,14,15]. In our analyses, no significant associations were found with the NF-κB pathway.

**Methods**

**Human cohort description**

The Normative Aging Study (NAS) is a longitudinal study of human aging in Eastern Massachusetts, established in 1963 by the Veterans Administration [13,33-36]. Community-dwelling men from the greater Boston metropolitan area were screened at entry and accepted into the study if they had no prior history of heart disease, hypertension, diabetes mellitus, cancer, peptic ulcer, gout, recurrent asthma, bronchitis, or sinusitis. Between 1963 and 1968, a total of 2,280 men were enrolled, ranging in age from 21 to 80 years (mean = 42 years) at entry. Since their enrollment, the participants have undergone comprehensive clinical examinations at 3–5 year intervals. As part of those examinations, many clinical measures have been obtained, ranging from blood pressure (systolic and diastolic) and smoking status to complete blood count (CBC) data, as described previously [37,38]. In the NAS, DNA has been extracted from leukocytes and stored in all visits since 1999. We conducted an epigenome-wide scan of the promoter regions of ~19,000 genes on 141 subjects from the NAS. The subjects were selected based on having sufficient DNA for the assay, while leaving DNA for subsequent studies.

Importantly, in our analyses, five blood cell proportions (obtained from the CBC data) were included as covariates in our models to account for any appreciable changes in blood composition: lymphocytes, neutrophils, monocytes, basophils, and eosinophils [37,38]. Moreover, to slightly increase our statistical power, blood cell proportions for two participants were derived by single imputation [39-41], given that Infinium HumanMethylation450 BeadChip array data (from another study [20]) were available in the absence of CBC data. Computationally, we used the minfi package function “estimateCellCounts(),” which applies the regression-calibration approach from Houseman et al. [39,42], using the flow-sorted end members from Reinius et al. [43]. Furthermore, to estimate those cell proportions that are not part of the Houseman et al. study—i.e., subdividing granulocytes into neutrophils, eosinophils, and basophils—we multiplied the estimated proportion of granulocytes obtained from the Houseman method by the mean proportion among NAS participants with measured proportions (e.g., % neutrophils among total granulocytes).

**Chromatin immunoprecipitation (ChIP) and DNA methylation microarray**

DNA samples were hybridized to the RefSeq 385K Promoter tiling array (Roche NimbleGen, Madison, WI) representing the promoter regions of all well-characterized genes in the RefSeq database (RefSeq genes with NM Prefix), in addition to all of the UCSC-annotated CpG islands. The array coverage is based on 50-75mer probes with approximately 100 bp spacing, depending on the sequence composition of the region. Sample immunoprecipitation, labeling, hybridization and data extraction were all performed according to standard procedures optimized by Roche-NimbleGen, as previously reported by Selzer et al. [44].

High-quality genomic DNA (~5 μg) was isolated from blood buffy coat using QiAmp DNA blood kits (QIAGEN, Hilden, Germany) and digested with 24U Mse I (5′-T▼CAATTAA) enzyme (New England BioLabs) to produce small fragments of approximately 200 bp–1 kb. This fragmented DNA was heat-denatured to produce single-stranded DNA, then immunoprecipitated using an anti-5mC (Abcam-ab10805) monoclonal mouse antibody. Methylated DNA immunoprecipitated (MeDIP) fragments were then heat-denatured for 10 min at 95°C and immediately cooled on ice. Immune complexes were captured with Protein-A agarose bead slurry (Invitrogen-15918-014) and washed to remove non-specifically-bounded material. Following elution of bound complexes, MeDIP samples were purified with phenol-chloroform:isoamyl alcohol and ethanol precipitation in a ~80°C freezer for 30 min. After centrifugation, the supernatant was carefully removed, and the pellet was washed with cold 70% ethanol and then centrifuged again to remove residual supernatant. MeDIP samples were completely air dried and re-suspended in 30 μl of 10 mM Tris HCl (pH 8.5). Fragments were amplified by whole-genome amplification (GenomePlex Complete Whole Genome Amplification [WGA2] Kit, Sigma-Aldrich). Experimental and total DNA samples were labeled using 9mer primers, with Cy3 and Cy5 dyes attached via Klenow labelling (50 units/μL, New England BioLabs). The labeled experimental IP and total DNAs were co-hybridized to the array for 16–20 hours, washed, and scanned by the Roche NimbleGen Service Laboratory (Reykjavik, Iceland). The intensity ratio of IP to total DNA was used to identify DNA methylation.

**Ambient air pollution modeling**

Continuous air pollution concentrations were measured at a Harvard School of Public Health monitoring site located on the Francis A. Countway Library of Medicine rooftop (10 Shattuck Street, Boston, MA), 1 km from the
clinical examination site. Average pollution measures for the month prior to the blood draw were computed. BC, a marker for traffic particles weighted toward diesel particles, was measured using an aethalometer (Magee Scientific, Berkeley, CA), and PM$_{2.5}$ was measured using a Tapered Element Oscillating Microbalance (model 1400A; Rupprecht & Pataschnick Co., East Greenbush, NY), operated at 50 degrees with two 4 liter per minute PM$_{2.5}$ impacters before the inlet. From September 25, 1999 to February 2, 2004, particulate sulfate was measured using the Harvard/EPA Denuder System (HEADS), which samples inorganic gaseous and particulate species in the air. From January 1, 2003 through 2007, daily particulate filter samples were analyzed, by X-ray fluorescence (XRF) spectroscopy, for elemental components. From these samples, we multiplied the mass of sulfur by three to obtain the mass of sulfate. For the days when both HEADS impacters and XRF were in operation, we used linear regression and determined that the measurements had a slope of 1 and $R^2 > 0.9$, indicating a high correlation between the two monitoring methods. XRF measurements were used during this period of overlap. These sulfate particles are secondary, long-range particles primarily from coal-burning power plants.

