ENVIRONMENTAL EPIGENETICS

Environmental Epigenetics, 2025, **11(1)**, 1–15

DOI: https://doi.org/10.1093/eep/dvaf013
Advance Access Publication 6 May 2025
Research Article

Prenatal ambient air pollution associations with DNA methylation in asthma- and allergy-relevant genes: findings from ECHO

Meredith Palmore[1,‡], Emma E. Thompson[2,‡], Fang Fang[3], Theresa M. Bastain[4], Carrie Breton[4], Scott Collingwood[5], Frank D. Gilliland[4], Diane R. Gold[6,7], Rima Habre[8], Tina Hartert[9], Gurjit K. Khurana Hershey[10,11], Daniel J. Jackson[12,13], Rachel Miller[14], Patrick Ryan[11,15], Lyndsey Shorey-Kendrick[16], Eliot R. Spindel[16], Joseph Stanford[17], James Gern[12,13], Chris McKennan[18], Carole Ober[2,†], Christine Ladd-Acosta,[1,†,*] for the ECHO Cohort Consortium,[**]

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, United States
[2]Department of Human Genetics, University of Chicago, Chicago, IL 60637, United States
[3]Genomics and Translational Research Center, RTI International, Research Triangle Park, NC 27709, United States
[4]Department of Population and Public Health Sciences, University of Southern California, Los Angeles, CA 90032, United States
[5]Department of Pediatrics, University of Utah, Salt Lake City, UT 84108, United States
[6]Department of Environmental Health, Harvard T. Chan School of Public Health, Boston, MA 02115, United States
[7]Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, United States
[8]Department of Environmental Health and Spatial Sciences, University of Southern California, Los Angeles, CA 90032, United States
[9]Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN 37232, United States
[10]Division of Asthma Research, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, United States
[11]Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45229, United States
[12]Department of Pediatrics, University of Wisconsin-Madison, Madison, WI 53792, United States
[13]Department of Medicine, University of Wisconsin-Madison, Madison, WI 53792, United States
[14]Division of Clinical Immunology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY 10029, United States
[15]Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH 45229, United States
[16]Division of Neuroscience, Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR 97006, United States
[17]Department of Family and Preventative Medicine, University of Utah, Salt Lake City, UT 84108, United States
[18]Department of Statistics, University of Pittsburgh, Pittsburgh, PA 15260, United States

[‡]Denotes equal first-author contribution.
[†]Denotes equal senior-authorship contribution.
[**]See the Supplementary Appendix for the full list of ECHO Cohort Consortium collaborators.
*Corresponding author. Department of Epidemiology (W6509), Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, Maryland 21205, United States. E-mail: claddac1@jh.edu

Abstract

Prenatal exposure to air pollution is an important risk factor for child health outcomes, including asthma. Identification of DNA methylation changes associated with air pollutant exposure can provide new intervention targets to improve children's health. The aim of this study is to test the association between prenatal air pollutant exposure and DNA methylation in developmental and asthma-/allergy-relevant biospecimens (placenta, buccal, cord blood, nasal mucosa, and lavage). A subset of 2294 biospecimens collected from 1906 child participants enrolled in the Environmental Influences on Child Health Outcomes program with prenatal air pollutant and high-quality Illumina Asthma&Allergy DNA methylation array measures ($n = 37\,197$ probes) were included. Prenatal ozone, nitrogen dioxide, and fine particulate matter were derived using residential history during pregnancy and spatiotemporal models. For each pollutant, biospecimen type, and prenatal exposure window, we estimated the effects of air pollution on gene DNA methylation levels. We compared results across pollutants, biospecimen types, and trimesters and tested for critical months of exposure using distributed lag models. DNA methylation levels at 154 out of 4746 tested genes were associated with air pollution; over 95% were exposure window, pollutant, and biospecimen-type specific. The fewest gene associations were detected in trimester 2, relative to other exposure windows. A variety of trends in methylation patterns were observed in response to lagged monthly pollution levels. Child DNA methylation changes at

Received 16 December 2024; Revised 17 April 2025; Accepted 1 May 2025

specific respiratory- and immune-relevant genes are associated with prenatal air pollutant exposures. Future studies should examine the relationship between these pollution-sensitive genes and child health.

Keywords: DNA methylation; Environmental influences on Child Health Outcomes (ECHO); prenatal air pollution

Introduction

Air pollution exposure is associated with adverse health outcomes [1] and shorter-life expectancy [2]. In particular, prenatal and early-life exposure to air pollution is an important risk factor for the development of asthma and allergic disease in childhood [3–7]. Policy efforts to improve air quality are urgently needed to reduce or eliminate these exposures. Additionally, identifying biological mechanisms through which pollution exposure affects health outcomes could provide a complementary approach to reduce the health impacts of air pollution, especially if exposure occurs and/or policies are not fully effective.

Epigenetic modifications likely mediate the effects of pollution exposure on health outcomes [8], and DNA methylation is the most studied epigenetic modification to date [9]. The effects of DNA methylation changes on clinical or phenotypic outcomes are mediated through their effects on gene regulation [10]. DNA methylation patterns in placenta, cord blood, and infant buccal cells and blood spots have been associated with prenatal exposures to maternal smoking, folate and iron levels, glycaemic index, chemicals, and air pollution [9]. Moreover, DNA methylation patterns at birth have been associated with a broad spectrum of health outcomes, including late-life health [11], adult cardiovascular traits [12], neurodevelopmental outcomes [13], and asthma and allergic phenotypes [14, 15].

Several studies have examined prenatal exposure to particulate matter with an aerodynamic diameter <10 mm (PM_{10}) and fine particulate matter <2.5 mm in diameter ($PM_{2.5}$), nitrogen dioxide (NO_2), ozone (O_3), or traffic (distance to major roadways) and DNA methylation patterns in placenta or cord blood using the 450k DNA methylation array (Illumina, Inc.). One focused on NO_2 specifically [16], seven focused on one or two pollutants [17–23], one included three pollutants [24], and one considered distance to heavily trafficked roadways [25]. Overall, few individual CpGs were associated with pollutant exposures (0–21 significant CpG associations across the studies; median = 4 CpG associations) among ~400 000 CpGs tested in sample sizes ranging from 128 to 1506 (in placenta and cord blood, respectively; median sample size = 384). In contrast, three studies performed analyses of differentially methylated regions (DMRs): one study reported nine significant DMRs with O_3 or NO_2 in 133 placenta or 175 cord blood samples [21]; one reported 42 and 41 significant DMRs with O_3 or NO_2, respectively, in 429 cord blood samples [23]; and another reported 18 DMRs associated with NO_2 in 128 cord blood samples [16]. Among the DMRs, there was negligible overlap of genes annotated to the DMRs across pollutants or tissues.

The paucity of associations between specific pollutants and DNA methylation patterns in cord blood or placenta cells across these studies could be due to the relatively small sample sizes, the accuracy of the measurements, or their focus on few biospecimen sources [9]. Moreover, all previous studies used the 450k array, which is enriched for hypomethylated (<20% methylation) or hypermethylated (>80%) CpGs. CpGs with methylation levels in those ranges tend to show lower reproducibility of DNA methylation measurements compared to those in the intermediate range [9], and intermediate methylated CpGs, which are under-represented on these arrays, are a signature of regions characterized by gene regulation [26].

To address this limitation of the 450k and Illumina MethylationEPIC arrays, we designed a 50k DNA methylation array that includes CpGs with (i) prior evidence of association with asthma or allergic diseases or locations at a genome-wide association study locus for asthma or allergic disease, and (ii) overlap with transcription start sites, predicted transcription factor-binding sites, or predicted enhancers [27]. The final content of the array, referred to as the Asthma&Allergy array, included probes for 37 863 CpGs, of which 94% overlapped with a predicted enhancer and 90% overlapped with a predicted transcription factor-binding site. Moreover, compared to the EPIC and 450k arrays, this array is enriched for intermediate methylated CpGs and depleted for hypomethylated and hypermethylated CpGs. Additional details on the design, characteristics, and performance of the Asthma&Allergy array can be found in Morin *et al.* [27].

In this study, we address gaps in earlier studies by using an array with high-value content, by including standardized assessments of daily prenatal exposures to $PM_{2.5}$, NO_2, and O_3, and by studying newborn and delivery tissues (cord blood, placenta, and newborn buccal swabs), as well as biospecimens collected through young adulthood (nasal mucosal cells, nasal lavage cells, and buccal swabs). In addition, we used a novel gene-based test that reduces the multiple testing burden and harnesses information across all CpGs assigned to each of the 4746 genes. We used this unique data set to explore tissue-specific responses, at multiple life stages, to prenatal exposures to $PM_{2.5}$, NO_2, and O_3 averaged over pregnancy, by trimester, and by gestational month among participants enrolled in the National Institutes of Health (NIH)-funded Environmental Influences on Child Health Outcomes (ECHO) program [28].

Materials and methods

Overall study population

This study included a subset of participants enrolled in the ECHO program, a national study of over 60 000 children and their caregivers from across the USA and Puerto Rico [28]. The goal of the program is to investigate the role of the environment, from preconception through adolescence, on perinatal and child health outcomes and development. A subset of ECHO participants participated in this study. Eligibility for this analysis required participants to have residential history data from the pregnancy time window and available DNA methylation measurements from the Asthma&Allergy array.

The ECHO Cohort protocol was approved by the ECHO single Institutional Review Board (IRB) and/or the cohorts' local IRBs. Written informed consent or parental/guardian permission with child assent was obtained prior to collecting data.

DNA methylation measures

DNA from child participants was extracted at each ECHO cohort site and sent to the University of Chicago Genomics Facility for sodium bisulphite conversion and hybridization to the Asthma&Allergy array v1.0 (Illumina [27]). Raw IDAT files for

2990 samples (1482 nasal mucosa, 960 nasal lavage, 218 buccal, 245 cord blood, and 85 placenta) and 43 661 probes were shared with the ECHO Data Analysis Center, where the data underwent quality control checks to remove poorly performing samples and probes. A flowchart describing our overall approach to implementing sample and probe quality control filters is shown in Supplementary Fig. S1. Samples with discrepancies between methylation-predicted sex and reported sex were removed ($n = 7$). After sex checks, probes located on the X chromosome were excluded, which led to the omission of one gene from analysis (*PPP1R3F*). Samples with low overall array intensity (methylated and unmethylated intensities <10), >1% of probes with detection *P*-values > .05 ($n = 24$), and technical replicates ($n = 6$) were removed. We removed probes that were not present in the 'minfi' Illumina Asthma&Allergy array v1.0 manifest file ($n = 5120$), with low quality (detection *P*-value > 0.05) reported in >1% of samples, cross-reactive probes, and probes that overlapped a single-nucleotide polymorphism ($n = 685$) or did not unambiguously map to the genome ($n = 659$). Samples that passed quality control were background corrected using the 'preprocessIllumina()' function and normalized within a sample using the 'preprocessNoob()' function in 'minfi' (version 1.36.0) [29]. Quantile normalization was applied using the 'ENmix' package in R version 4.1.2, stratified by tissue type to account for batch differences between samples [30]. Methylation *M*-values, calculated as the log2 ratio of proportion methylated over proportion unmethylated, were used in the analysis.

