Environment International 108 (2017) 127–136

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

DNA methylation and exposure to ambient air pollution in two prospective cohorts

Michelle Plusquin[a,b,c,1], Florence Guida[a,b,1], Silvia Polidoro[d], Roel Vermeulen[b,e], Ole Raaschou-Nielsen[f,g], Gianluca Campanella[a,b], Gerard Hoek[e], Soterios A. Kyrtopoulos[h], Panagiotis Georgiadis[h], Alessio Naccarati[d], Carlotta Sacerdote[i], Vittorio Krogh[j], H. Bas Bueno-de-Mesquita[a,k,l], W.M. Monique Verschuren[m,n], Sergi Sayols-Baixeras[o,p], Tommaso Panni[q], Annette Peters[q], Dennie G.A.J. Hebels[r,s], Jos Kleinjans[r], Paolo Vineis[a,b,d,2], Marc Chadeau-Hyam[a,b,e,*,2]

[a] Department of Epidemiology and Biostatistics, The School of Public Health, Imperial College London, London, United Kingdom
[b] Medical Research Council-Health Protection Agency Centre for Environment and Health, Imperial College London, London, United Kingdom
[c] Centre for Environmental Sciences, Hasselt University, Hasselt, Belgium
[d] IIGM, Italian Institute for Genomic Medicine, Turin, Italy
[e] Institute for Risk Assessment Sciences (IRAS), Division of Environmental Epidemiology, Utrecht University, Utrecht, The Netherlands
[f] Danish Cancer Society Research Center, Copenhagen, Denmark
[g] Department of Environmental Science, Aarhus University, Roskilde, Denmark
[h] National Hellenic Research Foundation, Institute of Biology, Medicinal Chemistry and Biotechnology, 48 Vas. Constantinou Ave., Athens 11635, Greece
[i] Unit of Cancer Epidemiology-CERMS, Department of Medical Sciences, University of Turin and Città Della Salute E Della Scienza Hospital, Turin, Italy
[j] Epidemiology Unit, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
[k] Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[l] Department of Social & Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia
[m] Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht, The Netherlands
[n] Centre for Nutrition, Prevention and Health Services, National Institute for Public Health and the Environment, Bilthoven, The Netherlands
[o] Cardiovascular Epidemiology and Genetics Research Group, IMIM (Hospital del Mar Medical Research Institute), 08003 Barcelona, Catalonia, Spain
[p] Universitat Pompeu Fabra (UPF), 08003 Barcelona, Catalonia, Spain
[q] Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany
[r] Department of Toxicogenomics, Maastricht University, The Netherlands
[s] Department of Cell Biology-Inspired Tissue Engineering, MERLN Institute, Maastricht University, Maastricht, The Netherlands

ARTICLE INFO

Keywords:
Air pollution
Epigenome-wide DNA methylation
Illumina 450 k human methylation array
Particulate matter
NO_x
EPIC

ABSTRACT

Long-term exposure to air pollution has been associated with several adverse health effects including cardiovascular, respiratory diseases and cancers. However, underlying molecular alterations remain to be further investigated. The aim of this study is to investigate the effects of long-term exposure to air pollutants on (a) average DNA methylation at functional regions and, (b) individual differentially methylated CpG sites. An assumption is that omic measurements, including the methylome, are more sensitive to low doses than hard health outcomes.

This study included blood-derived DNA methylation (Illumina-HM450 methylation) for 454 Italian and 159 Dutch participants from the European Prospective Investigation into Cancer and Nutrition (EPIC). Long-term air pollution exposure levels, including NO_2, NO_x, $PM_{2.5}$, PM_{coarse}, PM_{10}, $PM_{2.5\ absorbance}$ (soot) were estimated using models developed within the ESCAPE project, and back-extrapolated to the time of sampling when possible. We meta-analysed the associations between the air pollutants and global DNA methylation, methylation in

Abbreviations: CpG, 5′–C–phosphate–G–3′, a cytosine and guanine separated by one phosphate; CPMA, cross-phenotype meta-analysis; ELISA, enzyme-linked immunosorbent assay; ESCAPE, European Study of Cohorts for Air Pollution Effects; EPIC-Italy, The Italian part of the European Prospective Investigation into Cancer and Nutrition; EPIC-NL, The Dutch part of the European Prospective Investigation into Cancer and Nutrition; EpWAS, epigenome-wide association study; HPLC, high-performance liquid chromatography; LC-MS, liquid chromatography–mass spectrometry; LINE1, long interspersed nuclear elements; LUMA, luminometric methylation assay; NO_2, nitrogen dioxide; NO_x, nitrogen oxide; PM, particulate matter; $PM_{2.5}$, particulate matter with a diameter smaller than 2.5 μm; PM_{10}, particulate matter with a diameter smaller than 10 μm; PM_{coarse}, particulate matter with a diameter smaller than 10 μm and bigger than 2.5 μm; $PM_{2.5abs}$, absorbance of the $PM_{2.5}$ filter, a measure for soot

* Corresponding author at: MRC-PHE Centre for Environment and Health, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG, United Kingdom.
E-mail address: m.chadeau@imperial.ac.uk (M. Chadeau-Hyam).

[1] Joint first authors.
[2] Joint last authors.

http://dx.doi.org/10.1016/j.envint.2017.08.006
Received 30 January 2017; Received in revised form 9 August 2017; Accepted 9 August 2017

functional regions and epigenome-wide methylation. CpG sites found differentially methylated with air pollution were further investigated for functional interpretation in an independent population (EnviroGenoMarkers project), where (N = 613) participants had both methylation and gene expression data available.

Exposure to NO_2 was associated with a significant global somatic hypomethylation (p-value = 0.014). Hypomethylation of CpG island's shores and shelves and gene bodies was significantly associated with higher exposures to NO_2 and NO_x. Meta-analysing the epigenome-wide findings of the 2 cohorts did not show genome-wide significant associations at single CpG site level. However, several significant CpG were found if the analyses were separated by countries. By regressing gene expression levels against methylation levels of the exposure-related CpG sites, we identified several significant CpG-transcript pairs and highlighted 5 enriched pathways for NO_2 and 9 for NO_x mainly related to the immune system and its regulation.

Our findings support results on global hypomethylation associated with air pollution, and suggest that the shores and shelves of CpG islands and gene bodies are mostly affected by higher exposure to NO_2 and NO_x. Functional differences in the immune system were suggested by transcriptome analyses.

1. Introduction

Ambient air pollution includes gaseous components such as nitrogen oxides, benzene, and sulfur dioxide as well as particulate matter. The latter consists of acids, organic chemicals, metals and soil or dust particles of varying aerodynamic diameters. Because of their small size, these particles can be inhaled deeply into the lungs and deposited in the alveoli. The smallest particles can penetrate the lung epithelium and reach the blood stream (Brook et al., 2010). Exposure to air pollution is associated with adverse health outcomes including cardiovascular and respiratory diseases (Brook et al., 2010; Nawrot et al., 2011). There is increasing evidence that toxic substances in ambient air at levels that are at or below the current limits in the European Union may increase the risk of lung cancer. Analyses of large European cohorts showed that a 10 $\mu g/m^3$ increase in PM_{10} was associated with a 22% (95% confidence interval [CI]: 3 to 45%) greater risk for lung cancer, while a 5 $\mu g/m^3$ increase in $PM_{2.5}$ was associated with an 18% (95% CI: − 4 to 46%) increased risk (Raaschou-Nielsen et al., 2013). However, the potential modes of action of air pollutants are not well understood. Also, hard health outcomes such as cancer or cardiovascular diseases may not be sensitive enough (or require very large cohorts) to demonstrate low-dose effects.

Previous research suggests that inflammation, oxidative damage, and mitochondrial dysfunction (Demetriou et al., 2012; Mostafavi et al., 2015; Pettit et al., 2012) may be the underlying mechanisms leading from exposure to air pollution to health outcomes. Global DNA hypomethylation induces genomic instability, for instance through chromatin structure modelling (You and Jones, 2012), loss of imprinting, and increased activation of oncogenes (Feinberg and Tycko, 2004). Epigenetic studies report an inverse association between global methylation and long-term exposure to ambient air pollution (Baccarelli et al., 2009; De Prins et al., 2013; Janssen et al., 2013; Sanchez-Guerra et al., 2015; Tao et al., 2014), especially affecting 5-hydroxymethylcytosine methylation (Sanchez-Guerra et al., 2015). However, analyses of global methylation in these studies were based on the investigation of LINE-1 or Alu regions, or by means of HPLC, ELISA, LUMA or LC-MS, and analyses of functional regions in the genome have not yet been performed in relation to long-term exposure to air pollution.

Methylation levels at specific loci in genes including tissue factor F3, interferon gamma (*IFN-γ*), interleukin 6 (*IL-6*), toll-like receptor 2 (*TLR-2*), intercellular adhesion molecule 1 (*ICAM-1*), ten-eleven translocation (*TET1*) have been reported to be affected by exposure to air pollution (Bind et al., 2015; Bind et al., 2014; Lepeule et al., 2014; Somineni et al., 2015). Recent studies using Illumina Infinium Human Methylation 450 K technology (a) identified CpG sites whose methylation levels were affected by short and mid-term exposure to particulate matter (Panni et al., 2016); (b) investigated the association between long-term ambient air pollution exposure and DNA methylation (candidate sites and global) in monocytes of adults (Chi et al., 2016); (c) identified differential offspring DNA methylation in mitochondria-related genes in association with NO_2 exposure during pregnancy (Gruzieva et al., 2016). Also, long-term exposure to air pollution has been found associated with epigenetic aging measures (Ward-Caviness et al., 2016).

