Genome-wide DNA methylation in peripheral blood and long-term exposure to source-specific transportation noise and air pollution: the SAPALDIA study

Ikenna C. Eze\textsuperscript{1,2*, Ayoung Jeong\textsuperscript{1,2}, Emmanuel Schaffner\textsuperscript{1,2}, Faisal I. Rezwan\textsuperscript{3,4}, Akram Ghantous\textsuperscript{5}, Maria Foraster\textsuperscript{1,2,6,7,8,9}, Danielle Vienneau\textsuperscript{1,2}, Florian Kronenberg\textsuperscript{10}, Zdenko Herceg\textsuperscript{5}, Paolo Vineis\textsuperscript{11,12}, Mark Brink\textsuperscript{13}, Jean-Marc Wunderli\textsuperscript{14}, Christian Schindler\textsuperscript{1,2}, Christian Cajochen\textsuperscript{12}, Martin Röösli\textsuperscript{1,2}, John W. Holloway\textsuperscript{3}, Medea Imboden\textsuperscript{1,2}, Nicole Probst-Hensch\textsuperscript{1,2*}.

\textsuperscript{1}Swiss Tropical and Public Health Institute, Basel, Switzerland
\textsuperscript{2}University of Basel, Basel, Switzerland
\textsuperscript{3}Human Development and Health, Faculty of Medicine, University of Southampton, Southampton, UK
\textsuperscript{4}School of Water, Energy and Environment, Cranfield University, Cranfield, Bedfordshire, UK
\textsuperscript{5}Epigenetics Group, International Agency for Research on Cancer, Lyon, France
\textsuperscript{6}ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain
\textsuperscript{7}University Pompeu Fabra, Barcelona, Spain
\textsuperscript{8}CIBER Epidemiologia y Salud Publica, Madrid, Spain
\textsuperscript{9}Blanquerna School of Health Science, Universitat Ramon Llull, Barcelona, Spain
\textsuperscript{10}Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria
\textsuperscript{11}MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, UK.
\textsuperscript{12}Italian Institute for Genomic Medicine (IIGM), Turin, Italy.
\textsuperscript{13}Federal Office for the Environment, Bern, Switzerland
\textsuperscript{14}Empa Laboratory for Acoustics/Noise Control, Swiss Federal Laboratories for Material Science and Technology, Dübendorf, Switzerland
\textsuperscript{15}Center for Chronobiology, Psychiatric Hospital of the University of Basel, and Transfaculty Research Platform Molecular and Cognitive Neurosciences (MCN), Basel, Switzerland

*Corresponding Authors
Ikenna C. Eze; Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland. Tel: +41 61 284 8395. Email: ikenna.eze@swisstph.ch
Nicole Probst-Hensch; Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland. Tel: +41 61 284 8378. Email: nicole.probst@swisstph.ch
Acknowledgements

This work was supported by the Swiss National Science Foundation, SNF-SAPALDIA (grants numbers 33CS30-148470/1, 33CSCO-134276/1, 33CSCO-108796, 324730_135673, 3247BO-104283, 3247BO-104288, 3247BO-104284, 3247-065896, 3100-059302, 3200-052720, 3200-042532, 4026-028099, PMPDP3_129021/1, PMPDP3_141671/1); SNF-SiRENE (grant number CRSII3_147635); the European Commission Horizon 2020 research and innovation programme (ALEC study; grant number 633212); grant FP7 of the European Commission “Enhanced exposure 36 assessment and -omic profiling for high priority environmental exposures in Europe” (EXPOsOMICS study; grant number 308610).

SAPALDIA is also supported by the Swiss Federal Office for the Environment, Federal Office of Public Health, the Federal Office of Roads and Transport, the canton's government of Aargau, Basel-Stadt, Basel-Land, Geneva, Luzern, Ticino, Valais, and Zürich, the Swiss Lung League, the canton's Lung League of Basel Stadt/ Basel Landschaft, Geneva, Ticino, Valais, Graubünden and Zurich, Stiftung ehemals Bündner Heilstätten, SUVA, Freiwillige Akademische Gesellschaft, UBS Wealth Foundation, Talecris Biotherapeutics GmbH, Abbott Diagnostics, European Commission 018996 (GABRIEL), Wellcome Trust WT 084703MA. Maria Foraster has received support of the Beatriu de Pinós postdoctoral programme of the Government of Catalonia's Secretariat for Universities and Research of the Ministry of Economy and Knowledge.

Disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

Competing Financial Interests (CFI)

The authors declare they have no actual or potential competing financial interests.
ABSTRACT

Background: Few epigenome-wide association studies (EWAS) on air pollutants exist, but none on transportation noise exposures, which also contribute to environmental burden of disease.

Objective: We performed mutually independent EWAS on transportation noise and air pollution exposures.

Methods: We used data from two time-points of the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) from 1,389 participants contributing 2,542 observations. We applied multi-exposure linear mixed-effects regressions with participant-level random intercept to identify significant Cytosine-phosphate-Guanine (CpG) sites and differentially methylated regions (DMRs) in relation to one-year average aircraft, railway and road traffic day-evening-night noise (Lden), nitrogen dioxide (NO₂) and particulate matter with aerodynamic diameter <2.5 µm (PM₂.₅). We performed candidate (CpG-based; cross-systemic phenotypes, combined into “allostatic load”) and agnostic (DMR-based) pathway enrichment tests, and replicated previously reported air pollution EWAS signals.

Results: We found no statistically significant CpGs at false discovery rate <0.05. However, 14, 48, 183, 8 and 71 DMRs independently associated with aircraft, railway and road traffic Lden, NO₂ and PM₂.₅, respectively, with minimally overlapping signals. Transportation Lden and air pollutants tendentially associated with decreased and increased methylation, respectively. We observed significant enrichment of candidate DNA methylation related to C-reactive protein and body mass index (aircraft, road traffic Lden and PM₂.₅), renal function and “allostatic load” (all exposures). Agnostic functional networks related to cellular immunity, gene expression, cell growth/proliferation, cardiovascular, auditory, embryonic and neurological systems development were enriched. We replicated increased methylation in cg08500171 (NO₂) and decreased methylation in cg17629796 (PM₂.₅).

Conclusions: Mutually independent DNA methylation was associated with source-specific transportation noise and air pollution exposures, with distinct and shared enrichments for pathways related to inflammation, cellular development and immune responses. These contribute in clarifying the pathways linking these exposures and age-related diseases, but need further confirmation in the context of mediation analyses.
Transportation-related noise (including road traffic, railway and aircraft noise) and air pollution (including nitrogen dioxide (NO$_2$) and particulate matter with aerodynamic diameter <2.5 µm (PM$_{2.5}$)) both make the greatest contribution to the global environmental burden of disease (Fritschi et al. 2011; Hänninen et al. 2014; Vienneau et al. 2015). Both exposure groups have been linked to cross-systemic phenotypes including respiratory (Adam et al. 2015; Eze et al. 2018; Hoek et al. 2013), cardiovascular (Beelen et al. 2014; Fiorito et al. 2018; Foraster et al. 2017; Heritier et al. 2019; Kempen et al. 2018), and metabolic diseases (Eze et al. 2015; Zare Sakhvidi et al. 2018), cancers (Andersen et al. 2018; Hegewald et al. 2017; Raaschou-Nielsen et al. 2013) and neurological disturbances (Clark and Paunovic 2018; Stafojgia et al. 2014; Zhang et al. 2018). The potential mechanisms linking air pollution and these phenotypes include inflammatory, immune and oxidative stress responses following inhalation (Munzel et al. 2016), whereas noise is thought to act through annoyance reactions, sleep disturbances, stress and activation of the hypothalamic–pituitary–adrenal (HPA) axis and sympathetic nervous system, with subsequent release of stress hormones and inflammatory molecules (Daiber et al. 2019).

DNA methylation alterations may mediate part of the physiological and biochemical changes leading from these traffic-related exposures to their associated preclinical and clinical phenotypes. There has been increasing evidence linking short- and long-term air pollution exposure to various measures of DNA methylation in both children and adults (Abraham et al. 2018; de FC Lichtenfels et al. 2018; Gondalia et al. 2019; Gruzieva et al. 2017; Lee et al. 2019; Mostafavi et al. 2018; Panni et al. 2016; Plusquin et al. 2017; Plusquin et al. 2018; Sayols-Baixeras et al. 2019). Although findings from these studies were largely inconsistent with regard to single Cytosine-phosphate-Guanine (CpG) sites, the systematic review by Alfano and colleagues (Alfano et al. 2018) highlighted that air pollution was consistently linked to global hypomethylation in children (Breton et al. 2016; Cai et al. 2017; Janssen et al. 2015) and in adults (De Nys et al. 2018; De Prins et al. 2013), and accelerated epigenetic aging in adults (Nwanaji-Enwerem et al. 2016; Nwanaji-Enwerem et al. 2017; Ward-Caviness et al. 2016). Overarching epigenomic effects of air pollution identified across these studies include inflammation, mitochondrial and DNA damage responses, and accelerated biological aging (Alfano et al. 2018).

Transportation noise, like air pollution, may also influence health outcomes via DNA methylation. First, they might share mechanistic pathways, as there is growing evidence for the inflammatory and oxidative downstream effects of noise exposure (Daiber et al. 2019; Munzel et al. 2017). Second, oxidative DNA damage, a correlate of altered DNA methylation and gene expression (Russo et al. 2016) was linked to occupational noise (Bagheri Hosseinabadi et al. 2019; Nawaz and Hasnain 2013) and traffic noise (Hemmingsen et al. 2015). Third, DNA methylation changes were reported for noise-related stressors such as vehicular traffic (Commodore et al. 2018) and insufficient sleep (Gaine et al. 2018). Methylation changes were also reported in circadian rhythm genes in conditions of acute sleep loss (Cedernaes et al. 2018).
2015) and night-shift work (Zhu et al. 2011). A study investigating brain tissue DNA methylation in relation to noise exposure in animal models reported gene-specific methylation changes, which were in turn associated with metabolic health (Guo et al. 2017).

No transportation noise epigenome-wide association study (EWAS) has been performed so far in population-based studies, and previous air pollution EWAS studies did not account for concurrent noise exposure (Eze and Probst-Hensch 2018). This is a limitation as both exposures share common sources and might be potential mutual confounders (Tetreault et al. 2013). As demonstrated by previous studies (Franklin and Fruin 2017; Heritier et al. 2019; Tetreault et al. 2013), the effect estimates of air pollution might be overestimated if noise level is unaccounted for, and vice versa. In addition, a parallel consideration of both exposure groups might elucidate directly shared and independent pathways of individual exposures, towards an improved understanding of mechanisms linking these exposures to disease.