Analytic dataset derivation

Our final analytic sample consisted of ECHO child participants with (i) high-quality Asthma&Allergy array [27] DNA methylation data from a buccal, placenta, cord blood, nasal mucosa, or nasal lavage biospecimen; (ii) daily air pollutant exposure measures available from all three trimesters of pregnancy for at least one of the three air pollutants we examined: O_3, NO_2, and $PM_{2.5}$; and (iii) age at biospecimen collection. Within a biospecimen type, if a child had repeated DNA methylation measures, only the data from the sample with the earliest age at collection were included. Lastly, we excluded 12 siblings to minimize potential genetic and environmental correlation between related individuals. Data management was conducted using the 'tidyverse' and 'data.table' packages in R version 4.2.0 [31, 32].

Air pollutant estimates

We estimated ambient concentrations of O_3, NO_2, and $PM_{2.5}$ at the home address of participants during pregnancy by applying previously developed and validated spatiotemporal models. Daily estimates for all pollutants were available for the contiguous USA at a 1 km × 1 km spatial resolution from 1 January 2000 through 31 December 2016 from a neural network model (i.e. the 'Schwartz model') that incorporates satellite-based measurements, chemical transport model output, meteorology, and other data [33]. Because some participants were born prior to 2000 (or prenatal exposures occurred prior to 2000), we obtained monthly $PM_{2.5}$ estimates at a 6 km × 6 km spatial resolution from a previously developed and validated model (i.e. the 'Yanosky model') for the years 1988–2007 [34] and performed a previously described calibration step [4] to address differences in spatial and temporal resolution between the two models. Briefly, among participants with $PM_{2.5}$ estimates available from both models ($n = 1247$), a linear regression model was fitted with Schwartz monthly averages as the dependent variable and Yanosky monthly averages as the independent variable. Among participants with both Yanosky and Schwartz model estimates available, there was high concordance: 113 out of 125 calibrated Yanosky measures (90.4%) fell within 5 $\mu g/m^3$ of the Schwartz estimate, and there was a mean difference of 0.10 $\mu g/m^3$ between measures. Parameter estimates from the linear model were applied to estimates from the Yanosky model when only these were available (i.e. dates prior to 2000). These calibrations to the Yanosky model estimates were applied within ECHO cohorts. The Yanosky model provides estimates for $PM_{2.5}$ only, and therefore, three cohorts (CAS, COAST, and EHAAS; see full names in Supplementary Table S1) only contributed $PM_{2.5}$ measures, as participants in these studies were born prior to 2000. O_3 and NO_2 concentrations are provided as parts per billion (ppb), and $PM_{2.5}$ was estimated as micrograms per cubic meter ($\mu g/m^3$).

Monthly pollution estimates obtained from averages of daily pollution estimates were further averaged into pregnancy trimester and total pregnancy estimates, which were used in subsequent analyses. If prenatal air pollution data were available for a participant as both calibrated Yanosky estimates and Schwartz estimates, the Schwartz model prenatal estimate was prioritized for analysis due to the higher resolution of air pollution exposure it provides and greater data availability across cohorts and pollutants.

Mapping CpG probes to genes

Our statistical analyses focused on aggregate gene-level DNA methylation measures rather than single CpG positions. We were interested in identifying methylation changes associated with air pollutants at the gene level given the enrichment of probes on the Asthma&Allergy array in gene bodies and in or near genes relevant to respiratory and immunologic health and their regulatory regions [27]. Moreover, previous studies have focused on individual CpGs with few discoveries, and there is a gain in power in our moderate-sized sample due to a smaller number of statistical tests when performing gene-level analyses. As detailed further, we performed two types of statistical tests, including a 'signed' and 'unsigned' association. The latter was used to accommodate known differences in methylation levels across a gene—e.g. promoters and gene bodies often have inverse relationships between methylation levels and gene expression.

Gene-level analyses required first assigning individual CpGs measured on the Asthma&Allergy array to genes. CpG probes were mapped to the nearest gene that showed expression in any biospecimen type per the Ensembl database [35] using the GRCh37 build. The 'GenomicRanges' package in Bioconductor was used to merge array probe and gene positional information [36].

Statistical analyses

All analyses were conducted using Bioconductor and R versions 3.15 and 4.2.0. Air pollutant exposure estimates were modelled as a continuous measure in all analyses. For each pollutant (O_3, NO_2, and $PM_{2.5}$), biospecimen type (placenta, buccal, cord blood, nasal lavage, and nasal mucosa), and exposure window (throughout pregnancy, first trimester, second trimester, and third trimester), we estimated the effects of prenatal air pollution on gene methylation using unsigned and signed gene-based tests. Both test the null hypothesis that methylation levels are independent of pollution levels at every CpG mapped to the gene, although the signed and unsigned tests differ in their power to reject certain alternatives. The unsigned test will reject the null if methylation levels at ≥1 CpGs are truly related to pollution levels (regardless of the direction of association at each pollution-related CpG), while the signed test will reject if the directions of association at pollution-related CpGs are consistent. For test statistics, we let $z \sim N_p(0, \Sigma)$

be a length P-vector of z-scores for each of the P CpGs that map to a gene, where z-scores were defined to be the pollution beta coefficient estimates divided by their standard errors; these are marginally standardized and normally distributed under the null hypothesis. The joint distribution of z-scores under the null is therefore Σ, where the diagonal elements of $z^{\top}\Sigma^{-1}z$ are all one and the off-diagonal elements represent the spatial correlation between neighbouring CpGs (we used a Matérn covariance function to model this correlation; see the Supplementary Statistical Methods). The unsigned and signed test statistics were then taken to be $1_p^{\top}\Sigma^{-1}z\left(1_p^{\top}\Sigma^{-1}1_p\right)^{-1/2}$ and z, which are chi-squared with p degrees of freedom and standard normal, respectively, under the null. Mathematical derivations of these test statistics are given in the Supplementary Statistical Methods.

To correct for genomic inflation, we adjusted P-values using the empirical null distribution of z-scores calculated by the R package 'locfdr' (Comprehensive R Archive Network; https://CRAN.R-project.org/package=locfdr). For signed test results, we computed z-scores using the effect estimates and their corresponding P-values. For unsigned test results, z-scores were taken to be the inverse of the standard normal cumulative distribution function applied to P-values. P-values were recalculated by applying the standard normal cumulative distribution function to z-scores after dividing them by the standard deviation of the fitted empirical null distribution. To control for multiple testing, gene associations were considered significant if they reached a Benjamini–Hochberg false discovery rate q-value threshold of 0.05.

All statistical models were adjusted for child's sex (male/female), prenatal smoking exposure (yes/no), continuous gestational age (cord and placenta), or age at sample collection (nasal and buccal), an ECHO recruitment site indicator, season of birth (spring/summer/winter/fall), and DNA methylation latent factors (estimated separately for each pollutant, biospecimen, and exposure window; modelled continuously). Variances were corrected for CpG correlations within array SentrixIDs. Missing values for covariates, which ranged from 0% to 8%, were imputed with the mean value, conditional on a cohort of origin. Latent factors were computed using the 'FALCO' package in R version 3.4 to adjust for unwanted variation that was not related to air pollution exposure averages [37]. The relationships between the methylation latent factors, batch, array, and self-reported characteristics of the participant, such as race, ethnicity, and sex, were also explored. Cell composition is also an important covariate to consider. Based on prior studies showing that cell composition variability is captured by latent factors [38], we used these factors as covariates in these analyses.

Distributed lag analysis

Genes that were significant in gene tests were tested for critical time windows for air pollution exposure associations using distributed lag models. For buccal specimens, only those collected within 1 month of birth were included in this analysis. Monthly prenatal air pollution exposure averages were lagged for each month prior to birth. Since CpG sites were allowed to vary in their direction of effects in the trimester average gene tests, we modelled the mean of the absolute value of methylation levels at each CpG site as the dependent variable in distributed lag models. Models were adjusted for the same covariates mentioned earlier and up to the first 10 latent factors from the gene average tests. We selected a maximum of 10 latent factors to best balance having a statistical model that would converge and capturing a sufficient proportion of variance. Distributed lag models were fit with polynomials varying in degree from 1 to 3 using the 'dlnm' package in R version 3.4 [39]. The Akaike information criterion (AIC) was used to determine the model with the best fit among the varying polynomial models. Effect estimates and P-values were reported from the best-fitting model for each tissue and pollutant. Distributed lag model results were only selected for significance if there was a significant positive effect, indicating an increase in methylation changes in either direction (i.e. hypo- or hypermethylation), for any lagged month prior to birth. A flowchart of the number of genes included in each analysis is shown in Supplementary Fig. S2.

Results

Analytic sample derivation

The subset of ECHO cohorts that met our inclusion criteria differed in their original recruitment and eligibility criteria (Supplementary Table S1). Four cohorts enrolled families who had a parent with asthma and/or allergy (COAST, CCAAPS, EHAAS, and URECA), one cohort enrolled pregnant women who smoked for a clinical trial of vitamin C supplementation during pregnancy (VCSIP), and three cohorts specifically recruited underserved or low-income populations (CCCEH/M and N, URECA, and NuMoM2b-Utah). The remaining cohorts recruited mothers and children from the general population across multiple states (HOPEUtah, MADREs, NCS-ARS, NCS-IVS, TTP-Utah, CAS, CCCEH, INSPIRE, and WHEALS). Our analytic sample includes children born between 1988 and 2017. Most cohorts recruited children born over the span of 1–3 years. The largest spread of birth years covered by a cohort site (NCS-IVS) was 14 years (2001–2015), with an interquartile range of 4 years. The year of specimen collection was also determined by cohort site membership, with the maximum range of collection dates spanning 5 years. Most cohort sites had all three air pollutant measures; three cohorts (CAS, COAST, and EHAAS) had only $PM_{2.5}$ available for analyses.

