



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## ARCHIVIO ISTITUZIONALE DELLA RICERCA

### Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

DNA methylation in childhood asthma: An epigenome-wide meta-analysis

This is the submitted version (pre peer-review, preprint) of the following publication:

*Published Version:*

DNA methylation in childhood asthma: An epigenome-wide meta-analysis / Xu, Cheng-Jian; Söderhäll, Cilla; Bustamante, Mariona; Baiz, Nour; Gruzieva, Olena; Gehring, Ulrike; Mason, Dan; Chatzi, Leda; Basterrechea, Mikel; Llop, Sabrina; Torrent, Maties; Forastiere, Francesco; Fantini, Maria Pia; Carlsen, Karin C Lødrup; Haahtela, Tari; Morin, Andréanne; Kerkhof, Marjan; Merid, Simon Kebede; van Rijkom, Bianca; Jankipersadsing, Soesma A.; Bonder, Marc Jan; Ballereau, Stephane; Vermeulen, Cornelis J.; Aguirre, Geronimo; de Jongste, Johan C.; Smit, Henriette A.; Kumar, Ashish; Pershagen, Göran; Guerra, Stefano; Garcia-Añel, Judith; Gren, Darin; Buij, Lovisa; McEachan, Rosemary R.C.; Azad, Raf; Hovland, Vegard; Mowinckel, Petter; Alenius, Harri; Fyhrquist, Nanna; Lemonnier, Nathanaël; Pellet, Johann; Auffray, Charles; van der Vlies, Pieter; van Diemen, Cleo C.; Li, Yang; Wijmenga, Cisca; Netea, Mihai G.; Moffatt, Miriam S.; Gozdos, William O.C.M.; Anto, Josep M.; Bousquet, Jean; Laatikainen, Tiina; Laprise, Catherine; Carlsen, Kai-Håkon; Gori, Davide; Porta, Daniela; Iñiguez, Carmen; Bilbao, Jose Ramon; Kogevinas, Manolis; Wright, John; Brunekreef, Bert; Kere, Juha; Nawijn, Martijn C.; Annesi-Maesano, Isabella; Sunyer, Jordi; Melén, Erik; Koppelman, Gerard H.. - In: THE LANCET RESPIRATORY MEDICINE. - ISSN 2213-2600. - ELETTRONICO. - 6:5(2018), pp. 379-388. [10.1016/S2213-2600(18)30052-3]

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)



This is the pre-print version of:

Xu CJ, Söderhäll C, Bustamante M, Baiz N, Gruzieva O, Gehring U, Mason D, Chatzi L, Basterrechea M, Llop S, Torrent M, Forastiere F, Fantini MP, Carlsen KCL, Haahtela T, Morin A, Kerkhof M, Merid SK, van Rijkom B, Jankipersadsing SA, Bonder MJ, Ballereau S, Vermeulen CJ, Aguirre-Gamboa R, de Jongste JC, Smit HA, Kumar A, Pershagen G, Guerra S, Garcia-Aymerich J, Greco D, Reinius L, McEachan RRC, Azad R, Hovland V, Mowinckel P, Alenius H, Fyhrquist N, Lemonnier N, Pellet J, Auffray C; BIOS Consortium, van der Vlies P, van Diemen CC, Li Y, Wijmenga C, Netea MG, Moffatt MF, Cookson WOCM, Anto JM, Bousquet J, Laatikainen T, Laprise C, Carlsen KH, Gori D, Porta D, Iñiguez C, Bilbao JR, Kogevinas M, Wright J, Brunekreef B, Kere J, Nawijn MC, Annesi-Maesano I, Sunyer J, Melén E, Koppelman GH.

DNA methylation in childhood asthma: an epigenome-wide meta-analysis.

Lancet Respir Med. 2018 May;6(5):379-388.