In this paper, we aimed within the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) to:

(i) conduct a multi-exposure EWAS (single CpGs and genomic regions) involving long-term exposure to source-specific transportation noise (aircraft, railway and road traffic) and air pollution (NO$_2$ and PM$_{2.5}$);

(ii) assess in a candidate approach using a priori-curated CpGs, pathway enrichment for “allostatic load”-related cardio-metabolic, immunological and renal phenotypes (Johnson et al. 2017; McCrory et al. 2019; Seeman et al. 2010), given that both traffic-related exposures are chronic stressors of several physiological systems;

(iii) assess in an agnostic approach, functional pathway and network enrichment of EWAS-derived differentially methylated regions (DMRs); and

(iv) replicate previously reported CpGs associated with long-term exposure to NO$_2$ and PM$_{2.5}$ in the LifeLines (de FC Lichtenfels et al. 2018), EPIC-ITALY and EPIC-NL (Plusquin et al. 2017), and Korean COPD (Lee et al. 2019) cohorts.

METHODS

Study Population

The SAPALDIA cohort has been described elsewhere in detail (Ackermann-Liebrich et al. 2005; Martin et al. 1997). SAPALDIA is a population-based study that recruited 9,651 adults from eight geographically diverse Swiss areas (Aarau, Basel, Davos, Geneva, Lugano, Montana, Payerne, and Wald) in 1991 (SAP1) to investigate the respiratory effects of air pollution exposure. The first (SAP2) and second follow-up (SAP3) surveys occurred in 2002 and 2010, included 8,047 and 6,088 participants respectively, and additionally assessed cardio-metabolic and quality of life phenotypes. At each survey,
participants had physical examinations, health and lifestyle-related interviews. Measures of transportation noise and air pollution exposures were modeled at participants’ residences. At both SAP2 and SAP3, whole blood samples were collected, pre-processed and stored in a biobank (−80 °C) for biomarker assessments including DNA methylation. In the context two SAPALDIA nested studies—Aging Lungs in European Cohorts study (ALEC; www.alecstudy.org) and the EXPOsOMICS study (Vineis et al. 2017)—SAP2 and SAP3 repeat blood samples from 987 representative participants (ALEC), and an independent 405 non-smoking participants (for at least one year before SAP2; EXPOsOMICS), were applied toward DNA methylation assessment. The EXPOsOMICS sample was an asthma case-control sample of non-smokers (204 cases and 201 controls) where the asthma cases had methylation assessed only in SAP3 blood sample and the controls had methylation assessment in both SAP2 and SAP3 blood samples. Following exclusions due to methylation data quality and covariate data availability, we included SAP2 data from 972 ALEC and 198 EXPOsOMICS (asthma controls) participants, and SAP3 data from 970 ALEC and 402 EXPOsOMICS (204 asthma cases and 198 controls) participants.

The present study therefore includes 1,389 participants contributing 2,542 observations, with an average of 1.8 observations per participant. Figure 1 presents the details of participant selection and inclusion. The SAPALDIA study complies with the Helsinki declaration. All participants provided informed written consent before participating in any aspect of the SAPALDIA surveys and ethical approvals were obtained from the Swiss Academy of Medical Sciences and the Ethics committees of the participating cantons.

Assessment of DNA methylation

DNA methylation was measured in the ALEC and EXPOsOMICS samples separately in different laboratories, but using consistent methodology, using Illumina Infinium 450K BeadChip and processed as previously described (Imboden et al. 2019; Jeong et al. 2019). In brief, paired blood samples from the same participant were randomized across the arrays and placed next to each other on the arrays. Dye-bias correction (Triche et al. 2013) and absolute methylation level (β-values, defined as the ratio of methylation intensity over total intensity, with offset of 100) were computed using the minfi R-package (Aryee et al. 2014). Quality control criteria included call rate >95%, detection p-value <10^{-16}, sex consistency, autosomal chromosome location of probes, and restriction to probes identified in both ALEC and EXPOsOMICS samples. We applied beta mixture quantile normalization (BMQ) of the β-values to correct for the Illumina probe design bias (Teschendorff et al. 2013). We further excluded 37,882 CpGs that are cross-reactive or target polymorphic (Chen et al. 2013). For technical bias (batch effect) correction, these β-values were then regressed on the first 30 principal components derived from a principal component analysis of the control probes incorporated on the methylation chip (Lehne et al. 2015). Figure S1 shows the distributions of the BMQ-normalized beta values and technical bias-
corrected residuals. In line with previous studies (Imboden et al. 2019; Jeong et al. 2019), we applied these residuals (in place of the β-values) in subsequent EWAS covering 430,477 CpG sites that overlapped between ALEC and EXPOsOMICS samples.

To minimize bias due to extreme values (EVs) of the residuals—the dependent variables in the present EWAS—while retaining the distribution of DNA methylation at the non-skewed CpG sites, we performed “modified winsorization” of the CpG sites containing EVs. We defined EVs as values lying beyond three times the interquartile range (IQR) below and above the first (Q1) and third quartiles (Q3), respectively (Tukey 1977). We replaced each EV with the corresponding threshold of detection, in a CpG-specific manner, i.e.

If $EV < (Q1 - (3 \times IQR))$, then EV is replaced with the value of $Q1 - (3 \times IQR)$ at the CpG site;

If $EV > (Q3 + (3 \times IQR))$, then EV is replaced with the value of $Q3 + (3 \times IQR)$ at the CpG site.

Among the 430,477 included CpGs, 393,397 CpGs (91%) had at least one EV. The EVs comprised 0.04–22% of observations across the “winsorized” CpG sites. We used these “winsorized” data as the primary outcome variables in subsequent analyses (including EWAS, enrichment and replication), and only used the “non-winsorized” data for sensitivity testing of the top CpG signals identified in the EWAS.

Assessment of transportation noise exposure

As detailed elsewhere (Karipidis et al. 2014), annual average day-evening-night noise level (Lden, with respective 5 dB and 10 dB penalties for evening (19-23h) and night-time (23-07h; Lnight)) were calculated for 2001 and 2011 (corresponding to SAP2 and SAP3 respectively), for aircraft, railway and road traffic noise at the maximum-exposed façade of the residential floors of participants. This was done in the framework of the Short and Long Term Effects of Transportation Noise Exposure (SiRENE) project (Röösli et al. 2017). Aircraft noise was modeled with the FLULA2 software (Empa 2010) with input data covering four airports in Basel, Geneva, Zurich and Payerne. Railway noise emission and propagation were modeled using the sonRAIL (Thron and Hecht 2010) and the SEMIBEL (FOEN 1990) models respectively, whereas road traffic noise emission and propagation were separately modeled using sonROAD (Heutshi 2004) and StL-86 (FOEN 1987). Validation of noise calculations was done by comparing the calculated noise levels with measured levels from the field. Taking all measurements into account, a mean Lden difference of $1.6 \pm 5 \text{ dB}$ was observed (Schlatter et al. 2017). Lden and Lnight values were respectively truncated at 30 dB and 20 dB for railway and aircraft noise, and 35 dB and 25 dB for road traffic noise. Calculated Lden and Lnight values below these limits were replaced by the respective truncation values. Participants with truncated values were assigned a truncation indicator (yes/no) for subsequent modelling, in line with previous analyses in the context of the SiRENE study (Eze et al. 2017; Foraster et al. 2017). Given the high correlations of source-specific Lden and Lnight,
(Spearman rank correlation, \( r_{\text{aircraft}} = 0.76 \), \( r_{\text{railway}} = 0.97 \) and \( r_{\text{road traffic}} = 0.99 \)), and the generally lower \( \text{Lnight} \) levels (Table S1 and Figure S2); we focused on source-specific \( \text{Lden} \) as the noise parameters of interest in this EWAS.

Assessment of air pollution exposure

Traffic-related air pollutants—\( \text{NO}_2 \) and PM\(_{2.5}\)—were also modeled at participant’s residences at SAP2 and SAP3 as outdoor mean exposures. For SAP2, annual mean \( \text{NO}_2 \) was estimated (year 2003) using a hybrid model that regressed passive sampler measurements against dispersion model estimates, seasonal and climatic variables, as well as traffic and land use characteristics. Adjusted \( R^2 \) of the hybrid model was 0.8 (Liu et al. 2012). PM\(_{2.5}\) was derived from the Swiss PolluMap dispersion model (year 2000). Modeled values were validated against measured values with an \( R^2 \) of 0.9 (Heldstab et al. 2003). For SAP3, \( \text{NO}_2 \) and PM\(_{2.5}\) were modeled as average biennial exposures (2010/2011) using land-use regression models. The best models were used in each case: the combined PM\(_{2.5}\) with an adjusted \( R^2 \) of 0.6, and the area-specific \( \text{NO}_2 \) models with adjusted \( R^2 \)’s of 0.5-0.9 across SAPALDIA study areas (Eeftens et al. 2016). We included both pollutants in our analyses, examining their associations in both single and multi-exposure models.

Assessment of covariates

At the level of the individual, we considered groups of potential confounders measured at both SAP2 and SAP3, which might be associated with transportation noise, air pollution exposures, and DNA methylome. We considered sociodemographic factors including age (years), sex (male/female), and educational attainment (primary education/secondary education or apprenticeship/tertiary education)). We considered lifestyle factors such as smoking status (never/former/current), smoked pack years (calculated from cigarettes per day and duration of smoking), passive smoking (yes/no), alcohol consumption (beers, wines, liquors and spirits; \( \leq 1 \) glass per day/\( >1 \) glass per day), frequency of fruit intake (citrus or non-citrus fruits in any form; \( \leq 3 \) days per week/\( >3 \) days per week), vegetable intake (raw or cooked; \( \leq 3 \) days per week/\( >3 \) days per week), body mass index (BMI; kg/m\(^2\)) and sufficient moderate-to-vigorous physical activity (\( \geq 150 \) minutes per week of engagement in activities that makes one at least moderately sweat or breathless). To minimize the between-nested study differences, including residual batch effects and asthma status, we considered the contributing nested study (ALEC/EXPOsOMICS) and asthma status, defined as ever having a diagnosis of asthma. We also considered estimates of leukocyte composition for B cells, CD4T cells, CD8T cells, eosinophils, monocytes, natural killer cells and neutrophils (derived from the DNA methylation data using the “estimatecellcounts” function in the “minfi” R package (Houseman et al. 2012; Reinius et al. 2012)), to control the influence of cell proportions on methylation level (Adalsteinsson et al. 2012).
On the contextual scale, we considered some commonly indicated potential confounders of environmental health such as study area, a neighborhood index of socio-economic position (SEP; %) and greenness index within 1 km residential buffer. SEP was derived from a principal component analysis of household characteristics (education, occupation of household head, occupancy and median rent of household) based on 2001 census data (Panczak et al. 2012), and assigned to residential geo-coordinates at SAP2 and SAP3. Greenness index was calculated for 2014 as normalized difference vegetation index based on surface reflectance, and assigned to participants geo-coordinates at SAP2 and SAP3 (Vienneau et al. 2017).

