
Peer reviewed version

Link to published version (if available): 10.1183/13993003.01795-2018

Link to publication record in Explore Bristol Research

PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at https://erj.ersjournals.com/content/53/4/1801795. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore- bristol-research/ebr-terms/
Original article

Newborn DNA-methylation, childhood lung function, and the risks of asthma and COPD across the life course

Running head: DNA-methylation and respiratory disease

Herman T. den Dekker, M.D.¹,²,³,⁴, Kimberley Burrows, Ph.D.⁴, Janine F. Felix, M.D., Ph.D.¹,³,⁵, Lucas A. Salas, M.D., Ph.D.⁶-⁹, Ivana Nedeljkovic, M.D., M.Sc.³, Jin Yao, Ph.D.¹⁰, Sheryl L. Rifas-Shiman, M.P.H.¹¹, Carlos Ruiz-Arenas, M.Sc.⁶,⁸,⁹, N. Amin, Ph.D.³, Mariona Bustamante, Ph.D.⁶,⁸,⁹,¹², Dawn L. DeMeo, M.D., M.P.H.¹³, A. John Henderson, M.D.⁴, Caitlin G. Howe, Ph.D.¹⁰, Marie-France Hivert, M.D., M.Sc.¹¹, M. Arfan Ikram, M.D., Ph.D.³, Johan C. de Jongste, M.D., Ph.D.², Lies Lahousse, Pharm.D., Ph.D.³,¹⁴, Pooja R. Mandaviya, M.D.¹⁵,¹⁶, Joyce B. van Meurs, M.D., Ph.D.¹⁶, Mariona Pinart, Ph.D.⁶,⁸,⁹,¹⁷, Gemma C. Sharp, Ph.D.⁴, Lisette Stolk, Ph.D.¹⁶,¹⁸, André G. Uitterlinden, Ph.D.³,¹⁶,¹⁸, Josep M. Anto, M.D., Ph.D.⁶,⁸,⁹,¹⁷, Augusto A. Litonjua, M.D., M.P.H.¹³, Carrie V. Breton, Sc.D.¹⁰, Guy G. Brusselle, M.D., Ph.D.³,¹⁰,¹⁹, Jordi Sunyer, M.D., Ph.D.⁶,⁸,⁹,¹⁷, George Davey Smith, M.D., Ph.D.⁴, Caroline L. Relton, Ph.D.⁴,¹, Vincent W.V. Jaddoe, M.D., Ph.D.¹,³,⁵,¹, Liesbeth Duijts, M.D., Ph.D.²,²⁰,†

*These authors contributed equally
†These authors contributed equally

¹The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
²Department of Pediatrics, Division of Respiratory Medicine and Allergology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
³Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
⁴MRC Integrative Epidemiology Unit, University of Bristol, UK School of Social and Community Medicine, University of Bristol, UK
⁵Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
6 ISGLobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
7 Department of Epidemiology, Geisel School of Medicine at Dartmouth, Lebanon, NH, USA
8 Universitat Pompeu Fabra (UPF), Barcelona, Spain
9 CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
10 Department of Preventive Medicine, University of Southern California, Los Angeles, CA 90089
11 Obesity Prevention Program, Department of Population Medicine, Harvard Medical School and
Harvard Pilgrim Health Care Institute, Boston, MA 02115, USA
12 Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology,
Barcelona, Spain
13 Channing Division of Network Medicine, Brigham and Women’s Hospital, Boston, MA 02115, USA
14 Department of Respiratory Medicine, Ghent University Hospital, B-9000 Ghent, Belgium;
15 Department of Clinical Chemistry, Erasmus MC, University Medical Center Rotterdam, The
Netherlands
16 Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, the
Netherlands
17 IMIM (Hospital del Mar Medical Research Institute)
18 Netherlands Consortium for Healthy Ageing (NCHA), Erasmus MC, University Medical Center
Rotterdam, The Netherlands
19 Department of Respiratory Medicine, Erasmus MC, University Medical Center Rotterdam,
Netherlands
20 Department of Pediatrics, Division of Neonatology, Erasmus MC, University Medical Center
Rotterdam, Rotterdam, the Netherlands
Acknowledgements

A full description of acknowledgements and funding is provided in the Supplementary Appendix. The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. HD, KB, JF, CR, VJ, and LD contributed to the study design, data analysis plan, data collection, data and meta-analysis, data interpretation, writing, reviewing the manuscript critically and gave consent for submission. All other authors contributed equally to study design, data analysis plan, cohort specific data collection and analysis, reviewing the manuscript critically and gave consent for submission.

Corresponding author

Liesbeth Duijts, MD, PhD, Erasmus MC, University Medical Center Rotterdam, Sp-3435; PO Box 2060, 3000 CB, Rotterdam, The Netherlands.
Tel: *31 10 7036263, Fax: *31 10 7036811, E-mail: l.duijts@erasmusmc.nl

Abbreviations

DMR  Differentially Methylated Region
COPD  Chronic Obstructive Pulmonary Disease
FEV$_1$  Forced Expiratory Volume in 1 second
FEV$_1$/FVC  Forced Expiratory Volume in 1 second / Forced Vital Capacity
FEF$_{75}$  Forced Expiratory Flow at 75% of Forced Vital Capacity
ABSTRACT

Rationale We aimed to identify differentially methylated regions (DMRs) in cord blood DNA associated with childhood lung function, asthma and chronic obstructive pulmonary disease (COPD) across the life course.