The final published version is available online at

[https://doi.org/10.1016/S2213-2600\(18\)30052-3](https://doi.org/10.1016/S2213-2600(18)30052-3)

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

*This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)*

***When citing, please refer to the published version.***

## **DNA methylation in childhood asthma: an epigenome-wide meta-analysis**

Cheng-Jian Xu, Cilla Söderhäll, Mariona Bustamante, Nour Baïz, Olena Gruieva, Ulrike Gehring, Dan Mason, Leda Chatzi, Mikel Basteerechea, Sabrina Llop, Maties Torrent, Francesco Forastiere, Maria Pia Fantini, Karin C. Lodrup Carlsen, Tari Haahtela, Andréanne Morin, Marjan Kerkhof, Simon Kebede Merid, Bianca van Rijkom, Soesma A Jankipersadsing, Marc Jan Bonder, Stephane Ballereau, Cornelis J Vermeulen, Raul Aguirre-Gamboa, Johan C de Jongste, Henriette A Smit, Ashish Kumar, Göran Pershagen, Stefano Guerra, Judith Garcia-Aymerich, Dario Greco, Lovisa Reinius, Rosemary RC McEachan, Raf Azad, Vegard Hovland, Petter Mowinckel, Harri Alenius, Nanna Fyhrquist, Nathanaël Lemonnier, Johann Pellet, Charles Auffray, the BIOS Consortium, Pieter van der Vlies, Cleo C van Diemen, Yang Li, Cisca Wijmenga, Mihai G. Netea, Miriam F. Moffatt, William O.C.M Cookson, Josep M. Anto, Jean Bousquet, Tiina Laatikainen, Catherine Laprise, Kai-Håkon Carlsen, Davide Gori, Daniela Porta, Carmen Iñiguez, Jose Ramon Bilbao, Manolis Kogevinas, John Wright, Bert Brunekreef, Juha Kere, Martijn C Nawijn, Isabella Annesi-Maesano, Jordi Sunyer, Erik Melén, and Gerard H Koppelman

The authors' affiliations are as follows:

University of Groningen, University Medical Center Groningen, Department of Pulmonology, Groningen, the Netherlands. (C.J.X., C.J.V.) and University of Groningen, University Medical Center Groningen, GRIAC research institute Groningen, the Netherlands. (C.J.X., C.J.V., M.K., M.C.N., G.H.K) and University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands. (C.J.X., S.A.J., M.J.B., R.A-G., P.V., C.C.D., Y.L., C.W.) and Department of Women's and Children's Health, Karolinska Institutet, SE-171 77 Stockholm, Sweden. (C.S.) and Department of Biosciences and Nutrition, Karolinska Institutet, SE-141 83 Stockholm, Sweden. (C.S., L.R., J.K.) and ISGlobal, Centre for Research in Environmental Epidemiology (CREAL) and (M.Bu., S.G., J.G.-A., J.M.A., M.K., J.S. and Centre for Genomic Regulation (CRG) and The Barcelona Institute of Science and Technology.(M.Bu., J.M.A.) and Universitat Pompeu Fabra (UPF) (M.Bu., J.M.A., J.S.) and CIBER Epidemiología y Salud Pública (CIBERESP). (M.Bu., M.Ba., S.L., J.M.A., C.I., M.K., J.S.) and Epidemiology of Allergic and Respiratory Diseases Department (EPAR), Sorbonne Université, UPMC Univ Paris 06, INSERM, Pierre Louis Institute of Epidemiology and Public Health, Saint-Antoine Medical School, Paris, France. (N.B., I.A-M.) and Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. (O.G., A.K., S.K.M., G.P., H.A., M.F., E.M.) and Institute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands. (U.G., B.B.) and Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust. (D.M., R.R.M., R.A., J.W.) and Department of Social Medicine,