Using data from the Italian and Dutch components of the European Prospective Investigation into Cancer and Nutrition cohort study (EPIC), we investigate if differences in global DNA methylation or DNA methylation at certain functional regions could be induced by long-term exposure to air pollutants and could be used as a marker of low-dose effects. In addition, we adopt an epigenome-wide association study (EpWAS) approach to identify possible CpG sites whose methylation levels are affected by long-term exposure to air pollution. Using an independent population in which DNA methylation profiles and gene expression are available in the same participants (N = 613), we linked full-resolution gene expression levels data and methylation levels at exposure-related CpG sites to potentially characterise the functional consequences of these methylation alterations.

2. Methods

We first assessed the association between ambient air pollution estimates and global DNA methylation including functional regions in the two EPIC cohorts. In the second part of the manuscript, we performed an epigenome-wide association study. Finally, we investigated the transcripts associated with air pollution CpGs in the EnviroGenoMarkers study.

2.1. Study populations

EPIC is a multi-centre prospective cohort based on healthy, middle-aged subjects who agreed, following an active invitation, to participate in the study and to have their health status followed up for the rest of their lives. The present study included participants from two large population-based cohorts: the Italian and Dutch components of the EPIC study (EPIC-Italy, N = 47,749 and EPIC-Netherlands (EPIC-NL), N = 33,066); (Beulens et al., 2010; Palli et al., 2003).

- The Italian samples originate from two case-control studies on breast cancer (N = 231) (van Veldhoven et al., 2015) and colorectal cancer (N = 304), from two EPIC – Italy centres: Varese and Turin.
- The Dutch samples originate from a longitudinal study on weight change in healthy women nested within EPIC-NL (N = 170).

The rationale and design of the EPIC study have been described elsewhere (Riboli and Kaaks, 1997). The EPIC study protocol was approved by the ethical review boards of the International Agency for Research and Cancer (IARC) and by the local participating centres. For all participants, anthropometric measurements and lifestyle variables were collected at recruitment (1993–1998) through standardized

questionnaires, together with a blood sample. The Italian cancer cases were diagnosed with colon or breast cancer > 1 year after recruitment, and were healthy at the time of sampling.

2.2. Air pollution exposure assessment

Following the ESCAPE protocol (Beelen et al., 2013; Eeftens et al., 2012) addresses of the participants have been geocoded and air pollution concentrations at the baseline residential addresses of study participants were estimated by land-use regression models in a three-step, standardized procedure. First, particulate matter with an aerodynamic diameter of < 10 μm (PM_{10}), particulate matter with aerodynamic diameter of < 2.5 μm ($PM_{2.5}$), $PM_{2.5}$ absorbance ($PM_{2.5abs}$) determined by measurement of light reflectance (a marker for soot and black carbon), nitrogen oxides (NO_x), and nitrogen dioxide (NO_2) were measured in different seasons at residential locations for each cohort between October 2008, and April 2011 (Beelen et al., 2013; Eeftens et al., 2012). PM_{coarse} was calculated as the difference between PM_{10} and $PM_{2.5}$ (i.e., PM with diameter 2.5–10 μm). In Varese, only NO_2 and NO_x were available. Second, land-use regression models were developed for each pollutant in each study area, with the yearly mean concentration as the dependent variable and an extensive list of geographical attributes as possible predictors (Beelen et al., 2013; Eeftens et al., 2012). Generally, predictors for PM_{10}, $PM_{2.5}$, NO_x, and NO_2 were related to traffic and population density. Variables related to industry, proximity to a port, and altitude were also predictors in some models. The models generally explained a large fraction of measured spatial variation, the R^2 from leave-one-out cross-validation usually falling between 0.60 and 0.80. Finally, the models were used to assess exposure at the baseline address of each cohort member. Details on this procedure can be found on the website http://www.escapeproject.eu/manuals/. In EPIC-NL and the Turin component of EPIC Italy PM_{10}, PM_{coarse}, $PM_{2.5}$, $PM_{2.5\ absorbance}$, NO_x and NO_2 are available. In the Varese component of EPIC-Italy only NO_x and NO_2 are available.

Air pollution measurements were performed in 2008–2011, > 10 years after enrolment. We therefore extrapolated air pollution concentrations predicted by the land use regression models back to the time of enrolment using the absolute difference and the ratio between the two periods, based on data from routine background monitoring network site(s). For each study participant's home address, the back extrapolated concentration was obtained by either multiplying the modelled ESCAPE annual mean concentration with the ratio between average annual concentrations as derived from the routine monitoring site(s) for baseline and for the ESCAPE measurement period time or by adding to the modelled ESCAPE annual mean concentration the difference between the baseline and the ESCAPE measurement period time. The needed routine measurement data for back-extrapolation were only available for NO_2 and NO_x in EPIC-Italy and for NO_2, NO_x, PM_{10} and $PM_{2.5}$ absorbance in EPIC-NL. Details about the back-extrapolation method can be found on the following website http://www.escapeproject.eu/manuals/Procedure_for_extrapolation_back_in_time.pdf.

2.3. Epigenome-wide DNA methylation

Epigenome-wide DNA methylation profiles were obtained using the Illumina Infinium® HumanMethylation450 BeadChip. Epigenome-wide DNA methylation analyses were performed on samples using the Illumina Infinium HumanMethylation450 platform. Laboratory procedures were carried out according to manufacturers' protocols at the Human Genetics Foundation (Turin, Italy) and at ServiceXS BV (Leiden, The Netherlands) for the Italian and Dutch samples respectively. Buffy coats (leukocytes) stored in liquid nitrogen were thawed, and genomic DNA was extracted using the QIAGEN QIAsymphony DNA Midi Kit. DNA (500 ng) was bisulphite-converted using the Zymo Research EZ-96 DNA Methylation-Gold™ Kit, and hybridised to Illumina Infinium Human Methylation450K BeadChips. These were subsequently scanned using the Illumina HiScanSQ system, and sample quality was assessed using control probes present on the micro-arrays. Finally, raw intensity data were exported from Illumina GenomeStudio (version 2011.1). Data pre-processing was carried out using in-house software written for the R statistical computing environment. For each sample and each probe, measurements were set to missing if obtained by averaging intensities over less than three beads, or if averaged intensities were below detection thresholds estimated from negative control probes. Background subtraction and dye bias correction (for probes using the Infinium II design) were also performed. The probes targeting autosomal CpG loci (number of sex chromosomes = 11,648) were selected for further analyses. In these lists probes targeting non-specific CpGs (N = 40,590) were removed (Price et al., 2013). Probes detected in < 20% of the samples were excluded from the analyses leaving us with N = 398,372 and 339,810 for Italy and the Netherlands respectively for the analyses. Methylation levels at each CpG locus were expressed as the ratios of intensities arising from methylated CpGs over those arising from the sum of methylated and unmethylated CpGs (Beta values). The data was trimmed for outliers defined by > 3 interquartile ranges below the first quartile or above the fourth quartile.

2.4. Gene expression

Both blood-derived methylation (using the same Illumina 450 K technology) and genome-wide gene expression profiles (Agilent 4 × 44 K human whole genome microarray platform) were available in N = 613 participants from the EnviroGenomarkers study, which is described elsewhere (www.envirogenomarkers.net) (Georgiadis et al., 2016). Briefly, this study combines subjects from EPIC-ITALY (EPIC-Turin, EPIC-Florence, EPIC-Naples, EPIC-Varese, EPIC-Ragusa), which are different from the EPIC-Italy samples used for the air pollution EpWAS, and subjects from the Northern Sweden Health and Disease Study (NSHDS) (Bingham and Riboli, 2004; Hallmans et al., 2003). Both studies used population-based recruitment with standardized lifestyle and personal history questionnaires, anthropometric data and blood samples collected at recruitment (1993–1998 for EPIC-ITALY; 1990–2006 for NSHDS). The EnviroGenomarkers project and its associated studies and protocols were approved by the Regional Ethical Review Board of the Umea Division of Medical Research, as regards the Swedish cohort, and the Health Unit Local Ethical Committee, as regards the Italian cohort, and all participants gave written informed consent.

2.5. Statistical analyses

2.5.1. DNA methylation

After removing subjects that failed quality control for the DNA methylation arrays (N = 5 for EPIC-Italy and N = 11 for EPIC-NL) and those with missing values on air pollution estimates or confounders (N = 15 covariates and 61 air pollution for EPIC-Italy), 454 and 159 subjects were included in the analyses for EPIC-Italy and EPIC-NL, respectively.

Because of the strong air pollution contrast between Italy and the Netherlands, with lower air pollution estimates in the Netherlands not overlapping with those of Italy (Fig. S1), pooling data from both countries would hamper the possibility to disentangle country-specific from exposure effects, and would therefore penalize our results interpretability. For that reason, we meta-analysed the 2 EPIC cohorts and also adopted a stratified approach by showing the results of the two datasets separately. We considered three ways to define individual methylation profiles in concordance with a method described in a previous study (van Veldhoven et al., 2015): (i) global methylation, as measured by the arithmetic mean of beta-values across all somatic probes (N = 337,779) that were retained after quality control checks both in EPIC Italy and EPIC NL, (ii) methylation of seven categories of

CpG sites: physical characteristics (CpG islands, shelves, shores, or other) and functional characteristics (promoter, gene body or intergenic), and (iii) EpWAS: methylation fraction at each of the CpG sites assayed that passed quality control (N = 398,372 and 339,810 for Italy and the Netherlands). Irrespective of the DNA methylation metric, we investigated potential differences in relation to exposure to long-term air pollution estimates in EPIC-Italy and EPIC-NL.