Statistical analyses

EWAS

We described the characteristics of included participants by survey and by nested study, summarizing categorical variables as counts and proportions, and continuous variables as medians and interquartile ranges (IQR). Using the combined sample, we performed EWAS by linear mixed-effects regressions on the “winsorized” technical-bias corrected residuals of 430,477 CpGs, applying random intercepts at the level of participants. We used the “lmer” function of the “lme4” R package for the regressions (Bates et al. 2015). We performed single- as well as multi-exposure EWAS for aircraft, railway and road traffic Lden, NO\textsubscript{2} and PM\textsubscript{2.5}, adjusting for age, sex, education, smoking status, pack years, passive smoking, fruit, vegetable and alcohol intake, study area, SEP, greenness index, survey, nested study, asthma, Lden truncation indicator and Houseman estimates of leukocyte composition, in the main model. We identified genome-wide significant CpGs at false discovery rate (FDR) and Bonferroni-corrected p-value thresholds of 0.05 and 1.16E-07, respectively.

We tested sensitivity of the top 10 CpGs to further adjustment for the potential mediators BMI and physical activity, which have been associated with transportation noise (An et al. 2018b; Foraster et al. 2016; Roswall et al. 2017) and air pollution exposures (An et al. 2018a; An et al. 2018c). We further stratified these CpGs by nested study to assess consistency in effect direction, limited these models to participants who reported regular nighttime opening of windows, and assessed the robustness of top signals in models using the “non-winsorized” methylation data.

We tested for differentially methylated regions (DMRs) using the “dmrcate” function in the “DMRcate” R package (Peters et al. 2015) and the multi-exposure EWAS-derived parameters for each CpG, as input file. Thus, we performed the DMR analyses using the individual estimates of 430,477 CpGs from the main multi-exposure model. We defined DMRs as significant if they contained at least two CpGs, within at least 1000 base pairs, and had a minimum FDR p-value <0.05. Except when unannotated, all CpGs and DMRs are reported with gene annotation in parenthesis.
Pathway enrichment.

In a candidate pathway approach, we identified previously published EWAS signals for different physiological systems (captured by selected phenotypes) of potential relevance to transportation-related noise and air pollution effects. These systems (phenotypes) included (i) immunological (C-reactive protein, CRP); (ii) metabolic (glycaemia, insulin secretion/sensitivity, lipoprotein cholesterol, BMI, waist circumference/visceral adipose tissue mass, and metabolic syndrome); (iii) renal (estimated glomerular filtration rate, eGFR); and (iv) cardiovascular and autonomic nervous systems (blood pressure and cardiac autonomic responses, CAR). For all phenotypes, we curated CpGs from the EWAS atlas (Li et al. 2019), at a p-value threshold of 1.10E-05, except for CRP where only the genome-wide significant signals at p-value of 1.15E-07 were reported (Ligthart et al. 2016). CpGs that overlapped with those of the SAPALDIA study were finally curated for enrichment analyses (Excel Table S1).

We defined a global “allostatic load” pathway as the entirety of unique CpGs assigned to at least one of the constituent pathways (n = 1,626). We applied the Weighted Kolmogorov-Smirnov method (Charmpi and Ycart 2015) to test for enrichment of the candidate and “allostatic load” pathways using the absolute values of test statistics from multi-exposure EWAS. These test statistics from CpGs mapped to the pathway were compared to the empirical null distribution derived by 10,000 permutation samples. If the Kolmogorov-Smirnov test statistic was larger than the 90th percentile of test statistics obtained from the permutation samples after permutation-based multiple testing correction (van der Laan et al. 2005), we declared enrichment for the pathway.

In an agnostic approach, we assessed the functional pathways of exposure-specific differential methylation, using the “Core Analysis” of ingenuity pathway analysis (IPA; Ingenuity Systems, QIAGEN, CA, USA) on genes annotated to significant DMRs. We identified for each exposure, the functional pathways which were significantly enriched at p-value <0.05, as well as top disease and functional networks related to these pathways.

Replication of previously reported air pollution associated CpGs.

Using the SAPALDIA EWAS results, we looked up the single- and multi-exposure EWAS signals for validation of 22 and 10 previously reported genome-wide significant CpGs for long-term exposure to NO₂ and PM₂.₅, respectively. These include NO₂ signals from the LifeLines cohort (cg04908668, cg14938677, cg00344801, cg18379295, cg25769469, cg02234653, and cg08500171 (de FC Lichtenfels et al. 2018)), EPIC-ITALY (cg08120023, cg22856765, cg18164357, cg13918628, cg03870188, cg20939320, cg13420720, cg04914283, cg21156210, cg16205861, cg12790758, and cg18201392 (Plusquin et al. 2017)), and the Korean COPD cohort (cg05171937, cg06226567 and cg26583725 (Lee et al. 2019)). PM₂.₅-associated signals include in cg23890774 from EPIC-ITALY, and cg12575202, cg08630381, cg17629796, cg07084345, cg04319606, cg09568355, cg03513315, cg25489413,
RESULTS

The characteristics of the study sample at SAP2 and SAP3 are presented in Table 1. In general, participants tended to gain weight, smoke less with relatively lower pollutant exposures between surveys. ALEC and EXPOsOMICS participants mainly differed by design in their smoking habits and asthma status, but their environmental exposure profiles were comparable (Table S2). Figure S2 shows the overall distributions of the aircraft, railway and road traffic Lden and NO2 and PM2.5 exposures. The Spearman rank correlations (r) of the five exposures are shown on Table S1. Correlation of NO2 and PM2.5 was 0.65. NO2 had higher correlation with road traffic (r = 0.41) than railway (r = 0.23) and aircraft Lden (r = 0.10) whereas PM2.5 had higher correlation with aircraft (r = 0.23) than railway (r = 0.19) and road traffic Lden (r = 0.19). Correlation patterns were consistent across SAP2 and SAP3.

EWAS

In the single exposure models (Table S3), we observed one genome-wide significant (FDR = 0.040) PM2.5-associated CpG (cg26704043 (FARS2)) and one borderline-significant (FDR = 0.077) railway Lden-associated CpG (cg25201280 (ATPBD4)). There were no genome-wide significant signals in the multi-exposure model (Table 2). Here, the PM2.5-associated cg26704043 (FARS2) got considerably weaker (FDR = 0.180) whereas the railway Lden-associated cg25201280 (ATPBD4) remained borderline-significant (FDR = 0.075). The top 10 CpGs of the single exposure models were not entirely consistent with those of the multi-exposure models, with varying degrees of overlap across exposures. However, overlapping CpGs were directionally consistent (for all exposures), and the top CpG was positionally consistent (except for road traffic Lden) between the single and multi-exposure models (Tables 2 and S3).

In the multi-exposure model, aircraft Lden (cg02286155), railway Lden (cg25201280 (ATPBD4)) and road traffic Lden cg09129334 (ARHGEF7) were associated with reduced methylation, whereas NO2 (cg04337651 (ASB1)) and PM2.5 (cg26704043 (FARS2)) were associated with increased methylation, at their respective top CpG sites. Among the top 10 signals, we observed decreased methylation at nine (90%), eight (80%) and eight (80%) CpGs in relation to aircraft, railway and road traffic Lden, respectively, and increased methylation at five (50%) and 10 (100%) CpGs in relation to NO2 and PM2.5, respectively. These CpGs were generally robust to adjustment for BMI and physical activity (Table 2) and showed consistent effect direction when stratified by nested study (Table S4). Associations were also robust in the model limited to participants reporting regular nighttime window opening (Table S5). The “non-winsorized” model highlighted few CpGs where extreme values would potentially bias their estimates if unaddressed. These included directionally consistent, but considerably weaker effects on
Post hoc analysis of overlap among the top 100 exposure-specific CpGs identified from the multi-exposure EWAS showed minimal signal overlap between exposures (Figure 2A). We identified two overlapping signals between road traffic Lden and NO2 (cg12439232 and cg15590912 (CCSAP)). Yet, road traffic Lden was associated with decreased methylation whereas NO2 was associated with increased methylation at both sites (Excel Tables S4 and S5). Complete EWAS results from the single and multi-exposure main models of 430,477 CpGs in relation to aircraft, railway and road traffic Lden, NO2 and PM2.5 are deposited in the DRYAD public online depository (http://datadryad.org/) at the time of publication.

We identified independent DMRs (FDR <0.05) across all exposures in the main model (Table 3). There were 14 (10), 48 (39), 183 (189), 8 (8) and 71 (60) DMRs (genes) respectively associated with aircraft, railway, and road traffic Lden, NO2 and PM2.5. Among the top 10 CpGs identified in the multi-exposure model, two aircraft (cg10975000 and cg25462190 (N4BP3)), six railway (cg25201280 (ATPBD4), cg24653263 (EGFLAM), cg16825060 (LY6H), cg19270309 (ENPP7), cg23113715, and cg24047259), and five road traffic Lden-associated CpGs (cg17383236, cg23910243 (TGFB1II), cg06646021 (RAB44), cg03966094 (TMEM191A), and cg08351004 (DLX2)) were within the exposure-specific DMRs. Three NO2 (cg04337651 (ASB1), cg01746514 (LRRC16B) and cg26898336 (TEKT3)) and two PM2.5-associated CpGs (cg20099458 (WIP12) and cg26750893) were also within the exposure-specific DMRs. Top DMRs independently associated with exposures annotated to VTRNA2-1 (aircraft and railway Lden; Chr5:135415129–135416613), OXT (road traffic Lden; Chr20:3051954–3053196), ZSCAN31 (NO2; Chr6:28303923–28304451) and TRIM39, HCG18, and TRIM39-RPP21 (PM2.5; Chr6:30296689–30297941). Most of the CpGs within the DMRs associated with source-specific Lden showed decreased methylation (aircraft = 64%, railway = 69% and road traffic = 93%) whereas those associated with air pollution showed increased methylation (NO2 = 63% and PM2.5 = 93%). Excel Tables S7 and S8 show the results from the multi-exposure main model of DMRs associated with aircraft, railway and road traffic Lden, NO2 and PM2.5 at FDR <0.05.

