

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

Herman T. den Dekker\*, Kimberley Burrows\*, Janine F. Felix\*, Lucas A. Salas, Ivana Nedeljković, Jin Yao, Sheryl L. Rifas-Shiman, Carlos Ruiz-Arenas, N. Amin, Mariona Bustamante, Dawn L. DeMeo, A. John Henderson, Caitlin G. Howe, Marie-France Hivert, M. Arfan Ikram, Johan C. de Jongste, Lies Lahousse, Pooja R. Mandaviya, Joyce B. van Meurs, Mariona Pinart, Gemma C. Sharp, Lisette Stolk, André G. Uitterlinden, Josep M. Anto, Augusto A. Litonjua, Carrie V. Breton, Guy G. Brusselle, Jordi Sunyer, George Davey Smith, Caroline L. Relton<sup>#</sup>, Vincent W.V. Jaddoe<sup>#</sup>, Liesbeth Duijts<sup>#</sup>

*\*first authors contributed equally*

*<sup>#</sup>last authors contributed equally*

*This chapter is published in European Respiratory Journal, February 2019.*

*The supplemental material for this paper is available at:*

*<https://doi.org/10.1183/13993003.01795-2018>*

## ABSTRACT

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.

We meta-analysed 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 (FEV<sub>1</sub>), FEV<sub>1</sub>/Forced Vital Capacity (FVC), and Forced Expiratory Flow at 75% of FVC (FEF<sub>75</sub>) 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. 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.

Our findings suggest that the epigenetic status of the new-born 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.<sup>1</sup> Both diseases are characterized by airway obstruction, as indicated by a reduced Forced Expiratory Volume in 1 second (FEV<sub>1</sub>), FEV<sub>1</sub> to Forced Vital Capacity (FVC) ratio, and Forced Expiratory Flow at 75% of FVC (FEF<sub>75</sub>).<sup>2</sup> Childhood lung function predicts lung function and risks of asthma and COPD in later life.<sup>3</sup> An accumulating body of evidence suggests that asthma and COPD have at least part of their origins in fetal life.<sup>4,5</sup> 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.<sup>6</sup> Furthermore, adverse fetal exposures, such as maternal smoking and suboptimal diet, increase the risk of asthma and COPD.<sup>5</sup> The pathways linking genetic predisposition and environmental exposures in fetal life with life course respiratory disease may include epigenetic changes, including DNA-methylation.<sup>5</sup> Epigenetic changes are influenced by environmental exposures and could exert population effects much more rapidly than genetic mutations.<sup>6</sup> 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.<sup>5</sup> DNA-methylation may affect fetal development through effects on gene transcription and expression.<sup>7</sup> Recent studies assessing the associations between DNA-methylation and childhood respiratory health are mainly limited to candidate genes and small sample sizes.<sup>3,8</sup> 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.<sup>9</sup> 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.<sup>10,11</sup>

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 FEV<sub>1</sub>, FEV<sub>1</sub>/FVC

and FEF<sub>75</sub>. 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

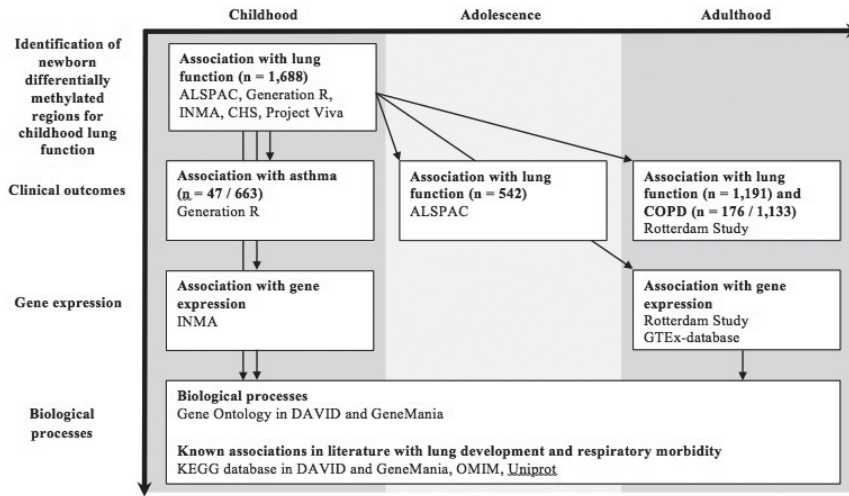
### Study design and data sources

We included population-based cohort studies participating in PACE consortium with data on epigenome-wide DNA-methylation at birth and lung function in childhood.<sup>12</sup> 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.<sup>13,14</sup>