As already proposed (Campanella et al., 2014), DNA methylation levels were modelled as dependent variable in a generalized linear model with beta-distributed response using the parameterization of Ferrari and Cribari-Neto (2004). The β value (0–1 scale) represents the difference in methylation percentage for every unit ($\mu g/m^3$) increase of air pollutant. To account for residual technical confounding, all models were adjusted for the position of the sample on the micro-array. Adjustment covariates included age at blood collection (continuous), smoking status (never, former, current), and sex; except for EPIC-NL which has only females and never and former smokers. In EPIC-Italy, models were also adjusted for the eventual prospective cancer diagnosis later in life (breast or colon cancer). All analyses were adjusted for blood cell composition estimated using an established de-convolution approach (Houseman et al., 2012). We employed a cross-phenotype meta-analysis approach (CPMA) (Cotsapas et al., 2011) based on the associations between global and regional methylation and air pollution and we used fixed effects models with inverse-variance weighting for individual CpG sites. In the latter, multiple testing was corrected for by using a stringent Bonferroni correction (ensuring a control of the family wise error rate below 0.05 in the epigenome-wide analyses significance level = 1.1e − 7). For completeness, we report CpG sites that are significant after FDR-correction (< 0.20).

2.5.2. Targeted integration of gene expression data, pathway analyses

We used linear regressions models relating all assayed transcripts (N = 29,662) to the methylation levels in a subset of CpG sites found to be associated with air pollution in at least one cohort. As already reported, (Chadeau-Hyam et al., 2014) we used a linear model approach to account for technical variation during the acquisition of gene expression and methylation profiles. Specifically, we ran a linear model setting the dates of the main experimental steps: RNA isolation, labelling, and hybridization, for gene expression data and array and position on array for methylation data. Removing from the observed values of gene expression and methylation levels the estimates of the random effects we obtained 'technically de-noised' profiles which were subsequently used in the linear models for targeted integration. CpG–transcript pairs were considered significant based on Bonferroni 5% significance level (per-test significance level $\alpha' = 0.05 / (N \times 29{,}662)$, where N denotes the number of CpG sites under investigation. Genes whose expression level was found involved in any CpG-transcript pairs were uploaded into an online overrepresentation analysis tool, DAVID 6.7 (https://david.ncifcrf.gov), to identify enriched pathways. Pathways with Bonferroni p-value smaller than 0.001 and fold change bigger than 1.5 were considered significant (Huang da et al., 2009). We additionally performed the same pathways analyses using genes expressed in human peripheral blood leukocyte (Palmer et al., 2006) as background.

3. Results

Population characteristics and air pollution estimates are reported in Tables 1, A.1 and Fig. A.1. Across all air pollution estimates, levels in Italy were much higher than those in the Netherlands, with limited overlap. The EPIC-NL study only consisted of female nonsmokers. Characteristics of the excluded subjects with missing values are compared with the study population in Supplementary Table A.2.

3.1. Global methylation

We observed inverse associations between NO_2 levels and global methylation on somatic chromosomes in EPIC-Italy (β = − 0.00007; SE = 0.00003; p-value = 0.014), in EPIC-NL (β = − 0.00388; SE = 0.00133; p-value = 0.031); and from the meta-analyses of the results of both cohorts (CPMA p-value = 0.014; Table 2). In association to NO_x, we observed a significant difference of − 0.01% in methylation for every unit ($\mu g/m^3$) increase in NO_X exposure in EPIC-Italy and a trend for lower methylation in the meta-analysis (CPMA p-value = 0.089).

A meta-analysis on exposure to NO_2 identified statistically significant average hypomethylation associated to NO_2 in all genetic and functional regions except for CpG-islands (CPMA p-values ranging from 0.010 to 0.029). Meta-analysing results from both studies, we found consistent average hypomethylation across shelves, shores, other non-CpG-island related regions and gene bodies (CPMA p-values ranging from 0.019 to 0.045) in relation to environmental exposure to NO_x. These associations were also detected in stratified analyses by cohort, but were weaker in EPIC-NL. We conducted meta-analyses for exposures to PM which did not result in any statistically significant finding. In EPIC-Italy, exposures to PM_{10} and PM_{coarse} were found significantly associated with hypermethylation on the promoter region of genes (PM_{10}: β = 0.0005; SE = 0.0002; p-value = 0.032; PM_{coarse}: β = 0.0007; SE = 0.0003; p-value = 0.035).

For the available back-extrapolated data, we found a significant or borderline significance global hypomethylation on somatic chromosomes in relation to exposure to NO_2 and NO_x in both countries and with the two back-extrapolation methods (Supplemental Table A.3 and A.4). As previously observed, almost all functional regions except CpG islands were hypomethylated and the associations were weaker in EPIC-NL. Back-extrapolated data for particulate matter were only available in EPIC-NL (PM_{10} and $PM_{2.5}$ absorbance). As previously observed, no significant associations were seen for $PM_{2.5}$ absorbance but a significant or borderline significance hypermethylation was seen for PM_{10} on somatic chromosomes and functional regions except CpG islands, shores and promoters (contrary to previously).

Sensitivity analyses excluding the subjects that developed cancer

Table 1
Population characteristics at sampling.

	EPIC-Italy	EPIC-Netherlands
N	454	159
Females	323 (71%)	159 (100%)
Age	54.2 ± 7.1	58.8 ± 5.6
Smoking		
Never	234 (52%)	93 (58%)
Former	111 (24%)	66 (42%)
Current	109 (24%)	–
BMI	25.5 ± 4.0	25.7 ± 4.2
Education		
Primary school	253 (56%)	29 (18%)
Secondary school	164 (36%)	107 (67%)
University level	37 (8%)	23 (15%)
Cancers cases n (%)		
Breast cancer	82 (18%)	NA
Colon cancer	133 (29%)	NA
Time to diagnosis (year)	6.10	3.36
Air pollution estimates ($\mu g/m^3$)		
NO_x	92.83 (34.46–131.00)	29.69 (24.25–42.58)
NO_2	50.00 (22.26–67.92)	19.98 (16.74–28.91)
PM_{10}	46.91 (38.24–53.23)	24.49 (23.96–26.16)
PM_{coarse}	16.75 (10.88–20.65)	8.05 (7.71–8.57)
$PM_{2.5}$	30.95 (26.74–32.73)	16.63 (16.40–17.12)
$PM_{2.5abs}$	3.38 (2.34–3.64)	1.19 (1.10–1.48)

Counts (percentages) and means ± standard deviation are reported for categorical and continuous variables, respectively. Air pollution estimates are reported as mean (5th–95th centile). NA: not applicable.

Table 2
Association between global methylation and air pollution estimates in EPIC-Italy and EPIC-NL.