As shown in Supplementary Fig. S3, of the 2990 biospecimens that underwent DNA methylation measurements on the Asthma&Allergy array, a total of 2953 samples and 37 197 CpG probes passed our methylation quality control filters. Participants without at least one air pollutant measured across all three pregnancy trimesters ($n = 518$), lacking age at time of sample collection ($n = 32$), and repeated samples or related individuals ($n = 109$) were removed. Our final analytic dataset contained 2294 biospecimens from 1906 participants (Supplementary Fig. S3). It consisted of 77 cord blood and 72 placenta birth specimens, 88 infant (0–1 year of age) buccal specimens, 1258 nasal mucosal cell specimens collected from ages 1 to 21 (median of 6.65) years, and 799 nasal lavage specimens collected from ages 11 to 30 (median of 14.7) years (Table 1). Multiple biospecimen types were contributed by 359 participants, as follows: 278 with nasal lavage and nasal mucosa (collected at different ages); 50 with cord blood, placenta, and buccal; 7 with cord blood and buccal; 14 with cord blood and placenta; and 10 participants with buccal and placenta.

Analytic sample characteristics

As shown in Table 1, gestational age at birth ranged from 30 to 43 weeks. Males and females were nearly equally represented (50.5% and 49.5%, respectively), with slight variations by specimen type. The most frequent parent-reported child races and ethnicities were White (49.8%), Black (34.6%), Hispanic (14.3%), and multiple (7.2%), with all other categories <7%. There were notable exceptions in nasal mucosa and lavage, with 30.9% and 49.4%, respectively, parent-reported Black race, and in cord blood and buccal specimens, where 22.1% and 18.2%, respectively, reported

Table 1. Descriptive characteristics of ECHO participants included in our analytic sample, by biospecimen type

	Placenta	Cord blood	Buccal	Nasal mucosa	Nasal lavage	Overall
	(*N* = 72)	(*N* = 77)	(*N* = 88)	(*N* = 1258)	(*N* = 799)	(*N* = 2294)
Age at specimen collection						
Days, median (min, max)	0 (0, 0)	0 (0, 0)	3.00 (0, 439)	2430 (504, 7760)	5370 (4040, 11 200)	4020 (0, 11 200)
Years, median (min, max)	—	—	0.01 (0, 1.2)	6.65 (1.4, 21.3)	14.7 (11.1, 30.7)	11.0 (0, 30.7)
Estimated gestational age at birth, weeks						
Median (min, max)	39.0 (36.0, 41.0)	39.0 (34.0, 41.0)	38.0 (30.0, 41.0)	39.0 (31.0, 43.0)	39.0 (30.0, 43.0)	39.0 (30.0, 43.0)
Missing	0 (0%)	0 (0%)	0 (0%)	8 (0.6%)	180 (22.0%)	188 (7.9%)
Child sex						
Female	34 (47.2%)	38 (49.4%)	43 (48.9%)	612 (48.6%)	408 (51.1%)	1135 (49.5%)
Male	38 (52.8%)	39 (50.6%)	45 (51.1%)	646 (51.4%)	391 (48.9%)	1159 (50.5%)
Parent-reported race of the child participant						
American Indian or Alaska Native	0 (0%)	0 (0%)	0 (0%)	<5 (<1%)	<5 (<1%)	<5 (<1%)
Asian	0 (0%)	0 (0%)	0 (0%)	<10 (<1%)	10 (1.3%)	<20 (<1%)
Black	<5 (<5%)	<5 (<2%)	6 (6.8%)	389 (30.9%)	395 (49.4%)	794 (34.6%)
Multiple identities	14 (19.4%)	14 (18.2%)	16 (18.2%)	78 (6.2%)	43 (5.4%)	165 (7.2%)
Native Hawaiian or other Pacific Islander	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Other[a]	0 (0%)	<5 (<3%)	0 (0%)	10 (0.8%)	6 (0.8%)	18 (0.8%)
White	55 (76.4%)	59 (76.6%)	66 (75.0%)	696 (55.3%)	266 (33.3%)	1142 (49.8%)
Missing	0 (0%)	<5 (<2%)	0 (0%)	76 (6.0%)	78 (9.8%)	155 (6.8%)
Ethnicity of the child participant						
Hispanic	11 (15.3%)	17 (22.1%)	16 (18.2%)	165 (13.1%)	118 (14.8%)	327 (14.3%)
Non-Hispanic	61 (84.7%)	59 (76.6%)	72 (81.8%)	1092 (86.8%)	673 (84.2%)	1957 (85.3%)
Missing	0 (0%)	<5 (<2%)	0 (0%)	<5 (<1%)	8 (1.0%)	10 (0.4%)
Ever used nicotine during pregnancy						
No	<17 (<23%)	<23 (<29%)	<20 (<20%)	1081 (85.9%)	663 (83.0%)	1798 (78.4%)
Yes	53 (73.6%)	51 (66.2%)	70 (79.5%)	146 (11.6%)	79 (9.9%)	399 (17.4%)
Missing	<5 (<5%)	<5 (<6%)	<5 (<3%)	31 (2.5%)	57 (7.1%)	97 (4.2%)

(continued)

Table 1. (Continued)

	Placenta (*N* = 72)	Cord blood (*N* = 77)	Buccal (*N* = 88)	Nasal mucosa (*N* = 1258)	Nasal lavage (*N* = 799)	Overall (*N* = 2294)
Education of biological mother						
Less than high school	6 (8.3%)	10 (13.0%)	9 (10.2%)	86 (6.8%)	81 (10.0%)	192 (8.4%)
High school degree, GED, or equivalent	19 (26.4%)	20 (26.0%)	30 (34.1%)	206 (16.4%)	119 (14.9%)	394 (17.2%)
Some college, no degree; Associate's degree (AA and AS); and trade school	25 (34.7%)	27 (35.1%)	29 (33.0%)	328 (26.1%)	231 (28.9%)	640 (27.9%)
Bachelor's degree (BA, BS)	14 (19.4%)	12 (15.6%)	12 (13.6%)	329 (26.2%)	167 (20.9%)	534 (23.3%)
Master's degree, professional, or Doctor's degree	8 (11.1%)	<10 (<10%)	8 (9.1%)	168 (13.4%)	107 (13.4%)	298 (13.0%)
Missing	0 (0%)	<5 (<2%)	0 (0%)	141 (11.2%)	94 (11.8%)	236 (10.3%)
Ever reported diagnosis of eczema						
No	47 (65.3%)	46 (59.7%)	54 (61.4%)	293 (23.3%)	287 (35.9%)	727 (31.7%)
Yes	25 (34.7%)	24 (31.2%)	34 (38.6%)	546 (43.4%)	383 (47.9%)	1012 (44.1%)
Missing	0 (0%)	7 (9.1%)	0 (0%)	419 (33.3%)	129 (16.1%)	555 (24.2%)
Ever reported diagnosis or symptoms of asthma						
No	67 (93.1%)	63 (81.8%)	77 (87.5%)	874 (69.5%)	452 (56.6%)	1533 (66.8%)
Yes	<5 (<6%)	5 (6.5%)	<11 (<12%)	361 (28.7%)	347 (43.4%)	727 (31.7%)
Missing	<5 (<2%)	9 (11.7%)	<5 (<2%)	23 (1.8%)	0 (0%)	34 (1.5%)
Ever reported diagnosis of allergic rhinitis						
No	19 (26.4%)	19 (24.7%)	17 (19.3%)	42 (3.3%)	78 (9.8%)	175 (7.6%)
Yes	12 (16.7%)	13 (16.9%)	19 (21.6%)	368 (29.3%)	245 (30.7%)	657 (28.6%)
Missing	41 (56.9%)	45 (58.4%)	52 (59.1%)	848 (67.4%)	476 (59.6%)	1462 (63.7%)
Season of birth						
Winter	22 (30.6%)	24 (31.2%)	23 (26.1%)	171 (13.6%)	195 (24.4%)	435 (19.0%)
Spring	17 (23.6%)	15 (19.5%)	23 (26.1%)	250 (19.9%)	199 (24.9%)	504 (22.0%)
Summer	15 (20.8%)	19 (24.7%)	25 (28.4%)	463 (36.8%)	194 (24.3%)	716 (31.2%)
Fall	18 (25.0%)	19 (24.7%)	17 (19.3%)	374 (29.7%)	211 (26.4%)	639 (27.9%)

Min = minimum; max = maximum; GED = General Educational Development Test; AA = associate of arts; AS = associate of science; BA = bachelor of arts; BS = bachelor of science.
Data shown are *n* (%) except where indicated.
[a]The participants reported something other than one of the categories available for selection or some combination of categories.

Table 2. Air pollution distributions and missingness by biospecimen type

	Placenta	Cord blood	Buccal	Nasal mucosa	Nasal lavage	Overall
	(*N* = 72)	(*N* = 77)	(*N* = 88)	(*N* = 1258)	(*N* = 799)	(*N* = 2294)
Pregnancy O_3 exposure (ppb)						
Median (Min, Max)	42.3 (30.2, 50.6)	41.6 (30.2, 51.7)	41.5 (30.2, 50.4)	38.0 (26.1, 62.9)	37.2 (24.9, 48.3)	38.0 (24.9, 62.9)
Missing	0 (0%)	0 (0%)	0 (0%)	38 (3.0%)	111 (13.9%)	149 (6.5%)
Pregnancy NO_2 exposure (ppb)						
Median (Min, Max)	18.1 (9.64, 28.6)	18.2 (8.58, 29.6)	18.0 (8.71, 30.9)	25.9 (2.33, 52.5)	33.3 (11.0, 51.7)	29.3 (2.33, 52.5)
Missing	0 (0%)	0 (0%)	0 (0%)	38 (3.0%)	111 (13.9%)	149 (6.5%)
Pregnancy $PM_{2.5}$ exposure ($\mu g/m^3$)						
Median (Min, Max)	7.99 (3.01, 13.2)	8.01 (3.01, 14.2)	8.40 (3.32, 13.2)	9.81 (3.31, 18.9)	13.8 (9.01, 21.3)	11.5 (3.01, 21.3)
Missing	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Column denominators are the total *N* from Table 1 per biospecimen type. Missing is computed across exposures within a biospecimen type.