Post hoc analysis of overlap among the gene-annotated significant DMRs also showed minimal overlap between exposures (Figure 2B). SLC27A3, B3GALT4, EN2 and AC008060.8 overlapped between railway and road traffic Lden, TRIM39, TRIM39-RPP21 and HCG18 overlapped between railway Lden and PM2.5, and HOXA2 overlapped between aircraft, road traffic Lden and PM2.5. Other overlapping genes include VTRNA2-1 (aircraft and railway Lden), ZFP57 (railway Lden and NO2) and ZSCAN31.
(road traffic Lden and NO$_2$) and PRRT1 (road traffic Lden and PM$_{2.5}$). Overlapping DMRs showed opposing effect direction between exposures, except for consistently increased methylation at TRIM39, TRIM39-RPP21, HCG18 (railway Lden and PM$_{2.5}$) and HOXA-2 (aircraft Lden and PM$_{2.5}$), and decreased methylation at SLC27A3, B3GALT4, EN2 and AC008060.8 (railway and road traffic Lden) (Excel Tables S7 and S8).

Pathway enrichment

For the candidate single phenotype and combined “allostatic load” pathways, multiple testing-corrected enrichment analyses showed varying enrichment across exposures (Table 4). Considering the results for single phenotypes, PM$_{2.5}$-related methylation was the most enriched whereas NO$_2$-related methylation was the least enriched. Single pathways (p-values) enriched for PM$_{2.5}$ included CRP (0.0004), glycaemia (0.0947), WC (0.0004), BMI (0.0004) and eGFR (0.0004). Aircraft and road traffic Lden-related methylation were enriched for CRP (0.0038; 0.0395), BMI (0.0007; 0.0008) and eGFR (0.0015; 0.0031). Railway Lden-related methylation was enriched for eGFR (0.0562) and CAR (0.0229), whereas NO$_2$–related methylation was only enriched for eGFR (0.0058). Considering the global “allostatic load” pathway, we found consistent enrichment across all exposures including aircraft (0.0004), railway (0.0871), and road traffic Lden (0.0004), PM$_{2.5}$ (0.0004), and NO$_2$ (0.0510).

Agnostic functional enrichment showed significantly enriched canonical pathways for railway, road traffic Lden and PM$_{2.5}$, with predominance of inflammatory and immune regulation-related pathways across exposures (Excel Table S9). Some of the enriched canonical pathways included type 2 diabetes signaling, tight junction signaling, mTOR signaling and lipopolysaccharide/IL-1 mediated inhibition of RXR function (railway Lden); Wnt/β-catenin signaling, cholecystokinin/gastrin-mediated signaling, G-protein coupled receptor signaling and Th1 and Th2 activation pathways (road traffic Lden); β-alanine, 4-aminobutyrate degradation; systematic lupus erythematosus signaling, and IL-4 signaling (PM$_{2.5}$). Network analyses showed the enrichment for disease mechanisms related to cellular signaling/interaction and embryonic/organ development (all exposures), cell-mediated immune/inflammatory responses (road traffic Lden and PM$_{2.5}$) and diseases related to connective tissue development/function (railway and road traffic Lden). Pathways related to cardiovascular function, gene expression and respiratory system disease (railway Lden); carbohydrate and small molecule transport/biochemistry, cancer and auditory function (road traffic Lden); hematological and nervous system function (PM$_{2.5}$) were enriched (Excel Table S10).

Replication of previously reported NO$_2$ and PM$_{2.5}$-associated CpGs

Using the single-exposure model, corresponding to the previous EWAS studies, we replicated the increased methylation in cg08500171 (B4T2) associated with NO$_2$ (p = 0.042). This replication remained robust in the multi-exposure model (p = 0.019). We observed a borderline significant
decreased methylation in cg17629796 associated with PM$_{2.5}$ (p = 0.076) in the single exposure model, which became statistically significant in the multi-exposure model (p = 0.033). Nine (41%) NO$_2$-associated CpGs and five (50%) PM$_{2.5}$-associated CpGs had the same direction of effect in the SAPALDIA study (Table 5).

DISCUSSION

In this first multi-exposure EWAS covering long-term aircraft, railway and road traffic noise as well as NO$_2$ and PM$_{2.5}$ exposures, we found genome-wide significant signals—at the level of genomic regions—associated with source-specific transportation noise and air pollution exposures. We demonstrated some mutual confounding of the associations of exposures with DNA methylation, where single exposure models slightly overestimated the observed methylation effect. Overall, methylation signals minimally overlapped, but showed common enrichment of inflammation and immune response-related pathways, across exposures. We also validated in this study, air pollution-related CpG signals identified in external EWAS.

The railway noise-related decrease in methylation at cg25201280 (transcription start site 200 of ATPBD4 gene on chromosome 15), and PM$_{2.5}$-related increase in methylation at cg26704043 (5' untranslated region of FARS2 gene on chromosome 6), were the strongest single CpG signals identified in the study. ATPBD4 (or DPH6) is a protein-coding gene that regulates ATP binding and diphthamide synthase activity, within the protein metabolism pathway (Chertow 1981; Young et al. 2004). Polymorphisms in this gene were associated—in various genome-wide association studies (GWAS)—with adiposity (Kichaev et al. 2019; Pulit et al. 2019), cognitive function (Lee et al. 2018; Li et al. 2015) and Crohn’s disease progression (O’Donnell et al. 2019). FARS2 is a protein-coding gene that regulates the translation of mitochondrial proteins and gene expression (Bullard et al. 1999). Variants of this gene were associated—in GWAS—with extreme obesity (Wheeler et al. 2013) and acute myeloid leukemia (Lv et al. 2017), whereas cg26704043 was associated—in EWAS—with systemic lupus erythematosus (Imingenberg-Kreuze et al. 2018) and Grave’s disease (Chen et al. 2019). These highlight the role of DNA methylation changes in these genes as potential mechanisms by which railway noise and PM$_{2.5}$ contribute to the resulting inflammatory and neurological phenotypes.

Our observation of numerous independent DMRs (for all five exposures) despite paucity of single CpG signals highlights the relevance of concurrent investigation of genomic regions in EWAS. Analyses of genomic regions contextualizes CpGs to identify DMRs with possible functional relevance in regulation of gene transcription (Rakyan et al. 2011). Interestingly, transportation noise showed mostly decreased methylation whereas air pollution showed mostly increased methylation at the respective DMRs. This difference in direction of association was also observed at the levels overlapping CpGs and annotated DMRs. The relevance of these findings remains to be clarified as DNA hypermethylation is often associated with transcriptional gene repression, while hypomethylation is associated with a chromatin
configuration that allows transcription (Sawalha 2008). However, concordant hypermethylation in
TRIM39 (railway noise and PM_{2.5}) and HOXA2 (PM_{2.5} and aircraft noise and PM_{2.5}), and
hypomethylation in SLC27A3 and EN2 (railway and road traffic noise) protein-coding genes indicate
some synergistic pathways between exposures. TRIM39 activates the apoptotic signaling pathway, plays
a role in cellular signaling and response to stimuli (Zhang et al. 2012), and its hypermethylation was
associated with depression (Crawford et al. 2018). HOXA2 encodes a transcriptional regulator that
controls cellular differentiation during development (Akin and Nazarali 2005). Hypermethylation at this
gene was associated transcriptional suppression in various cancers (Li et al. 2013). SLC27A3 is involved
in lipid metabolism and brain development (Stahl 2004), and hypomethylation in this gene was
associated with recurrent endometrial carcinoma (Hsu et al. 2013). EN2 functions in the development
of the central nervous system and is commonly associated with autism (Lupu et al. 2018; Márquez-
Valadez et al. 2018). Aberrant methylation in EN2 was recently linked to clear cell renal carcinoma (Lai
et al. 2017). While these highlight certain distinct and shared pattern of associations, they demonstrate
the complex network in the mechanisms linking these exposures to disease. The integration of gene
expression or transcriptomic data in future studies will improve our understanding of the mechanisms
going from divergence in directions of associations to convergence on pathway level across these
exposures.

The results demonstrating enrichment for DNA methylation associated with phenotypes of “allostatic
load” for both, air pollution and noise exposure are novel, but in line with previously hypothesized
mechanisms. Chronic low-level exposure to air pollution and noise are established risk factors for
cardio-metabolic disease and were previously linked to the single phenotypes making up the allostatic
load pathway in this paper (Munzel and Daiber 2018; Thomson 2019). Recent experimental evidence
specifically linked particulate matter and ozone pollution as well as transportation noise to alterations
in the HPA axis (Jafari et al. 2017). The findings of “allostatic load” enrichment agrees with the
functional enrichment of pathways related to oxidative stress and immune responses across all
exposures. Furthermore, the enriched disease networks identified for PM_{2.5} overlaps the networks that
recently reported for PM_{10} in another study (Lee et al. 2019). Taken together, the evidence especially
supports the recent links of transportation noise to markers of inflammation and oxidative stress (Bagheri
Hosseinabadi et al. 2019; Munzel et al. 2016; Munzel et al. 2017; Munzel and Daiber 2018). However,
our findings on the inflammatory pathway might have been influenced by the higher prevalence of
asthma in this sample (22%) compared to the entire SAPALDIA study (12%), but our models contained
asthma status, in addition to smoking and pack years, and should minimize this potential bias. In line
with recent evidence on the early life methylome effects of air pollution (Cai et al. 2017; Gruzijeva et al.
2017) and potentially noise exposures, we observed enrichment of pathways related to embryonic and
organ development for all exposures. Altered DNA methylation due to these exposures might therefore
explain the reported associations of PM and poor birth outcomes (Liu et al. 2019; Smith et al. 2020), as
well as the early life theory of age-related disease (Walhovd et al. 2016). This is even more relevant if
these methylation changes are heritable, and when confirmed, improves our understanding of the relevance of this pathway in relation to these exposures.