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.<sup>15-17</sup> 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**.



**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.

## Respiratory outcome assessment

Lung function measures comprised pre-bronchodilator  $FEV_1$ ,  $FEV_1/FVC$  and  $FEF_{75}$ , which were converted into sex-, age-, height- and ethnicity-adjusted z-scores.<sup>18</sup> Physician-diagnosed asthma was obtained by questions adapted from the International Study on Asthma and Allergy in Childhood.<sup>19</sup> COPD was defined as pre-bronchodilator  $FEV_1/FVC < 0.70$  in the absence of asthma, or a doctor diagnosis.<sup>20</sup>

## 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  $FEV_1$ ,  $FEV_1/FVC$  and

FEF<sub>75</sub>. Analyses were adjusted for maternal age, socio-economic status, smoking during pregnancy, parity, asthma or atopy, technical covariates and estimated cell counts.<sup>21</sup> 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.<sup>9,22</sup> Regions were defined as a minimum of 2 probes within a window size of 500 bases with an FDR-threshold <0.05.<sup>9</sup> 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.<sup>23</sup> 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).<sup>24</sup> 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.<sup>25</sup>

### ***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, FEV<sub>1</sub>, FEV<sub>1</sub>/FVC and FEF<sub>75</sub> 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.<sup>9,26</sup> 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  $\pm 250$ kb 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.<sup>26, 27</sup> Additionally, we explored whether the annotated and differentially expressed genes were expressed in human lung specimens of the Genotype-Tissue Expression (GTEx) database.<sup>15</sup>

### ***Exploration biological processes***

The Gene Ontology database implemented in DAVID and Genemania was used to examine gene function in biological processes.<sup>16, 17</sup> 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.<sup>28</sup> 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 FEV<sub>1</sub>, FEV<sub>1</sub>/FVC and FEF<sub>75</sub>, 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  $I^2 < 50$ : FEV<sub>1</sub> 140/163 (86%), FEV<sub>1</sub>/FVC 82/89 (92%) and FEF<sub>75</sub> 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 FEV<sub>1</sub> 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 FEV<sub>1</sub>/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 FEF<sub>75</sub> 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

**Table 1.** Characteristics of Cohorts and Their Participants.

|                           | No. of participants | Type of blood sample for DNA-methylation | No. of available CpGs | No. of subjects with expression data | Age at lung function measurement | Asthma |          | COPD  |          |
|---------------------------|---------------------|------------------------------------------|-----------------------|--------------------------------------|----------------------------------|--------|----------|-------|----------|
|                           |                     |                                          |                       |                                      | Years (SD)                       | Cases  | Controls | Cases | Controls |
|                           |                     |                                          |                       |                                      |                                  |        |          |       |          |
| Primary analyses          |                     |                                          |                       |                                      |                                  |        |          |       |          |
| ALSPAC (UK)               | 654                 | Cord blood                               | 471,193               | NA                                   | 8·6 (0·2)                        | NA     | NA       | NA    | NA       |
| Generation R (NL)         | 643                 | Cord blood                               | 436,013               | NA                                   | 9·8 (0·3)                        | 47     | 663      | NA    | NA       |
| INMA (Spain)              | 140                 | Cord blood                               | 439,306               | 107                                  | 6·9 (0·3)                        | NA     | NA       | NA    | NA       |
| CHS (USA)                 | 75                  | Cord blood                               | 383,857               | NA                                   | 13·3 (0·6)                       | NA     | NA       | NA    | NA       |
| Project Viva (USA)        | 176                 | Cord blood                               | 470,870               | NA                                   | 7·9 (0·7)                        | NA     | NA       | NA    | NA       |
| Secondary analyses        |                     |                                          |                       |                                      |                                  |        |          |       |          |
| ALSPAC (UK)               | 542                 | Cord blood                               | NA                    | NA                                   | 15·4 (0·2)                       | NA     | NA       | NA    | NA       |
| Rotterdam Study – I (NL)  | 488                 | Peripheral blood                         | NA                    | 488                                  | 64·0 (6·3)                       | NA     | NA       | 63    | 425      |
| Rotterdam Study – II (NL) | 703                 | Peripheral blood                         | NA                    | 703                                  | 67·5 (5·9)                       | NA     | NA       | 92    | 611      |