	β	EPIC-Italy		β	EPIC-NL		CPMA
		SE	p-value		SE	p-value	p-value
NO_2							
Somatic probes	− 7.37e − 5	3.00e − 5	0.014	− 2.88e − 3	1.33e − 3	0.031	0.014
Relation to CpG island							
Island	− 4.04e − 5	2.80e − 5	0.140	− 1.63e − 3	1.23e − 3	0.185	0.359
Shelves	− 1.48e − 4	5.95e − 5	0.013	− 4.67e − 3	2.40e − 3	0.051	0.020
Shores	− 8.50e − 5	3.36e − 5	0.011	− 2.62e − 3	1.17e − 3	0.025	0.010
Other	− 1.28e − 4	5.11e − 5	0.012	− 3.68e − 3	1.92e − 3	0.055	0.020
Position on gene							
Promoter	− 6.23e − 5	2.67e − 5	0.019	− 2.46e − 3	1.16e − 3	0.033	0.019
Gene body	− 8.05e − 5	3.14e − 5	0.010	− 2.82e − 3	1.30e − 3	0.029	0.010
Intergenic	− 9.17e − 5	3.86e − 5	0.017	− 3.03e − 3	1.61e − 3	0.060	0.029
NO_x							
Somatic probes	− 1.37e − 4	6.64e − 5	0.039	− 9.17e − 4	5.76e − 4	0.112	0.089
Relation to CpG island							
Island	− 2.37e − 5	6.19e − 5	0.702	− 5.05e − 4	5.31e − 4	0.341	0.653
Shelves	− 3.24e − 4	1.31e − 4	0.014	− 2.05e − 3	1.03e − 3	0.045	0.019
Shores	− 1.59e − 4	7.44e − 5	0.032	− 1.15e − 3	5.03e − 4	0.022	0.021
Other	− 2.87e − 4	1.13e − 4	0.011	− 1.54e − 3	8.22e − 4	0.061	0.020
Position on gene							
Promoter	− 9.54e − 5	5.91e − 5	0.106	− 7.61e − 4	5.01e − 4	0.129	0.216
Gene body	− 1.57e − 4	6.93e − 5	0.024	− 9.89e − 4	5.60e − 4	0.077	0.045
Intergenic	− 1.91e − 4	8.51e − 5	0.025	− 1.08e − 3	6.95e − 4	0.119	0.067
PM_{10}							
Somatic probes	2.87e − 4	2.53e − 4	0.257	− 2.22e − 3	4.77e − 3	0.642	0.885
Relation to CpG island							
Island	3.98e − 4	2.17e − 4	0.067	-3.15e − 3	4.36e − 3	0.470	0.394
Shelves	3.12e − 4	4.94e − 4	0.529	− 3.04e − 3	8.52e − 3	0.721	0.358
Shores	4.12e − 4	2.69e − 4	0.125	− 1.63e − 3	4.19e − 3	0.698	0.773
Other	2.73e − 4	4.17e − 4	0.513	− 1.64e − 3	6.82e − 3	0.811	0.306
Position on gene							
Promoter	4.56e − 4	2.12e − 4	0.032	− 2.46e − 3	4.14e − 3	0.553	0.260
Gene body	2.47e − 4	2.62e − 4	0.346	− 2.04e − 3	4.63e − 3	0.659	0.685
Intergenic	3.14e − 4	3.17e − 4	0.322	− 1.42e − 3	5.74e − 3	0.805	0.601
PM_{coarse}							
Somatic probes	4.88e − 4	3.86e − 4	0.206	− 1.11e − 2	8.83e − 3	0.209	0.489
Relation to CpG island							
Island	5.37e − 4	3.32e − 4	0.105	− 1.35e − 2	8.07e − 3	0.095	0.172
Shelves	6.13e − 4	7.54e − 4	0.416	− 1.12e − 2	1.58e − 2	0.478	0.771
Shores	6.33e − 4	4.11e − 4	0.123	− 1.29e − 2	7.74e − 3	0.095	0.192
Other	5.04e − 4	6.36e − 4	0.428	− 1.15e − 2	1.26e − 2	0.364	0.918
Position on gene							
Promoter	6.85e − 4	3.24e − 4	0.035	− 1.31e − 2	7.64e − 3	0.087	0.067
Gene body	4.17e − 4	3.99e − 4	0.297	− 9.50e − 3	8.59e − 3	0.269	0.729
Intergenic	5.45e − 4	4.84e − 4	0.260	− 1.06e − 2	1.06e − 2	0.320	0.749
$PM_{2.5}$							
Somatic probes	− 4.10e − 4	5.87e − 4	0.486	− 5.61e − 3	5.86e − 3	0.338	0.887
Relation to CpG island							
Island	1.36e − 5	5.07e − 4	0.979	− 7.52e − 3	6.17e − 3	0.223	0.712
Shelves	− 6.25e − 4	1.15e − 3	0.586	− 1.11e − 2	1.04e − 2	0.285	0.878
Shores	− 4.24e − 5	6.27e − 4	0.946	− 6.58e − 3	5.46e − 3	0.228	0.720
Other	− 5.95e − 4	9.68e − 4	0.539	− 7.69e − 3	8.27e − 3	0.353	0.799
Position on gene							
Promoter	2.92e − 6	4.96e − 4	0.995	− 6.03e − 3	5.39e − 3	0.263	0.596
Gene body	− 4.08e − 4	6.07e − 4	0.502	− 5.45e − 3	5.65e − 3	0.335	0.874
Intergenic	− 3.90e − 4	7.36e − 4	0.597	− 5.72e − 3	7.06e − 3	0.418	0.627
$PM_{2.5abs}$							
Somatic probes	− 1.21e − 3	2.63e − 3	0.647	− 2.94e − 2	2.09e − 2	0.161	0.857
Relation to CpG island							
Island	1.94e − 3	2.27e − 3	0.392	-2.89e − 2	2.21e − 2	0.190	0.699
Shelves	− 3.51e − 3	5.14e − 3	0.494	− 6.04e − 2	3.69e − 2	0.102	0.543
Shores	2.91e − 5	2.80e − 3	0.998	− 3.30e − 2	1.95e − 2	0.090	0.784
Other	− 3.80e − 3	4.33e − 3	0.380	− 4.25e − 2	2.95e − 2	0.150	0.588
Position on gene							
Promoter	1.09e − 3	2.22e − 3	0.624	− 2.68e − 2	1.92e − 2	0.164	0.596
Gene body	− 1.39e − 3	2.72e − 3	0.609	− 3.06e − 2	2.01e − 2	0.128	0.874
Intergenic	− 2.32e − 3	3.29e − 3	0.481	− 3.21e − 2	2.52e − 2	0.202	0.627

The different measures for particulate matter (PM_{10}, $PM_{2.5}$, $PM_{2.5}$ absorbance and PM_{coarse}) are available in 297 subjects of the Turin centre while NO_2 and NO_x are available for 454 subjects. β = regression coefficient; SE = standard error for regression coefficient. β value (0–1 scale) represents the difference in methylation for every unit (μg/m^3) increase of air pollutant. The cross phenotype meta-analysis p-value is shown as CPMA p-value.

later in life were conducted, and provided similar results for NO_x exposure. Although the direction of the association remained the same, we observed weaker, non-significant associations for NO_2 (Supplemental Table A.5). Sensitivity analyses not correcting for cancer case status provided similar results, though with stronger associations for NO_2 (Table A.6.).

3.2. Epigenome-wide analyses

The meta-analyses of EPIC-NL and EPIC-Italy showed that no individual CpG sites were significantly associated to the air pollution components and indicated high heterogeneity between the site-specific methylation levels (Table A.7).

By stratifying site-specific methylation differences by study, the analyses of NO_2 (Table 3, Table A.8) revealed 12 significant associations in EPIC-Italy (Fig. A.3) and none in EPIC-NL (Fig. A.4). Of the 12 significant CpG sites in EPIC-Italy, the vast majority (11 CpG sites) showed hypomethylation. The epigenome-wide association study of NO_x identified 1 and 6 differentially methylated CpG in EPIC-Italy and EPIC-NL, respectively (Table 3, Table A.8). Of the differentially methylated sites found in EPIC-NL, 5 CpG sites were hypomethylated.

In EPIC-Italy, we identified only 1 differentially methylated CpG site in relation to $PM_{2.5}$ (Table 4, Table A.8) but no significant CpG sites with the other PM metrics. More associations were found in EPIC-NL, where we identified 4, 9, and 7 differentially methylated CpG loci, for PM_{10}, $PM_{2.5}$, and $PM_{2.5abs}$ respectively (Table 4; Table A.8). Among those, 2 CpG sites were found associated with the 3 exposures with the same direction: cg17629796 (hypomethylated) and cg03513315 (hypermethylated). Four additional CpG sites overlapped for $PM_{2.5}$ and $PM_{2.5abs}$: cg07084345 and cg00005622 (hypomethylated), cg04319606 and cg09568355 (hypermethylated). The use of back-extrapolated exposure estimates affected the associations only marginally, however CpG cg03513315 and cg17629796 were no longer significantly associated with PM_{10} in EPIC-NL (Tables A.9, A.10).

In total, we identified 32 unique CpG sites differentially methylated at an epigenome-wide level with at least one of the 5 air pollution components in at least 1 cohort. Of these, 21 CpG sites were annotated. FDR significant CpG sites with an adjusted p-value of maximum 0.2 are presented in Table A.11 and A.11 for respectively EPIC-Italy and EPIC-NL.

Concerning the cancer cases in the EPIC-Italy we performed 2 sensitivity analyses: i) when excluding the subjects that developed cancer later in life, all identified CpG sites remained nominally significant ($p < 0.05$), 12 being Bonferroni significant ($p < 0.05/14 = 3.75e-03$) except for cg22856765 ($p = 0.0575$) (Table A.13); ii) when not correcting for cancer case status, all associations remained significant ($p < 3.75e-06$) (Table A.14).

3.3. Investigation of the CpG-transcript pairs

To study the association between the 21 exposure-related and annotated CpG sites (Tables 3 and 4) and transcriptomic profiles, we assessed the association of the $21 \times 29{,}662 = 622{,}902$ CpG–transcript pairs in the EnviroGenoMarkers study. We identified 2135 significant CpG–transcript pairs, corresponding to 620 unique genes, that were associated to CpG sites related to NO_2, 598 genes for NO_x related CpGs, 41 for PM_{10}, 174 for $PM_{2.5}$ and 173 unique genes for $PM_{2.5abs}$. The number of positive and inverse associations is shown is Table A.15. Gene enrichment analyses using the default background (all genes expressed in *Homo sapiens*) and based on these sets of transcripts identified significantly enriched pathways (setting the EASE score below 0.01 and the minimal number of genes per group to 5) as summarised in Table 5. We found 5 enriched pathways for NO_2 and 9 for NO_x mainly relating to the immune system and its regulation. No significant enriched pathway was found for the particulate matter exposures. Changing the background list of genes to those found expressed in leukocytes from human peripheral blood (Palmer et al., 2006), we found no significant enriched pathway after correction for multiple testing. However, several nominally significant pathways were found: 26, 42, 5, and 4 for NO_2, NO_x, $PM_{2.5}$ and $PM_{2.5abs}$ respectively (Table A.16). Several of the pathways found for NO_2 and NO_x relate to the immune system and its regulation.

Table 3
CpG sites whose methylation is significantly associated with exposure to NO_2 or NO_X for EPIC-Italy and EPIC-NL at Bonferroni significance level ($p < 1.1e-7$).