The relevance of confounding by leukocyte composition, on the association between exposures and peripheral blood DNA methylation has been of interest (Heiss and Brenner 2017; van Rooij et al. 2019). We demonstrated—in post hoc analyses—the general weakening of top CpGs (with some signals not attaining nominal significance) when leukocyte composition was not considered. However, main effect sizes remained robust (Table S7). Source-specific transportation noise and air pollution exposures were also associated with various leukocyte types (Table S8). These indicate that changes in leukocyte composition probably drive methylation at certain sites, and should be considered potential confounders in this and similar studies. Our findings in this EWAS (adjusted for leukocyte composition) therefore reflect more of actual DNA methylation changes rather than underlying leukocyte composition changes, in relation to transportation noise and air pollution exposure. This would otherwise be difficult to interpret given the observed associations of exposures with leukocyte types.

The strengths of our study include being the first EWAS to assess mutually independent effects of source-specific transportation noise and air pollution exposures. Our study considered in parallel, the independent effects of these exposures on DNA methylation at single CpG sites and in genomic regions. The availability of individually assigned, source-specific noise and air pollution data, derived from validated models, allowed the exploration of mutual confounding of these exposures on DNA methylation. We were able to control for potential confounders given the detailed characterization of the participants in the SAPALDIA study. The multi-exposure approach also allowed the investigation of the independent pathway enrichment using both candidate and agnostic approaches, to improve mechanistic understanding of transportation noise and air pollution exposures. We used a novel approach to investigate indirectly, the effect of these exposures on physiological stress systems. We could also explore the influence of leukocyte composition, among other sensitivity analyses. The availability of genome-wide methylation data allowed the validation of previously reported air pollution signals, contributing to their external validity in the SAPALDIA study.

Our study is limited in its cross-sectional design, which precludes causal inferences and differentiating short-term vs. long-term exposure effects. Our noise and air pollution estimates may have been biased by errors in input data, which could be exposure-specific, and potentially accounting for variations in observed effects. Such bias, however, would most likely be non-systematic and non-differential. The air pollution and noise levels in the SAPALDIA study are relatively low and may have reduced statistical power of identifying signals, and limited the replication of previous findings in settings of higher exposure. Unlike road traffic noise, which was ubiquitous, railway and aircraft noise in the SAPALDIA study were less common, with only about 45% exposure to these sources. Nevertheless, we made some significant observations. We could not identify any EWAS on stress hormones, which directly captured
alterations in the HPA axis, thus, our definition of allostatic load may have been biased. However, we expect any bias to be minimal given our inclusion of CAR, and other downstream biomarkers of physiological stress. As demonstrated in a recent review (Johnson et al. 2017), there is yet no consensus on what markers best capture allostatic load. Interestingly, allostatic load score was always associated with negative health outcomes regardless of constituent biomarkers or phenotypes (Castagné et al. 2018; Johnson et al. 2017; Ribeiro et al. 2019).

In conclusion, DNA methylation was independently associated with transportation-related noise and air pollution exposures in the SAPALDIA study, with enrichment for pathways related to inflammation and immune response. Differential methylation due to these exposures may therefore explain the link between these exposures and several age-related outcomes. More EWAS with combined exposures and gene expression or transcriptome data are needed, especially from more polluted areas (including low- and middle-income countries), to corroborate present findings, and capture better, the full extent and relevance of DNA methylation changes associated with these exposures. In particular, it remains to be seen if DNA methylation related to the identified pathways mediates the association between transportation noise and air pollution exposures and incident age-related phenotypes.
REFERENCES


Table 1. Summary of the included SAPALDIA sample

<table>
<thead>
<tr>
<th>Categorical variables, n (%)</th>
<th>SAPALDIA2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SAPALDIA3&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>623 (53)</td>
<td>737 (54)</td>
</tr>
<tr>
<td>Formal education, ≤9 years</td>
<td>56 (5)</td>
<td>62 (4)</td>
</tr>
<tr>
<td>Formal education, 10-12 years</td>
<td>759 (65)</td>
<td>887 (65)</td>
</tr>
<tr>
<td>Formal education, &gt;12 years</td>
<td>355 (30)</td>
<td>423 (31)</td>
</tr>
<tr>
<td>Smoking status, never</td>
<td>539 (46)</td>
<td>659 (48)</td>
</tr>
<tr>
<td>Smoking status, former</td>
<td>355 (30)</td>
<td>496 (36)</td>
</tr>
<tr>
<td>Smoking status, current</td>
<td>276 (24)</td>
<td>217 (16)</td>
</tr>
<tr>
<td>Passive smoke exposure</td>
<td>289 (25)</td>
<td>159 (12)</td>
</tr>
<tr>
<td>Alcohol intake &gt;1 glass per day</td>
<td>456 (39)</td>
<td>529 (38)</td>
</tr>
<tr>
<td>Fruit intake ≤3days/week</td>
<td>326 (28)</td>
<td>280 (20)</td>
</tr>
<tr>
<td>Vegetable intake ≤3days/week</td>
<td>86 (7)</td>
<td>100 (7)</td>
</tr>
<tr>
<td>Urban area</td>
<td>689 (59)</td>
<td>897 (60)</td>
</tr>
<tr>
<td>Prevalent asthma</td>
<td>161 (14)</td>
<td>398 (21)</td>
</tr>
<tr>
<td>MVPA &lt;150 minutes per week</td>
<td>330 (28)</td>
<td>354 (26)</td>
</tr>
<tr>
<td>Regular nighttime opening of windows</td>
<td>971 (83)</td>
<td>1,131 (82)</td>
</tr>
<tr>
<td>Nested study, ALEC</td>
<td>972 (83)</td>
<td>970 (71)</td>
</tr>
<tr>
<td>Nested study, EXPOsOMICS</td>
<td>198 (17)</td>
<td>402 (29)</td>
</tr>
</tbody>
</table>

Continuous variables, median (IQR)

<table>
<thead>
<tr>
<th>Variable</th>
<th>SAPALDIA2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SAPALDIA3&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>50 (18)</td>
<td>58 (18)</td>
</tr>
<tr>
<td>Body mass index (kg/m&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>24.9 (5)</td>
<td>25.8 (6)</td>
</tr>
<tr>
<td>Smoking pack years</td>
<td>0.4 (15)</td>
<td>0 (14)</td>
</tr>
<tr>
<td>Neighborhood index of socio-economic position (%)</td>
<td>64.6 (13)</td>
<td>64.8 (13)</td>
</tr>
<tr>
<td>Greenness index within 1 km buffer</td>
<td>0.61 (0.2)</td>
<td>0.62 (0.2)</td>
</tr>
<tr>
<td>Aircraft Lnight (dB)</td>
<td>20 (2)</td>
<td>20.1 (5)</td>
</tr>
<tr>
<td>Railway Lnight (dB)</td>
<td>22.9 (14)</td>
<td>20 (10)</td>
</tr>
<tr>
<td>Road traffic Lnight (dB)</td>
<td>44.9 (111)</td>
<td>45.1 (11)</td>
</tr>
<tr>
<td>Aircraft Lden (dB)</td>
<td>30 (9)</td>
<td>32.7 (8)</td>
</tr>
<tr>
<td>Railway Lden (dB)</td>
<td>30 (11)</td>
<td>30 (7)</td>
</tr>
<tr>
<td>Road traffic Lden (dB)</td>
<td>53.7 (11)</td>
<td>53.9 (11)</td>
</tr>
<tr>
<td>NO&lt;sub&gt;2&lt;/sub&gt; (µg/m&lt;sup&gt;3&lt;/sup&gt;)</td>
<td>20.2 (14)</td>
<td>16.7 (10)</td>
</tr>
<tr>
<td>PM&lt;sub&gt;2.5&lt;/sub&gt; (µg/m&lt;sup&gt;3&lt;/sup&gt;)</td>
<td>14.3 (5)</td>
<td>12.9 (2)</td>
</tr>
</tbody>
</table>