Methods We meta-analyzed epigenome-wide data of 1688 children from five cohorts to identify cord blood DMRs and their annotated genes, in relation to Forced Expiratory Volume in 1 second (FEV1), FEV1/Forced Vital Capacity (FVC), and Forced Expiratory Flow at 75% of FVC (FEF75) at ages 7 to 13 years. Identified DMRs were explored for associations with childhood asthma, adult lung function and COPD, gene expression and involvement in biological processes.

Results We identified 59 DMRs associated with childhood lung function, of which 18 were associated with childhood asthma and 9 with COPD in adulthood. Genes annotated to the top ten identified DMRs were HOXA5, PAOX, LINC00602, ABCA7, PER3, CLCA1, VENTX, NUDT12, PTPRN2 and TCL1A. Differential gene expression in blood was observed for 32 DMRs in childhood and 18 in adulthood. Genes related with 16 identified DMRs were associated with respiratory developmental or pathogenic pathways.

Interpretation Our findings suggest that the epigenetic status of the newborn affects respiratory health and disease across the life course.
INTRODUCTION

Asthma and chronic obstructive pulmonary disease (COPD) have become major global health problems in the last decades (1). Both diseases are characterized by airway obstruction, as indicated by a reduced Forced Expiratory Volume in 1 second (FEV₁), FEV₁ to Forced Vital Capacity (FVC) ratio, and Forced Expiratory Flow at 75% of FVC (FEF₇₅) (2). Childhood lung function predicts lung function and risks of asthma and COPD in later life (3). An accumulating body of evidence suggests that asthma and COPD have at least part of their origins in fetal life (4, 5). Genetics alone fail to explain the quickly altering prevalence of allergies and chronic respiratory diseases in the past decades, because any mutation would require multiple generations to occur on a population level (6). Furthermore, adverse fetal exposures, such as maternal smoking and suboptimal diet, increase the risk of asthma and COPD (5). The pathways linking genetic predisposition and environmental exposures in fetal life with life course respiratory disease may include epigenetic changes, including DNA-methylation (5). Epigenetic changes are influenced by environmental exposures and could exert population effects much more rapidly than genetic mutations (6). DNA-methylation is currently the best understood epigenetic mechanism, and techniques have been developed to assess epigenome-wide DNA-methylation patterns in large population-based studies. Fetal development is characterized by high rates of DNA-methylation changes and rapid organ development (5). DNA-methylation may affect fetal development through effects on gene transcription and expression (7). Recent studies assessing the associations between DNA-methylation and childhood respiratory health are mainly limited to candidate genes and small sample sizes (3, 8). We focused on identification of differential DNA-methylated regions (DMRs) because regional methylation of CpGs controls cell-type-specific transcription. Also, the use of DMRs increases statistical power and minimizes the effects of genetic variants in the methylation analyses (9). Identification of genomic regions with altered DNA-methylation levels related to lung function and respiratory diseases across the life course is important to understand mechanisms underlying associations of environmental and genetic factors with the development of lower lung function and risk of chronic respiratory diseases. We hypothesized that fetal differential DNA-methylation reflected in cord blood DNA of newborns affect gene expression and subsequent respiratory tract development, and predispose individuals for obstructive airway diseases in later life (10, 11).
We meta-analyzed five epigenome-wide association studies using data from 1,688 children participating in prospective cohort studies to identify differential DNA-methylated regions (DMRs) of newborns associated with childhood FEV₁, FEV₁/FVC and FEF₇₅. Identified top DMRs were subsequently explored for their associations with childhood asthma, lung function in adolescence and adulthood, and COPD in adulthood, and explored for association with gene expression and involvement in biological processes.

METHODS

Study Design and Data Sources
We included population-based cohort studies participating in PACE consortium with data on epigenome-wide DNA-methylation at birth and lung function in childhood (12). We used data from 1,688 Caucasian children aged 7 to 13 years participating in the Avon Longitudinal Study of Parents and Children (ALSPAC) (United Kingdom), Generation R (Netherlands), INfancia y Medio Ambiente Study (INMA) (Spain), Children’s Health Study (CHS) and Project Viva (both from the U.S.A.). These data were used for the primary discovery epigenome-wide meta-analysis to identify DMRs of newborns related to childhood lung function. We aimed to identify DMRs instead of single CpGs while differences at any individual CpG may be small, and the use of DMRS might minimize the effects of genetic variants in the methylation analysis (13, 14).