Lung function was obtained by spirometry and sex-, age-, height- and ethnicity-adjusted Z-scores were calculated. FEV<sub>1</sub>: Forced Expiratory Volume in 1 second; FVC: Forced Vital Capacity; FEF<sub>75</sub>: Forced Expiratory Flow at 75% of FVC; NA: not applicable. UK: United Kingdom. NL: the Netherlands. USA: United States of America.

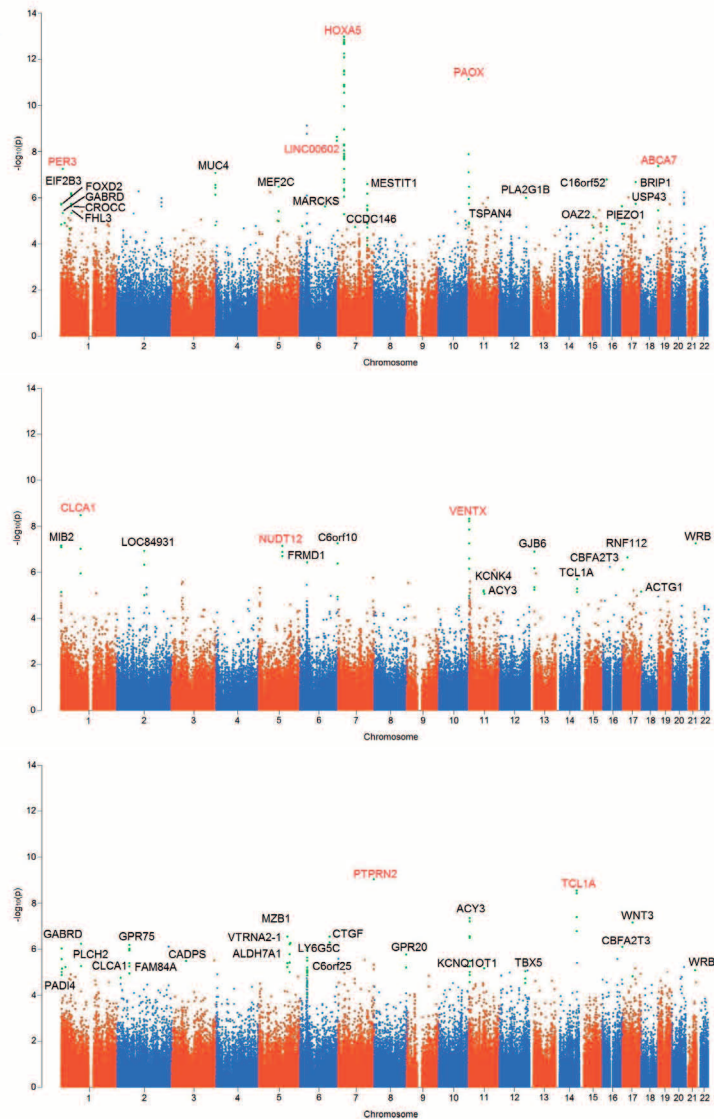
of the 59 previously identified DMRs still contained  $\geq 2$  CpGs (**supplementary appendix Table 4**). Of these 41 DMRs, 54% (n=22) remained to be associated with childhood lung function (**Supplementary appendix Table 5**).