SAPALDIA: Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults. MVPA: moderate to vigorous physical activity. ALEC: Aging Lungs in European Cohorts. PM<sub>2.5</sub>: particulate matter with aerodynamic diameter <2.5 µm. NO<sub>2</sub>: nitrogen dioxide. ALEC and EXPOsOMICS are European cohort consortia in which SAPALDIA participates. <sup>a</sup> Population included in the analysis was limited to participants with complete methylome, exposure, and covariate data.
Table 2. Top ten CpGs independently associated with source-specific transportation noise and air pollution in the SAPALDIA study, multi-exposure models.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>CpG ID</th>
<th>CHR</th>
<th>Location</th>
<th>Gene</th>
<th>Feature</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
<th>P (FDR)</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft Lden</td>
<td>cg02286155</td>
<td>5</td>
<td>176826262</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.007</td>
<td>0.001</td>
<td>1.48E-06</td>
<td>0.637</td>
<td>-0.007</td>
<td>0.001</td>
<td>2.13E-06</td>
</tr>
<tr>
<td></td>
<td>cg15065350</td>
<td>2</td>
<td>17716941</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.009</td>
<td>0.002</td>
<td>8.14E-06</td>
<td>0.659</td>
<td>-0.009</td>
<td>0.002</td>
<td>7.80E-06</td>
</tr>
<tr>
<td></td>
<td>cg16218477</td>
<td>7</td>
<td>1066167</td>
<td>C7orf50</td>
<td>Body</td>
<td>-0.004</td>
<td>0.001</td>
<td>8.53E-06</td>
<td>0.659</td>
<td>-0.004</td>
<td>0.001</td>
<td>7.05E-06</td>
</tr>
<tr>
<td></td>
<td>cg21602842</td>
<td>18</td>
<td>46291908</td>
<td>KIAA0427</td>
<td>Body</td>
<td>-0.005</td>
<td>0.001</td>
<td>9.84E-06</td>
<td>0.659</td>
<td>-0.005</td>
<td>0.001</td>
<td>9.84E-06</td>
</tr>
<tr>
<td></td>
<td>cg09042449</td>
<td>10</td>
<td>44064225</td>
<td>ZNF239</td>
<td>5’UTR</td>
<td>-0.002</td>
<td>0.0004</td>
<td>1.06E-05</td>
<td>0.659</td>
<td>-0.002</td>
<td>0.0004</td>
<td>1.18E-05</td>
</tr>
<tr>
<td></td>
<td>cg10975000</td>
<td>13</td>
<td>28371375</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.002</td>
<td>0.0005</td>
<td>1.12E-05</td>
<td>0.659</td>
<td>-0.002</td>
<td>0.0005</td>
<td>1.16E-05</td>
</tr>
<tr>
<td></td>
<td>cg06220958</td>
<td>17</td>
<td>10452851</td>
<td>MTH2</td>
<td>5’UTR</td>
<td>0.011</td>
<td>0.002</td>
<td>1.24E-05</td>
<td>0.659</td>
<td>0.010</td>
<td>0.002</td>
<td>1.43E-05</td>
</tr>
<tr>
<td></td>
<td>cg25462190</td>
<td>5</td>
<td>17754067</td>
<td>N4BP3</td>
<td>Body</td>
<td>-0.007</td>
<td>0.002</td>
<td>1.32E-05</td>
<td>0.659</td>
<td>-0.007</td>
<td>0.002</td>
<td>1.06E-05</td>
</tr>
<tr>
<td></td>
<td>cg11944797</td>
<td>13</td>
<td>99135711</td>
<td>STK24</td>
<td>Body</td>
<td>-0.001</td>
<td>0.0003</td>
<td>1.38E-05</td>
<td>0.659</td>
<td>-0.001</td>
<td>0.0003</td>
<td>1.70E-05</td>
</tr>
<tr>
<td></td>
<td>cg04635504</td>
<td>11</td>
<td>2829241</td>
<td>KCNQ1</td>
<td>Body</td>
<td>-0.005</td>
<td>0.001</td>
<td>1.61E-05</td>
<td>0.664</td>
<td>-0.005</td>
<td>0.001</td>
<td>1.62E-05</td>
</tr>
<tr>
<td>Railway Lden</td>
<td>cg25201280</td>
<td>15</td>
<td>35838552</td>
<td>ATPBD4</td>
<td>TSS200</td>
<td>-0.001</td>
<td>0.0002</td>
<td>1.74E-07</td>
<td>0.075</td>
<td>-0.001</td>
<td>0.0002</td>
<td>1.99E-07</td>
</tr>
<tr>
<td></td>
<td>cg24653263</td>
<td>5</td>
<td>38258335</td>
<td>EGFLAM</td>
<td>TSS200</td>
<td>-0.003</td>
<td>0.001</td>
<td>8.97E-07</td>
<td>0.193</td>
<td>-0.003</td>
<td>0.001</td>
<td>1.01E-06</td>
</tr>
<tr>
<td></td>
<td>cg16825060</td>
<td>8</td>
<td>14442342</td>
<td>LY6H</td>
<td>TSS1500</td>
<td>-0.003</td>
<td>0.001</td>
<td>2.34E-06</td>
<td>0.256</td>
<td>-0.003</td>
<td>0.001</td>
<td>2.56E-06</td>
</tr>
<tr>
<td></td>
<td>cg23468045</td>
<td>5</td>
<td>12669584</td>
<td>N.A.</td>
<td>N.A.</td>
<td>0.001</td>
<td>0.0002</td>
<td>2.37E-06</td>
<td>0.256</td>
<td>0.001</td>
<td>0.0002</td>
<td>2.19E-06</td>
</tr>
<tr>
<td></td>
<td>cg19270309</td>
<td>17</td>
<td>77712853</td>
<td>ENPP7</td>
<td>3’UTR</td>
<td>0.002</td>
<td>0.0005</td>
<td>3.25E-06</td>
<td>0.266</td>
<td>0.002</td>
<td>0.0005</td>
<td>3.56E-06</td>
</tr>
<tr>
<td></td>
<td>cg07461273</td>
<td>7</td>
<td>99697172</td>
<td>MCM7</td>
<td>Body</td>
<td>-0.004</td>
<td>0.001</td>
<td>3.78E-06</td>
<td>0.266</td>
<td>-0.004</td>
<td>0.001</td>
<td>3.70E-06</td>
</tr>
<tr>
<td></td>
<td>cg13013145</td>
<td>7</td>
<td>27153580</td>
<td>HOXA3</td>
<td>5’UTR</td>
<td>-0.003</td>
<td>0.001</td>
<td>4.60E-06</td>
<td>0.266</td>
<td>-0.003</td>
<td>0.001</td>
<td>4.32E-06</td>
</tr>
<tr>
<td></td>
<td>cg23113715</td>
<td>22</td>
<td>25800663</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.004</td>
<td>0.001</td>
<td>5.79E-06</td>
<td>0.266</td>
<td>-0.004</td>
<td>0.001</td>
<td>5.73E-06</td>
</tr>
<tr>
<td></td>
<td>cg13402217</td>
<td>1</td>
<td>15158437</td>
<td>SNX27</td>
<td>TSS1500</td>
<td>-0.003</td>
<td>0.001</td>
<td>6.03E-06</td>
<td>0.266</td>
<td>-0.003</td>
<td>0.001</td>
<td>5.50E-06</td>
</tr>
<tr>
<td></td>
<td>cg24047259</td>
<td>14</td>
<td>65347275</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.001</td>
<td>0.0003</td>
<td>7.35E-06</td>
<td>0.266</td>
<td>-0.001</td>
<td>0.0003</td>
<td>4.90E-06</td>
</tr>
<tr>
<td>Road traffic Lden</td>
<td>cg09129334</td>
<td>13</td>
<td>11183767</td>
<td>ARHGEF7</td>
<td>Body</td>
<td>-0.007</td>
<td>0.001</td>
<td>1.73E-06</td>
<td>0.384</td>
<td>-0.007</td>
<td>0.001</td>
<td>2.12E-06</td>
</tr>
<tr>
<td></td>
<td>cg17383236</td>
<td>7</td>
<td>10001750</td>
<td>N.A.</td>
<td>N.A.</td>
<td>-0.002</td>
<td>0.0005</td>
<td>2.76E-06</td>
<td>0.384</td>
<td>-0.002</td>
<td>0.0005</td>
<td>2.74E-06</td>
</tr>
<tr>
<td></td>
<td>cg01066220</td>
<td>6</td>
<td>31696240</td>
<td>DDAT2</td>
<td>Body</td>
<td>0.001</td>
<td>0.0001</td>
<td>3.48E-06</td>
<td>0.384</td>
<td>0.001</td>
<td>0.0001</td>
<td>4.29E-06</td>
</tr>
<tr>
<td></td>
<td>cg23910243</td>
<td>16</td>
<td>31484618</td>
<td>TGBF11</td>
<td>Body</td>
<td>-0.002</td>
<td>0.0005</td>
<td>4.32E-06</td>
<td>0.384</td>
<td>-0.002</td>
<td>0.0005</td>
<td>7.13E-06</td>
</tr>
<tr>
<td></td>
<td>cg06646021</td>
<td>1</td>
<td>22940652</td>
<td>RAB4A</td>
<td>TSS1500</td>
<td>-0.005</td>
<td>0.001</td>
<td>4.47E-06</td>
<td>0.384</td>
<td>-0.005</td>
<td>0.001</td>
<td>4.78E-06</td>
</tr>
<tr>
<td></td>
<td>cg03066594</td>
<td>20</td>
<td>10415919</td>
<td>C2orf94</td>
<td>TSS200</td>
<td>0.0005</td>
<td>0.0001</td>
<td>6.00E-06</td>
<td>0.386</td>
<td>0.0004</td>
<td>0.0001</td>
<td>7.82E-06</td>
</tr>
<tr>
<td></td>
<td>cg03966094</td>
<td>22</td>
<td>21058792</td>
<td>TMEM191A</td>
<td>Body</td>
<td>-0.003</td>
<td>0.001</td>
<td>7.06E-06</td>
<td>0.386</td>
<td>-0.003</td>
<td>0.001</td>
<td>6.73E-06</td>
</tr>
</tbody>
</table>
CpG: Cytosine-phosphate-Guanine. SAPALDIA: Swiss cohort study on air pollution and lung and heart diseases in adults. CHR: chromosome. SE: standard error. Lden: day-evening-night noise level. NO$_2$: nitrogen dioxide. PM$_{2.5}$: particulate matter with aerodynamic diameter <2.5 µm. Beta coefficients represent increase or decrease in DNA methylation per two-fold increase in aircraft, railway or road traffic Lden or 10 µg/m$^3$ increase in NO$_2$ or PM$_{2.5}$. All estimates were from multi-exposure epigenome-wide linear mixed-effects models, with random intercept at the level of participant. Multi-exposure models included all five exposures (Aircraft, railway, road traffic Lden and respective truncation indicators, NO$_2$ and PM$_{2.5}$) at the same time. In a preliminary step, DNA methylation $\beta$-values were regressed on the Illumina control probe-derived first 30 principal components to correct for correlation structures and technical bias, and residuals of these regressions were used as the technical bias-corrected methylation level at the CpG sites. Extreme values of the residuals (lying beyond three times the interquartile range below the first quartile and above the third quartile at each CpG site) were replaced with their corresponding detection threshold value (“modified winsorization”). The “winsorized” data were then used as the dependent variables in the EWAS. Model 1: adjusted for age, sex, educational level, area, neighborhood socio-economic status, greenness index, smoking status, smoking pack years, exposure to passive smoke, consumption of fruits, vegetables and alcohol, nested study, asthma status, noise truncation indicators, survey and leukocyte composition (main model). Model 2: Model 1 + body mass index and physical activity. N.A.: not annotated.
Table 3. Summary of EWAS-derived differentially methylated regions and enrichment in relation to transportation noise and air pollution exposure in the SAPALDIA study