University of Crete, Greece. (L.C.) and Department of Genetics & Cell Biology, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands. (L.C.) and Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, USA. (L.C.) and Health Research Institute Biodonostia, San Sebastián, Spain. (M.Ba.) and Public Health Department of Gipuzkoa, San Sebastián, Spain. (M.Ba.) and Epidemiology and Environmental Health Joint Research Unit, FISABIO–Universitat Jaume I–Universitat de València, Valencia, Spain. (S.L.) and ib-salut, Area de Salut de Menorca. (M.T.) and Department of Epidemiology Lazio Regional Health Service, Rome, Italy. (F.F.) and Department of Biomedical and Neuromotor sciences, University of Bologna, Bologna, Italy. (M.P.F.) and Department of Paediatrics, Oslo University Hospital. (K.C.L.C., V.H.,P.M., K.H.C.) and University of Oslo. (K.C.L.C.) Skin and Allergy Hospital, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. (T.H.) and Department of Human Genetics, McGill University and Genome Quebec, Innovation Centre, 740 Dr. Penfield Avenue, Montréal, Québec H3A 1A5, Canada. (A.M.) and European Institute for systems Biology and Medicine, Campus Charles Mérieux - Université de Lyon, CIRI CNRS UMR5308, CNRS - ENS - UCBL – ENS. (S.B., N.L., J.P., C.A.,) and Institute for Advanced Biosciences, UGA-INSERM U1209-CNRS UMR 5309, Site Santé, Allée des Alpes, 38700 La Tronche (N.L.) and Cancer Research UK Cambridge Institute, University of Cambridge, Robinson Way, Cambridge CB2 0RE, UK. (S.B.) and Department of Pediatrics, Erasmus University Hospital/Sophia Children’s Hospital, Rotterdam, The Netherlands. (J.C.J.) and Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands. (H.A.S.) Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland. (A.K.) and University of Basel, Basel, Switzerland. (A.K.) and Asthma and Airway Disease Research Center, University of Arizona, Tucson, AZ, USA (S.G.) Faculty of Medicine and Life Sciences & Institute of Biosciences and Medical Technology (BioMediTech), University of Tampere, Finland (D.Gr.) and Institute for Advanced Biosciences, UGA-INSERM U1209-CNRS UMR5309, Site Santé, Allée des Alpes, 38700 La Tronche, France (N.L.) and HZPC Research BV, Metslawier, The Netherlands. (P.V.) and National Heart and Lung Institute, Imperial College, London SW3 6LY, UK (M.F.M., W.O.C.M.C.) and Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, 6525 HP Nijmegen, the Netherlands (M.G.N.) and University Hospital, Montpellier, France (J.B.) and National Institute for Health and Welfare, Helsinki, Finland (T.L.) and Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Helsinki, Finland (T.L.) and Centre intégré universitaire de santé et de services sociaux du Saguenay–Lac-Saint-Jean, 930, rue Jacques-Cartier Est, Chicoutimi (Québec) G7H 7K9 Canada (C.L.) and Department of Biomedical and Neuromotor sciences, University of Bologna, Bologna, Italy (D.Go.) and Department of Epidemiology Lazio Regional Health Service, Rome, Italy (D.P.) and Epidemiology and Environmental Health Joint Research Unit, FISABIO–Universitat Jaume I–Universitat de València, Valencia, Spain (C.I.) and University of the Basque Country UPV-EHU, Department of Genetics, Physical Anthropology and

Animal Physiology (J.R.B.) and BiorCruces Health Research Institute (J.R.B.) and Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands (B.B) and King's College London, Department of Medical and Molecular Genetics, London, UK (J.K.) and University of Groningen, University Medical Center Groningen, Department of Pathology and Medical Biology, Experimental Pulmonology and Inflammation Research, Groningen, the Netherlands. (M.C.N.) and IMIM (Hospital del Mar Medical Research Institute (J.S.) and Sachs Children's Hospital, Stockholm, Sweden (E.M.) and Centre for Occupational and Environmental Medicine, Stockholm County Council, Stockholm, Sweden (E.M.) and University of Groningen, University Medical Center Groningen, Department of Pediatric Pulmonology and Pediatric Allergy, Beatrix Children's Hospital, Groningen, the Netherlands (G.H.K.)

Dr. Melén and Dr. Koppelman contributed equally to this article.

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Koppelman at PO Box 30.001, CA43, 9700 RB Groningen, Beatrix Children's Hospital, Department of Pediatric Pulmonology and Pediatric Allergology, University Medical Center Groningen, Groningen, the Netherlands or at [g.h.koppelman@umcg.nl](mailto:g.h.koppelman@umcg.nl)

## **Abstract**

**Background** DNA methylation profiles associated with childhood asthma may provide novel insights into disease pathogenesis.

**Methods** Epigenome-wide DNA methylation was investigated in whole blood DNA of 392 asthma cases and 1156 controls at age 4-8 years from four European birth cohorts within the Mechanisms of the Development of ALLergy (MeDALL) project. Twenty-seven differentially methylated CpG sites (CpGs) were further analyzed in 247 cases and 2,949 controls at age 4-16 from six additional European cohorts, and meta-analyzed. We addressed cell-type specificity in eosinophils and nasal epithelial cells. Whole blood transcriptional profiles associated with replicated CpGs were annotated using RNA-seq data of FACS-sorted subsets of peripheral blood mononuclear cells.