### Identified DMRs and lung function and respiratory diseases across the life course

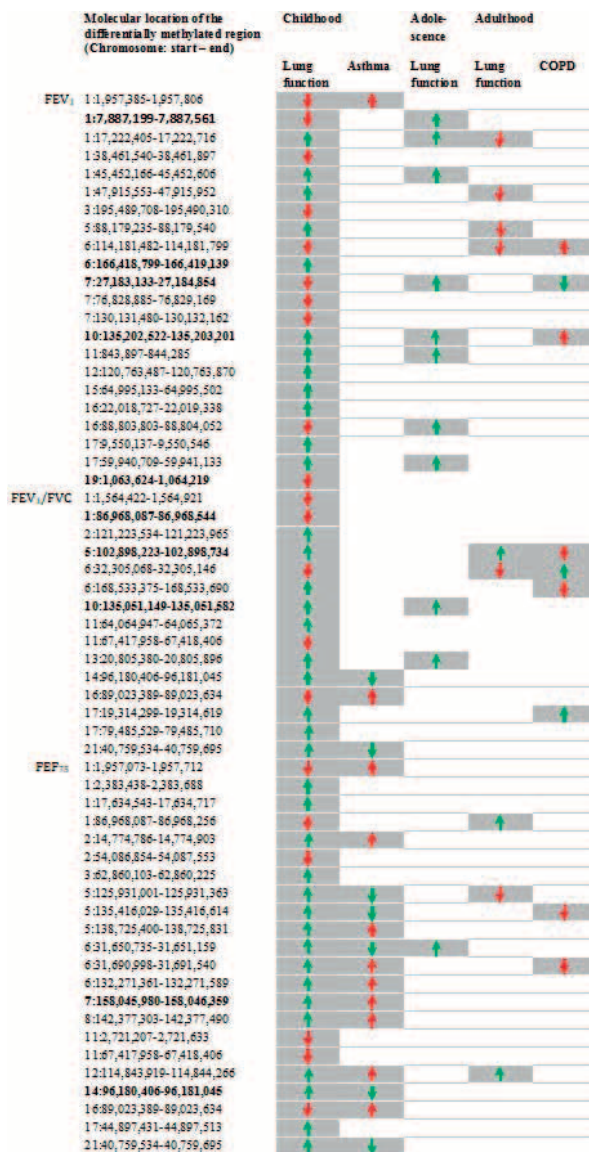
Of all 59 identified DMRs related with childhood lung function, 18 (31%) were associated with childhood asthma (**Figure 3, Supplementary appendix Table 6**).

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 FEV<sub>1</sub> and COPD, but not with childhood asthma or adult lung function. The DMRs annotated to *PER3* and *VENTX* were associated with childhood and adolescence FEV<sub>1</sub> and FEV<sub>1</sub>/FVC, respectively. The DMR annotated to *NUDT12* was associated with childhood FEV<sub>1</sub>/FVC and COPD. The DMRs annotated to *PTPRN2* and *TCL1A* were associated with childhood FEF<sub>75</sub> and asthma. The DMRs annotated to *LINC00602*, *ABCA7* and *CLCA1* were associated with childhood lung function but not with other outcomes.





**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 (A) Forced Expiratory Volume in 1 second (FEV<sub>1</sub>), (B) FEV<sub>1</sub>/ Forced Vital Capacity (FVC) and (C) Forced Expiratory Flow at 75% of FVC (FEF<sub>75</sub>). P-values of DMRs ranged from  $3.05 \times 10^{-14}$  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.



**Figure 3.** Identified Differentially Methylated Regions and Their Location, and Direction of Associations with Childhood Lung Function, Childhood Asthma, Adolescent Lung Function, and Adult Lung Function and COPD. Results present identified differentially methylated regions (DMRs) from association-analyses of DNA-methylation at birth with childhood Forced Expiratory Volume in 1 second (FEV<sub>1</sub>), FEV<sub>1</sub>/ Forced Vital Capacity (FVC) and Forced Expiratory Flow at 75% of FVC (FEF<sub>75</sub>), 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).