	CpG	Gene	CHR	Localisation on Gene	Localisation on CGI	β	SE	p-Value
NO_2	EPIC-Italy							
	cg08120023	C1orf203	1:116947203	–	Body	−0.00321	0.00054	3.02e−9
	cg18201392	RNF2	1:185023741	–	5UTR	−0.00483	0.00090	8.02e−8
	cg04914283	EPHB2	1:23181832	–	Body	−0.00464	0.00086	5.85e−8
	cg18164357	C11orf67	11:77534497	South shelf	5UTR	−0.00892	0.00155	9.61e−9
	cg16205861	–	12:54146572	South shore	–	−0.00387	0.00072	6.57e−8
	cg03870188	MCF2L	13:113717830	North shelf	Body	−0.00429	0.00075	1.02e−8
	cg12790758	MEIS2	15:37369914	–	Body	−0.00439	0.00081	7.06e−8
	cg20939320	NCRNA00119	3:132563279	–	Body	−0.00616	0.00112	3.49e−8
	cg21156210	RG9MTD2	4:100485208	Island	TSS1500	0.01002	0.00185	6.03e−8
	cg13420207	CACNA2D1	7:81666278	–	Body	−0.00983	0.00181	5.56e−8
	cg22856765	THAP1	8:42693384	–	3UTR	−0.00818	0.00139	4.27e−9
	cg13918628	CD72	9:35610380	–	3UTR	−0.01169	0.00220	1.02e−7
NO_x	EPIC-Italy							
	cg05036212	–	2:119609691	North shore	–	0.00370	0.00067	3.60e−8
	EPIC-NL							
	cg08509991	COL17A1	10:105845720	–	TSS200	0.05303	0.00988	8.01e−8
	cg09487985	–	12:34484805	North Shelf	–	−0.02267	0.00413	4.01e−8
	cg18059012	PATL2	15:44969129	–	TSS200	−0.02653	0.00495	8.54e−8
	cg18351711	ODZ3	4:183243982	–	TSS1500	−0.03160	0.00572	3.28e−8
	cg12232118	TRIM15	6:30131458	–	5UTR	−0.02979	0.00488	1.03e−9
	cg09499965	–	7:151220664	South Shelf	–	−0.02527	0.00408	5.97e−10

CpGs sites that are significant after correction for multiple testing are shown (Bonferroni p-value 1.1e−7).
The number of subjects in the analyses in EPIC-Italy is 454.
β value (0–1 scale) represents the difference in methylation for every unit ($\mu g/m^3$) increase of air pollutant. TSS = transcription start site, UTR = untranslated region.
Column headers: Gene = UCSC annotated gene; CHR = chromosome and Chromosomal position; localisation on Gene = UCSC gene region feature category; Localisation on CGI = UCSC relation to CpG islands; β = regression coefficient; SE = standard error for regression coefficient.

Table 4
CpG sites whose methylation is significantly associated with exposure to PM_{10}, $PM_{2.5}$, and $PM_{2.5abs}$ for EPIC-Italy and EPIC-NL at Bonferroni significance level ($p < 1.1e-7$).

	CpG	Gene	CHR	Localisation on gene	Localisation on CGI	β	SE	p-Value
PM_{10}	EPIC-NL							
	cg17629796	–	13:30707265	–	–	−0.38612	0.07252	1.01e−7
	cg03025825	SMG6	17:1975245	–	Body	−0.39554	0.07483	1.25e−7
	cg03513315	PES1	22:30988383	Island	TSS1500	0.22314	0.04245	1.47e−7
	cg21232615	C9orf11	9:27292606	–	Body	−0.34525	0.06082	1.37e−8
$PM_{2.5}$	EPIC-NL							
	cg12575202	–	10:133331128	–	–	−0.46704	0.08005	5.40e−9
	cg08630381	–	13:100612277	Island	–	0.46095	0.07291	2.58e−10
	cg17629796	–	13:30707265	–	–	−0.56342	0.09407	2.11e−9
	cg07084345	–	15:61972967	–	–	−0.51254	0.07480	7.26e−12
	cg04319606	C2orf70	2:26785290	Island	TSS200	0.26098	0.06824	1.31e−7
	cg09568355	–	2:45228633	Island	–	0.26098	0.04958	1.41e−7
	cg03513315	PES1	22:30988383	Island	TSS1500	0.30682	0.05820	1.35e−7
	cg25489413	ZMIZ2	7:44794343	–	TSS1500	−0.36547	0.06762	6.48e−8
	cg00005622	–	8:145180403	North shore	–	−0.39818	0.06408	5.16e−10
	EPIC-Italy							
	cg23890774	–	19:36618841	Island	–	0.07774	0.01385	1.98e−8
$PM_{2.5abs}$	EPIC-NL							
	cg17629796	–	13:30707265	–	–	−2.18426	0.36365	1.90e−9
	cg16608593	MTA1	14:105912068	Island	Body	−1.07134	0.18704	1.02e−8
	cg07084345	–	15:61972967	–	–	−1.84644	0.28999	1.92e−10
	cg04319606	C2orf70	2:26785290	Island	TSS200	1.37346	0.25425	6.59e−8
	cg09568355	–	2:45228633	Island	–	0.99171	0.18423	7.32e−8
	cg03513315	PES1	22:30988383	Island	TSS1500	1.18511	0.21738	4.99e−8
	cg00005622	–	8:145180403	North shore	–	−1.32860	0.25132	1.25e−7

CpGs sites that are significant after correction for multiple testing are shown (Bonferroni p-value: 1.1e−7).
The number of subjects in the analyses in EPIC-Italy is 297.
β value (0–1 scale) represents the difference in methylation for every unit ($\mu g/m^3$) increase of air pollutant. TSS = transcription start site, UTR = untranslated region.
Column headers: Gene = UCSC annotated gene; CHR = chromosome and Chromosomal position; localisation on Gene = UCSC gene region feature category; Localisation on CGI = UCSC relation to CpG islands; β = regression coefficient; SE = standard error for regression coefficient.

4. Discussion

By affecting the epigenome and downstream processes, DNA methylation has the ability to biologically mediate the effect of environmental exposures. The main finding of this study was that lower DNA methylation levels in several functional regions on the genome including CpG island's shores and shelves and gene bodies were associated with exposure to higher ambient outdoor concentrations of NO_2 and NO_x.

In EPIC-Italy, global hypomethylation was found associated with higher annual averages of NO_2 and NO_x. The association with NO_x did not reach statistical significance in EPIC-NL, which could be attributed to a lower statistical power, to the overall lower exposure levels observed in the Netherlands, or to other population differences between the 2 cohorts. Future studies should further investigate whether the methylome is actually sensitive to low levels of exposure. Reduced global DNA methylation in blood quantified by high-performance liquid chromatography has been previously reported after mid-term exposure of 60 days to ambient NO_2 (De Prins et al., 2013), while a study with annual NO_x exposures did not find this association investigating Alu and LINE repeats (Chi et al., 2016). Since a hallmark of cancer is genome instability (Hanahan and Weinberg, 2011), lower DNA methylation resulting from exposure to NO_2 and NO_x could enable the development of cancers. This hypothesis is reinforced in a recent meta-analysis showing evidence for a relationship between NO_2, as a proxy for traffic-related air pollution, and lung cancer (Hamra et al., 2015). Indeed, global DNA hypomethylation has been suggested to play an important role in environment-related cancers (Cao, 2015; Koestler et al., 2012; van Veldhoven et al., 2015).

Our study investigated epigenome-wide DNA methylation and long-

Table 5
Significant pathways associated with the CpG-transcript pairs.

Database	Term	Count	p-Value	Fold enrichment	Bonferroni-corrected p-Value (5%)
NO_2					
GOTERM_BP_FAT	GO:0002694~regulation of leukocyte activation	23	9.50e−9	4.44	2.24e−5
GOTERM_BP_FAT	GO:0050865~regulation of cell activation	23	2.53e−8	4.21	5.99e−5
GOTERM_BP_FAT	GO:0051249~regulation of lymphocyte activation	20	1.64e−7	4.33	3.87e−4
INTERPRO	IPR001849:Pleckstrin homology	27	7.16e−7	3.09	6.42e−4
GOTERM_CC_FAT	GO:0044459~plasma membrane part	108	7.63e−7	1.57	2.58e−4
NO_x					
GOTERM_BP_FAT	GO:0046649~lymphocyte activation	28	1.52e−10	4.45	3.54e−7
GOTERM_BP_FAT	GO:0045321~leukocyte activation	30	6.40e−10	3.92	1.49e−6
GOTERM_BP_FAT	GO:0050865~regulation of cell activation	24	6.70e−9	4.33	1.56e−5
GOTERM_BP_FAT	GO:0051249~regulation of lymphocyte activation	22	7.71e−9	4.70	1.79e−5
GOTERM_BP_FAT	GO:0002694~regulation of leukocyte activation	23	1.23e−8	4.38	2.86e−5
GOTERM_BP_FAT	GO:0050863~regulation of T cell activation	19	2.54e−8	5.13	5.91e−5
GOTERM_BP_FAT	GO:0001775~cell activation	30	3.20e−8	3.30	7.45e−5
GOTERM_BP_FAT	GO:0002684~positive regulation of immune system process	26	1.38e−7	3.45	3.22e−4
SP_PIR_KEYWORDS	B-cell	7	1.03e−6	17.89	4.40e−4

term air pollution using the Illumina 450 K technology, which enabled us to study the overall methylation across the main functional genomic regions. In subjects exposed to higher levels of NO_2 and NO_x, we observed lower methylation levels in CpG island's shores and shelves and gene bodies, however not in CpG-islands. CpG island's shores are defined as the 2 kb of sequences flanking a CpG island; they have been reported to be more dynamic than the island itself (Bibikova et al., 2011; Irizarry et al., 2009) and are thought to have a functional role in the regulation of gene expression (Tao et al., 2014; Vanderkraats et al., 2013). Flanking the shores, shelves are defined as the neighbouring 2 kb genomic regions. In the assembly of shores and shelves there are various levels of stability corresponding to different levels of control on gene expression (Edgar et al., 2014). Although the method we used to determine global methylation averages the results of a high-resolution epigenome-wide array, we find a stable signal in shore and shelve regions. Describing the molecular pathways leading to global or regional hypo- or hypermethylation is speculative since the process has not been completely elucidated yet. However, a central role for reactive oxygen species has been put forward (Wu and Ni, 2015).