We used several resources for the secondary analyses. For clinical outcomes, we used childhood asthma data (Generation R, mean age 6 years), lung function data from adolescents (ALSPAC, mean age 15 years) and adults (Rotterdam Study, mean age 66 years, The Netherlands), and COPD data in adults (Rotterdam Study) (Figure 1). For gene expression, we used blood samples from children (INMA, at birth and mean age 4 years) and adults (Rotterdam Study). Last, we used publicly available resources to relate identified DMRs with biological processes (15-17). Parents, legal representatives or participants provided informed consent in accordance with local ethics policies. Detailed information about the study design and cohorts is provided in the Supplementary Appendix.

DNA-methylation
All cohorts extracted DNA from blood samples and used the EZ-96 DNA Methylation kit (Zymo Research Corporation, Irvine, USA) for bisulfite conversion. Samples were processed with the Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, USA) followed by cohort-specific quality control, probe exclusion and data normalization. Detailed information on cohort-specific data acquisition and quality control is provided in the Supplementary Appendix.

Respiratory Outcome Assessment

Lung function measures comprised pre-bronchodilator FEV₁, FEV₁/FVC and FEF₇⁵, which were converted into sex-, age-, height- and ethnicity-adjusted z-scores (18). Physician-diagnosed asthma was obtained by questions adapted from the International Study on Asthma and Allergy in Childhood (19). COPD was defined as pre-bronchodilator FEV₁/FVC <0.70 in the absence of asthma, or a doctor diagnosis (20).

Statistical Analyses

Primary Meta-analysis on Childhood Lung Function A detailed description of applied methods is presented in the Supplementary Appendix. Individual cohorts used robust linear regression models to examine the associations of DNA-methylation levels of CpGs with childhood FEV₁, FEV₁/FVC and FEF₇⁵. Analyses were adjusted for maternal age, socio-economic status, smoking during pregnancy, parity, asthma or atopy, technical covariates and estimated cell counts (21). Results were combined using inverse variance-weighted fixed-effect meta-analyses. Results from unadjusted models were similar to fully adjusted models (Supplementary Appendix Table 1). Using p-values obtained from the meta-analyses, we identified DMRs using the software-tool Comb-p, which is the most robust tool to identify DMRs with small effect sizes (9, 22). Regions were defined as a minimum of 2 probes within a window size of 500 bases with an FDR-threshold <0.05 (9). Comb-p uses unadjusted p-values for each probe as input, and calculates adjusted p-values for each probe that account for the correlation with nearby CpGs (23). Next, the SLK p-values were adjusted for multiple testing and adjusted into q-values. Comb-p finds DMRs based on these q-values and calculates p-values for these DMRs based on the original p-values. Finally, the DMR p-values were adjusted for multiple testing using the Šidák-correction based on the size of the region and number of possible regions of that size. A sliding window identifies a DMR without any predefined regional borders, and therefore
(theoretically) does not have a maximum number of windows and DMRs. A more extensive description of the identification of DMRS is provided in the Supplementary Material. Annotation of the genes located nearest to the DMRs was performed using Peak Annotation and Visualization (PAVIS) (24). We limited annotation to a region of 500kb (250kb upstream, 250kb downstream of the beginning and end of the region, respectively). All annotations were based on human GRCh37/hg19 assembly. Because genetic variants in Infinium probes could result in spurious methylation measurements, we performed a sensitivity analysis in a subset of high-quality probes (n=294,834) without SNPs, insertions or deletions, repeats, polymorphic probes and bisulfite induced reduced genomic complexity (25).

**Secondary Analyses on Later Life Lung Function and Respiratory Diseases** We used linear and logistic regression models to examine the associations of CpGs within identified DMRs with asthma in childhood, FEV₁, FEV₁/FVC and FEF₇₅ in adolescence and adulthood, and COPD in adulthood. Single CpG p-values were used to reconstruct the identified DMRs with Comb-p, applying identical parameter settings as in the discovery meta-analyses including false discovery rate (FDR)-correction (9, 26). We did not apply Šidák-correction because analyses were limited to the identified DMRs.

**Gene Expression Analyses** We assessed the associations of CpGs within identified DMRs with gene expression in a region of ±250kb in blood samples from children and adults. P-values of CpGs associated with gene expression were combined for each DMR using a modified generalized Fisher method and FDR-correction (26, 27). Additionally, we explored whether the annotated and differentially expressed genes were expressed in human lung specimens of the Genotype-Tissue Expression (GTEx) database (15).

**Exploration Biological Processes** The Gene Ontology database implemented in DAVID and Genemania was used to examine gene function in biological processes (16, 17). We examined pathways for all genes annotated to the DMRs and for genes with differential expression in association with the identified DMRs. We used the Kyoto Encyclopedia of Genes and Genomes (KEGG) database in DAVID and Genemania, the OMIM database and the Universal Protein Resource (UniProt) to explore whether annotated or expressed genes have been related to respiratory development or diseases (28). We used the Ensemble Genome browser to visualize the genomic structure of the identified DMRs.
RESULTS

Meta-analysis of Epigenome-wide Association Studies on Childhood Lung Function

Characteristics of the participating cohorts are given in Table 1 and Supplementary Appendix Table 2.