Normalization and pre-processing of DNA methylation data
We normalized the raw methylation intensities (log2 green vs. red channel ratio) for each probe by subtracting the overall median and then dividing by the probe’s GC-content specific standard deviation, which is the standard deviation of all the probes whose sequence has the same number of G and C nucleotides as the target probe. We then smoothed the normalized scores using a local linear kernel smoother, as in Fan and Gijbels (1996), over the probe locations. For a given gene, its methylation score was then calculated by taking the area under the smoothed curve, truncated at zero, over a 500 bp window around the transcription start site of the gene, and then dividing by the percentage of CpG dinucleotides in the DNA sequence within the window and the number of probes having positive scores within the window.

Biostatistical methods and pathway analysis
We identified the genes associated with the MAP kinase and NF-kB signaling pathways using the BioCarta reference website. Since the outcome (gene methylation) is high dimensional, we employed a canonical-correlation-analysis (CCA)-based approach, which is a type of extension of the usual regression model for multiple outcomes. Under this approach, to adjust for confounders, we first regressed each of the exposures of interest, and the outcomes of interest, on the set of confounders, and then used the residuals from these regressions in the analysis.

For each gene pathway, we performed three analyses. We studied the association of the gene methylations in the pathway with (i) BC exposure alone, adjusting for age, blood pressure, smoking status, blood cell composition, and sulfate exposure as confounders, (ii) the association of gene methylation with sulfate alone, adjusting for the aforementioned confounders and BC exposure, and (iii) joint association modeling in which we studied the association of the pathway with both BC and sulfate exposure jointly (while adjusting for all other confounders). Our goal was to identify exposure-specific effects, but also to identify the effect of air pollution in the more realistic scenario in which people are exposed to both pollutants.

As mentioned above, we used a sparse forward stepwise-CCA method to identify specific genes that contribute to the association between the exposure and the DNA methylation status in the gene pathway. In short, the exposure set is held fixed, while at each step, the gene methylation score that contributes most to the association (between the gene set and the exposure) is selected. Genes are added to the set until a score is maximized. The association between the exposure and methylations is measured using the canonical correlation between the two sets, namely X and Y. The canonical correlation is given by $\text{corr}(Xa, Yb)$, where a and b are weight vectors (also called loading vectors), with lengths representing the number of measures in the exposures (pollutants) as “set X,” and the outcomes (gene methylations) as “set Y,” which are calculated to maximize the canonical correlation under the constrain $a^\top \text{cov}(X)a = b^\top \text{cov}(Y)b = 1$. Each entry in a is a weight corresponding to a specific pollutant (exposure), and each entry in b is a weight corresponding to a specific gene methylation measure. The larger the weight (in absolute value), the larger the influence of the variable it represents on the canonical correlation.

The score used as a criterion to select genes was the empirical CIC (Correlation Information Criterion), which takes the correlation between the identified set of genes and the exposures and removes the 99th percentile of this distribution under the null. This distribution was determined by 1000 random samples for each combination of 1 or 2 exposures (depending on the particular exposure model of interest) and any number of “outcomes” (1, 2, 3, 4, ..., number of genes in the pathway under study).

To test the significance of the canonical correlation between the set selected (set of “outcomes” and the set of exposures) we used the Wilks’ Lambda tests statistic, applied with a permutation procedure. The Wilks’ Lambda given by the ratio $\text{det}(\text{cov}(X,Y))/(\text{det}(\text{cov}(X))\text{det}(\text{cov}(Y)))$, where $\text{det}(C)$ is the determinant of a matrix C, and $(X,Y)$ is the data matrix of both the exposures X and the outcomes Y, is used to test the null hypothesis of no association between two data matrices X and Y. It cannot be used on the selected data because the variable selection...
method (stepwise-CCA) was applied, changing the null distribution of the correlation between the outcomes and the exposures. Thus we permuted the exposure data while holding the methylation scores for all genes in the pathway, fixed 3000 times, then performed the stepwise analysis on the permuted data, and finally computed the Wilks’ Lambda statistic for the exposure and genes identified by the stepwise-CCA method. The P-value for the true exposure is the proportion of permuted exposures with a lower Wilks’ Lambda P-value than the one for the true exposure.

Bioinformatic visualization of an integrated MAPK network
Pathways from the hiPathDb [45] Integrated database were selected to link all BioCarta MAPK genes; these pathways included the KEGG [46,47] ErbB signaling and MAPK signaling pathways (path:hsa04010 and path:hsa0401, respectively), in addition to the BioCarta MAPK signaling pathway (pid p 100113 mapkpathway). The integrated pathways were downloaded in XML format and imported into Cytoscape [48]. All non-gene nodes were removed by hand, and the network was restricted to second neighbors of the BioCarta MAP kinase genes. Supernodes with only a single connection or only single direction edges were removed, and redundant supernode-to-supernode connections were simplified whenever possible; self-directed loops and redundant edges of the same direction were also removed. Nodes were annotated by fill-color with methylation coefficients on a truncated scale of −1.5 to 1.5, as shown in the legend.

Disease ontologies of selected genes
The MAPK (epi)gene hits were analyzed for disease association without regard to enrichment with the GeneAnswers library [49]. Selected genes (11 total) comprised those with non-zero methylation coefficients for any of the three exposure conditions (black carbon, sulfate, and multi-pollutant carbon and sulfate). GeneAnswers associates genes with disease using DOLite [50], a database based on the Disease Ontology [51], an open-source ontology for the semantic integration of biomedical data associated with human disease.

Results and discussion
Novel (epi)gene pathway-exposure assessment
We hypothesized that short-term exposure to environmental air pollution components, singly or in a multi-pollution paradigm, would be associated with blood DNA methylation alterations in known inflammation-linked gene networks, specifically the MAPK pathway, and possibly in a downstream target, NF-κB. To study this hypothesis, we implemented a multidisciplinary strategy with an established sparse stepwise canonical correlation analysis (stepwise-CCA) method [52,53], by which we were able to evaluate genes in the abovementioned pathways with respect to air pollution-related epigenetic alterations.