Cancer-associated DNA hypomethylation as well as hypermethylation occur, and these two types of epigenetic abnormalities usually seem to affect different DNA regions (Ehrlich, 2002). In this study, we also found a positive association (hypermethylation) between DNA methylation on promoter regions and PM_{10} and PM_{coarse} exposures. Other studies reported an inverse association between the global methylome (Alu repeats) and short time PM-exposure (Bellavia et al., 2013). We only found this association in the Italian cohort, possibly because of its higher exposure levels or larger size. At the gene-specific level, increased DNA methylation of tumour suppressor genes has been reported in carcinogenesis (Feinberg and Tycko, 2004) and particulate matter is classified by IARC as a group 1 carcinogen (Loomis et al., 2014). Further studies will be needed to confirm whether tumour suppressor genes drive the effects of higher methylation in relation to particulate matter exposure.

Based on the cohort-specific EpWASes, we identified several novel single CpG sites not previously associated to air pollution. For example, the CpG site cg03513315 located on the PES1 gene (Pescadillo Ribosomal Biogenesis Factor) was positively associated to PM_{10}, $PM_{2.5}$ and $PM_{2.5abs}$ in EPIC-NL. This gene seems to play a role in breast cancer proliferation and tumorigenesis. Also, the CpG cg18351711 located on the promoter region of the gene ODZ3 also known as Teneurin Transmembrane Protein 3 (TENM3) was inversely associated to NO_x in EPIC-NL and is involved in regulation of neuronal development. Using meta-analyses to combine the epigenome-wide findings in EPIC-Italy and EPIC-NL did not provide evidence of robust associations between epigenome-wide methylation and air pollution components. Since several CpG loci were associated with NO_2, NO_x, $PM_{2.5}$, PM_{10} and $PM_{2.5abs}$ exposures in EPIC-Italy or EPIC-NL, we attempted to further generalize these study-specific results from EPIC-Italy and EPIC-NL to other studies, including the German KORA F3, F4 studies (Holle et al., 2005; Wichmann et al., 2005) and the Spanish REGICOR study (Grau et al., 2007). Of the CpG sites we found differentially methylated in relation to exposure in either EPIC-Italy or EPIC-NL, none reached statistical significance either by a replication approach (taking into account multiple testing) in either KORA or REGICOR, or when performing a meta-analysis combining all the cohorts (Appendix p3, Table A.3). A translation of the CpGs associated with at least one of the exposures in at least one of the cohorts to the transcriptomic level resulted in significant pathways for NO_X and NO_2 exposure mainly in the immune system. One important mechanism of air pollution induced health effects is the induction of a persistent inflammatory state mediated by the immune system (Patel et al., 2013). A study in older men of the Normative Aging study used a targeted approach to show that exposure to particles is associated to the methylation state of inflammatory genes (Bind et al., 2015).

A study on short-term to mid-term exposure to $PM_{2.5}$ including KORA F3 and F4 meta-analysed data from 450 K human methylation Infinium arrays in 3 cohorts and did identify several significant CpG sites (Panni et al., 2016), as did a study in cord blood that meta-analysed 4 cohorts (Gruzieva et al., 2016). Possible explanations for the negative (not replicated) results of our epigenome-wide analyses of long-term exposure can involve (i) differences in the study populations; however, sensitivity analyses stratifying for sex or smoking status in our studies did not result in more replication (results not shown); (ii) differences in exposure levels, with Italy being the highest exposed cohort, the Netherlands having a narrow exposure range, and with only minor overlap between the 2 cohorts; (iii) air pollution being a mixture of pollutants which may have multiple and synergistic effects. In the case of PM, either the compounds absorbed to the particles can cause damage or the particles themselves can cause inflammatory reactions (Gong et al., 2014). Hence, the complexity of toxicological pathways involved could be related to exposure levels, and at an epigenome-wide level this might be an important reason why we observe study-specific results. Using single probe analyses to identify strong replicable signals for a complex exposure such as long-term air pollution should be addressed in larger populations.

Despite the large number of tests performed in epigenome-wide association studies, previous studies have been able to find and replicate strong exposure-triggered signals, for instance in relation to smoking (Guida et al., 2015; Zhang et al., 2014). Air pollution occurs at levels that are much lower than other exposures such as smoking and this can explain differences in the magnitude of health effects or the molecular changes underlying them.

Future studies can explore the epigenome-wide signals of exposure to long-term air pollution in the general population by either pooling the results of several cohorts with well-matched laboratory conditions and accurate air pollution exposure models or by performing larger meta-analyses including several studies. These additional analyses should also include non-linear models to account for possible complex dose-response relationship.

Our study has specific strengths and limitations. We used GIS (Geographical Information System) and land use regression modelling based on actual measurements in the designated study areas combined with predictor variables for nearby traffic intensity, population/household density and land use models, which have been generated and validated in the ESCAPE study (Beelen et al., 2013; Eeftens et al., 2012). Several studies have shown that the spatial distribution of air pollution is stable over 10-years periods (Eeftens et al., 2012; Gulliver et al., 2011) and can be applied successfully to estimate air pollution concentrations several years forwards or backwards in time. Although the absolute level of exposure to air pollution has generally decreased over time, measured and predicted NO_2 values from LUR models, from samples collected > 10 years apart, had good agreement and showed spatial stability (Cesaroni et al., 2012). Also in this study, we observed associations with global and regional methylation both when exposure was assessed at the historical home address (with data on air pollution dating from > 15 years later than the population sampling), and with back-extrapolated estimates. Since we have methylation information only at a single time point, we cannot describe the effect of exposure to air pollutants over time. For regional methylation, we did not correct for multiple testing as we only look at the results of these analyses marginally and never compare them or either conclude on the joint associations. Nevertheless, all of the associations would survive correction for 5 tests which correspond to the effective number of tests performed among the 7 actual tests.

Due to the nested case–control design of this study population cancer cases have been included in EPIC-Italy. These individuals were diagnosed with colon or breast cancer > 1 year after recruitment. However, these cancers are not known to be related to air pollution. In order to account for potential bias, we corrected for disease status in our models and performed sensitivity analyses by restricting the study population to healthy controls and by not adjusting for the cancer case

status. In these analyses the sign and estimates of the effect sizes were consistent, and most of the relevant associations were found nominally significant.

5. Conclusions

In our study, lower DNA methylation levels in CpG island's shores and shelves and gene bodies in adults were associated with exposure to higher ambient outdoor concentrations of NO_2 and NO_x. These differences in methylation may point to a mechanism underlying the association between air pollution and health outcomes. Our results may indicate a specific response to low dose exposure which should be further studied to test whether the methylome can be used as a low-dose marker. Exposure to air pollutants, however, was not associated with differential DNA methylation in single probes in the combined analyses of the two cohorts.

Funding

EPIC-Italy was financially supported by the Italian Association for Cancer Research (AIRC). Genome-wide DNA methylation profiling of EPIC-Italy samples was financially supported by the Human Genetics Foundation (HuGeF) and Compagnia di San Paolo. EPIC-Netherlands was financially supported by the Dutch Ministry of Public Health, Welfare, and Sports (VWS), by the Netherlands Cancer Registry, by LK Research Funds, by Dutch Prevention Funds, by the Netherlands Organisation for Health Research and Development (ZON), and by the World Cancer Research Fund (WCRF). Genome-wide DNA methylation profiling of EPIC-Netherlands samples was financially supported by internal Imperial College funds. The REGICOR study was supported by the Spanish Ministry of Economy and Innovation through the Carlos III Health Institute [Red HERACLES RD12/0042, PI12/00232, PI09/90506], European Funds for Development (ERDF-FEDER), and by the Catalan Research and Technology Innovation Interdepartmental Commission [SGR 1195]. Michelle Plusquin was supported by the People Program (Marie Curie Actions) of the European Union's Seventh Framework Program FP7/2007-2013/ under REA grant agreement n [628858]. Florence Guida was supported by the COLT foundation. Support for this work was also provided by the project EXPOSOMICS, grant agreement 308610-FP7 European Commission. Sergi Sayols-Baixeras was funded by a contract from Instituto de Salud Carlos III FEDER [IFI14/00007].

Acknowledgements

The Centre for Environment and Health is supported by the Medical Research Council and Public Health England (MR/L01341X/1). Marc Chadeau-Hyam acknowledges support from Cancer Research UK, Population Research Committee Project grant ('Mechanomics' project #22184 to MC-H).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.envint.2017.08.006.

References

Baccarelli, A., Wright, R.O., Bollati, V., Tarantini, L., Litonjua, A.A., Suh, H.H., Zanobetti, A., Sparrow, D., Vokonas, P.S., Schwartz, J., 2009. Rapid DNA methylation changes after exposure to traffic particles. Am. J. Respir. Crit. Care Med. 179, 572–578.

Beelen, R., Hoek, G., Vienneau, D., Eeftens, M., Dimakopoulou, K., Pedeli, X., Tsai, M.Y., Kunzli, N., Schikowski, T., Marcon, A., Eriksen, K.T., Raaschou-Nielsen, O., Stephanou, E., Patelarou, E., Lanki, T., Yli-Toumi, T., Declercq, C., Falq, G., Stempfelet, M., Birk, M., Cyrys, J., von Klot, S., Nador, G., Varro, M.J., Dedele, A., Grazuleviciene, R., Molter, A., Lindley, S., Madsen, C., Cesaroni, G., Ranzi, A., Badaloni, C., Hoffmann, B., Nonnemacher, M., Kraemer, U., Kuhlbusch, T., Cirach, M., de Nazelle, A., Nieuwenhuijsen, M., Bellander, T., Korek, M., Olsson, D., Stromgren, M., Dons, E., Jerrett, M., Fischer, P., Wang, M., Brunekreef, B., de Hoogh, K., 2013. Development of NO_2 and NOx land use regression models for estimating air pollution exposure in 36 study areas in Europe - the ESCAPE project. Atmos. Environ. 72, 10–23.