We identified 22, 15 and 22 DMRs associated with FEV₁, FEV₁/FVC and FEF₇₅, respectively (Figure 2, Supplementary Appendix Tables 3 and 4). A higher mean methylation of CpGs located within 37 (63%) of the identified DMRs was associated with higher lung function measures, and within 22 (37%) of identified DMRs with lower lung function measures. We observed a high homogeneity across the included studies (CpGs with I² <50: FEV₁ 140/163 (86%), FEV₁/FVC 82/89 (92%) and FEF₇₅ 139/148 (87%)) (Supplemental Tables 4a-c).

Of the top ten significant DMRs associated with childhood lung function, the 5 DMRs and their annotated genes for FEV₁ were located at chr7:27,183,133-27,184,854 (HOXA5), chr10:135,202,522-135,203,201 (PAOX), chr6:166,418,799-166,419,139 (LINC00602), chr19:1,063,624-1,064,219 (ABCA7) and chr1:7,887,199-7,887,561 (PER3). Three DMRs and their annotated genes for FEV₁/FVC were located at chr1:86,968,087-86,968,544 (CLCA1), chr10:135,051,149-135,051,582 (VENTX), and chr5:102,898,223-102,898,734 (NUDT12). Two DMRs for FEF₇₅ and their annotated genes were located at chr7:158,045,980-158,046,359 (PTPRN2) and chr14:96,180,406-96,181,045 (TCL1A). After exclusion of potentially problematic probes containing genomic variants, 41 of the 59 previously identified DMRs still contained ≥2 CpGs (supplementary appendix Table 4). Of these 41 DMRs, 54% (n=22) remained to be associated with childhood lung function (Supplementary appendix Table 5).

Identified DMRs and Lung Function and Respiratory Diseases Across the Life Course

Of all 59 identified DMRs related with childhood lung function, 18 (31%) were associated with childhood asthma (Figure 3, Supplementary appendix Table 6). Furthermore, 11 (19%) and 9 (15%) DMRs were associated with lung function in adolescence and adulthood, respectively, and 9 (15%) were associated with COPD. The DMRs annotated to HOXA5 and PAOX were associated with childhood and adolescence FEV₁ and COPD, but not with childhood asthma or adult lung function.
The DMRs annotated to PER3 and VENTX were associated with childhood and adolescence FEV₁ and FEV₁/FVC, respectively. The DMR annotated to NUDT12 was associated with childhood FEV₁/FVC and COPD. The DMRs annotated to PTPRN2 and TCL1A were associated with childhood FEF₇₅ and asthma. The DMRs annotated to LINC00602, ABCA7 and CLCA1 were associated with childhood lung function but not with other outcomes.

**Identified DMRs and Gene Expression**

Of the 59 identified DMRs, 32 (54%) DMRs at birth were associated with gene expression at age 4 years, and 18 (31%) DMRs with gene expression in adulthood (Supplementary Appendix Table 7). The DMR annotated to HOXA5 was associated with differential expression of several genes of the HOX-family (Table 2). The DMRs annotated to PER3, VENTX, NUDT12 and TCL1A were associated with differential expression of their respective genes. The DMRs annotated to PAOX, LINC00602, ABCA7, CLCA1 and PTPRN2 were not associated with expression of their corresponding genes. Genes annotated to 28 (47%) of all identified DMRs were expressed in adult lung tissue, including the top significant DMRs annotated to PAOX, ABCA7, CLCA1, VENTX and NUDT12 (Supplementary Appendix Table 8).

**Identified DMRs and related biological processes**

Of all 59 identified DMRs, 43 were annotated to genes not previously associated with lung function or respiratory morbidity (Supplementary Appendix Table 7). Of the genes annotated to the top ten significant DMRs, HOXA5, CLCA1, TCL1A and NUDT12 were previously associated with respiratory development including alveogenesis, respiratory diseases and cellular immunity (Table 2). Genes related to the identified DMRs, including HOXA5, PER3, CLCA1, NUDT12 and PTPRN2, were located in pathways related to regionalization, DNA- and RNA-regulation and embryonic development (Supplementary Appendix Tables 9). The genes HLA-DRB4 and HLA-DRB5 were enriched in processes including asthma. These genes were associated with the DMR located at chr6:32,305,068-32,305,146, which was related with childhood and adulthood FEV₁/FVC and COPD.

Of the top ten significant DMRs, the DMRs annotated to HOXA5, CLCA1 and TCL1A contained CTCF-binding sites (Supplementary Figure 2). The DMRs annotated to HOXA5, PAOX, PER3 and
NUDT12 were located in promotor regions of their respective genes. The DMR annotated to ABCA7 was located in a CpG-island.

DISCUSSION

We identified 59 DMRs in neonatal cord blood associated with childhood lung function. Eighteen (31%) of all identified DMRs were also associated with childhood asthma, 11 (19%) and 9 (15%) with adolescent and adult lung function, respectively, and 9 (15%) with COPD. Differential gene expression was observed for 32 (54%) DMRs in childhood and 18 (31%) DMRs in adulthood. Multiple genes related to the identified DMRs have previously been associated with respiratory development and morbidity, and many identified DMRs were located within known regulatory elements for gene expression.