We first performed epigenome-wide scans of promoter regions for ~19,000 genes from 141 participants from the

Table 1 The 84 MAPK pathway-linked genes considered in our analyses

| Gene     | Gene     | Gene     |
|----------|----------|----------|
| 01 ATF2  | 38 MAP3K5| 73 TRADD |
| 02 CREB1 | 39 MAP3K6| 74 TGFBI |
| 03 CEBPA | 40 MAP3K7| 75 TGFBI |
| 04 CHUK  | 41 MAP3K8| 76 TGFBI |
| 05 DAIX  | 42 MAP3K9| 77 TGFBR1|
| 06 ELK1  | 43 MAP4K1| 78 FOS   |
| 07 GRB2  | 44 MAP4K2| 79 HRAS  |
| 08 IRKB  | 45 MAP4K3| 80 MYC   |
| 09 JUN   | 46 MAP4K4| 81 ARAF  |
| 10 MKNK1 | 47 MAP4K5| 82 BRAF  |
| 11 MKNK2 | 48 MAP4K6| 83 RAF1  |
| 12 MAPK1 | 49 MAP4K7| 84 RELA  |
| 13 MAPK10| 50 MAP4K8|
| 14 MAPK11| 51 MAP4K9|
| 15 MAPK12| 52 MAP5K5|
| 16 MAPK13| 53 MAP5K6|
| 17 MAPK14| 54 MAP5K7|
| 18 MAPK3 | 55 MAP5K8|
| 19 MAPK4 | 56 MAP5K9|
| 20 MAPK6 | 57 MAP6D |
| 21 MAPK7 | 58 NFKB1 |
| 22 MAPK9 | 59 NFKBA |
| 23 MAP2K1| 60 PAK2  |
| 24 MAP2K2| 61 PAK1  |
| 25 MAP2K3| 62 RAC1  |
| 26 MAP2K4| 63 RAC2  |
| 27 MAP2K5| 64 RPS6KB1|
| 28 MAP2K6| 65 RPS6KB2|
| 29 MAP2K7| 66 RPS6KA1|
| 30 MAP3K10| 67 RPS6KA2|
| 31 MAP3K11| 68 RPS6KA3|
| 32 MAP3K12| 69 RPS6KA4|
| 33 MAP3K13| 70 RPS6KA5|
| 34 MAP3K14| 71 RPS6KA6|
| 35 MAP3K2 | 72 RPS6KA7|
| 36 MAP3K3 | 73 RPS6KA8|
| 37 MAP3K4 | 74 RPS6KA9|

To our knowledge, this is the first human study to systematically evaluate MAPK gene-promoter methylation using a wide collection of upstream and downstream pathway components with respect to ambient air pollution exposures.
Normative Aging Study [37,54,55]. Briefly, our previously validated workflow (described in Methods) included DNA fragmentation, methylated DNA immunoprecipitation (MeDIP) capture, DNA purification, hybridization to the RefSeq 385K Promoter tiling array for methylomic detection, and (epi)gene clustering by methylation status via the stepwise-CCA algorithm [53]. Using the BioCarta reference database [56], we identified methylation changes in the promoters of 84 genes from the MAP kinase-signaling pathway, which are listed in Table 1. In humans, this large gene network has very rarely been evaluated within the context of air pollution-related effects in vivo.

We normalized the raw methylation intensities for each array probe separately, smoothed these normalized intensities for each probe, and finally constructed gene-specific methylation scores based on the intensities of neighboring probes around the transcription start site (TSS) of each gene, as previously described [53]. We used these gene-specific methylation scores, or “weights,” to investigate the association between DNA methylation and environmental exposures.

The gene methylation weights found by the stepwise-CCA algorithm are provided throughout all of the data tables: A large weight in the absolute value is interpreted as a stronger effect. A positive value indicates increased methylation (hypermethylation) with increased exposure to BC (or sulfate, etc.), and a negative weight implies decreased methylation (hypomethylation) with increased exposure. These weighted coefficients, however, should not be interpreted individually—but rather as a combined (epi)gene cluster “hit,” constituting a set of bi-directional epigenetic marks (+ or −) among pathway genes associated with an exposure model.

**Human cohort characteristics**

Table 2 shows a summary of the characteristics of the subgroup of NAS participants from our methylation study for which sulfate measures were available (n = 90) vs. the larger set. The characteristics reported for this subgroup—BC, age, and other measures—were almost identical to those calculated using information from the entire participant list (n =141). The air pollution measures are averages of the 30 days of ambient BC and/or sulfate concentration prior to the day that blood was collected for methylomic analysis. This time window was selected to balance the evidence of effects of short-term air pollution exposure on the cardiovascular and respiratory systems, with the common understanding that at least some potential environmental effects on DNA methylation require days (or weeks) to become apparent [57-59].

Our data included 84 MAPK pathway genes (listed in Table 1), from which we identified exposure-specific (epi) gene sets (Table 3) based on three pollution models. We identified 11 genes whose methylation status was associated with BC exposure (P-value 0.04) after adjusting for relevant confounders: age, sulfate exposure, blood-cell-type

| NAS cohort description | All participants | With sulfate measures |
|------------------------|------------------|----------------------|
| Age (years)            | n = 141 people   | n = 90 people        |
|                        | Median 73; Range 56–88; SD ±6.8 | Median 73; Range 58–88; SD ±6.6 |
| Blood pressure (mmHg)  | n = 141          | n = 90               |
| Systolic               | Median 130; Range 87–188; SD ±16.13 | Median 129.5; Range 87–188; SD ±15.65 |
| Diastolic              | Median 79; Range 54–98; SD ±9.1 | Median 78; Range 55–96; SD ±9.11 |
| Exposure (μg/m³)       | n = 141          | n = 90               |
| BC, 30-days averaged   | Mean 0.84; SD ±0.16 | Mean 0.83; SD ±0.15 |
| Sulfate, 30-days averaged | N/A             | Mean 3.06; SD ±0.79 |
| Smoking status (n)     | n = 141          | n = 90               |
| Never                  | 53               | 32                   |
| Ever                   | 7                | 5                    |
| Former                 | 81               | 53                   |
| Blood count (%)        | n = 136          | n = 88               |
| Lymphocytes            | Median 26; Range 6–39; SD ±6.66 | Median 26; Range 6–38; SD ±6.92 |
| Neutrophils            | Median 62; Range 45–86; SD ±7.71 | Median 63; Range 45–86; SD ±7.78 |
| Monocytes              | Median 9; Range 4–14; SD ±1.92 | Median 8; Range 4–14; SD ±1.96 |
| Basophils              | Median 1; Range 0–2; SD ±0.51 | Median 1; Range 0–2; SD ±0.52 |
| Eosinophils            | Median 3; Range 0–11; SD ±1.9 | Median 3; Range 0–10; SD ±1.83 |