## 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 FEV<sub>1</sub>/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

**Table 2.** Associations of the top ten significant identified DMRs with gene expression and related respiratory outcomes.

| Molecular location of the DMR (Chromosome: start - end) | Lung function         | Annotated gene*  | Expressed gene†                                                                                       | Gene expression in children‡ | Gene expression in adults§ | Previously associated with lung development or respiratory morbidity                         |
|---------------------------------------------------------|-----------------------|------------------|-------------------------------------------------------------------------------------------------------|------------------------------|----------------------------|----------------------------------------------------------------------------------------------|
| chr1: 7,887,199 - 7,887,561                             | FEV <sub>1</sub>      | <i>PER3</i>      | <i>PER3</i> , <i>RP3-46711.4</i> , <i>RNA5SP23</i> , <i>RP4-726F1.1</i><br><i>RP11-431K24.1</i>       | ↓<br>-                       | -<br>↑                     |                                                                                              |
| chr1: 86,968,087 - 86,968,544                           | FEV <sub>1</sub> /FVC | <i>CLCA1</i>     | no expression                                                                                         | -                            | -                          | Lung development, asthma, COPD                                                               |
| chr5: 102,898,223 - 102,898,734                         | FEV <sub>1</sub> /FVC | <i>NUDT12</i>    | <i>NUDT12</i><br><i>CMBL</i>                                                                          | ↓<br>↓                       | -<br>-                     | Smoking behavior in COPD                                                                     |
| chr6: 166,418,799 - 166,419,139                         | FEV <sub>1</sub>      | <i>LINC00602</i> | no expression                                                                                         | -                            | -                          |                                                                                              |
| chr7: 27,183,133 - 27,184,854                           | FEV <sub>1</sub>      | <i>HOXA5</i>     | <i>HOXA1</i> , <i>HOTTIP</i><br><i>EVX1</i> , <i>HOXA4</i> , <i>HOXA7</i>                             | ↓<br>↓                       | ↓<br>-                     | Lung development, FEV <sub>1</sub> , FEV <sub>1</sub> /FVC<br>Lung development, asthma, COPD |
| chr7: 158,045,980 - 158,046,359                         | FEF <sub>75</sub>     | <i>PTPRN2</i>    | no expression                                                                                         | -                            | -                          |                                                                                              |
| chr10: 135,202,522 - 135,203,201                        | FEV <sub>1</sub>      | <i>PAOX</i>      | no expression                                                                                         | -                            | -                          |                                                                                              |
| chr10: 135,051,149 - 135,051,582                        | FEV <sub>1</sub> /FVC | <i>VENTX</i>     | <i>TUBGCP2</i> , <i>RP11-122K13.12</i><br><i>VENTX</i> , <i>ECHS1</i><br><i>SPRN</i><br><i>ZNF511</i> | ↓<br>↑<br>↑<br>-             | -<br>-<br>↓<br>↑           |                                                                                              |
| chr14: 96,180,406 - 96,181,045                          | FEF <sub>75</sub>     | <i>TCL1A</i>     | <i>TCL1A</i> , <i>CCDC85C</i>                                                                         | ↓                            | -                          | Asthma                                                                                       |
| chr19: 1,063,624 - 1,064,219                            | FEV <sub>1</sub>      | <i>ABCA7</i>     | no expression                                                                                         | -                            | -                          |                                                                                              |

Results present identified differentially methylated regions (DMRs) from association-analyses of DNA-methylation at birth with childhood Forced Expiratory Volume in 1 second (FEV<sub>1</sub>), FEV<sub>1</sub>/ Forced Vital Capacity (FVC) and Forced Expiratory Flow at 75% of FVC (FEF<sub>75</sub>). \*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.