**Findings** Fourteen of 27 CpGs from the discovery analyses were replicated and passed genome-wide significance ( $P < 1.14 \times 10^{-7}$ ) after meta-analysis. Consistently lower methylation levels were observed at all associated loci across childhood from age 4 to 16 years, but not at birth. All 14 CpGs were significantly associated with asthma in a second replication study using whole blood DNA, and were strongly associated with asthma in purified eosinophils. Whole blood transcriptional signatures associated with these CpGs indicated increased activation of eosinophils, effector/memory CD8-T and NK cells, and reduced naïve T cells. Five of the 14 CpGs were associated with asthma in respiratory epithelial cells indicating cross-tissue epigenetic effects.

**Interpretation** Reduced whole blood DNA methylation at 14 CpGs acquired after birth was strongly associated with childhood asthma. These CpGs and their associated transcriptional profiles indicate activation of eosinophils and cytotoxic T cells in childhood asthma.

## Introduction

Asthma is a heterogeneous chronic inflammatory airway disease, characterized by variable respiratory symptoms and reversible airflow limitation. Worldwide, over 300 million patients suffer from asthma, with significant morbidity, reduced quality of life of patients and their families, and substantial

healthcare costs<sup>1</sup>. Asthma has its origins predominantly in early childhood<sup>2</sup>. According to twin studies, about 50-60% of asthma susceptibility is explained by genetic factors<sup>3</sup>. Large scale genome-wide association studies (GWAS) have uncovered several genetic variants related to asthma, some of which are specific to childhood onset asthma<sup>4,5</sup>. The dramatic increase in asthma prevalence in the past half-century suggests an important role for environmental exposures early in life<sup>6,7</sup>, which can affect epigenetic patterns leading to changes in gene regulation, and consequently to altered biological responses contributing to disease susceptibility<sup>8,9</sup>. In support of this, epigenetic changes have been observed at specific genomic loci for asthma, including DNA methylation CpG sites (CpGs)<sup>10,11,12,13,14,15</sup> non-coding RNAs<sup>16</sup> and microRNA<sup>17</sup>. Specifically, previous candidate and epigenome-wide studies in peripheral whole blood (or peripheral blood mononuclear cells) have found methylation associations with immune-related genes, such as *FOXP3*<sup>18</sup> and *RUNX3*<sup>12</sup>, involved in T-cell regulation and maturation. Eosinophil-associated genes have further been implicated, exemplified by *IL5RA*<sup>19</sup> and *IL13*<sup>12</sup>, the latter being differentially methylated also in airway epithelial cells following IL13-stimulation<sup>15</sup>. In studies on total IgE levels, pathways relevant to eosinophil activation<sup>10</sup> and lipid / fatty acid metabolism<sup>20</sup> with suggested involvement in asthma pathogenesis are highlighted.

Despite the recent progress in asthma epigenetics, it remains largely unknown which DNA methylation sites are robustly associated with asthma across the genome; whether these asthma associated methylation patterns are already present at birth or develop during childhood<sup>9</sup>, and whether age-specific patterns can be observed in children from different countries with varying environmental exposures and lifestyles. To answer these questions, we performed a large-scale epigenome-wide association study (EWAS) within the Mechanisms of the Development of ALLergy (MeDALL) project<sup>21</sup> of whole blood DNA in children at preschool age (4-5 years) and school age (8 years) in four birth cohorts together with a validation and replication study in children (0-16 years) and adults (1-79 years). We subsequently investigated cell-type specific methylation of the identified CpGs in eosinophils and epithelial cells and their related gene expression signatures.

## Methods

A full description of the methods is provided in the online supplement. Medical Ethical Committees of all institutions approved this study. Written informed consent was obtained from parents or legal guardians of all participating children.



Asthma definition: We defined asthma as agreed upon by experts as described previously in the MeDALL study<sup>22</sup>, with two of three criteria present: 1) doctor diagnosis of asthma ever, 2) use of asthma medication in the past 12 months and 3) wheezing/breathing difficulties in the past 12 months.