Relevant units are supplied in the left-hand side for each characteristic, with median (or mean), range, standard deviation, and sample size indicated.
proportions (derived from the CBC data), smoking status, and blood pressure (see Methods). Association analysis between sulfate exposure and DNA methylation in the MAPK pathway identified 12 genes after adjusting for BC exposure and the other aforementioned confounders, yet this group fell below the level of statistical significance (P-value 0.10) likely due to a smaller sample size (n = 90).

Finally, multi-pollutant analysis of DNA methylation associated jointly with sulfate and BC exposure yielded 14 significant genes (P-value 0.01).

Importantly, adjusting for blood cell proportions barely influenced the effect estimates, so in Table 3, we report both P-values, with and without the cell proportion adjustment (while including all of the other confounders). In Additional file 1, A-D, we present clustered heatmaps of the correlations between the full list of 84 MAPK pathway genes (from Table 1) vs. the cluster of 27 gene hits (Table 3), before and after adjusting for all confounders.

To aid best in visualizing all possible relationships amongst our various MAPK hits from Table 3, we next constructed an annotated Venn diagram (Figure 1). Interestingly, no genes occupied the union of all 3 exposure models, and minor overlap was observed across any 2 given exposure combinations. For example, only two genes—MAPK13 and MAPKAPK2—overlapped between BC and sulfate, yet their methylation status reversed directionality. Strikingly, it is evident that the genes associated with the multi-pollutant model (BCS) are not simply the aggregate collection of hits found in the BC and sulfate models. In fact, it behaves like an entirely novel exposure combination, suggesting that multi-pollutant exposures may impact the epigenome in disparate ways, unlike their single-exposure counterparts.

This key observation is also consistent with our past work to identify new methylomic changes in gene promoters related to the asthma pathway: Only a single allergy-specific receptor gene, FCER1G, exhibited methylation changes associated with both BC and sulfate exposure [53]. All other significant immune-gene clusters were either specific to BC exposure alone (6 genes) or to sulfate exposure (4 genes). Functionally, however, all of these genes did share a common biological network across immune cells and the bronchus, which could be easily visualized.

Visualization of MAPK pathway components
In the present work, however, our 27 (epi)gene hits in the MAPK pathway do not share any obvious physiological link, collectively, within some unified cardiopulmonary pathway relevant to air pollution—thus perhaps constituting a novel epigenetic/signaling “crossroads” of exposure-associated genes relevant to the processing of environmental PM signals in vivo. To help visualize this dynamic system, we next used an integrated bioinformatic approach to overlay our DNA methylation coefficients (from Table 3) onto an expanded MAPK signaling map. In Figure 2, we can best appreciate the following points: (i.) our initial set of 84 MAPK genes (dark circles) map widely across the broadened MAPK signaling system, thereby ensuring that most branches of the network were aptly queried via our stepwise-CAA method; (ii.) the distribution of exposure-

| ID       | Weighted coefficients by exposure model | Black Carbon (BC) | Sulfates (S) | Multi (BCS) |
|----------|-----------------------------------------|------------------|--------------|-------------|
| VISCA2   | 0                                      | 0.33             | 0            |
| MAPK10   | 0                                      | 0.33             | -0.22        |
| MAPK21   | 0                                      | 0.42             | -0.3         |
| MAPK5    | 0                                      | 0.22             | -0.3         |
| MAPK11   | 0                                      | -0.4             | 0            |
| MAPK14   | 0                                      | 0.37             | -0.1         |
| MAPK2    | -0.25                                  | 0                | 0            |
| MAPK6    | -0.37                                  | 0                | 0            |
| MAPK7    | 0                                      | 0.5              | 0.1         |
| MAPK1    | 0                                      | -0.19            | 0.31         |
| MAPK3    | 0.42                                   | 0                | 0            |
| MAPK4    | 0                                      | 0.57             | -0.61        |
| MAPKAPK2 | -1.36                                  | 0.62             | 0            |
| MEF2A    | 0.32                                   | 0                | 0.33         |
| PAK1     | 0                                      | 0                | -0.57        |
| RPS6KB1  | 0                                      | -0.28            | 0.09         |
| RPS6KB2  | 0.27                                   | 0                | 0            |
| RPS6KB3  | 0                                      | -0.33            | 0            |
| SHC1     | 0.33                                   | 0                | 0            |
| STAT1    | 0                                      | -0.49            | 0.43         |
| TGFBI    | 0                                      | 0.28             | 0            |
| MYC      | 0                                      | 0                | 0.23         |
| RELA     | 0.27                                   | 0                | 0.48         |
| Canonical correlation | 0.67 | 0.73 | 0.73 |
| P-value* | 0.02                                   | 0.04             | 0.05         |
| P-value** | 0.04         | 0.10             | 0.01         |

We identified a cluster of 27 MAPK gene hits, and their corresponding weights (as coefficients after 3K permutation tests) are shown. At the bottom of the table, P-values for each model are labeled as either adjusted for all covariates, with two asterisks (**), or for all covariates except blood cell proportions, with one asterisk (*).
specific MAPK hits across the multi-pollution model (Figure 2c) is not simply an aggregate of signaling components found in the two other models, BC (2a) and sulfates (2b).