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.<sup>10, 29</sup> 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.<sup>30</sup> Epigenetic mechanisms could link environmental exposures with the unexplained heritability for childhood asthma.<sup>31, 32</sup> 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 FEV<sub>1</sub>.<sup>8</sup> 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.<sup>33</sup> 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.<sup>34</sup> 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*).<sup>28</sup> 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.<sup>28</sup> The DMR annotated to *HOXA5* was associated with childhood and adolescent FEV<sub>1</sub>, COPD and differential expression of *HOXA1*, *HOXA4* and *HOXA7*. One DMR associated with childhood FEV<sub>1</sub>/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 (lncRNA) 602) was linked to childhood FEV<sub>1</sub>. 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.<sup>35,36</sup> The DMR annotated to *PAOX* was linked to childhood and adolescence FEV<sub>1</sub> 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 FEV<sub>1</sub> 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.<sup>37</sup> *PER3*, annotated to a DMR associated with FEV<sub>1</sub> in children and adolescents, is a key element in the endogenous circadian rhythm. The DMR linked to *CLCA1* was associated with childhood FEV<sub>1</sub>/FVC and FEF<sub>75</sub> 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.<sup>38-40</sup> *NUDT12*, annotated to a DMR associated with childhood FEV<sub>1</sub>/FVC and COPD, is involved in intracellular biochemical reactions. *NUDT12* is associated with smoking behavior in COPD.<sup>41</sup>

*PTPRN2*, annotated to a DMR associated with childhood FEF<sub>75</sub> 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.<sup>42</sup> *TCL1A*, annotated to a DMR associated with childhood FEV<sub>1</sub>/FVC, FEF<sub>75</sub> and asthma, is specific to developing lymphocytes when expressed and is associated with asthma.<sup>34</sup> 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.<sup>12</sup> Genes related with four identified DMRs associated with childhood lung function were differentially methylated in association with maternal folate levels during pregnancy.<sup>43</sup> 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.<sup>44, 45</sup> 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.<sup>46, 47</sup> 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.<sup>48</sup> 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.<sup>49</sup> 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, inser-

tions 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.<sup>25,50</sup> 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.<sup>51</sup> 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.<sup>52</sup> 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.<sup>53</sup> 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.<sup>54</sup> 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

1. Chronic obstructive pulmonary disease (COPD): World Health Organisation; 2015 [Available from: <http://www.who.int/mediacentre/factsheets/fs315/en/>].
2. McDonough JE, Yuan R, Suzuki M, Seyednejad N, Elliott WM, Sanchez PG, et al. Small-airway obstruction and emphysema in chronic obstructive pulmonary disease. *N Engl J Med*. 2011; 365(17):1567-75.
3. McGeachie MJ, Yates KP, Zhou X, Guo F, Sternberg AL, Van Natta ML, et al. Patterns of Growth and Decline in Lung Function in Persistent Childhood Asthma. *New Engl J Med*. 2016;374(19): 1842-52.
4. Martinez FD. Early-Life Origins of Chronic Obstructive Pulmonary Disease. *N Engl J Med*. 2016;375(9):871-8.
5. Duijts L, Reiss IK, Brusselle G, de Jongste JC. Early origins of chronic obstructive lung diseases across the life course. *Eur J Epidemiol*. 2014;29(12):871-85.
6. Begin P, Nadeau KC. Epigenetic regulation of asthma and allergic disease. *Allergy Asthma Clin Immunol*. 2014;10(1):27.
7. Krauss-Etschmann S, Meyer KF, Dehmel S, Hylkema MN. Inter- and transgenerational epigenetic inheritance: evidence in asthma and COPD? *Clin Epigenetics*. 2015;7:53.
8. Yang IV, Pedersen BS, Liu A, O'Connor GT, Teach SJ, Kattan M, et al. DNA methylation and childhood asthma in the inner city. *J Allergy Clin Immunol*. 2015;136(1):69-80.
9. Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. *Bioinformatics*. 2012;28(22):2986-8.
10. Postma DS, Rabe KF. The Asthma-COPD Overlap Syndrome. *N Engl J Med*. 2015;373(13):1241-9.
11. Fu JJ, McDonald VM, Baines KJ, Gibson PG. Airway IL-1 beta and Systemic Inflammation as Predictors of Future Exacerbation Risk in Asthma and COPD. *Chest*. 2015;148(3):618-29.
12. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet*. 2016;98(4):680-96.
13. Lister R, Pelizzola M, Downen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009; 462(7271):315-22.
14. Bock C. Analysing and interpreting DNA methylation data. *Nat Rev Genet*. 2012;13(10):705-19.
15. Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank*. 2015; 13(5):311-9.
16. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*. 2009;37(1):1-13.
17. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research*. 2010;38(Web Server issue):W214-20.
18. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. *Eur Respir J*. 2012;40(6):1324-43.
19. Asher MI, Keil U, Anderson HR, Beasley R, Crane J, Martinez F, et al. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *Eur Respir J*. 1995;8(3): 483-91.