Hispanic ethnicity. Prenatal exposure to smoking was more prevalent in cord blood (66.2%), placenta (73.6%), and buccal specimens (79.5%) compared to nasal specimens. This is expected, since these three biospecimen types originated from a study that recruited mothers who smoked (VCSIP). Most participant mothers had at least a high school degree (81.3%). Lastly, 19.0% of samples were collected from children born in winter (defined as 20 December through 19 March), 22.0% in the spring (20 March through 19 June), 31.2% in the summer (20 June through 20 September), and 27.9% in the fall (21 September through 19 December).

Prenatal air pollution exposure distributions

As shown in Table 2, the median pregnancy O_3 exposure level was 38.0 ppb, ranging from 24.9 to 62.9 ppb. There were higher median levels of pregnancy O_3 exposure for placenta, cord blood, and buccal than those for nasal mucosa (38.0 ppb; range: 26.1–62.9 ppb) and nasal lavage specimens. The median pregnancy NO_2 exposure was 29.3 ppb across biospecimens, ranging from 2.33 to 52.5 ppb. The median NO_2 exposure levels for placenta, cord blood, and buccal specimens were lower than those for nasal mucosa and lavage specimens. Nasal biospecimens were missing both O_3 and NO_2 data at a rate of 3.0% for mucosa and 13.9% for lavage. Median pregnancy $PM_{2.5}$ levels were slightly lower for placenta, cord blood, and buccal than for nasal mucosa and lavage specimens.

Gene methylation associations with air pollutants

Using the signed association tests, we identified a total of 25 unique genes that were associated with at least one of the pollutants and exposure time periods that we examined in any of the biospecimen types. All but five of these genes (*RP11-108P20.2, CTD-2384B11.2, RP5-1028K7.2, AL662890.3*, and *CITF22-92A6.1*) were also identified by the unsigned test. Thus, we focused subsequent analyses, and results reported below, on the unsigned test results. All 25 genes that reached significance for the signed gene association tests, their effect estimates, and *P*-values are listed in Supplementary Table S2. The effect sizes for each of these 25 gene associations ranged from a 0.19% to 9.98% change in DNA methylation per 1 ppb or $\mu g/m^3$ of pollution (or a 0.95% to 49.9% change in DNA methylation per 5-unit increase in pollution exposure). For example, as shown in Supplementary Table S2, *AHRR* showed a 1.90% increase in methylation for each 1 $\mu g/m^3$ increase in $PM_{2.5}$ exposure among nasal lavage biospecimens. Methylation decreases were also observed—e.g. *GLRX* showed a percent methylation decrease of 6.58 for each 1 ppb increase in pregnancy O_3 exposure in placenta biospecimens (Supplementary Table S2).

The unsigned test revealed 154 unique genes (a false discovery rate of <0.05) with methylation associations with at least one air pollutant, exposure window, and biospecimen type. Our quantile–quantile plots and lambda values show reasonable control of potential genomic inflation, with λ_{GC} ranging from 0.98 to 1.08, depending on the specific pollutant and biospecimen examined (Supplementary Figs S4–S7). All significant unsigned gene associations, their beta coefficient estimates, and *P*-values for each biospecimen type and each exposure window are shown in Supplementary Table S3.

O_3-associated gene methylation

Among the set of 154 genes associated with any exposure using the unsigned association test, 77 (50%) were associated with O_3 levels (Fig. 1a, Supplementary Fig. S8), including 8 that were detected in two exposure windows. The greatest number of genes was discovered for third-trimester O_3 exposure ($n = 31$), followed by the first trimester ($n = 26$). We also observed differences in the number of genes identified by biospecimen type: 29 for placenta, 19 for cord blood, 17 for nasal mucosa, 8 for nasal lavage, and 6 for buccal. Two genes (*HIST1H3PS1* and *KIAA1875*) were associated with O_3 in more than one biospecimen type (Supplementary Table S4). Each exposure window showed a different pattern in the number of gene methylation associations detected within a biospecimen type; i.e., there were no biospecimens with consistently higher numbers of genes detected across all exposure windows. For example, the biospecimen with the largest number of gene associations in the first trimester was cord blood ($n = 11$ genes), the second trimester was nasal mucosal ($n = 7$), and the third trimester was placenta ($n = 18$).

NO_2–gene methylation associations

Among the 154 genes associated with any exposure using the unsigned association test, 36 (23%) were associated with NO_2 exposure (Fig. 1b, Supplementary Fig. S8), including 12 genes detected in more than one exposure window (Fig. 1b). The greatest number of gene methylation associations with NO_2 were observed in the third trimester ($n = 22$). The number of associations differed by biospecimen type, with most observations in cord blood ($n = 14$), followed by buccal ($n = 9$), placenta ($n = 5$), nasal lavage ($n = 5$), and nasal mucosa ($n = 3$). Two genes (*ADCY9* and *SERPINB6*) were associated with NO_2 in more than one biospecimen type (Supplementary Table S5). Similar to O_3, we observed differences in the number of associations by biospecimen type across exposure windows (Fig. 1b).

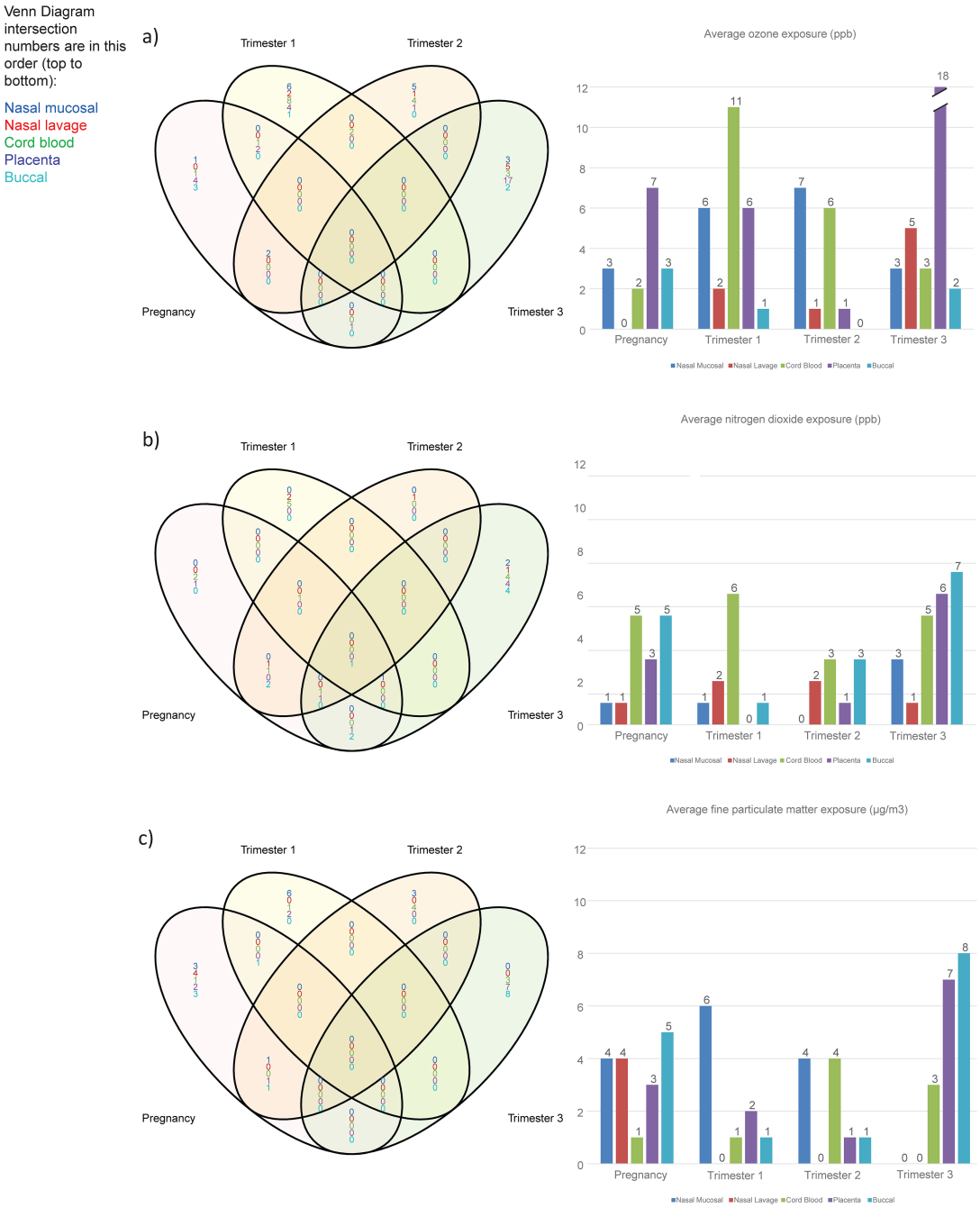

Figure 1. Total number of genes in each exposure window showing significant methylation changes associated with prenatal exposure to (a) O_3, (b) NO_2, and (c) $PM_{2.5}$ levels. The right side of each panel shows the number of gene associations with the pollutant in bar plots. Genes in the bar plots are not mutually exclusive across time periods. Patterns in the bar charts denote each biospecimen source. On the left of each panel is a Venn diagram comparing overlap across time periods within a biospecimen type.

$PM_{2.5}$-associated gene methylation

Among the set of 154 genes associated with any exposure by the unsigned association test, 49 (32%) were associated with $PM_{2.5}$ levels (Fig. 1c, Supplementary Fig. S8), including four genes detected in two exposure windows (Fig. 1c). The third trimester showed the greatest number of associations ($n = 18$) followed closely by pregnancy ($n = 17$). Differences in the number of associations detected differed by biospecimen type, with 13 in buccal, 13 in nasal mucosa, 12 in placenta, 9 in cord blood, and 4 in nasal lavage. Two genes (*FBXO17* and *RTEL1*) were associated with $PM_{2.5}$ in more than one biospecimen type (Supplementary Table S6).

Consistent with O_3 and NO_2, there were no biospecimens with consistently higher numbers of genes detected across all exposure windows (Fig. 1c).

Comparison of associations across biospecimen types

We were interested in determining which, if any, genes showed pollutant-specific methylation changes across different biospecimen types. For O_3, the results of the cross-tissue comparison revealed that only 2/77 (2.6%) genes were associated with O_3 exposure in more than one biospecimen type. *HIST1H3PS1* was

associated with first-trimester O_3 exposure levels in both buccal and cord blood specimens, and *KIAA1875* was associated with buccal and cord blood for the third- and first-trimester exposure windows, respectively (Supplementary Table S4).