In light of these data, therefore, we next hypothesized that perhaps some of these genes (across all 3 exposure models) would be linked to relevant human disease outcomes already known to be exacerbated/modulated by air pollution—e.g., heart disease [60,61], atherosclerosis [62-64], stroke [65,66], cancers [67-69], etc.—among various other disease contexts. Indeed, biocomputational profiling (Figure 3) confirmed that 11 of our 27 hits (~41%) were in fact highly associated with a wealth of PM-linked pathophysiological conditions (and to many other varied diseases), thereby underscoring that these genes may serve as an epigenetic/signaling nexus of exposure-related signals in vivo, a role not fully appreciated among their other biological functions.

Importantly, some individual connections between our MAPK gene hits and various air pollution exposure models do exist, upon examining a handful of in vitro cell-based and/or animal-exposure studies. This literature helps to further confirm the biological relevance of the MAPK pathway within air pollution models and paradigms. For example, the MAP2K1 protein (also known as MEK1, MAPKK1, or MKK1) was previously found to be important for induction of COX-2 protein expression via vanadate exposure in vitro using a human lung carcinoma cell line (A549) [70]. Moreover, another group reported that primary human pulmonary cells exposed to cigarette smoke extract (CSE), especially normal human bronchial smooth muscle cells, required MAPKAPK2/MK2 protein activation to promote pathological expression of pro-inflammatory Interlukin-8 [71]. Lastly, two key downstream transcription factors associated with MAPK signaling merit special consideration, STAT-1 and MYC. Vehicular-derived airborne nano-sized particulate matter (from Los Angeles, CA) that was re-aerosolized and administered to C57Bl/6 J male mice led to increased MYC expression in brain (cerebellum), liver, and lung tissue [72], and vanadium compounds present in PM, such as vanadium pentoxide, induced STAT-1 activation in lung myofibroblasts, which required upstream hydrogen peroxide and MAP kinase signaling activity [73].

In our study, the associations between the abovementioned MAPK pathway genes and our exposure models were so specific (and reproducible) that further attempts to find other significant (epi)gene clusters with the NF-κB
pathway, a downstream network often coupled to MAPK signaling, failed consistently, despite a ~36% overlap in gene sets across both pathways, as noted in Tables 4 and 5. We do recognize, however, that perhaps a higher sample size could have helped to facilitate an association, which may not have been possible given our current number of participants. Moreover, we also acknowledge other key limitations inherent to our work, given that the NAS is a male-only cohort and that we analyzed blood-cell DNA.

---

**Figure 2** Methylation coefficients of our (epi)gene hits within the broader MAPK signaling system. A nexus integrating all BioCarta MAP kinase genes to other previously studied MAPK networks is diagrammed, wherein nodes representing the genes within the BioCarta MAP kinase pathway (84 total) are outlined and labeled in dark black. Arrows indicate known direction of action. Methylation coefficients (from Table 3) are represented here in a scale from blue (negative values), to white (zero), to orange (positive values). For simplicity, both unmeasured values and zero are represented in white. Exposure-specific MAPK coefficients are shown across all three of our models: (a) black carbon; (b) sulfates; and (c) the multi-pollutant paradigm.

**Figure 3** Biocomputational profiling of disease-linked MAPK (epi)gene hits. A heatmap (left) of methylation coefficients for the three pollution paradigms—black carbon (BC), sulfates (S), and BC with sulfate (BCS)—and a corresponding disease ontology table (right) are shown. Only genes with an annotation within the disease ontology [51] are shown: i.e., 11 out of 27 MAPK genes (~41%). Heatmap colors represent the methylation coefficients, with negative values in blue, zero in white, and positive values in red. Rows within both the heatmap and the concept-map represent individual genes as noted. Columns within the disease ontology concept-map represent individual diseases within the disease ontology. Disease category hierarchies were collapsed, so diseases may represent subcategories of other diseases represented (e.g., “Breast cancer” and “Ovarian cancer” are both sub-categories of “Cancer”). Dots within the disease ontology table denote an association of a gene with a given pathology. Alternative shading is used to help demarcate columns.
**Table 4 Complete list of the NF-κB signaling pathway and associated genes (22 total) considered in this study**

| Gene   | Gene   |
|--------|--------|
| 01     | CHUK   |
| 02     | FADD   |
| 03     | IKKβ   |
| 04     | IKBKG  |
| 05     | IL1R1  |
| 06     | IL1A   |
| 07     | IRAK1  |
| 08     | MAP3K7B |
| 09     | MAP3K7T2 |
| 10     | TAB1   |
| 11     | MYD88  |
| 12     | NFκB1  |
| 13     | NFκBIA |
| 14     | RIPK1  |
| 15     | TRAF6  |
| 16     | TRADD  |
| 17     | TLR4   |
| 18     | TNFRSF1A |
| 19     | TNFRSF1B |
| 20     | TNFAIP3 |
| 21     | RELA   |
| 22     | RELB   |

| Weighted coefficients by exposure model | BC       | Sulfates (S) | Multi (BCS) |
|---------------------------------------|----------|--------------|-------------|
| Black Carbon (BC)                     |          |              |             |
| S                                    | 0        | 0.87         | 1.08        |
| MYD88                                | 0        | 0.96         | 0           |
| RELA                                 | 0        | 0.52         | 0           |
| TNFRSF1A                             | 0        | 0.82         | 0           |
| TNFRSF1B                             | 1.15     | 0            | 0           |
| Canonical correlation                 | 0.23     | 0.28         | 0.24        |
| P-value*                             | 0.08     | 0.61         | 0.81        |
| P-value**                            | 0.98     | 0.54         | 0.84        |

P-values for each model are indicated as either adjusted for all covariates (**) or all covariates except blood cell proportions (*).
prevention when it is most effective, i.e. in early stages or, among exposed individuals, even before any subclinical cardiopulmonary impairment is detectable. As methods for accurate epigenomic profiling become increasingly available and affordable, these approaches may allow for better multi-pollutant exposure assessment to be brought to numerous environmental health studies, as well as to preventive settings where exposure data are lacking or where funds and opportunities for expensive personal monitoring are limited. As the age of the US population increases, such efforts will have the potential to help millions of individuals in the prevention of air pollution-related pathophysiological outcomes and their sequelae, particularly among vulnerable people, and to narrow health disparities and promote equity.