### *Discovery*

Whole blood DNA from children from four MeDALL birth cohorts was investigated in BAMSE (Sweden), EDEN (France), INMA (Sabadel [SAB], Spain) and PIAMA (the Netherlands)<sup>22</sup>. Given the strong effects of age on whole blood methylation patterns in childhood<sup>9</sup>, we considered the possibility of age specific methylation in asthma and performed discovery analysis at preschool and school age. We thus examined epigenome-wide DNA methylation using Illumina Infinium HumanMethylation450 BeadChips (450K) in 207 asthma cases and 610 controls at age 4-5 years, and 185 asthma cases and 546 controls at age 8 years using a cross-sectional case-control design (Figure 1).

### *Replication*

All blood DNA samples of children not participating in the discovery phase were selected for replication from seven MeDALL birth cohorts: BIB (UK), ECA (Norway), Karelia (Finland), PIAMA, RHEA (Greece), Robbic (Italy; subsets Bologna and Rome), INMA (Spain; subsets Gipuzkoa [GIP], Menorca [MEN] and Valencia [VAL]) (Table S1). We used an independent technology, iPLEX (Agena Biosciences) to validate and replicate selected CpGs after quality control (QC) of the assays. In total, 3,196 samples from 6 birth cohorts were selected for our main meta-analysis. Top CpGs at both age 4-5 and 8 years were selected for replication (Table S2-S3). We next investigated whether replicated CpGs in cord blood predict later asthma in 1,316 children participating in BIB, EDEN, INMA-VAL and Robbic (Table S4).

We subsequently studied cell type specificity of the asthma association of the replicated CpGs in 455 respiratory epithelial cell samples, collected by nasal brushing of 16-year-old children as well as in DNA isolated from blood eosinophils (16 cases, 8 controls (age 2-56 years) and compared this to whole blood DNA samples of 167 subjects (age 1-79 years) from families from the Saguenay-Lac-Saint-Jean (SLSJ) region in Canada<sup>23</sup> (Table S4).

### *Statistical analyses*

Asthma associated differentially methylated CpGs were identified by fitting a robust linear regression method corrected for sex, cohort and technical co-variables. We used models that do not correct for blood cell-type as main model for discovery and replication, and report cell-type corrected models by inclusion of the estimated cell type proportions as covariates as sensitivity analysis.

We performed inverse variance-weighted fixed-effects meta-analyses with METAL<sup>19</sup>, or random effects meta-analysis using Metafor<sup>25</sup> in case of a heterogeneity  $\chi^2$  test  $P$  value  $< 0.05$ . Our replicated CpGs were those that were significantly associated in the meta-analysis of replication samples (Bonferroni correction,  $P < 0.0019$ , 27 tests) and passed epigenome-wide significance using Bonferroni correction ( $P < 1.14 \times 10^{-7}$ , 439,306 tests) after meta-analysis of all studies.

### *Functional genomic analysis*

The association of CpG methylation with single nucleotide polymorphisms (SNPs) was studied in 496 children, and with matched regional whole blood gene expression in blood in 119 children (4 years, INMA), and 260 children (16 years, BAMSE) using Affymetrix Human Transcriptome Array 2.0 Genechips<sup>26</sup>. Association between CpG methylation with whole genome expression by RNA-seq was performed in the BIOS consortium dataset. CpGs were annotated by GREAT<sup>27</sup>. Functional enrichment analysis was done by overlapping our replicated CpGs with histone marks and chromatin states of 27 blood cell-types in the Roadmap epigenome project<sup>28</sup>. Gene expression sets associated with replicated CpGs were interrogated for cell-type specificity using the 500FG dataset<sup>29</sup>.

## **Results**

### **EWAS discovery and replication in whole blood DNA**

At age 4-5 years, we identified one genome-wide significant CpG annotated to *ETV6* and *BCL2L14* ( $P$  value =  $7.96 \times 10^{-8}$ ) (Figure 2a and S2a) (Table 1). Therefore, we used a looser threshold (FDR adjusted  $P < 0.10$ ) to select eleven CpGs for replication (Table S2) as we considered age-specific effects on asthma. Next, we identified 26 CpGs that passed genome-wide significance threshold ( $P < 1.14 \times 10^{-7}$ ) at age 8 years (Figures 2b and S2b, Table S3), with the strongest association ( $P = 8.82 \times 10^{-10}$ ) found for cg03695871 in *AP5B1* (*Adaptor related protein complex 5 beta 1 subunit*), a gene associated with eczema<sup>27</sup>. Cg11456013 and cg1362844 were identified at both ages and in total 35 CpGs were selected for replication.