Bellavia, A., Urch, B., Speck, M., Brook, R.D., Scott, J.A., Albetti, B., Behbod, B., North, M., Valeri, L., Bertazzi, P.A., Silverman, F., Gold, D., Baccarelli, A.A., 2013. DNA hypomethylation, ambient particulate matter, and increased blood pressure: findings from controlled human exposure experiments. J. Am. Heart Assoc. 2, e000212.

Beulens, J.W., Monninkhof, E.M., Verschuren, W.M., van der Schouw, Y.T., Smit, J., Ocke, M.C., Jansen, E.H., van Dieren, S., Grobbee, D.E., Peeters, P.H., Bueno-de-Mesquita, H.B., 2010. Cohort profile: the EPIC-NL study. Int. J. Epidemiol. 39, 1170–1178.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, L., Schroth, G.P., Gunderson, K.L., Fan, J.B., Shen, R., 2011. High density DNA methylation array with single CpG site resolution. Genomics 98, 288–295.

Bind, M.A., Lepeule, J., Zanobetti, A., Gasparrini, A., Baccarelli, A., Coull, B.A., Tarantini, L., Vokonas, P.S., Koutrakis, P., Schwartz, J., 2014. Air pollution and gene-specific methylation in the Normative Aging Study: association, effect modification, and mediation analysis. Epigenetics 9, 448–458.

Bind, M.A., Coull, B.A., Peters, A., Baccarelli, A.A., Tarantini, L., Cantone, L., Vokonas, P.S., Koutrakis, P., Schwartz, J.D., 2015. Beyond the mean: quantile regression to explore the association of air pollution with gene-specific methylation in the normative aging study. Environ. Health Perspect. 123, 759–765.

Bingham, S., Riboli, E., 2004. Diet and cancer—the European prospective investigation into cancer and nutrition. Nat. Rev. Cancer 4, 206–215.

Brook, R.D., Rajagopalan, S., Pope 3rd, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith Jr., S.C., Whitsel, L., Kaufman, J.D., American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism, 2010. Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. Circulation 121, 2331–2378.

Campanella, G., Polidoro, S., Di Gaetano, C., Fiorito, G., Guarrera, S., Krogh, V., Palli, D., Panico, S., Sacerdote, C., Tumino, R., Elliott, P., Matullo, G., Chadeau-Hyam, M., Vineis, P., 2014. Epigenetic signatures of internal migration in Italy. Int. J. Epidemiol.

Cao, Y., 2015. Environmental pollution and DNA methylation: carcinogenesis, clinical significance, and practical applications. Front. Med. 9, 261–274.

Cesaroni, G., Porta, D., Badaloni, C., Stafoggia, M., Eeftens, M., Meliefste, K., Forastiere, F., 2012. Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. Environ. Health-Glob. 11.

Chadeau-Hyam, M., Vermeulen, R.C.H., Hebels, D.G.A.J., Castagne, R., Campanella, G., Portengen, L., Kelly, R.S., Bergdahl, I.A., Melin, B., Hallmans, G., Palli, D., Krogh, V., Tumino, R., Sacerdote, C., Panico, S., de Kok, T.M.C.M., Smith, M.T., Kleinjans, J.C.S., Vineis, P., Kyrtopoulos, S.A., Project, E., 2014. Prediagnostic transcriptomic markers of chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis. Ann. Oncol. 25, 1065–1072.

Chi, G.C., Liu, Y., MacDonald, J.W., Barr, R.G., Donohue, K.M., Hensley, M.D., Hou, L., McCall, C.E., Reynolds, L.M., Siscovick, D.S., Kaufman, J.D., 2016. Long-term outdoor air pollution and DNA methylation in circulating monocytes: results from the Multi-Ethnic Study of Atherosclerosis (MESA). Environ. Health 15, 119.

Cotsapas, C., Voight, B.F., Rossin, E., Lage, K., Neale, B.M., Wallace, C., Abecasis, G.R., Barrett, J.C., Behrens, T., Cho, J., De Jager, P.L., Elder, J.T., Graham, R.R., Gregersen, P., Klareskog, L., Siminovitch, K.A., van Heel, D.A., Wijmenga, C., Worthington, J., Todd, J.A., Hafler, D.A., Rich, S.S., Daly, M.J., Consortia, F.O.N.O., 2011. Pervasive sharing of genetic effects in autoimmune disease. PLoS Genet. 7, e1002254.

De Prins, S., Koppen, G., Jacobs, G., Dons, E., Van de Mieroop, E., Nelen, V., Fierens, F., Int Panis, L., De Boever, P., Cox, B., Nawrot, T.S., Schoeters, G., 2013. Influence of ambient air pollution on global DNA methylation in healthy adults: a seasonal follow-up. Environ. Int. 59, 418–424.

Demetriou, C.A., Raaschou-Nielsen, O., Loft, S., Moller, P., Vermeulen, R., Palli, D., Chadeau-Hyam, M., Xun, W.W., Vineis, P., 2012. Biomarkers of ambient air pollution and lung cancer: a systematic review. Occup. Environ. Med. 69, 619–627.

Edgar, R., Tan, P.P., Portales-Casamar, E., Pavlidis, P., 2014. Meta-analysis of human methylomes reveals stably methylated sequences surrounding CpG islands associated with high gene expression. Epigenetics Chromatin 7, 28.

Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C., Dedele, A., Dons, E., de Nazelle, A., Dimakopoulou, K., Eriksen, K., Falq, G., Fischer, P., Galassi, C., Grazuleviciene, R., Heinrich, J., Hoffmann, B., Jerrett, M., Keidel, D., Korek, M., Lanki, T., Lindley, S., Madsen, C., Molter, A., Nador, G., Nieuwenhuijsen, M., Nonnemacher, M., Pedeli, X., Raaschou-Nielsen, O., Patelarou, E., Quass, U., Ranzi, A., Schindler, C., Stempfelet, M., Stephanou, E., Sugiri, D., Tsai, M.Y., Yli-Tuomi, T., Varro, M.J., Vienneau, D., Klot, S., Wolf, K., Brunekreef, B., Hoek, G., 2012. Development of Land Use Regression models for PM(2.5), PM(2.5) absorbance, PM(10) and PM(coarse) in 20 European study areas; results of the ESCAPE project. Environ. Sci. Technol. 46, 11195–11205.

Ehrlich, M., 2002. DNA methylation in cancer: too much, but also too little. Oncogene 21, 5400–5413.

Feinberg, A.P., Tycko, B., 2004. The history of cancer epigenetics. Nat. Rev. Cancer 4, 143–153.

Ferrari, S., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. J. Appl. Stat. 31, 799–815.

Georgiadis, P., Hebels, D.G., Valavanis, I., Liampa, I., Bergdahl, I.A., Johansson, A., Palli, D., Chadeau-Hyam, M., Chatziioannou, A., Jennen, D.G., Krauskopf, J., Jetten, M.J., Kleinjans, J.C., Vineis, P., Kyrtopoulos, S.A., EnviroGenomarkers consortium, 2016. Omics for prediction of environmental health effects: Blood leukocyte-based cross-omic profiling reliably predicts diseases associated with tobacco smoking. Sci Rep 6,