Additional file

Additional file 1: Clustered heatmaps of the observed correlations between our full MAPK gene set and those shown to be associated with air pollution. (A) Unadjusted DNA methylation coefficients were used to cluster the 84 genes listed in Table 1; (B) clustering of adjusted methylation coefficients, after accounting for all relevant confounders included in this study—age, blood pressure, smoking status, blood cell proportions, etc. (described in Methods); similarly, the same is shown for the 27 MAPK gene hits from Table 3, prior to adjusting for all confounders (C) and afterwards (D).

Abbreviations

PM: Particulate matter; BC: Black carbon; S: Sulfate; BCS: Black carbon and sulfate; CBC: Complete blood count; MAPK: Mitogen-activated protein kinase; stepwise-CCA: Stepwise-canonical correlation analysis; NF-κB: Nuclear factor kappa-light-chain-enhancer of activated B cells; SmcC: S-methyl-cytosine; NAS: Normative Aging Study.

Competing interests

The authors declare that they have no competing interests or other relevant affiliations within organizations or other entities having a financial interest that could conflict with the subject matter and/or materials discussed in this manuscript. Neither the USEPA nor NIEHS expressly endorse the purchase of any commercial products and/or services mentioned/used in this publication.

Authors’ contributions

JC, TS, and JT participated in the experiments, and JC wrote the manuscript with intellectual and editorial assistance from AAB, JS, and AM. Together, JC and TS analyzed the data, and JC revised and submitted the manuscript. XL, LC, and BC provided additional biostatistical expertise, and PV designed the human cohort study. All authors read and approved the final version of the manuscript.

Acknowledgements

The Veterans Affairs (VA) Normative Aging Study (NAS) is supported by the Cooperative Studies Program/Epidemiology Research and Information Centers of the US Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research Information Center (MAVERIC), Boston, MA. The National Institute of Environmental Health Science (R01ES015172, R01ES021733, and P30ES00022 for AAB & JS, R00ES01774 for AM, and T32 ES07069 and 1F32ES024068-01 for JC) and the National Cancer Institute (NCI R37CA076404 & P01CA134294 for TS, AM, and XL) supported this work. Moreover, this publication was also made possible by a USEPA-grant (RD 83479801) and a Harvard-NIEHS Center for Environmental Health Pilot Grant (ES050002 to AM). All authors explicitly state that they obtained appropriate Institutional Review Board (IRB) approval and have followed the principles outlined in the Declaration of Helsinki for all human experimental investigations; informed consent was obtained from all study participants. The authors thank Edgar Antonio Carmona for helping us to create, edit, and format the various images used throughout this manuscript.

Author details

1. Laboratory of Human Environmental Epigenetics, Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA. 2. Exposure, Epidemiology, and Risk Program, Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA. 3. Program in Quantitative Genomics, Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA. 4. Center for Health Bioinformatics, Harvard School of Public Health, Boston, MA, USA. 5. Department of Clinical Sciences and Community Health, Universit degli Studi di Milano, Milan, Italy. 6. Department of Statistics, North Carolina State University, Raleigh, NC, USA. 7. VA Normative Aging Study, Veterans Affairs Boston Healthcare System and the Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA.

Received: 2 May 2014 Accepted: 6 October 2014
Published: 13 November 2014

References

1. Anderson JD, Thundiyil JG, Stoolbach A: Clearing the air: a review of the effects of particulate matter air pollution on human health. J Med Toxicol 2012, 8:166–175.
2. WHO: Burden of disease from Ambient Air Pollution for 2012. 2014. http://www.who.int/phe/health_topics/outdoorair/databases/ AAP_Bd_results_March2014.pdf.
3. Kuehn BM: WHO: More than 7 million air pollution deaths each year. JAMA 2014, 311:1486.
4. Delfino RJ, Staimer N, Vazni JD: Air pollution and circulating biomarkers of oxidative stress. Air Qual Atmos Health 2011, 4:37–52.
5. Mordukhovich J, Wilker E, Suh H, Wright R, Sparrow D, Vokonas PS, Schwartz J: Black carbon exposure, oxidative stress genes, and blood pressure in a repeated-measures study. Environ Health Perspect 2009, 117:1767–1772.
6. O’Neill MG, Veyes A, Zanobetti A, Samat JA, Gold DR, Economides PA, Horton ES, Schwartz J: Diabetes enhances vulnerability to particulate air pollution-associated impairment in vascular reactivity and endothelial function. Circulation 2005, 111:2913–2920.
7. McConnell R, Islam T, Shankardass K, Jerrett M, Lummann F, Gilliland F, Gauderman J, Avol E, Kunzli N, Yao L, Peters J, Berhane K: Childhood incident asthma and traffic-related air pollution at home and school. Environ Health Perspect 2010, 118:1021–1026.
8. Carlsten C, Dybunndo A, Becker A, Chan-Yeung M, Brauer M: Traffic-related air pollution and incident asthma in a high-risk birth cohort. Occup Environ Med 2011, 68:291–295.
9. Peters A, Dockery DW, Heinrich J, Wichmann HE: Medication use modifies the health effects of particulate sulfate air pollution in children with asthma. Environ Health Perspect 1997, 105:450–455.
10. Fann N, Bell ML, Walker K, Hubbell B: Improving the linkages between air pollution epidemiology and quantitative risk assessment. Environ Health Perspect 2011, 119:1671–1675.
11. Weidenhan D, Plass C: The evolving epigenome. Hum Mol Genet 2013, 221:R6–R16.
12. Holloway JW, Savarimuthu Francis S, Feng KM, Yang JA: Genomics and the respiratory effects of air pollution exposure. Respirology 2012, 17:590–600.
13. Baccarelli A, Wright RO, Bollati V, Tarantini L, Litonjua AA, Suh HH, Zanobetti A, Sparrow D, Vokonas PS, Schwartz J: Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 2009, 179:572–578.
14. Baccarelli A, Bollati V: Epigenetics and environmental chemicals. Clin Opin Pediatr 2009, 21:243–251.
15. Breton CV, Marutani AN: Air pollution and epigenetics: recent findings. Curr Envi Health Rep 2014, 1:35–45.
16. Wright R, Saul PA: Epigenetics and primary care. Pediatr Res 2013, 73:216–223.
17. Wright RJ, Brunst KJ: Programming of respiratory health in childhood: influence of outdoor air pollution. Curr Opin Pediatr 2013, 25:232–239.
18. Blatter A, Farnham PJ: Cross-talk between site-specific transcription factors and DNA methylation states. J Biol Chem 2013, 288:34287–34294.
19. Felsenfeld G: A brief history of epigenetics. Cold Spring Harb Perspect Biol 2014, 6: http://cshperspectives.cshlp.org/content/6/1/a018200.
20. Carmona JJ, Iiz R, Just AC, Burpal J, Binder AM, Hutchinson J, Hoffmann O, Schwartz J, Baccarelli A, Michels KB: Comparison of multiplexed reduced representation bisulfite sequencing (mRRBS) with the 450 K Illumina Human BeadChip: from concordance to practical applications for methylomic profiling in epidemiologic epigenetic studies. Epigenetics Chromatin 2013, 6: http://www.epigenetichuman.com/content/6/5/3/36.
21. Palacios OM, Carmona JJ, Michan S, Chen KY, Manabe Y, Ward JL 3rd, Goodyear LJ, Tong Q: Diet and exercise signals regulate SIRT3 and activate AMPK and PGC-1alpha in skeletal muscle. Aging (Albany NY) 2009, 1:771–783.