<table>
<thead>
<tr>
<th>Exposure</th>
<th>DMRs (Genes), n</th>
<th>Average effect on DMRs</th>
<th>Top DMR (Gene)</th>
<th>CpGs in top DMR, n</th>
<th>FDR p-value, top DMR</th>
<th>Top enriched canonical pathway (TECP)</th>
<th>Genes in TECP</th>
<th>Top enriched disease networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aircraft Lden</td>
<td>14 (10)</td>
<td>↓methylation (64%)</td>
<td>Chr5:135415129–135416613 (VTRNA2-I)</td>
<td>19</td>
<td>8.04E-14</td>
<td>N.A.</td>
<td>N.A.</td>
<td>Cell Cycle, Embryonic Development, Organismal Development</td>
</tr>
<tr>
<td>Railway Lden</td>
<td>48 (39)</td>
<td>↓methylation (69%)</td>
<td>Chr5:135415129–135416613 (VTRNA2-I)</td>
<td>19</td>
<td>3.61E-05</td>
<td>Type II Diabetes Mellitus Signaling; Diphthamide Biosynthesis</td>
<td>PRKAA1, SLC27A3, TNFRSF11B, DPH6</td>
<td>Cardiovascular System Development and Function, Gene Expression, Organ Development</td>
</tr>
<tr>
<td>Road traffic Lden</td>
<td>183 (189)</td>
<td>↓methylation (93%)</td>
<td>Chr20:3051954–3053196 (OXT)</td>
<td>13</td>
<td>3.93E-06</td>
<td>Wnt/β-catenin Signaling; Cholecystokinin/Gastrin-mediated Signaling</td>
<td>CDH1, CSNK1E, SOX2, SOX8, WNT16, MEF2D, PXN, SHC1, TNF</td>
<td>Cell-mediated Immune Response, Cell-To-Cell Signaling and Interaction, Cellular Movement</td>
</tr>
<tr>
<td>NO₂</td>
<td>8 (8)</td>
<td>↑methylation (63%)</td>
<td>Chr6:28303923–28304451 (ZSCAN31)</td>
<td>11</td>
<td>2.57E-06</td>
<td>N.A.</td>
<td>N.A.</td>
<td>Cell Cycle, Cell-To-Cell Signaling and Interaction, Post-Translational Modification</td>
</tr>
<tr>
<td>PM₂.₅</td>
<td>71 (60)</td>
<td>↑methylation (93%)</td>
<td>Chr6:30296689–30297941 (TRIM39, HCG18, TRIM39-RPP21)</td>
<td>14</td>
<td>4.28E-08</td>
<td>β-alanine and 4-aminobutyrate Degradation I; Systemic Lupus Erythematosus signaling</td>
<td>ABAT1, PRPF31, PRPF8, PTPN6</td>
<td>Cell Cycle, Nervous System Development and Function, Organismal Injury and Abnormalities</td>
</tr>
</tbody>
</table>

EWAS: epigenome-wide association study. Lden: day-evening-night noise level. NO₂: nitrogen dioxide. PM₂.₅: particulate matter with aerodynamic diameter <2.5 µm. FDR: false discovery rate.

SAPALDIA: Swiss cohort study on air pollution and lung and heart diseases in adults. CpG: Cytosine-phosphate-Guanine. Each DMR analysis had the corresponding multi-exposure EWAS-derived parameters as input. Multi-exposure EWAS derived from linear mixed-effects models, with random intercept at the level of participant, and adjusted for age, sex, educational level, area, neighborhood socio-economic status, greenness index, smoking status, smoking pack years, exposure to passive smoke, consumption of fruits, vegetables and alcohol, nested study, asthma status, survey, noise truncation indicators and leukocyte composition. In a preliminary step, DNA methylation β-values were regressed on the Illumina control probe-derived first 30 principal components to correct for correlation structures and technical bias, and residuals of these regressions covering 430,477 CpGs were used as the technical bias-corrected methylation level at the CpG sites. Extreme values of the residuals (lying beyond three times the interquartile range below the first quartile and above the third quartile at each CpG site) were replaced with their corresponding detection threshold value (“modified winsorization”). The “winsorized” data were then used as the dependent variables in the EWAS. Significant (FDR <0.05) and annotated DMRs were used for canonical pathway and network enrichment in the Ingenuity Pathway Analysis software (Ingenuity Systems, Redwood City, CA, USA). ↓ denotes a decrease in methylation whereas ↑ denotes an increase in methylation. N.A.: not applicable due to few significant and annotated DMRs, limiting statistical power to detect enriched canonical pathways.
Table 4. Pathway enrichment tests (p-values) for transportation noise and air pollution exposures based on curated CpGs reported for selected cross-systemic outcomes.

<table>
<thead>
<tr>
<th>Exposure</th>
<th>CRP</th>
<th>Metabolic syndrome</th>
<th>Lipids</th>
<th>FG/HbA1c</th>
<th>Insulin</th>
<th>WC</th>
<th>BMI</th>
<th>Blood pressure</th>
<th>eGFR</th>
<th>CAR</th>
<th>“Allostatic load”</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (CpGs)</td>
<td>256</td>
<td>10</td>
<td>14</td>
<td>9</td>
<td>158</td>
<td>168</td>
<td>893</td>
<td>20</td>
<td>296</td>
<td>7</td>
<td>1626</td>
</tr>
<tr>
<td>Aircraft Lden</td>
<td>0.0038</td>
<td>0.5915</td>
<td>0.6326</td>
<td>0.4545</td>
<td>0.6630</td>
<td>0.2925</td>
<td>0.0007</td>
<td>0.1020</td>
<td>0.0015</td>
<td>0.9096</td>
<td>0.0004</td>
</tr>
<tr>
<td>Railway Lden</td>
<td>0.3163</td>
<td>0.4943</td>
<td>0.9533</td>
<td>0.5778</td>
<td>0.6284</td>
<td>0.3001</td>
<td>0.9023</td>
<td>0.6773</td>
<td>0.0562</td>
<td>0.0229</td>
<td>0.0871</td>
</tr>
<tr>
<td>Road traffic Lden</td>
<td>0.0395</td>
<td>0.2434</td>
<td>0.9969</td>
<td>0.3879</td>
<td>0.3461</td>
<td>0.5361</td>
<td>0.0008</td>
<td>0.2858</td>
<td>0.0031</td>
<td>0.8013</td>
<td>0.0004</td>
</tr>
<tr>
<td>NO2</td>
<td>0.2117</td>
<td>0.1237</td>
<td>0.7831</td>
<td>0.5733</td>
<td>0.3697</td>
<td>0.6873</td>
<td>0.8818</td>
<td>0.5355</td>
<td>0.0058</td>
<td>0.9728</td>
<td>0.0510</td>
</tr>
<tr>
<td>PM2.5</td>
<td>0.0004</td>
<td>0.4517</td>
<td>0.1263</td>
<td>0.0947</td>
<td>0.3451</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.3759</td>
<td>0.0004</td>
<td>0.3814</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

CpG: Cytosine-phosphate-Guanine. CRP: C-reactive proteins. FG: fasting glucose. HbA1c: glycated hemoglobin. WC: waist circumference. BMI: body mass index. eGFR: estimated glomerular filtration rate. CAR: cardiac autonomic response. Lden: day-evening night noise level; NO2: nitrogen dioxide; PM2.5: particulate matter <2.5 microns in diameter. Lipids includes triglycerides, high-, low- and very low-density lipoprotein cholesterol. Insulin includes measures of insulin secretion and resistance. WC also includes central obesity and adiposity. BMI also includes general obesity. Blood pressure includes systolic and diastolic blood pressure. eGFR also includes impaired renal function. CAR includes acceleration and deceleration capacity. “Allostatic load” combines all the phenotypes. Pathway enrichment p-values derived from Weighted Kolmogorov-Smirnov method using the absolute values of test statistics from multi-exposure epigenome-wide association studies (EWAS), and comparing the EWAS-derived CpGs mapped to each pathway to the empirical null distribution derived by 10,000 permutation samples. The overall procedure included permutation-based multiple testing correction. EWAS was done using linear mixed-effects models, with random intercept at the level of participant, and adjusted for age, sex, educational level, area, neighborhood socio-economic status, greenness index, smoking status, smoking pack years, exposure to passive smoke, consumption of fruits, vegetables and alcohol, nested study, asthma status, noise truncation indicators, survey and leukocyte composition. In a preliminary step, DNA methylation β-values were regressed on the Illumina control probe-derived first 30 principal components to correct for correlation structures and technical bias, and residuals of these regressions covering 430,477 CpGs were used as the technical bias-corrected methylation level at the CpG sites. Extreme values of the residuals (lying beyond three times the interquartile range below the first quartile and above the third quartile at each CpG site) were replaced with their corresponding detection threshold value (“modified winsorization”). The “winsorized” data were then used as the dependent variables in the EWAS.
Table 5. Replication of previously reported EWAS signals for long-term exposure to NO$_2$ and PM$_{2.5}$, in the SAPALDIA study, single and multi-exposure models