Reduced lung function in childhood is associated with reduced lung function and increased risks of asthma and COPD many decades later (10, 29). Pathways of environmental exposures in early life, such as tobacco smoke exposure or lack of breastfeeding, that affect lung development and risk of chronic obstructive respiratory diseases in later life might be modified by genetic susceptibility. Vice versa, genetic susceptibility could partly explain the difference in adverse effects of early environmental exposures on the risk of chronic obstructive respiratory diseases in later life. Identified genetic variants associated with childhood asthma in large-scale GWA studies only account for up to 7.5% of the explained variance (30). Epigenetic mechanisms could link environmental exposures with the unexplained heritability for childhood asthma (31, 32). Studies that examined associations of DNA-methylation with lung function, asthma or COPD are scarce, limited to candidate genes or high-risk population and lack replication. An epigenome-wide study among 97 asthmatics and 97 healthy children aged 6-12 years identified 81 DMRs associated with asthma, of which 16 DMRs were also associated with FEV₁ (8). Of these 81 DMRs, 19 were located within 500kb of our identified DMRs and may affect the same genes. Another epigenome-wide study in 1,454 adults identified 349 CpGs associated with COPD (33). Four annotated genes in this adult study (CBFA2T3, PADI4, LST1, KCNQ1) were replicated in our study of children. Multiple genes associated with the identified DMRs have previously been related with asthma and COPD in genome-wide association studies. TCL1A has been identified as asthma-susceptibility gene (34). Nine (15%) of the 59 DMRs we identified were associated with adult lung function, and annotated to, or associated with differential expression of 11
genes. Nine of these genes were previously linked with pulmonary structures (CROCC, CLCA1), immunity (MARCKS, FOXD2, MEF2C, CMBL, CLCA1), asthma (MARCKS, HCG23, CLCA1), COPD (MARCKS, TBX5, CLCA1), and smoking behavior in COPD (NUDT12) (28). This suggests that genes associated with respiratory diseases could be influenced by differential DNA-methylation from early life onwards.

We explored the biological processes of the top significant DMRs for development of respiratory morbidity (28). The DMR annotated to HOXA5 was associated with childhood and adolescent FEV1, COPD and differential expression of HOXA1, HOXA4 and HOXA7. One DMR associated with childhood FEV1/FVC was annotated to VENTX, which is a member of the HOX-gene family. The DMR annotated to LINC00602 (Long Intergenic Non-Protein Coding RNA (IncRNA) 602) was linked to childhood FEV1. LncRNAs influence gene-specific epigenetic regulation and interact amongst others with the transcription of HOX-genes. HOX-genes are critical for segmental fetal development, and especially HOXA5 is required for embryonic respiratory tract morphogenesis (35, 36). The DMR annotated to PAOX was linked to childhood and adolescence FEV1 and COPD. PAOX is involved in the regulation of intracellular polyamine, which is essential for protein synthesis. The DMR linked to ABCA7 was associated with FEV1 in childhood and adolescence. ABCA7 is involved in the lipid homeostasis in the cellular immune system and is essential for phagocytosis of apoptotic cells by alveolar macrophages (37). PER3, annotated to a DMR associated with FEV1 in children and adolescents, is a key element in the endogenous circadian rhythm. The DMR linked to CLCA1 was associated with childhood FEV1/FVC and FEF75, and expressed in adult lung tissue. CLCA1 affects IL-13 driven mucus production in human airway epithelial cells and is associated with asthma and COPD (38-40). NUDT12, annotated to a DMR associated with childhood FEV1/FVC and COPD, is involved in intracellular biochemical reactions. NUDT12 is associated with smoking behavior in COPD (41).

PTPRN2, annotated to a DMR associated with childhood FEF75 and asthma is member of a gene family regulating cell growth and differentiation, and is involved in vesicle-mediated secretory processes. DNA-methylation of PTPRN2 differentiates between lung cancer, pulmonary fibrosis and COPD (42). TCL1A, annotated to a DMR associated with childhood FEV1/FVC, FEF75 and asthma, is specific to developing lymphocytes when expressed and is associated with asthma (34). Thus, many
of the genes annotated to the top significant DMRs are involved in respiratory development, cellular immunity and respiratory morbidity, which warrant further studies.

This is the largest study to date evaluating the associations of newborn epigenome-wide DNA-methylation with lung function and respiratory disease in children and adults, and it provides new insights into the epigenetic changes in fetal life that increase the risk of life-time respiratory morbidity.