22. Yang H, Yang T, Baur JA, Perez E, Matsui T, Carmona JJ, Lamming DW, Souza-Pinto NC, Sohr VA, Rosenzweig A, de Cabo R, Sause AA, Sinclair DA: Nutrient-sensitive mitochondrial NAD+ levels dictate cell survival. Cell 2007, 130:1095–1107.

23. Shukla A, Timblin C, BeruBe K, Gordon T, McKinney W, Driscoll K, Vacek P, Yang H, Yang T, Baur JA, Perez E, Matsui T, Carmona JJ, Lamming DW, Palacios OM, Carmona JJ, Michan S, Chen KY, Manabe Y, Ward JL 3rd, Carmona et al. Environmental Health Perspectives 2002, 102:200–208.

24. Mondal K, Stephen Haskill J, Becker S: Adhesion and pollution particle-induced oxidant generation is neither necessary nor sufficient for cytokine induction in human alveolar macrophages. Am J Respir Cell Mol Biol 2002, 26:200–208.

25. Shukla A, Timblin C, BeruBe K, Gordon T, McKinney W, Driscoll K, Vacek P, Mossman BT: Inhaled particulate matter causes expression of nuclear factor (NF)-kappaB-related genes and oxidant-dependent NF-kappaB activation in vitro. Am J Respir Cell Mol Biol 2000, 23:182–187.

26. Chung A, Xie C, Wang X, Vincent R, Wang RD: Air pollution particles activate NF-kappaB on contact with airway epithelial cell surfaces. Toxicol Appl Pharmacol 2005, 208:37–45.

27. Mossman BT, Lounsbury KM, Reddy SP: Oxidants and signaling by mitogen-activated protein kinases in lung epithelium. Am J Respir Cell Mol Biol 2006, 34:666–669.

28. Huang YC, Ghio AJ: Vascular effects of ambient pollutant particles and metals. Curr Vasc Pharmacol 2006, 4:199–203.

29. Kyriakis JM, Avruch J: Mammalian MAPK signal transduction pathways activated by stress and inflammation: a 10-year update. Physiol Rev 2012, 92:699–737.

30. Wu W, Bromberg PA, Samet JM: Zinc ions as effectors of environmental oxidative lung injury. Free Radic Biol Med 2013, 65:57–69.

31. Plotnikov A, Zehrai E, Procaccia S, Sieger R: The MAPK cascades: signaling components, nuclear roles and mechanisms of nuclear translocation. Biochim Biophys Acta 2013, 1831:1619–1633.

32. Bandypadhyay S, Chiang CY, Silvastava J, Gersten M, White S, Bell R, Kurschner C, Martin C, Smith M, Sahasrabudhe S, Barber DL, Chanda SK, Ideker T: A human MAP kinase interactome. Nat Methods 2010, 7:801–805.

33. Baja ES, Schwartz JD, Coull BA, Wellenius GA, Vokonas PS, Suh HH: Structural equation modeling of parasympathetic and sympathetic function decline and COPD. DeMeo DL, Litonjua AA, Silverman EK, Connett JE, Anthonisen NR, Wise RA, Sunyer J, Tittanen P, Wojtyniak B, Zmirou D, Schwartz J: Associations of IL6 polymorphisms with lung function in cities in the APHEA2 project. Environ Health Perspect 2009, 117:982–987.

34. Wilker EH, Alexeef SE, Poon A, Litonjua AA, Sparrow D, Vokonas PS, Suh HH: The temporal pattern of respiratory and heart disease mortality in response to airborne particulate matter associated with methylation pattern in the asthma pathway. Epigenomics 2013, 5:147–154.

35. Lepeule J, Bind MA, Baccarelli AA, Kantrakis P, Tarantini L, Lijonja 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 2012, 4:819–336.

36. Baccarelli AA, Cantrone L, Coull BA, Baccarelli A, Schwartz JD, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.

37. Lijonja A, Schwartz J, Tarantini L, Vokonas PS, Koutrakis P, Schwartz J: From disease ontology to disease-ontology lite: statistical methods to adapt a general-purpose ontology for the test of gene-ontology associations. Environ Health Perspect 2012, 120:658–663.

38. Houseman EA, Accomando WP, Korten DC, Chistensm BC, Marst CJ, Stallings RL: Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chromosomes Cancer 2005, 44:305–319.

39. Yu N, Seo J, Plo K, Yang J, Park J, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.

40. Kanehisa M, Goto S: KEGG: kyoto encyclopedia of genes and genomes. Nucl Acids Res 2000, 28:270–37.

41. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: Cytoscape: a software environment for integrated models of molecular interaction networks. Genome Res 2003, 13:2498–2504.