20. Terzikhan N, Verhamme KM, Hofman A, Stricker BH, Brusselle GG, Lahousse L. Prevalence and incidence of COPD in smokers and non-smokers: the Rotterdam Study. *Eur J Epidemiol*. 2016.
21. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC bioinformatics*. 2012;13:86.
22. Peters TJ, Buckley MJ, Statham AL, Pidsley R, Samaras K, Lord RV, et al. De novo identification of differentially methylated regions in the human genome. *Epigenet Chromatin*. 2015;8.
23. Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, et al. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet*. 2006;38(12):1378-85.
24. Huang W, Loganantharaj R, Schroeder B, Fargo D, Li L. PAVIS: a tool for Peak Annotation and Visualization. *Bioinformatics*. 2013;29(23):3097-9.
25. Naeem H, Wong NC, Chatterton Z, Hong MK, Pedersen JS, Corcoran NM, et al. Reducing the risk of false discovery enabling identification of biologically significant genome-wide methylation status using the HumanMethylation450 array. *BMC Genomics*. 2014;15:51.
26. Benjamini Y, Hochberg Y. Controlling for False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society, Series B*. 1995;57:289-300.
27. Dai H, Leeder JS, Cui Y. A modified generalized Fisher method for combining probabilities from dependent tests. *Front Genet*. 2014;5:32.
28. UniProt C. UniProt: a hub for protein information. *Nucleic acids research*. 2015;43(Database issue):D204-12.
29. McGeachie MJ, Yates KP, Zhou X, Guo F, Sternberg AL, Van Natta ML, et al. Patterns of Growth and Decline in Lung Function in Persistent Childhood Asthma. *N Engl J Med*. 2016;374(19):1842-52.
30. Soler Artigas M, Loth DW, Wain LV, Gharib SA, Obeidat M, Tang W, et al. Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat Genet*. 2011;43(11):1082-90.
31. Cookson W, Moffatt M, Strachan DP. Genetic risks and childhood-onset asthma. *J Allergy Clin Immunol*. 2011;128(2):266-70; quiz 71-2.
32. English S, Pen I, Shea N, Uller T. The information value of non-genetic inheritance in plants and animals. *PLoS One*. 2015;10(1):e0116996.
33. Qiu W, Baccarelli A, Carey VJ, Boutaoui N, Bacherman H, Klanderman B, et al. Variable DNA methylation is associated with chronic obstructive pulmonary disease and lung function. *Am J Respir Crit Care Med*. 2012;185(4):373-81.
34. George BJ, Reif DM, Gallagher JE, Williams-DeVane CR, Heidenfelder BL, Hudgens EE, et al. Data-driven asthma endotypes defined from blood biomarker and gene expression data. *PLoS One*. 2015;10(2):e0117445.
35. Golpon HA, Geraci MW, Moore MD, Miller HL, Miller GJ, Tudor RM, et al. HOX genes in human lung: altered expression in primary pulmonary hypertension and emphysema. *Am J Pathol*. 2001;158(3):955-66.
36. Mandeville I, Aubin J, LeBlanc M, Lalancette-Hebert M, Janelle MF, Tremblay GM, et al. Impact of the loss of Hoxa5 function on lung alveogenesis. *American Journal of Pathology*. 2006; 169(4):1312-27.