Of the 36 unique genes identified for NO_2, 2 (5.5%) were associated in more than one biospecimen type (Supplementary Table S5). One of these genes, *ADCY9*, was associated in the third-trimester exposure window for both placenta and cord blood. *SERPINB6* was associated in pregnancy and all three trimesters in placenta and cord blood samples. *SERPINB6* is particularly notable due to its consistent detection across birth and nasal biospecimen types and across exposure time windows.

Of the 49 genes associated with $PM_{2.5}$, 2 (4.1%) were detected in multiple biospecimen types (Supplementary Table S6). *FBXO17* was associated in cord blood and buccal consistently in the first and second trimesters, and also in pregnancy for buccal specimens. *RTEL1* was associated in the third trimester for both cord blood and placenta.

Comparison of Associations across air pollutants

Overall, 146 (94%) of the genes we identified were specific to one pollutant: 70 genes were associated with O_3, 31 with NO_2, and 45 with $PM_{2.5}$ (Supplementary Fig. S8). One gene was associated with both NO_2 and $PM_{2.5}$, four genes with both O_3 and NO_2, and three genes with both O_3 and $PM_{2.5}$ air pollutant exposures. No genes were associated with all three air pollutants. Genes that were associated with more than one pollutant are reported in Supplementary Table S7.

Placenta association findings

As shown in Fig. 1, a total of 55 genes were identified in placenta: 32 genes for O_3, 10 for NO_2, and 13 genes for $PM_{2.5}$ exposures. Of those, only six genes (10.9%) were identified in more than one exposure time interval. This included two genes identified in pregnancy and first-trimester exposure windows for O_3 (*ENO1* and *HPN*); one gene identified in pregnancy and third-trimester exposure windows for O_3 (*GLRX*); one gene in pregnancy, second, and third trimesters for NO_2 (*SERPINB6*); one gene (*C21orf119*) identified in both pregnancy and third trimester for NO_2; and one gene (*HIST1H1A*) identified in pregnancy and second trimester for $PM_{2.5}$.

Cord blood association findings

For cord blood, we identified a total of 50 gene associations, including 22 for O_3, 19 for NO_2, and 9 for $PM_{2.5}$. Of these, six (12%) were associated in more than one exposure time interval. For O_3 exposure, we detected associations with one gene in both pregnancy and first trimester (*TMEM116*), and two genes (*JARID2* and *KIAA0513*) overlapped pregnancy and second trimester (Fig. 1a, left panel). NO_2 exposure was associated with one gene (*MIR4309* and *EGLN3*) in pregnancy and second trimester; one gene for pregnancy, first, and second trimesters (*PLEKHF1*); and one gene across pregnancy, second, and third trimesters (*SERPINB6*) (Fig. 1b, left panel). No genes were associated with $PM_{2.5}$ in more than a single time window in cord blood specimens (Fig. 1c, left panel).

Buccal association findings

Methylation levels at a total of 37 genes were associated with O_3 ($n = 6$), NO_2 ($n = 16$), and $PM_{2.5}$ ($n = 15$) prenatal exposure levels. Of these, seven genes (18.9%) were associated in more than one exposure window. No genes were associated with O_3 in more than a single time window (Fig. 1a, left panel). Associations with NO_2 exposure levels included one gene (*RP11-59H1.1*) in every exposure window—i.e. pregnancy and all trimesters. Additionally, for NO_2, two genes were associated in pregnancy and second-trimester exposure windows (*RPSAP2* and *ARID5B*), and two genes (*PHF2* and *ARFRP1*) were associated in both pregnancy and third-trimester exposure windows (Fig. 1b, left panel). One gene was associated with $PM_{2.5}$ in pregnancy and first-trimester exposure windows (GRK4), and another (*FBXO17*) was associated in both pregnancy and second-trimester exposure windows in buccal specimens (Fig. 1c, left panel).

Nasal lavage association findings

A total of 18 genes showed methylation associations with prenatal air pollution, including eight associations for O_3, six for NO_2, and four for $PM_{2.5}$. Figure 1a and c shows that none of the genes associated with O_3 or $PM_{2.5}$ exposure levels were detected in multiple exposure windows. One NO_2-associated gene (*MXRA8*) was identified in both the second-trimester and pregnancy windows (Fig. 1b, left panel).

Nasal mucosal findings

In nasal mucosal cells, 31 total genes were associated with either O_3 ($n = 19$), NO_2 ($n = 5$), or $PM_{2.5}$ ($n = 14$). Of those, associations were observed across time intervals for four genes (13%). Two genes (*PLEKHA2* and *PRPF38AP1*) were associated with O_3 in pregnancy and the second trimester (Fig. 1a, left panel); one gene (*RERE*) with NO_2 in pregnancy, first, and third trimesters (Fig. 1b, left panel); and one gene with $PM_{2.5}$ (*HIST1H2AK*) in pregnancy and second-trimester exposure windows (Fig. 1c, left panel).

Evaluation of critical months of exposure

We used distributed lag models to determine whether there were specific critical months of exposure associated with DNA methylation patterns. We classified the distributed lag model results into trend types based on the best-fitting model, as determined by the AIC, and the slope for polynomial coefficients of its line of best fit. A conceptual overview of the trend types can be found in Fig. 2.

This analysis revealed 21 gene associations with at least one lagged exposure month in the positive direction, indicating increased risk for changes in DNA methylation, for a given month, in response to a specific pollutant (Table 3). Of these, three were in buccal (Supplementary Fig. S9), three in cord blood (Supplementary Fig. S10), six in nasal mucosa (Supplementary Fig. S11), four in nasal lavage (Supplementary Fig. S12), and five were in placenta tissue (Supplementary Fig. S13; Supplementary Table S8).

Eight of the 21 genes best fit a third-order polynomial distributed lag relationship between air pollutant exposure and methylation levels. As shown in Table 3, two genes (*AL662890.3* and *TNFRSF14*) were classified as trend type A ('negative serpentine') and six genes (*EGLN3, MIR4309, CCDC88C, ADAT3, PPP1R1B,* and *VPS45*) were classified as trend type B ('positive serpentine'). These results showed that the associations between monthly air pollution exposure and DNA methylation at these genes vary throughout pregnancy, with two or more periods of increased effect. This may indicate multiple critical windows of exposure at various time periods in pregnancy.

Six genes best fit a nonlinear model, as determined by the AIC. Of these, two genes (*PCSK7* and *SERPINB6*) were classified as trend type C ('downward U-shaped') and four additional genes (*BAIAP2L1, GNG7, MALT1,* and *MXRA8*) were classified as trend type D ('upward U-shaped'; Table 3). For these genes, the association between monthly air pollution exposure and DNA methylation at the gene was significant at one or two distinct periods. This

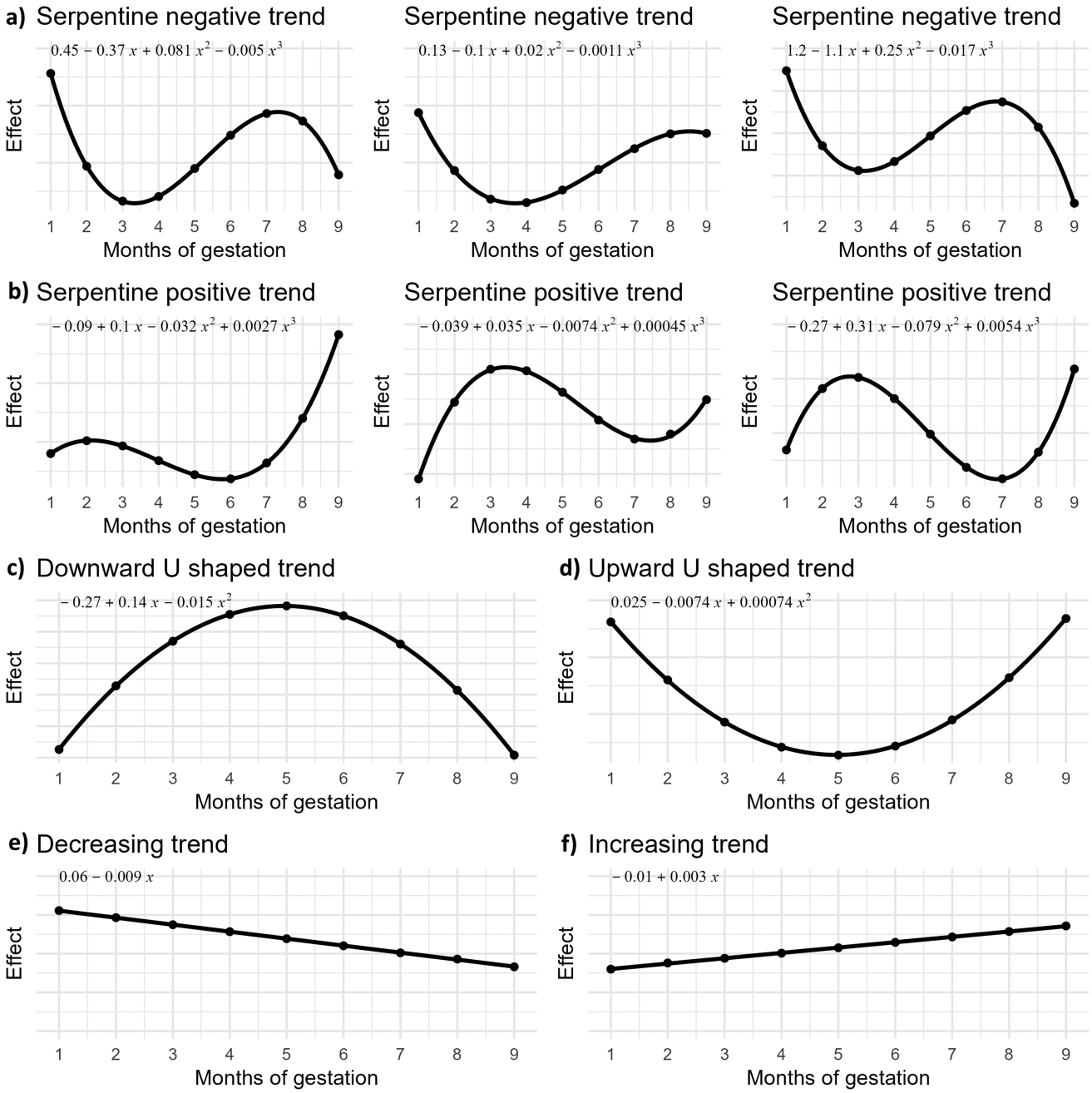

Figure 2. Conceptual overview of how genes were categorized into distributed lag result types. (a) Serpentine trend where the coefficient for the third-order polynomial was negative. (b) Serpentine trend where the coefficient for the third-order polynomial was positive. (c) U-shaped trend where the coefficient for the second-order polynomial was negative. (d) U-shaped trend where the coefficient for the second-order polynomial was positive. (e) Linear trend where the coefficient for the first-order polynomial was negative. (f) Linear trend where the coefficient for the first-order polynomial was positive.

may indicate critical windows of exposure in the middle of pregnancy (downward U-shaped) or at the very beginning and end of pregnancy (upward U-shaped).