<table>
<thead>
<tr>
<th>Air pollutant; Cohort (Reference)</th>
<th>CpG ID</th>
<th>CHR</th>
<th>Location</th>
<th>Gene</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
<th>Beta</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO$_2$; LifeLines (de FC Lichtenfels et al. 2018)</td>
<td>cg04908668</td>
<td>6</td>
<td>32823941</td>
<td>PSMB9</td>
<td>-0.012</td>
<td>0.002</td>
<td>7.94E-09</td>
<td>0.0003</td>
<td>0.0003</td>
<td>0.333</td>
<td>0.0004</td>
<td>0.0004</td>
<td>0.322</td>
</tr>
<tr>
<td></td>
<td>cg14938677</td>
<td>7</td>
<td>127231698</td>
<td>ARF5</td>
<td>0.023</td>
<td>0.004</td>
<td>1.05E-08</td>
<td>0.0004</td>
<td>0.001</td>
<td>0.619</td>
<td>0.0004</td>
<td>0.001</td>
<td>0.663</td>
</tr>
<tr>
<td></td>
<td>cg00344801</td>
<td>22</td>
<td>46685728</td>
<td>TTC38</td>
<td>-0.028</td>
<td>0.005</td>
<td>2.38E-08</td>
<td>0.001</td>
<td>0.001</td>
<td>0.209</td>
<td>0.001</td>
<td>0.001</td>
<td>0.282</td>
</tr>
<tr>
<td></td>
<td>cg18379295</td>
<td>14</td>
<td>52326155</td>
<td>GNG2</td>
<td>0.020</td>
<td>0.004</td>
<td>3.50E-08</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.204</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.367</td>
</tr>
<tr>
<td></td>
<td>cg25769469</td>
<td>5</td>
<td>71643841</td>
<td>PTCD2</td>
<td>0.035</td>
<td>0.006</td>
<td>3.69E-08</td>
<td>0.001</td>
<td>0.002</td>
<td>0.547</td>
<td>0.002</td>
<td>0.002</td>
<td>0.237</td>
</tr>
<tr>
<td></td>
<td>cg02234653</td>
<td>2</td>
<td>224625080</td>
<td>AP1S3</td>
<td>-0.017</td>
<td>0.003</td>
<td>4.70E-08</td>
<td>0.0005</td>
<td>0.001</td>
<td>0.553</td>
<td>0.0003</td>
<td>0.001</td>
<td>0.752</td>
</tr>
<tr>
<td></td>
<td>cg08500171</td>
<td>6</td>
<td>31590674</td>
<td>BAT2</td>
<td>0.024</td>
<td>0.004</td>
<td>9.81E-08</td>
<td>0.002</td>
<td>0.001</td>
<td>0.042</td>
<td>0.003</td>
<td>0.001</td>
<td>0.199</td>
</tr>
<tr>
<td>NO$_2$; EPIC-ITALY (Plusquin et al. 2017)</td>
<td>cg08120023</td>
<td>1</td>
<td>116947203</td>
<td>C1orf203</td>
<td>-0.003</td>
<td>0.0005</td>
<td>3.02E-09</td>
<td>0.001</td>
<td>0.001</td>
<td>0.586</td>
<td>0.002</td>
<td>0.001</td>
<td>0.179</td>
</tr>
<tr>
<td></td>
<td>cg22856765</td>
<td>8</td>
<td>42693384</td>
<td>THAP1</td>
<td>-0.008</td>
<td>0.001</td>
<td>4.27E-09</td>
<td>-0.00001</td>
<td>0.0002</td>
<td>0.960</td>
<td>-0.00001</td>
<td>0.0002</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td>cg18164357</td>
<td>11</td>
<td>77534497</td>
<td>C11orf67</td>
<td>-0.009</td>
<td>0.001</td>
<td>9.61E-09</td>
<td>0.00005</td>
<td>0.0003</td>
<td>0.042</td>
<td>0.001</td>
<td>0.0003</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>cg13918628</td>
<td>9</td>
<td>35610380</td>
<td>CD72</td>
<td>-0.012</td>
<td>0.002</td>
<td>1.02E-08</td>
<td>-0.0003</td>
<td>0.0002</td>
<td>0.184</td>
<td>-0.0005</td>
<td>0.0003</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>cg03870188</td>
<td>13</td>
<td>113717830</td>
<td>MCF2L</td>
<td>-0.004</td>
<td>0.001</td>
<td>1.02E-08</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.670</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.650</td>
</tr>
<tr>
<td></td>
<td>cg20939320</td>
<td>3</td>
<td>132563279</td>
<td>NCRNA00119</td>
<td>-0.006</td>
<td>0.001</td>
<td>3.49E-08</td>
<td>0.001</td>
<td>0.001</td>
<td>0.091</td>
<td>0.001</td>
<td>0.001</td>
<td>0.185</td>
</tr>
<tr>
<td></td>
<td>cg13420207</td>
<td>7</td>
<td>81666278</td>
<td>CACNA2D1</td>
<td>-0.010</td>
<td>0.002</td>
<td>5.56E-08</td>
<td>-0.00004</td>
<td>0.001</td>
<td>0.947</td>
<td>-0.00003</td>
<td>0.001</td>
<td>0.973</td>
</tr>
<tr>
<td></td>
<td>cg04914283</td>
<td>1</td>
<td>23181832</td>
<td>EPHB2</td>
<td>-0.005</td>
<td>0.001</td>
<td>5.85E-08</td>
<td>0.001</td>
<td>0.001</td>
<td>0.298</td>
<td>0.001</td>
<td>0.001</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>cg21156210</td>
<td>4</td>
<td>100485208</td>
<td>RGMMD2</td>
<td>0.010</td>
<td>0.002</td>
<td>6.03E-08</td>
<td>-0.00002</td>
<td>0.0002</td>
<td>0.941</td>
<td>0.0001</td>
<td>0.0003</td>
<td>0.731</td>
</tr>
<tr>
<td></td>
<td>cg16205861</td>
<td>12</td>
<td>54146572</td>
<td>N.A.</td>
<td>-0.004</td>
<td>0.001</td>
<td>6.57E-08</td>
<td>-0.00001</td>
<td>0.0005</td>
<td>0.788</td>
<td>-0.0001</td>
<td>0.0005</td>
<td>0.882</td>
</tr>
<tr>
<td></td>
<td>cg12790758</td>
<td>15</td>
<td>37369914</td>
<td>MEIS2</td>
<td>-0.004</td>
<td>0.001</td>
<td>7.06E-08</td>
<td>0.00006</td>
<td>0.0005</td>
<td>0.287</td>
<td>-0.0001</td>
<td>0.001</td>
<td>0.898</td>
</tr>
<tr>
<td>NO2: Korean COPD Cohort (Lee et al. 2019)</td>
<td>PM2.5: EPIC-ITALY (Plusquin et al. 2017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg18201392 1 185023741  RNF2  -0.005  0.001  8.02E-08  0.0002  0.0001  0.118  0.0002  0.0001  0.133</td>
<td>cg12575202 10 133331128  N.A.  -0.467  0.080  5.40E-09  -0.001  0.003  0.786  -0.002  0.003  0.529</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg05171937 12 27396765  STK38L  0.010  0.002  1.10E-08  0.003  0.001  0.068  0.002  0.002  0.125</td>
<td>cg08630381 13 100612277  N.A.  0.461  0.073  2.58E-10  -0.001  0.001  0.415  -0.001  0.001  0.394</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg06226567 20 22559676  C20orf56  0.003  0.001  3.50E-08  0.0002  0.0001  0.863  0.0001  0.0002  0.514</td>
<td>cg17629796a 13 30707265  N.A.  -0.563  0.094  2.11E-09  -0.002  0.001  0.076  -0.003  0.001  0.033</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg25838725 13 110397643  N.A.  -0.001  2.3E-04  4.90E-08  0.0001  0.0001  0.183  0.0001  0.0001  0.228</td>
<td>cg07084345 15 61972967  N.A.  -0.512  0.075  7.26E-12  0.002  0.008  0.816  0.002  0.008  0.769</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg04319606 2 26785290  C2orf70  0.261  0.068  1.31E-07  0.0002  0.002  0.893  -0.00002  0.002  0.989</td>
<td>cg04319606 2 26785290  C2orf70  0.261  0.068  1.31E-07  0.0002  0.002  0.893  -0.00002  0.002  0.989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg09568355 2 45228633  N.A.  0.261  0.049  1.41E-07  0.003  0.002  0.286  0.004  0.003  0.179</td>
<td>cg03513315 2 30988383  PES1  0.307  0.058  1.35E-07  -0.0003  0.001  0.631  -0.0001  0.001  0.837</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg03513315 2 30988383  PES1  0.307  0.058  1.35E-07  -0.0003  0.001  0.631  -0.0001  0.001  0.837</td>
<td>cg25489413 7 44794343  ZMIZ2  -0.365  0.068  6.48E-08  0.001  0.002  0.712  0.0002  0.002  0.933</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg00005622 8 145180403  N.A.  -0.398  0.064  5.16E-10  -0.004  0.001  0.109  -0.004  0.003  0.136</td>
<td>cg05171937 12 27396765  STK38L  0.010  0.002  1.10E-08  0.003  0.001  0.068  0.002  0.002  0.125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

EWAS: epigenome-wide association study. CpG: Cytosine-phosphate-Guanine. CHR: chromosome. SE: standard error. All SAPALDIA estimates were derived from linear mixed-effects EWAS models, with random intercept at the level of participant, adjusted for age, sex, educational level, area, neighborhood socioeconomic status, greenness index, smoking status, smoking pack years, exposure to passive smoke, consumption of fruits, vegetables and alcohol, nested study, asthma status, noise truncation indicators, survey and leukocyte composition. In a preliminary step, DNA methylation β-values were regressed on the Illumina control probe-derived first 30 principal components to correct for correlation structures and technical bias, and residuals of these regressions covering 430,477 CpGs were used as the technical bias-corrected methylation level at the CpG sites. Extreme values of the residuals (lying beyond three times the interquartile range below the first quartile and above the third quartile at each CpG site) were replaced with their corresponding detection threshold value (“modified winsorization”). The “winsorized” data were then used as the dependent variables in the present EWAS. The multi-exposure model contained all five exposures at the same time. a Validated in the SAPALDIA study. b SAPALDIA estimates derived from never-smoker sample comparable to the EPIC-NL never-smoker estimates. N.A.: not annotated.
Figure 1. Selection of participants included in the present study

SAPALDIA: Swiss cohort study on air pollution and lung and heart diseases in adults. ALEC: Aging lungs in European cohorts. ALEC and EXPOsOMICS are European cohort consortia in which SAPALDIA participates. DNA methylation was measured using Illumina Infinium 450K BeadChip and processed in the same manner across the ALEC and EXPOsOMICS samples, to derive the residuals of the beta values (corrected for technical bias) of overlapping 430,477 CpGs, which were subsequently applied to the present epigenome-wide association study.

Figure 2. Overlap of top 100 CpG signals (A) and genes annotated to significant differentially methylated regions (B) in relation to aircraft, railway, and road traffic Lden, NO2 and PM2.5 identified from multi-exposure EWAS in the SAPALDIA study

Lden: day-evening-night noise level. NO2: nitrogen dioxide; PM2.5: particulate matter <2.5 microns in diameter. EWAS: epigenome-wide association study. SAPALDIA: Swiss cohort study on air pollution and lung and heart diseases in Adults. CpGs were identified by multi-exposure EWAS using multivariable linear mixed-effects models with random intercepts at the level of participants, and adjusted for age, sex, educational level, area, neighborhood socio-economic status, greenness index, smoking status, smoking pack years, exposure to passive smoke, consumption of fruits, vegetables and alcohol, nested study, asthma status, noise truncation indicators, survey and leukocyte composition. In a preliminary step, DNA methylation β-values were regressed on the Illumina control probe-derived first 30 principal components to correct for correlation structures and technical bias, and residuals of these regressions covering 430,477 CpGs were used as the technical bias-corrected methylation level at the CpG sites. Extreme values of the residuals (lying beyond three times the interquartile range below the first quartile and above the third quartile at each CpG site) were replaced with their corresponding detection threshold value (“modified winsorization”). The “winsorized” data were then used as the dependent variables in the present EWAS. CpGs (annotated gene) intersecting at the level of road traffic Lden and NO2 were cg12439232 and cg15590912 (CCSAP). Genes (DMR) intersecting at the level of road traffic Lden and PM2.5 was PRRT1 (chr6:32115964–32117401); and at the level of aircraft, road traffic Lden and PM2.5 was HOXA2 (chr7:27141774–27143806). VTRNA2-1 (chr5:135415129–135416613) intersected between aircraft and railway Lden, ZFP57 (chr6:29648161–29649084) between railway Lden and NO2, and ZSCAN31 (chr6:28303923–28304451) between road traffic Lden and NO2. TRIM39, TRIM39-RPP21 and HCG18 (chr6:30296689–30297941) intersected between railway Lden and PM2.5, whereas SLC27A3 (chr1:153746588–153747856), B3GALT4 (chr6:33244976–33246185), EN2 and AC008060.8 (chr7:155249398–155251925) intersected between railway and road traffic Lden.