To the best of our knowledge, no other cohort studies with data on cord blood DNA-methylation and childhood lung function are available. We aimed to strengthen our results using public databases on gene expression and biological pathways, which added additional support for the observed associations. Ideally, the presence of identified DMRs would be replicated in lung cells. However, in cohort studies, this is ethically not done. It is unknown whether nasal cells, which are easier to acquire, have a high enough correlation in DNA-methylation with lung tissue. Therefore, further studies should aim to examine whether DNA-methylation in nasal cells is a good proxy for lung tissue and data on DNA-methylation and phenotypes should be shared in consortia to increase the statistical power to identify DNA-methylation patterns affecting respiratory health across the life course. These results cannot currently be used as predictors of disease in individuals, but are important from an etiological perspective. Genes associated with 29 of the identified DMRs, including HOXA5, PAOX, VENTX, PTPRN2 and TCL1A, have been reported to be differentially methylated in relation with maternal smoking during pregnancy (12). Genes related with four identified DMRs associated with childhood lung function were differentially methylated in association with maternal folate levels during pregnancy (43). This supports the hypothesis that adverse exposures in fetal life may impact DNA-methylation at birth, gene expression and subsequent respiratory development in the child, predisposing individuals for obstructive airway diseases. Further experimental or Mendelian randomization studies on the identified DMRs and associated genes might inform strategies in early life to improve lung function and lower the lifetime risk of obstructive respiratory diseases.

Some limitations should be discussed. We measured DNA-methylation in blood because it is easily accessible in large cohort studies. Blood DNA-methylation does not necessarily reflect lung epithelial DNA-methylation. However, asthma and COPD have systemic manifestations, characterized by increased inflammatory blood markers (44, 45). Although the analyses were adjusted for estimated cell counts, we cannot rule out residual confounding due to alterations in cell type distribution.

Recently, two new reference sets for cell type adjustment in cord blood samples were published (46,
These reference sets are currently being validated, and future studies will shed light on the differences between reference panels. In our secondary analyses, we assessed whether the identified DMRs were associated with lung function measured in adolescence and adulthood, similar to the associations identified between cord blood DNA methylation and childhood lung function. DNA methylation patterns and expression of genes vary depending on the developmental stage, and these changes could be non-linear (48). We were not able to assess the stability of DNA methylation in the identified DMRs in the same individuals from birth to adulthood. In a recent study addressing DNA methylation changes in early life, significantly reduced or increased methylation of single CpGs between ages 0 to 4 years and 4 to 8 years occurred in <4% of all CpGs, suggesting only a minor global change in DNA methylation in childhood (49). Longitudinal changes in DNA methylation from early life until adulthood in relation to respiratory morbidity have not been studied yet. We observed similar associations between DNA methylation of a specific genetic region with lung function, asthma or COPD observed in early life and adulthood, and this strengthens our hypothesis that specific DNA methylation patterns affect respiratory health across the life course. Further studies in longitudinal cohort studies with repeated measures of DNA methylation from birth into adulthood in the same individuals are needed to confirm this.

We presented our primary results including all probes, and provided results of analyses excluding potentially problematic probes, namely those with SNPs, insertions or deletions, repeats and bisulfite induced reduced genomic complexity. These underlying variants may affect probe binding and as such, affect the identified associations. In our stringent sensitivity analyses, we observed similar size and direction of the effect estimates in 54% of the identified DMRs associated with childhood lung function. However, the true exact impact of potentially problematic probes on the measurement of DNA methylation in our analyses remains unknown (25, 50). Discarding probes a priori may discard information. Therefore, we present all results of the main and sensitivity analysis.

Genetic variation as opposed to environmental variation might be influencing the DMRs associated with respiratory health. A recent study in two ethnic diverse adult cohorts in 557 subjects showed that DNA methylation of airway epithelium plays a central role in mediating the effects of SNPs and gene expression on asthma risk and its clinical course (51). Another study in 115 subjects participating in an adult cohort study reported a potential mediating effect of DNA methylation of single CpGs on the associations between SNPs located at chromosomal locus 17q21 and asthma (52). The study
identified 6 CpGs associated with gene expression of ORMDL3 and GSDMB. The authors did not assess the associations of DNA methylation with asthma or lung function. We did not identify any DMR located near the 17q21 locus. This could be explained by the young age of our study subjects or differences in main respiratory outcomes measurements. The previous studies stepwise assessed the effect of DNA methylation with the gene expression, and associations of SNPs with asthma, whereas our study focused on the direct associations between DNA methylation and respiratory outcomes. Further research is needed to assess this potential biological pathway.

Several identified DMRs were associated with gene expression other than the nearest and therefore annotated gene, which limits the potential biological effect of the annotated genes. The genomic inflation factor for the primary analyses ranged from 1.07 to 1.21 (Supplementary Figure 1).

Recently, it was shown that the genomic inflation factor provides an invalid estimate of test-statistic inflation when the outcome of interest is associated with many, small genetic effects (53). Furthermore, estimating the inflation factor using the genomic inflation factor results both in an overestimation of the actual inflation and in imprecise estimates contributing to the previously unexplained, high variability across studies. This might explain the genomic inflation in our analyses. The statistical steps in Comb-P limit the final number of DMRs identified, and genomic inflation in the identification of DMRs could not be tested. Further studies are needed to develop statistical tools dealing with genomic inflation in epigenome-wide studies.