Eight genes best fit a linear model as determined by the AIC. Of these, three genes (*SRSF10P1*, *HPN*, and *ZNF184*) were classified as trend type E ('decreasing'), and four genes (*CCNL2*, *RP4-536B24.3*, *RP5-1028K7.2*, and *ENO1*) were classified as trend type F ('increasing'). These trends reflected a gradual increase or decrease in the association between monthly air pollution exposure and DNA methylation at the gene throughout gestation and may indicate additive or decreasing effects of exposure across pregnancy.

Discussion

We conducted one of the largest investigations to date on DNA methylation associations with prenatal exposure to ambient air pollutants in US children. This study examined an enriched set of genes where methylation changes were previously associated with asthma and allergic sensitization and also showed evidence for having functional relevance. Thus, the DNA methylation changes we identified to be associated with air pollutant exposure levels are also directly relevant to functional biological changes and have been previously linked to asthma and allergy outcomes. Unlike prior studies, we focused on regions of the genome given high

Table 3. Trends of air pollution exposure

Trend	Gene symbols	Interpretation
A	*AL662890.3* *TNFRSF14*	Methylation changes in response to air pollution may start early, but rapidly become more stable. There are fewer methylation changes in midpregnancy. Later in pregnancy, methylation changes may increase rapidly again and taper off leading up to birth. This may indicate 'critical windows' in months where methylation changes are more responsive to air pollution exposure.
B	*EGLN3* *MIR4309* *CCDC88C* *ADAT3* *PPP1R1B* *VPS45*	Methylation changes in response to air pollution may not start immediately but may rapidly increase in the following months. There are relatively fewer methylation changes in midpregnancy. Later in pregnancy, methylation changes may increase rapidly again. This may indicate 'critical windows' in months where methylation changes are more responsive to air pollution exposure.
C	*PCSK7* *SERPINB6*	There are few methylation changes in response to air pollution in the early months and later months. Most changes take place midpregnancy. This may indicate a 'critical window' in midpregnancy where methylation changes are more responsive to air pollution exposure.
D	*BAIAP2L1* *GNG7* *MALT1* *MXRA8*	Methylation changes in response to air pollution are frequent at the beginning of pregnancy and rapidly become more stable midpregnancy. Changes may rapidly increase again later in pregnancy. This may indicate 'critical windows' in early and late pregnancy where methylation changes are more responsive to air pollution exposure.
E	*SRSF10P1* *HPN* *ZNF184*	Methylation changes in response to air pollution start at the beginning of pregnancy, but methylation gradually becomes more stable as pregnancy continues.
F	*CCNL2* *RP4-536B24.3* *RP5-1028K7.2* *ENO1*	Fewer methylation changes in response to air pollution occur at the beginning of pregnancy, but changes gradually increase as pregnancy continues. This could also be indicative of cumulative effects of air pollution on methylation change over pregnancy.

Linear and nonlinear regression models were fit to the coefficients from the distributed lag model to classify the functional form of the trends. Trends were classified into A: 'Serpentine negative', B: 'Serpentine positive', C: 'Downward U-shaped', D: 'Upward U-shaped', E: 'Decreasing', or F: 'Increasing' based on the directions of slopes of the polynomial coefficients in the best-fitting model, determined by the lowest AIC. 'Increasing' and 'Decreasing' trends were further confirmed by a Mann–Kendall test. Figure 2 contains a conceptual illustration of how the trends were determined for each gene.

priority for contributing to asthma and allergy risk, evaluated critical months of exposure, and compared findings across air pollutant and biospecimen types to identify shared and exposure- and biospecimen-specific effects. We identified 154 genes showing methylation associations with O_3 ($n = 77$), NO_2 ($n = 36$), or $PM_{2.5}$ ($n = 49$). Most of the associations were pollutant-specific (95%) and biospecimen-specific (96%), with a variety of trends in critical windows of exposure. The magnitude of methylation differences varied by air pollutant and tissue type, ranging from ~1% up to 50% for each one- to five-unit increase in exposure. While the specific gene methylation probes we examined here have not been studied in prior air pollutant studies, the magnitudes of effect we observed exceeded or were consistent with those reported in prior studies.

We observed differences in the number of genes showing methylation-exposure associations by trimester, with the largest number of associations in the first and third trimesters. Previous studies have supported both early and late pregnancy critical windows for impacts of O_3 exposure on adverse birth outcomes, including one study [40] that even identified an increased risk of low birth weight and preterm related to O_3 exposure in the final week of pregnancy [41, 42]. Interestingly, some genes associated with O_3 also followed a pattern of abrupt increased risk of methylation change in the last month of gestation in distributed lag models (*LZTS2* and *CCDC88C*). NO_2 exposure late in pregnancy was associated with adverse birth outcomes, but exposure during all of pregnancy has been associated with asthma, allergic outcomes, and child lung function trajectories [4, 6, 7, 42–45]. Particulate matter exposure during the third trimester, in particular, has been associated with a wide range of adverse offspring health outcomes [43, 46]. Overall, our results are consistent with prior literature showing that exposure to air pollutants at critical windows in early pregnancy, late pregnancy, and all of pregnancy impacts a wide range of child health outcomes (e.g. respiratory, immune, and birth) and provides potential clues into the corresponding biology relevant to these associations. Nasal mucosal and lavage biospecimens followed a pattern of a larger number of pregnancy average or early-to-mid-pregnancy gene associations with fewer third-trimester associations. Distributed lag models further elucidated critical months of exposure associated with DNA methylation at 21 genes, which were observed in each biospecimen type. Most of the associations with these 21 genes were with two or more critical windows during pregnancy, while only a few showed linear effects with either a consistent increase or decrease in exposure effects on methylation with each increasing month of pregnancy. Our results provide insights into the effects of prenatal exposure to ambient air pollutants on child biology and highlight high-priority genes in which DNA methylation may mediate the effects of prenatal air pollutant exposure on health outcomes in offspring.

While most of the associations we identified were pollutant-specific, some were with multiple pollutants. In particular, one gene, SERPINB6, was associated with NO_2 exposure in every trimester for placenta and cord blood, as well as third trimester O_3 for nasal lavage biospecimens, suggesting a biologic target of ambient pollution that is shared across cell types and persistent throughout pregnancy. This gene, which is highly expressed in mast cells [47], is a member of the serpin superfamily, which is critical to regulating airway physiology [48] and has also been linked to pulmonary diseases [49] and inflammatory response regulation, as well as asthma [50–52]. This is a strong candidate for future studies to investigate whether DNA methylation at this locus mediates the risk effects observed with prenatal air pollutant exposure on respiratory health outcomes in offspring.

There were also multiple genes associated with other pairwise combinations of exposures, including O_3–NO_2 ($n = 4$), O_3–$PM_{2.5}$ ($n = 3$), and NO_2–$PM_{2.5}$ ($n = 1$). Given the inverse correlation between NO_2 and O_3 levels, due to atmospheric chemical reactions, it is not surprising that some of the methylation changes are detected for both exposures. The large number of pollutant-specific changes suggests distinct biological impacts of different components of ambient air pollutants.

The age at biospecimen collection varied from birth (placenta and cord blood) through childhood and early adulthood (buccal and nasal). Overall, we observed fewer associated genes in the nasal and buccal biospecimens related to ambient air pollution exposures compared to the birth specimens. It is possible that this reflects differences in the effects of exposure by cell type, that stronger DNA methylation effects are observed with more recent exposure, or that some prenatal exposure effects on methylation do not persist into later childhood. Future longitudinal studies are needed to clarify these possibilities. The first- and second-trimester specificity of the exposure timing and methylation changes in nasal lavage and nasal mucosal swabs, which are primarily composed of immune cells, is particularly intriguing, given that human immune cells undergo development, maturation, and transitions to compartmentalization in organs between pregnancy weeks 4 and 20 [53]. Air pollution exposure during this critical window may shape DNA methylation patterns in immune cells that migrate to the airway and persist into later childhood. Another notable exception is for $PM_{2.5}$ exposure. In particular, $PM_{2.5}$ exposure was associated with a larger number of genes in nasal mucosal and buccal swabs relative to other biospecimen types.

Previous reports identified up to 21 CpGs or genomic regions associated with prenatal O_3, NO_2, or $PM_{2.5}$ exposures. From our search of the literature regarding methylation changes associated with prenatal ambient air pollutants, one gene was observed in our study and previous studies in different paediatric population samples. A DMR in *NUDT12*, a gene that was associated with O_3 exposure in placenta tissue in our study, was previously discovered in cord blood and also associated with O_3 exposure [23]. Replication of this gene result in the ECHO program cohorts further supports O_3 associations with methylation changes in *NUDT12* and extends prior results by showing changes in the placenta, another developmentally important tissue. In addition, a DMR annotated to this gene was previously reported in cord blood and associated with lung function measures (FEV_1/FVC) and with the expression of *NUDT12* in childhood [54]. Overall, we were limited in our ability to replicate prior findings in this large independent sample of US children and across biospecimen types due to the different CpG content on the Asthma&Allergy array. Nonetheless, this array allowed us to identify novel DNA methylation associations between specific genes and prenatal ambient air pollution exposure, and overall many more associations than reported previously despite testing a much smaller number of CpGs (~45 000 vs. 450 000 or 800 000 on the 450K or EPIC arrays, respectively). The enrichment of CpGs near genes involved in immune responses on the Asthma&Allergy array may have contributed to this finding. Our results showing pollutant-specific methylation effects are also largely consistent with prior literature and add new insights into critical windows of exposure- and biospecimen-specific and shared methylation changes.