37. Jehle AW, Gardai SJ, Li S, Linsel-Nitschke P, Morimoto K, Janssen WJ, et al. ATP-binding cassette transporter A7 enhances phagocytosis of apoptotic cells and associated ERK signaling in macrophages. *J Cell Biol.* 2006;174(4):547-56.
38. Poole A, Urbanek C, Eng C, Schageman J, Jacobson S, O'Connor BP, et al. Dissecting childhood asthma with nasal transcriptomics distinguishes subphenotypes of disease. *J Allergy Clin Immunol.* 2014;133(3):670-8 e12.
39. Hegab AE, Sakamoto T, Uchida Y, Nomura A, Ishii Y, Morishima Y, et al. CLCA1 gene polymorphisms in chronic obstructive pulmonary disease. *J Med Genet.* 2004;41(3):e27.
40. Alevy YG, Patel AC, Romero AG, Patel DA, Tucker J, Roswit WT, et al. IL-13-induced airway mucus production is attenuated by MAPK13 inhibition. *J Clin Invest.* 2012;122(12):4555-68.
41. Siedlinski M, Cho MH, Bakke P, Gulsvik A, Lomas DA, Anderson W, et al. Genome-wide association study of smoking behaviours in patients with COPD. *Thorax.* 2011;66(10):894-902.
42. Wielscher M, Vierlinger K, Kegler U, Ziesche R, Gsur A, Weinhausel A. Diagnostic Performance of Plasma DNA Methylation Profiles in Lung Cancer, Pulmonary Fibrosis and COPD. *EBioMedicine.* 2015;2(8):929-36.
43. Joubert BR, den Dekker HT, Felix JF, Bohlin J, Ligthart S, Beckett E, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat Commun.* 2016;7:10577.
44. Barnes PJ. Inflammatory mechanisms in patients with chronic obstructive pulmonary disease. *J Allergy Clin Immunol.* 2016;138(1):16-27.
45. Nadif R, Siroux V, Boudier A, le Moual N, Just J, Gormand F, et al. Blood granulocyte patterns as predictors of asthma phenotypes in adults from the EGEA study. *Eur Respir J.* 2016.
46. Bakulski KM, Feinberg JI, Andrews SV, Yang J, Brown S, S LM, et al. DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics.* 2016;11(5):354-62.
47. Gervin K, Page CM, Aass HC, Jansen MA, Fjeldstad HE, Andreassen BK, et al. Cell type specific DNA methylation in cord blood: a 450K-reference data set and cell count-based validation of estimated cell type composition. *Epigenetics.* 2016:0.
48. Feng L, Wang J, Cao B, Zhang Y, Wu B, Di X, et al. Gene expression profiling in human lung development: an abundant resource for lung adenocarcinoma prognosis. *PLoS One.* 2014; 9(8):e105639.
49. Xu CJ, Bonder MJ, Soderhall C, Bustamante M, Baiz N, Gehring U, et al. The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics.* 2017;18(1):25.
50. Lehne B, Drong AW, Loh M, Zhang WH, Scott WR, Tan ST, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies (vol 16, 37, 2015). *Genome Biol.* 2016;17.
51. Kothari PH, Qiu W, Croteau-Chonka DC, Martinez FD, Liu AH, Lemanske RF, Jr., et al. Role of local CpG DNA methylation in mediating the 17q21 asthma susceptibility gasdermin B (GSDMB)/ORMDL sphingolipid biosynthesis regulator 3 (ORMDL3) expression quantitative trait locus. *J Allergy Clin Immunol.* 2018;141(6):2282-6 e6.
52. Nicodemus-Johnson J, Myers RA, Sakabe NJ, Sobreira DR, Hogarth DK, Naureckas ET, et al. DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight.* 2016;1(20):e90151.
53. van IJterson M, van Zwet EW, Consortium B, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biol.* 2017;18(1):19.

54. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet.* 2006;38(2):209-13.