There were no cohort studies available for replication analyses. We included all available cohorts in the meta-analysis to obtain the largest possible power to detect new associations. A previously published study has shown that in (epi)genome-wide association analyses a meta-analysis of all participating cohorts rather than a split sample analysis with a properly selected level of (epi)genome-wide significance is the most powerful approach to identify new associations (54). The high between-study homogeneity observed for the vast majority of CpGs in our meta-analysis (Supplementary Tables 4a-c) also provides support for the stability of the reported associations. Nevertheless, confirmatory studies are needed.

In conclusion, we identified 59 DMRs in cord blood that were associated with childhood lung function. Multiple DMRs were additionally related with childhood asthma, adolescent and adult lung function, or adult COPD. Also, multiple DMRs were associated with differential gene expression of genes involved in embryonic and respiratory tract development, or were located in regulatory elements for gene
expression. These findings suggest that epigenetic changes during fetal life might modify the risk of respiratory diseases across the life course.
REFERENCES


Table 1. Characteristics of Cohorts and Their Participants.

<table>
<thead>
<tr>
<th>Primary analyses</th>
<th>No. of participants</th>
<th>Type of blood sample for DNA-methylation</th>
<th>No. of available CpGs</th>
<th>No. of subjects with expression data</th>
<th>Age at lung function measurement</th>
<th>Asthma Cases</th>
<th>Controls</th>
<th>COPD Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSPAC (UK)</td>
<td>654</td>
<td>Cord blood</td>
<td>471,193</td>
<td>n.a.</td>
<td>8·6 (0·2)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Generation R (NL)</td>
<td>643</td>
<td>Cord blood</td>
<td>436,013</td>
<td>n.a.</td>
<td>9·8 (0·3)</td>
<td>47</td>
<td>663</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>INMA (Spain)</td>
<td>140</td>
<td>Cord blood</td>
<td>439,306</td>
<td>107</td>
<td>6·9 (0·3)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>CHS (USA)</td>
<td>75</td>
<td>Cord blood</td>
<td>383,857</td>
<td>n.a.</td>
<td>13·3 (0·6)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Project Viva (USA)</td>
<td>176</td>
<td>Cord blood</td>
<td>470,870</td>
<td>n.a.</td>
<td>7·9 (0·7)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
</tbody>
</table>

Secondary analyses
<table>
<thead>
<tr>
<th>Study</th>
<th>Sample Size</th>
<th>Sample Type</th>
<th>n.a.</th>
<th>n.a.</th>
<th>Z-score (SD)</th>
<th>n.a.</th>
<th>n.a.</th>
<th>n.a.</th>
<th>n.a.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALSPAC (UK)</td>
<td>542</td>
<td>Cord blood</td>
<td>n.a.</td>
<td>n.a.</td>
<td>15.4 (0.2)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>Rotterdam Study – I (NL)</td>
<td>488</td>
<td>Peripheral blood</td>
<td>n.a.</td>
<td>488</td>
<td>64.0 (6.3)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>63</td>
<td>425</td>
</tr>
<tr>
<td>Rotterdam Study – II (NL)</td>
<td>703</td>
<td>Peripheral blood</td>
<td>n.a.</td>
<td>703</td>
<td>67.5 (5.9)</td>
<td>n.a.</td>
<td>n.a.</td>
<td>92</td>
<td>611</td>
</tr>
</tbody>
</table>

Lung function was obtained by spirometry and sex-, age-, height- and ethnicity-adjusted Z-scores were calculated. FEV₁: Forced Expiratory Volume in 1 second; FVC: Forced Vital Capacity; FEF₇₅: Forced Expiratory Flow at 75% of FVC; n.a.: not applicable. UK: United Kingdom. NL: the Netherlands. USA: United States of America.
Table 2. Associations of the Top Ten Significant Identified Differentially Methylated Regions with Gene Expression and Related Respiratory Outcomes.