The strengths of this study include our inclusion of a large cohort of children from the USA, ability to assess the effects of prenatal ambient air pollutant exposures on multiple tissue types, exposure windows, and child life stages, ranging from infancy through adolescence. Furthermore, our focus on CpGs enriched for characteristics reflecting gene-regulatory properties and for regions of the genome relevant to asthma and allergic diseases that are not currently measured on other platforms provides novel biologic insights and potential biologic targets of environmental risks for asthma and allergy, including air pollution, examined here. Despite these strengths, we recognize there are some limitations. First, postnatal air pollutant exposure could influence DNA methylation in nasal lavage and nasal mucosal samples collected in childhood. However, this is unlikely to be a major concern for the following reasons: (i) for postnatal exposure to impact our nasal results, it would have to be highly correlated with prenatal exposure levels, and (ii) strong evidence showing prenatal-specific changes can persist from birth through adulthood and are independent of postnatal exposure has been shown for other exposure domains [55–57]. This is not a concern in the buccal, cord blood, or placenta results because there was no postnatal exposure at the time of collection (birth). This suggests the mode of exposure (i.e. maternal prenatal vs. postnatal) may influence DNA methylation patterns differently. Additionally, this distinguishable pattern of methylation from prenatal exposure may be identifiable later in life. A second limitation is the possible influence of unmeasured confounders in our analyses. For example, we did not have estimates of relative humidity or the temperature at the time of the air pollution measurement across cohorts available to include as covariates and instead opted for a seasonal indicator. Other variables we investigated, like maternal diet, education, income, and insurance, were either mostly or completely missing in several cohort sites, so we could not reasonably impute missing values. The second cycle of data collection in the ECHO program attempts to address these limitations by ensuring cohort sites collect the same set of measures [58]. Also, there are likely environmental and health exposures that alter childhood DNA methylation, which have not yet been investigated in the CpG sites covered by the Asthma&Allergy array. Since pollution exposure is the focus of this study, we did not delve into other environmental exposures. The novel array will bring many opportunities for future investigation. Third, it is possible that a lack of complete adjustment occurred in our analyses. However, we carefully considered our modelling choices to account for confounding without introducing additional bias. One consideration was the downward trend of air pollution exposure levels over time. Long-term temporal differences can be addressed by model adjustments for cohort site because cohorts recruited pregnancies over the course of shorter time windows. We ultimately decided not to include birth year as an additional covariate because adjusting for variables that are predictive of exposures (i.e. air pollution) but not outcomes (i.e. DNA methylation) can induce bias [59], and there is limited evidence to support interannual differences in DNA methylation within shorter time windows that individual cohorts collected data. Other differences across cohort sites, such as DNA extraction methods and years of sample collection, could impact variability in DNA methylation measures. Beyond adjusting for cohort site, we took multiple steps to reduce technical variation. We applied between-sample normalization methods, corrected models for within-array correlations of CpGs, and included latent factors that adjusted for unobserved variation induced by different collection and extraction methodologies, such as cell-type composition differences. Finally, different cohorts provided different specimen sources, which may limit comparability of associations across biospecimen types. To extend our results and enable direct cross-specimen comparisons as well as to assess longitudinal persistence of air pollution-associated methylation changes, future

cohort studies will need to collect multiple biospecimen types from the same participants across multiple life stages.

Consistent with the developmental origins of health and disease hypothesis [60], DNA methylation changes in specific genes, enriched for their relevance to asthma and allergy health outcomes, may be mechanistically involved in prenatal air pollutant exposure effects on offspring health outcomes in a tissue-, exposure window-, and pollutant-specific manner.

Acknowledgements

The authors wish to thank our ECHO colleagues; the medical, nursing, and program staff; and the children and families participating in the ECHO cohort. We also thank Chris Reyes for providing administrative and management support for a subset of cohort sites that contributed to this project.

Author contributions

Conceptualization: J.G. [equal], C.O. [equal], and C.L.-A [equal]. F.F. [supporting], T.M.B.[supporting], C.B. [supporting], S.C.[supporting], F.D.G.[supporting], D.G.[supporting], R.H.[supporting], T.H.[supporting], G.K.K.H.[supporting], D.J.J.[supporting], R.M.[supporting], P.R.[supporting], L.S.-K.[supporting], E.R.S.[supporting], J.S.[supporting], J.G.[supporting], and C.M.[supporting], and Data acquisition: J.G, C.O., F.D.G., D.G., R.H., and P.R. Formal analysis: M.P., E.E.T., and F.F., and C.M. Software: C.M.,. Supervision: C.O., C.L-A, C.M. Writing - original draft: M.P., E.E.T, C.O., C.L-A. Writing - review and editing: F.F., T.M.B., C.B., S.C., F.D.G., D.G., R.H., T.H., G.K.K.H., D.J.J., R.M., P.R., L.S.-K., E.R.S., J.S., J.G., and C.M.

Supplementary data

Supplementary data is available at *EnvEpig* online.

Conflict of interest. Dr C.L.-A. reports consulting fees from the University of Iowa for providing expertise on autism epigenetics, outside the scope of this work. All other authors have no conflicts to declare.

Funding

Research reported in this publication was supported by the Environmental influences on Child Health Outcomes (ECHO) Program, Office of the Director, National Institutes of Health (NIH), under Award Numbers U2COD023375 (Coordinating Center), U24OD023382 (Data Analysis Center), U24OD023319 with cofunding from the Office of Behavioral and Social Science Research (Measurement Core), U24OD035523 (Lab Core), ES0266542 (Human Health Exposure Analysis Resource; HHEAR), U24ES026539 (HHEAR Barbara O'Brien), U2CES026533 (HHEAR Lisa Peterson), U2CES026542 (HHEAR Patrick Parsons, Kannan Kurunthacalam), U2CES030859 (HHEAR Manish Arora), U2CES030857 (HHEAR Timothy R. Fennell, Susan J. Sumner, Xiuxia Du), U2CES026555 (HHEAR Susan L. Teitelbaum), U2CES026561 (HHEAR Robert O. Wright), U2CES030851 (HHEAR Heather M. Stapleton, P. Lee Ferguson), UG3/UH3OD023251 (Akram Alshawabkeh), UH3OD023320 and UG3OD035546 (Judy Aschner), UH3OD023332 (Clancy Blair, Leonardo Trasande), UG3/UH3OD023253 (Carlos Camargo), UG3/UH3OD023248 and UG3OD035526 (Dana Dabelea), UG3/UH3OD023313 (Daphne Koinis Mitchell), UH3OD023328 (Cristiane Duarte), UH3OD023318 (Anne Dunlop), UG3/UH3OD023279 (Amy Elliott), UG3/UH3OD023289 (Assiamira Ferrara), UG3/UH3OD023282 (James Gern), UH3OD023287 (Carrie Breton), UG3/UH3OD023365 (Irva Hertz-Picciotto), UG3/UH3OD023244 (Alison Hipwell), UG3/UH3OD023275 (Margaret Karagas), UH3OD023271 and UG3OD035528 (Catherine Karr), UH3OD023347 (Barry Lester), UG3/UH3OD023389 (Leslie Leve), UG3/UH3OD023344 (Debra MacKenzie), UH3OD023268 (Scott Weiss), UG3/UH3OD023288 (Cynthia McEvoy), UG3/UH3OD023342 (Kristen Lyall), UG3/UH3OD023349 (Thomas O'Connor), UH3OD023286 and UG3OD035533 (Emily Oken), UG3/UH3OD023348 (Mike O'Shea), UG3/UH3OD023285 (Jean Kerver), UG3/UH3OD023290 (Julie Herbstman), UG3/UH3OD023272 (Susan Schantz), UG3/UH3OD023249 (Joseph Stanford), UG3/UH3OD023305 (Leonardo Trasande), UG3/UH3OD023337 (Rosalind Wright), UG3OD035508 (Sheela Sathyanarayana), UG3OD035509 (Anne Marie Singh), UG3OD035513 and UG3OD035532 (Annemarie Stroustrup), UG3OD035516 and UG3OD035517 (Tina Hartert), UG3OD035518 (Jennifer Straughen), UG3OD035519 (Qi Zhao), UG3OD035521 (Katherine Rivera-Spoljaric), UG3OD035527 (Emily S. Barrett), UG3OD035540 (Monique Marie Hedderson), UG3OD035543 (Kelly J. Hunt), UG3OD035537 (Sunni L. Mumford), UG3OD035529 (Hong-Ngoc Nguyen), UG3OD035542 (Hudson Santos), UG3OD035550 (Rebecca Schmidt), UG3OD035536 (Jonathan Slaughter), and UG3OD035544 (Kristina Whitworth). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Role of funder: The sponsor, National Institutes of Health (NIH), participated in the overall design and implementation of the ECHO Program, which was funded as a cooperative agreement between NIH and grant awardees. The sponsor approved the Steering Committee-developed ECHO protocol and its amendments, including COVID-19 measures. The sponsor had no access to the central database, which was housed at the ECHO Data Analysis Center. Data management and site monitoring were performed by the ECHO Data Analysis Center and Coordinating Center. All analyses for scientific publication were performed by the study statistician, independently of the sponsor. The lead author wrote all drafts of the manuscript and made revisions based on coauthors and the ECHO Publications Committee (a subcommittee of the ECHO Operations Committee) feedback without input from the sponsor. The study sponsor did not review or approve the manuscript for submission to the journal.

Data availability

Select de-identified data from the ECHO Program are available through NICHD's Data and Specimen Hub (DASH). Information on study data not available on DASH, such as some Indigenous datasets, can be found on the ECHO study DASH webpage.