42. Feng G, Du P, Kibbe WA, Lin S: GeneAnswers, integrated interpretation of genes. Bioconductororg 2014. http://www.bioconductor.org/packages/release/bioc/vignettes/GeneAnswers/inst/doc/geneAnswers.pdf.

43. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen SE, Greco D, Stallings RL: Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chromosomes Cancer 2005, 44:305–319.

44. Yu N, Seo J, Plo K, Yang J, Park J, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.

45. Kanehisa M, Goto S: KEGG: kyoto encyclopedia of genes and genomes. Nucl Acids Res 2000, 28:270–37.

46. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: Cytoscape: a software environment for integrated models of molecular interaction networks. Genome Res 2003, 13:2498–2504.

47. Sofer T, Baccarelli A, Schwartz JD, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.

48. Baja ES, Schwartz JD, Coull BA, Wellenius GA, Vokonas PS, Suh HH: Structural equation modeling of parasympathetic and sympathetic function decline and COPD. DeMeo DL, Litonjua AA, Silverman EK, Connett JE, Anthonisen NR, Wise RA, Sunyer J, Tittanen P, Wojtyniak B, Zmirou D, Schwartz J: Associations of IL6 polymorphisms with lung function in cities in the APHEA2 project. Environ Health Perspect 2009, 117:982–987.

49. Wilker EH, Alexeef SE, Poon A, Litonjua AA, Sparrow D, Vokonas PS, Suh HH: The temporal pattern of respiratory and heart disease mortality in response to airborne particulate matter associated with methylation pattern in the asthma pathway. Epigenomics 2013, 5:147–154.

50. Lepeule J, Bind MA, Baccarelli AA, Kantrakis P, Tarantini L, Lijonja 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 2012, 4:819–336.

51. Baccarelli AA, Cantrone L, Coull BA, Baccarelli A, Schwartz JD, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.

52. Sofer T, Baccarelli A, Cantrone L, Coull BA, Baccarelli A, Schwartz JD, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.

53. Baccarelli AA, Cantrone L, Coull BA, Baccarelli A, Schwartz JD, Kim WK, Lee S: NPATHDB: a human-integrated pathway database with facile visualization. Nucleic Acids Res 2012, 40:7979–D802.
64. Rivera M, Basagana X, Aguilera I, Foraster M, Agis D, de Groot E, Perez L, Mendez MA, Bousu L, Targa J, Ramos R, Sala J, Marrugat J, Elosua R, Kardil N: The association between air pollution and subclinical atherosclerosis: Rivera et al. respond. Environ Health Perspect 2014, 122:A8–A9.

65. Chiu HF, Yang CY: Short-term effects of fine particulate air pollution on ischemic stroke occurrence: a case-crossover study. J Toxicol Environ Health A 2013, 76:1188–1197.

66. Xiang H, Mertz KJ, Arena VC, Brink LL, Xu X, Bi Y, Talbott EO: Estimation of short-term effects of air pollution on stroke hospital admissions in Wuhan. China PloS One 2013, 8:e61168.

67. Vigotti MA, Mataloni F, Bruni A, Mincer C, Gianicolo EA: Mortality analysis by neighbourhood in a city with high levels of industrial air pollution. Int J Public Health 2014, 59:654–653.

68. Fajerstajn L, Veras M, Barrozo LV, Saldiva P: Air pollution: a potentially modifiable risk factor for lung cancer. Nat Rev Cancer 2013, 13:674–678.

69. Hu H, Dailey AB, Kan H, Xu X: The effect of atmospheric particulate matter on survival of breast cancer among US females. Breast Cancer Res Treat 2013, 139:217–226.

70. Chien PS, Mak OT, Huang HJ: Induction of COX-2 protein expression by vanadate in A549 human lung carcinoma cell line through EGF receptor and p38 MAPK-mediated pathway. Biochem Biophys Res Commun 2006, 339:562–568.

71. Moretto N, Bertolini S, Iadicicco C, Marchini G, Kaur M, Volpi G, Patacchini R, Singh D, Facchinetti F: Cigarette smoke and its component acrolein augment IL-8/ 0XCL8 mRNA stability via p38 MAPK/MK2 signaling in human pulmonary cells. Am J Physiol Lung Cell Mol Physiol 2012, 303i:1929–1938.

72. Zhang H, Liu H, Davies KJ, Sioutas C, Finch CE, Morgan TE, Forman HJ: Nrf2-regulated phase II enzymes are induced by chronic ambient nanoparticle exposure in young mice with age-related impairments. Free Radic Biol Med 2012, 52:2038–2046.

73. Wang YZ, Ingram JL, Walters EM, Rice AB, Santos JH, Van Houten B, Bonner JC: Vanadium-induced STAT-1 activation in lung myofibroblasts requires H2O2 and P38 MAP kinase. Free Radic Biol Med 2003, 35:845–855.

74. Lin CM, Kuo HW: Sex-age differences in association with particulate matter and emergency admissions for cardiovascular diseases: a hospital-based study in Taiwan. Public Health 2013, 127:828–833.

75. Faustini A, Staffogga M, Colais P, Berti G, Bisanti L, Cadum E, Cernigliaro A, Mallone S, Scamato C, Forastiere F, EpiAir Collaborative G: Air pollution and multiple acute respiratory outcomes. Eur Respir J 2013, 42:304–313.

76. Schwela D: Air pollution and health in urban areas. Rev Environ Health 2000, 15:13–42.

77. Salvi S, Blomberg A, Rudell B, Kelly F, Sandstrom T, Holgate ST, Frew A: Acute inflammatory responses in the airways and peripheral blood after short-term exposure to diesel exhaust in healthy human volunteers. Am J Respir Crit Care Med 1999, 159:702–709.

78. Xu X, Deng F, Guo X, Lv P, Zhong M, Liu C, Wang A, Tzan K, Jiang SY, Lippmann M, Rajagopalan S, Qu Q, Chen LC, Sun Q: Association of systemic inflammation with marked changes in particulate air pollution in Beijing in 2008. Toxicol Lett 2012, 212:147–156.