<table>
<thead>
<tr>
<th>Molecular location of the</th>
<th>Lung function</th>
<th>Annotated gene*</th>
<th>Expressed gene†</th>
<th>Gene expression in children‡</th>
<th>Gene expression in adults§</th>
<th>Previously associated with lung development or respiratory morbidity‖</th>
</tr>
</thead>
<tbody>
<tr>
<td>differently methylated region (Chromosome: start – end)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr1: 7,887,199 - 7,887,561</td>
<td>FEV₁</td>
<td>PER3</td>
<td>PER3, RP3-467l1.4, RNA5SP23, RP4-726F1.1</td>
<td>↓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>RP11-431K24.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr1: 86,968,087 – 86,968,544</td>
<td>FEV₁/FVC</td>
<td>CLCA1</td>
<td>no expression</td>
<td>-</td>
<td>-</td>
<td>Lung development, asthma, COPD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr5: 102,898,223 - 102,898,734</td>
<td>FEV₁/FVC</td>
<td>NUDT12</td>
<td>NUDT12</td>
<td>↓</td>
<td>-</td>
<td>Smoking behavior in COPD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CMBL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr6: 166,418,799 – 166,419,139</td>
<td>FEV₁</td>
<td>LINC00602</td>
<td>no expression</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>chr7: 27,183,133 - 27,184,854</td>
<td>FEV₁</td>
<td>HOXA5</td>
<td>HOXA1, HOTTIP</td>
<td>↓</td>
<td>↓</td>
<td>Lung development, FEV₁, FEV₁/FVC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr7: 158,045,980 – 158,046,359</td>
<td>FEF₇₅</td>
<td>PTPRN2</td>
<td>no expression</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>chr10: 135,202,522 – 135,203,201</td>
<td>FEV₁</td>
<td>PAOX</td>
<td>no expression</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr10: 135,051,149 –</td>
<td>FEV₁/FVC</td>
<td>VENTX</td>
<td>TUBGCP2, RP11-122K13.12</td>
<td>↓</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Chromosome</td>
<td>Start</td>
<td>End</td>
<td>Gene 1</td>
<td>Gene 2</td>
<td>Expression</td>
<td>Direction 1</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>-----</td>
<td>--------</td>
<td>--------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>14</td>
<td>96,180,406</td>
<td>96,181,045</td>
<td>VENTX</td>
<td>ECHS1</td>
<td>↑</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SPRN</td>
<td></td>
<td>↑</td>
<td>↓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ZNF511</td>
<td></td>
<td>-</td>
<td>↑</td>
</tr>
<tr>
<td>19</td>
<td>1,063,624</td>
<td>1,064,219</td>
<td>FEF75</td>
<td>TCL1A</td>
<td>↓</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TCL1A, CCDC85C</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Results present identified differentially methylated regions (DMRs) from association-analyses of DNA-methylation at birth with childhood Forced Expiratory Volume in 1 second (FEV₁), FEV₁/FV (FVC) and Forced Expiratory Flow at 75% of FVC (FEF₇₅). *DMRs were annotated to their nearest gene. †: Identified DMRs at birth were associated with gene expression in: ‡: childhood (INMA; mean age 4 years) and §: adulthood (the Rotterdam Study, mean age 66 years). Gene expressions levels were assessed limited to 250kb up- and downstream of the outer border of the DMR. Directions of associations are marked ↓ if a higher methylation of the DMR was associated with a decreased expression of the specific gene, ↑ if a higher methylation of the DMR was associated with an increased expression of the specific gene, and - if no direction of associations were observed. ‖: Associations of expressed genes with lung development and respiratory morbidity were explored in previous published studies the OMIM database and UniProt.
Figure 1. Overall Study Design.

Epigenome-wide meta-analyses were performed to identify methylated CpGs associated with lung function in children using data from 1,689 children participating in ALSPAC, Generation R, INMA, CHS and Project Viva. Identified differentially methylated regions (DMRs) were annotated to their nearest gene using PAVIS. Next, we examined if identified DMRs were associated with asthma in children participating in Generation R, lung function in adolescents and adults participating in the ALSPAC or Rotterdam Study, or COPD in adults participating in Rotterdam Study, and with gene expression levels in children participating in INMA, adults in Rotterdam Study, and the GTEx-database. We further explored biological processes and associations with lung development and respiratory morbidity using publicly available resources (DAVID, GeneMania, OMIM and UniProt). N = x: number of participants included for the analysis. N = x/x: number of cases / total number of participants included in the analysis.
Figure 2. Manhattan Plots of Associations of CpGs located in Differentially Methylated Regions with Childhood Lung Function Outcomes.

Green dots represent p-values from associations of CpGs located in differentially methylated regions (DMRs) at birth with childhood Forced Expiratory Volume in 1 second (FEV₁), FEV₁/Forced Vital Capacity (FVC) and Forced Expiratory Flow at 75% of FVC (FEF₇₅). P-values of DMRs ranged from 3.05E-14 to 0.031, and details are provided in Supplementary Appendix Table 3. Nearest annotated genes of DMRs are provided. The genes annotated to the top ten significant DMRs associated with childhood lung function are written in red. Single CpGs are presented as red and blue dots, corrected for correlations with neighboring CpGs.
Results present identified differentially methylated regions (DMRs) from association-analyses of DNA-methylation at birth with childhood Forced Expiratory Volume in 1 second (FEV$_1$), FEV$_1$/ Forced Vital Capacity (FVC) and Forced Expiratory Flow at 75% of FVC (FEF$_{75}$), their location, and their direction of association with childhood lung function, childhood asthma, adolescent lung function, and adult lung function and COPD. Molecular locations of the top ten significant DMRs are presented in bold. Identified DMRs associated with childhood lung function and other respiratory outcomes are marked in grey, and if not associated with respiratory outcomes in white. Directions of associations are marked ↓ if a higher mean methylation of the DMRs was associated with a lower z-score for lung function or lower risk of asthma or COPD, and marked ↑ if a higher mean methylation of the DMRs was associated with a higher z-score of lung function or higher risk of asthma or COPD. Red colored arrows represent disadvantageous effect estimates (lower lung function, increased risk of asthma or COPD), and green colored arrows beneficial effect estimates (higher lung function, lower risk of asthma or COPD).