References

1. Brunekreef B, Holgate ST. Air pollution and health. *Lancet* 2002;**360**:1233–42.
2. Abbasi-Kangevari M, Malekpour MR, Masinaei M *et al.*; GBD 2019 North Africa and the Middle East Air Pollution Collaborators. Effect of air pollution on disease burden, mortality, and life expectancy in North Africa and the Middle East: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Planet Health* 2023;**7**:e358–69.
3. Sbihi H, Tamburic L, Koehoorn M *et al.* Perinatal air pollution exposure and development of asthma from birth to age 10 years. *Eur Respir J* 2016;**47**:1062–71.
4. Zanobetti A, Ryan PH, Coull BA *et al.*; ECHO Children's Respiratory and Environmental Workgroup; Early-life exposure to air

pollution and childhood asthma cumulative incidence in the ECHO CREW Consortium. *JAMA Network Open* 2024;**7**: e240535.
5. Neophytou AM, Lutzker L, Good KM *et al*. Associations between prenatal and early-life air pollution exposure and lung function in young children: exploring influential windows of exposure on lung development. *Environ Res* 2023;**222**:115415.
6. Lu C, Zhang Y, Li B *et al*. Interaction effect of prenatal and postnatal exposure to ambient air pollution and temperature on childhood asthma. *Environ Int* 2022;**167**:107456.
7. Garcia E, Rice MB, Gold DR. Air pollution and lung function in children. *J Allergy Clin Immunol* 2021;**148**:1–14.
8. Ho SM, Johnson A, Tarapore P *et al*. Environmental epigenetics and its implication on disease risk and health outcomes. *ILAR J* 2012;**53**:289–305.
9. Bakulski KM, Blostein F, London SJ. Linking prenatal environmental exposures to lifetime health with epigenome-wide association studies: state-of-the-science review and future recommendations. *Environ Health Perspect* 2023;**131**:126001.
10. Dhar GA, Saha S, Mitra P *et al*. DNA methylation and regulation of gene expression: guardian of our health. *Nucleus (Calcutta)* 2021;**64**:259–70.
11. Li S, Ye Z, Mather KA *et al*. Early life affects late-life health through determining DNA methylation across the lifespan: a twin study. *EBioMedicine* 2022;**77**:103927.
12. Cameron VA, Jones GT, Horwood LJ *et al*. DNA methylation patterns at birth predict health outcomes in young adults born very low birthweight. *Clin Clin Epigenet* 2023;**15**:47.
13. Rygiel CA, Dolinoy DC, Bakulski KM *et al*. DNA methylation at birth potentially mediates the association between prenatal lead (Pb) exposure and infant neurodevelopmental outcomes. *Environ Epigenet* 2021;**7**:dvab005.
14. Everson TM, Lyons G, Zhang H *et al*. DNA methylation loci associated with atopy and high serum IgE: a genome-wide application of recursive Random Forest feature selection. *Genome Med* 2015;**7**:89.
15. Rathod A, Melaram R, Zhang H *et al*. The association of DNA methylation at birth with adolescent asthma is mediated by atopy. *Clin Exp Allergy* 2023;**53**:1226–9.
16. Lee S, Sbihi H, MacIsaac JL *et al*. Persistent DNA methylation changes across the first year of life and prenatal NO_2 exposure in a Canadian prospective birth study. *Environ Health Perspect* 2024;**132**:47004.
17. Breton CV, Gao L, Yao J *et al*. Particulate matter, the newborn methylome, and cardio-respiratory health outcomes in childhood. *Environ Epigenet* 2016;**2**:dvw005.
18. Gruzieva O, Xu CJ, Breton CV *et al*. Epigenome-wide meta-analysis of methylation in children related to prenatal NO_2 air pollution exposure. *Environ Health Perspect* 2017;**125**:104–10.
19. Gruzieva O, Xu CJ, Yousefi P *et al*. Prenatal particulate air pollution and DNA methylation in newborns: an epigenome-wide meta-analysis. *Environ Health Perspect* 2019;**127**:57012.
20. Isaevska E, Fiano V, Asta F *et al*. Prenatal exposure to PM_{10} and changes in DNA methylation and telomere length in cord blood. *Environ Res* 2022;**209**:112717.
21. Ladd-Acosta C, Feinberg JI, Brown SC *et al*. Epigenetic marks of prenatal air pollution exposure found in multiple tissues relevant for child health. *Environ Int* 2019;**126**:363–76.
22. Plusquin M, Chadeau-Hyam M, Ghantous A *et al*. DNA methylome marks of exposure to particulate matter at three time points in early life. *Environ Sci Technol* 2018;**52**:5427–37.
23. Starling AP, Wood C, Liu C *et al*. Ambient air pollution during pregnancy and DNA methylation in umbilical cord blood, with potential mediation of associations with infant adiposity: The Healthy Start study. *Environ Res* 2022;**214**:113881.
24. Song AY, Feinberg JI, Bakulski KM *et al*. Prenatal exposure to ambient air pollution and epigenetic aging at birth in newborns. *Front Genet* 2022;**13**:929416.
25. Peng C, den Dekker M, Cardenas A *et al*. Residential proximity to major roadways at birth, DNA methylation at birth and mid-childhood, and childhood cognitive test scores: Project Viva(Massachusetts, USA). *Environ Health Perspect* 2018;**126**:97006.
26. Elliott G, Hong C, Xing X *et al*. Intermediate DNA methylation is a conserved signature of genome regulation. *Nat Commun* 2015;**6**:6363.
27. Morin A, Thompson EE, Helling BA *et al*.; program collaborators for Environmental Influences on Child Health Outcomes and Children's Respiratory and Environmental Workgroup. A functional genomics pipeline to identify high-value asthma and allergy CpGs in the human methylome. *J Allergy Clin Immunol* 2023;**151**:1609–21.
28. Knapp EA, Kress AM, Parker CB *et al*.; on behalf of program collaborators for Environmental Influences on Child Health Outcomes. The Environmental Influences on Child Health Outcomes (ECHO)-Wide Cohort. *Am J Epidemiol* 2023;**192**:1249–63.
29. Aryee MJ, Jaffe AE, Corrada-Bravo H *et al*. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 2014;**30**:1363–9.
30. Pidsley R, Yw CC, Volta M *et al*. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* 2013;**14**:293.
31. Wickham H, Averick M, Bryan J *et al*. Welcome to the tidyverse. [Computer program]. Journal of Open Source Software. Version 2.0.0. 2019.
32. Barrett T, Dowle M,Srinivasan A *et al*. data.table: Extension of 'data.frame' https://Rdatatable.gitlab.io/data.table, https://github.com/Rdatatable/data.table, https://r-datatable.com. R package version 1.17.99, 2025
33. Di Q, Amini H, Shi L *et al*. An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. *Environ Int* 2019;**130**:104909.
34. Yanosky JD, Paciorek CJ, Laden F *et al*. Spatio-temporal modeling of particulate air pollution in the conterminous United States using geographic and meteorological predictors. *Environ Health* 2014;**13**:63.
35. Cunningham F, Allen JE, Allen J *et al*. Ensembl 2022. *Nucleic Acids Res* 2022;**50**:D988–95.
36. Lawrence M, Huber W, Pages H *et al*. Software for computing and annotating genomic ranges. *PLoS Comput Biol* 2013;**9**:e1003118.
37. McKennan C. Factor analysis in high dimensional biological data with dependent observations. preprint arXiv:200911134. 2020.
38. Morin A, McKennan CG, Pedersen CT *et al*. Epigenetic landscape links upper airway microbiota in infancy with allergic rhinitis at 6 years of age. *J Allergy Clin Immunol* 2020;**146**:1358–66.
39. Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *J Stat Softw* 2011;**43**:1–20.
40. Zhang C, Yang J, Wei J *et al*. Individual ambient ozone exposure during pregnancy and adverse birth outcomes: exploration of the potentially vulnerable windows. *J Hazard Mater* 2024;**464**:132945.
41. Rappazzo KM, Nichols JL, Rice RB *et al*. Ozone exposure during early pregnancy and preterm birth: a systematic review and meta-analysis. *Environ Res* 2021;**198**:111317.
42. Ju L, Li C, Yang M *et al*. Maternal air pollution exposure increases the risk of preterm birth: evidence from the meta-analysis of cohort studies. *Environ Res* 2021;**202**:111654.
43. Liao J, Zhang Y, Yang Z *et al*. Identifying critical windows of air pollution exposure during preconception and gestational period on birthweight: a prospective cohort study. *Environ Health* 2023;**22**:71.

44. Usemann J, Decrue F, Korten I *et al*. Exposure to moderate air pollution and associations with lung function at school-age: a birth cohort study. *Environ Int* 2019;**126**:682–9.
45. Tingskov Pedersen CE, Eliasen AU, Ketzel M *et al*. Prenatal exposure to ambient air pollution is associated with early life immune perturbations. *J Allergy Clin Immunol* 2023;**151**:212–21.
46. Johnson NM, Hoffmann AR, Behlen JC *et al*. Air pollution and children's health—a review of adverse effects associated with prenatal exposure from fine to ultrafine particulate matter. *Environ Health Prev Med* 2021;**26**:72.
47. Strik MC, Wolbink A, Wouters D *et al*. Intracellular serpin SERPINB6 (PI6) is abundantly expressed by human mast cells and forms complexes with beta-tryptase monomers. *Blood* 2004;**103**:2710–7.
48. Kelly-Robinson GA, Reihill JA, Lundy FT *et al*. The serpin superfamily and their role in the regulation and dysfunction of serine protease activity in COPD and other chronic lung diseases. *Int J Mol Sci* 2021; **12**:22.
49. Uemasu K, Tanabe N, Tanimura K *et al*. Serine protease imbalance in the small airways and development of centrilobular emphysema in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2020;**63**:67–78.
50. Burgener SS, Leborgne NGF, Snipas SJ *et al*. Cathepsin G inhibition by serpinb1 and serpinb6 prevents programmed necrosis in neutrophils and monocytes and reduces GSDMD-driven inflammation. *Cell Rep* 2019;**27**:3646–3656e5.
51. Woodruff PG, Boushey HA, Dolganov GM *et al*. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci U S A* 2007;**104**:15858–63.
52. Pividori M, Schoettler N, Nicolae DL *et al*. Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies. *Lancet Respir Med* 2019;**7**:509–22.
53. Park JE, Jardine L, Gottgens B *et al*. Prenatal development of human immunity. *Science* 2020;**368**:600–3.
54. den Dekker HT, Burrows K, Felix JF *et al*. Newborn DNA-methylation, childhood lung function, and the risks of asthma and COPD across the life course. *Eur Respir J* 2019;**4**:53.
55. Ladd-Acosta C, Fallin MD. DNA methylation signatures as biomarkers of prior environmental exposures. *Curr Epidemiol Rep* 2019;**6**:1–13.
56. Richmond RC, Suderman M, Langdon R *et al*. DNA methylation as a marker for prenatal smoke exposure in adults. *Int J Epidemiol* 2018;**47**:1120–30.
57. Richmond RC, Simpkin AJ, Woodward G *et al*. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet* 2015;**24**:2201–17.
58. Blackwell CK, Cella D, Adair L *et al*. Extending the Environmental influences on Child Health Outcomes (ECHO) Cohort through 2030: rationale and study protocol. *PLoS One* 2024; e0312677.
59. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019;**34**:211–9.
60. Gillman MW. Developmental origins of health and disease. *N Engl J Med* 2005;**353**:1848–50.

Environmental Epigenetics, 2025, 11(1), dvaf013 , DOI: https://doi.org/10.1093/eep/dvaf013, Advance Access Publication 6 May 2025, Research Article
Received 16 December 2024; Revised 17 April 2025; Accepted 1 May 2025