20544.
Gong, J., Zhu, T., Kipen, H., Wang, G., Hu, M., Guo, Q., Ohman-Strickland, P., Lu, S.E., Wang, Y., Zhu, P., Rich, D.Q., Huang, W., Zhang, J., 2014. Comparisons of ultrafine and fine particles in their associations with biomarkers reflecting physiological pathways. Environ. Sci. Technol. 48, 5264–5273.
Grau, M., Subirana, I., Elosua, R., Solanas, P., Ramos, R., Masia, R., Cordon, F., Sala, J., Juvinya, D., Cerezo, C., Fito, M., Vila, J., Covas, M.I., Marrugat, J., 2007. Trends in cardiovascular risk factor prevalence (1995-2000-2005) in northeastern Spain. Eur. J. Cardiovasc. Prev. Rehabil. 14, 653–659.
Gruzieva, O., Xu, C.J., Breton, C.V., Annesi-Maesano, I., Anto, J.M., Auffray, C., Ballereau, S., Bellander, T., Bousquet, J., Bustamante, M., Charles, M.A., de Kluizenaar, Y., den Dekker, H.T., Duijts, L., Felix, J.F., Gehring, U., Guxens, M., Jaddoe, V.V., Jankipersadsing, S.A., Merid, S.K., Kere, J., Kumar, A., Lemonnier, N., Lepeule, J., Nystad, W., Page, C.M., Panasevich, S., Postma, D., Slama, R., Sunyer, J., Soderhall, C., Yao, J., London, S.J., Pershagen, G., Koppelman, G.H., Melen, E., 2016. Epigenome-wide meta-analysis of methylation in children related to prenatal NO_2 air pollution exposure. Environ. Health Perspect. http://dx.doi.org/10.1289/EHP36.
Guida, F., Sandanger, T.M., Castagne, R., Campanella, G., Polidoro, S., Palli, D., Krogh, V., Tumino, R., Sacerdote, C., Panico, S., Severi, G., Kyrtopoulos, S.A., Georgiadis, P., Vermeulen, R.C., Lund, E., Vineis, P., Chadeau-Hyam, M., 2015. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. Hum. Mol. Genet. 24, 2349–2359.
Gulliver, J., Morris, C., Lee, K., Vienneau, D., Briggs, D., Hansell, A., 2011. Land use regression modeling to estimate historic (1962–1991) concentrations of black smoke and sulfur dioxide for Great Britain. Environ. Sci. Technol. 45, 3526–3532.
Hallmans, G., Agren, A., Johansson, G., Johansson, A., Stegmayr, B., Jansson, J.H., Lindahl, B., Rolandsson, O., Soderberg, S., Nilsson, M., Johansson, I., Weinehall, L., 2003. Cardiovascular disease and diabetes in the northern Sweden health and disease study cohort - evaluation of risk factors and their interactions. Scand. J. Public Health Suppl. 61, 18–24.
Hamra, G.B., Laden, F., Cohen, A.J., Raaschou-Nielsen, O., Brauer, M., Loomis, D., 2015. Lung cancer and exposure to nitrogen dioxide and traffic: a systematic review and meta-analysis. Environ. Health Perspect. 123, 1107–1112.
Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: the next generation. Cell 144, 646–674.
Holle, R., Happich, M., Lowel, H., Wichmann, H.E., Group, M.K.S, 2005. KORA—a research platform for population based health research. Gesundheitswesen 67 (Suppl. 1), S19–25.
Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., Kelsey, K.T., 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 13.
Huang da, W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.
Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., Ji, H., Potash, J.B., Sabunciyan, S., Feinberg, A.P., 2009. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat. Genet. 41, 178–186.
Janssen, B.G., Godderis, L., Pieters, N., Poels, K., Kicinski, M., Cuypers, A., Fierens, F., Penders, J., Plusquin, M., Gyselaers, W., Nawrot, T.S., 2013. Placental DNA hypomethylation in association with particulate air pollution in early life. Part Fibre Toxicol. 10, 22.
Koestler, D.C., Marsit, C.J., Christensen, B.C., Accomando, W., Langevin, S.M., Houseman, E.A., Nelson, H.H., Karagas, M.R., Wiencke, J.K., Kelsey, K.T., 2012. Peripheral blood immune cell methylation profiles are associated with nonhematopoietic cancers. Cancer Epidemiol. Biomark. Prev. 21, 1293–1302.
Lepeule, J., Bind, M.A., Baccarelli, A.A., Koutrakis, P., Tarantini, L., Litonjua, A., Sparrow, D., Vokonas, P., Schwartz, J.D., 2014. Epigenetic influences on associations between air pollutants and lung function in elderly men: the normative aging study. Environ. Health Perspect. 122, 566–572.
Loomis, D., Huang, W., Chen, G., 2014. The International Agency for Research on Cancer (IARC) evaluation of the carcinogenicity of outdoor air pollution: focus on China. Chin. J. Cancer 33, 189–196.
Mostafavi, N., Vlaanderen, J., Chadeau-Hyam, M., Beelen, R., Modig, L., Palli, D., Bergdahl, I.A., Vineis, P., Hoek, G., Kyrtopoulos, S., Vermeulen, R., 2015. Inflammatory markers in relation to long-term air pollution. Environ. Int. 81, 1–7.
Nawrot, T.S., Vos, R., Jacobs, L., Verleden, S.E., Wauters, S., Mertens, V., Dooms, C., Hoet, P.H., Van Raemdonck, D.E., Faes, C., Dupont, L.J., Nemery, B., Verleden, G.M., Vanaudenaerde, B.M., 2011. The impact of traffic air pollution on bronchiolitis obliterans syndrome and mortality after lung transplantation. Thorax 66, 748–754.
Palli, D., Berrino, F., Vineis, P., Tumino, R., Panico, S., Masala, G., Saieva, C., Salvini, S., Ceroti, M., Pala, V., Sieri, S., Frasca, G., Giurdanella, M.C., Sacerdote, C., Fiorini, L., Celentano, E., Galasso, R., Decarli, A., Krogh, V., Italy, E., 2003. A molecular epidemiology project on diet and cancer: the EPIC-Italy prospective study. Design and baseline characteristics of participants. Tumori 89, 586–593.
Palmer, C., Diehn, M., Alizadeh, A.A., Brown, P.O., 2006. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. BMC Genomics 7, 115.
Panni, T., Mehta, A.J., Schwartz, J.D., Baccarelli, A.A., Just, A.C., Wolf, K., Wahl, S., Cyrys, J., Kunze, S., Strauch, K., Waldenberger, M., Peters, A., 2016. A genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. Environ. Health Perspect.
Patel, M.M., Chillrud, S.N., Deepti, K.C., Ross, J.M., Kinney, P.L., 2013. Traffic-related air pollutants and exhaled markers of airway inflammation and oxidative stress in New York City adolescents. Environ. Res. 121, 71–78.
Pettit, A.P., Brooks, A., Laumbach, R., Fiedler, N., Wang, Q., Strickland, P.O., Madura, K., Zhang, J., Kipen, H.M., 2012. Alteration of peripheral blood monocyte gene expression in humans following diesel exhaust inhalation. Inhal. Toxicol. 24, 172–181.
Price, M.E., Cotton, A.M., Lam, L.L., Farre, P., Emberly, E., Brown, C.J., Robinson, W.P., Kobor, M.S., 2013. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin 6, 4.
Raaschou-Nielsen, O., Andersen, Z.J., Beelen, R., Samoli, E., Stafoggia, M., Weinmayr, G., Hoffmann, B., Fischer, P., Nieuwenhuijsen, M.J., Brunekreef, B., Xun, W.W., Katsouyanni, K., Dimakopoulou, K., Sommar, J., Forsberg, B., Modig, L., Oudin, A., Oftedal, B., Schwarze, P.E., Nafstad, P., De Faire, U., Pedersen, N.L., Ostenson, C.G., Fratiglioni, L., Penell, J., Korek, M., Pershagen, G., Eriksen, K.T., Sorensen, M., Tjonneland, A., Ellermann, T., Eeftens, M., Peeters, P.H., Meliefste, K., Wang, M., Bueno-de-Mesquita, B., Key, T.J., de Hoogh, K., Concin, H., Nagel, G., Vilier, A., Grioni, S., Krogh, V., Tsai, M.Y., Ricceri, F., Sacerdote, C., Galassi, C., Migliore, E., Ranzi, A., Cesaroni, G., Badaloni, C., Forastiere, F., Tamayo, I., Amiano, P., Dorronsoro, M., Trichopoulou, A., Bamia, C., Vineis, P., Hoek, G., 2013. Air pollution and lung cancer incidence in 17 European cohorts: prospective analyses from the European Study of Cohorts for Air Pollution Effects (ESCAPE). Lancet Oncol. 14, 813–822.
Riboli, E., Kaaks, R., 1997. The EPIC Project: rationale and study design. European prospective investigation into cancer and nutrition. Int. J. Epidemiol. 26 (Suppl. 1), S6–14.
Sanchez-Guerra, M., Zheng, Y., Osorio-Yanez, C., Zhong, J., Chervona, Y., Wang, S., Chang, D., McCracken, J.P., Diaz, A., Bertazzi, P.A., Koutrakis, P., Kang, C.M., Zhang, X., Zhang, W., Byun, H.M., Schwartz, J., Hou, L., Baccarelli, A.A., 2015. Effects of particulate matter exposure on blood 5-hydroxymethylation: results from the Beijing truck driver air pollution study. Epigenetics 10, 633–642.
Somineni, H.K., Zhang, X., Biagini Myers, J.M., Kovacic, M.B., Ulm, A., Jurcak, N., Ryan, P.H., Khurana Hershey, G.K., Ji, H., 2015. Ten-eleven translocation 1 (TET1) methylation is associated with childhood asthma and traffic-related air pollution. J. Allergy Clin. Immunol.
Tao, M.H., Zhou, J., Rialdi, A.P., Martinez, R., Dabek, J., Scelo, G., Lissowska, J., Chen, J., Boffetta, P., 2014. Indoor air pollution from solid fuels and peripheral blood DNA methylation: findings from a population study in Warsaw. Poland Environ. Res. 134, 325–330.
Vanderkraats, N.D., Hiken, J.F., Decker, K.F., Edwards, J.R., 2013. Discovering high-resolution patterns of differential DNA methylation that correlate with gene expression changes. Nucleic Acids Res. 41, 6816–6827.
van Veldhoven, K., Polidoro, S., Baglietto, L., Severi, G., Sacerdote, C., Panico, S., Mattiello, A., Palli, D., Masala, G., Krogh, V., Agnoli, C., Tumino, R., Frasca, G., Flower, K., Curry, E., Orr, N., Tomczyk, K., Jones, M.E., Ashworth, A., Swerdlow, A., Chadeau-Hyam, M., Lund, E., Garcia-Closas, M., Sandanger, T.M., Flanagan, J.M., Vineis, P., 2015. Epigenome-wide association study reveals decreased average methylation levels years before breast cancer diagnosis. Clin. Epigenetics 7, 67.
Ward-Caviness, C.K., Nwanaji-Enwerem, J.C., Wolf, K., Wahl, S., Colicino, E., Trevisi, L., Kloog, I., Just, A.C., Vokonas, P., Cyrys, J., Gieger, C., Schwartz, J., Baccarelli, A.A., Schneider, A., Peters, A., 2016. Long-term exposure to air pollution is associated with biological aging. Oncotarget. http://dx.doi.org/10.18632/oncotarget.12903.
Wichmann, H.E., Gieger, C., Illig, T., MONICA/KORA Study Group, 2005. KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67 (Suppl. 1), S26–30.
Wu, Q., Ni, X., 2015. ROS-mediated DNA methylation pattern alterations in carcinogenesis. Curr. Drug Targets 16, 13–19.
You, J.S., Jones, P.A., 2012. Cancer genetics and epigenetics: two sides of the same coin? Cancer Cell 22, 9–20.
Zhang, Y., Yang, R., Burwinkel, B., Breitling, L.P., Brenner, H., 2014. F2RL3 methylation as a biomarker of current and lifetime smoking exposures. Environ. Health Perspect. 122, 131–137.