IPlex assays for 27 out of 35 CpGs were available after QC. Fourteen of the 27 CpGs were significantly associated with asthma across childhood after meta-analysis of the six MeDALL replication cohorts and passed the genome-wide significance threshold in the meta-analysis of replication and discovery cohorts

(Table 1, Tables S5-S6). Expression and function of genes annotated to replicated CpGs are described in Table S7, and genomic annotations of all CpGs and regional co-methylation plots are provided in Figures S3-S5. The most significant association ( $P$  value =  $2.55 \times 10^{-17}$ ) was observed for cg01901579 in *DICER1* (Figure S6). Specific replication at age 4 is provided in Table S8, but could not be performed at age 8 years due to low power. Methylation levels in asthma cases were consistently lower than those in controls for all 14 replicated CpGs, with small effects in whole blood in the discovery cohorts (Figure S7). To evaluate potential SNP effects within the probe for the 14 replicated CpGs, the beta value distributions were visually assessed in the discovery cohorts (Figure S8), and no bimodal distribution was detected that could reveal an underlying SNP within the probe effect. In addition, the residuals were normally distributed. Homogeneity testing revealed consistent effects across childhood for all replicated sites ( $P_{\text{het}} > 0.05$  and  $I^2 < 25\%$ ) (Figures S9-S12). Strikingly, all 14 CpGs (100%) significantly associated with asthma in a second replication study using whole blood DNA of Canadian families (age 1-79) (Table 2). However, methylation of the 14 asthma-associated CpGs in 1,316 cord blood DNA samples did not significantly predict asthma development at age 3-4 years in four different cohorts (Table S9).

### **Cell type heterogeneity and cell type specific analysis**

In our discovery samples, white blood cell composition was measured in INMA (Table S10) and estimated in all four cohorts (Table S11). Asthma was significantly associated with lower estimated CD4 and CD8 positive T cell subsets in whole blood, but not with blood eosinophil counts. However, we observed that the association of CpG methylation with asthma became less significant after eosinophil and neutrophil correction ( $P$  value after correction ranged from  $1.23 \times 10^{-2}$  to  $4.41 \times 10^{-6}$  versus  $1.17 \times 10^{-6}$  to  $2.79 \times 10^{-9}$  before correction; Table 1, Table S5-6). Using Reinius' blood cell-type specific methylation data<sup>31</sup>, we found that 11 of the 14 asthma associated CpGs displayed relatively lower methylation levels in eosinophils (Table S12) ( $P$  value =  $1.92 \times 10^{-24}$ , Fisher exact test).

We next investigated whether CpG methylation within eosinophils was also associated with asthma in participants of the SLSJ study, and found all 14 CpGs to be significantly associated in subjects with a very broad age range of 2-56 years (Table S4). CpG methylation in eosinophils was on average 18% lower in asthmatics (Table 2, Figure S13). Furthermore, we found significant differential methylation in relation to asthma for five out of 14 CpGs in nasal respiratory epithelial cells<sup>13</sup> from 16-year-old children (enrichment  $P$  value =  $3.8 \times 10^{-4}$ ) (Table 2).

## Functional and mediation analysis

Four of 14 CpGs associated with SNPs within 250 kb (*cis*-MeQTL) (Table S13), including *rs9425436* that associated with CpG methylation of cg01770400 in *SERPINC1* ( $p = 5.6 \times 10^{-4}$ ). One of these SNPs was previously related to asthma: *rs9425436* was in strong LD ( $r^2=0.95$ ) with *rs4652298* in *ZBTB37* that has previously shown evidence for association with asthma ( $P = 2.7 \times 10^{-3}$ ) in the GABRIEL GWAS<sup>4</sup>. Mediation analysis revealed that 14.4% of the SNP *rs9425436* effect on asthma was due to CpG methylation of cg01770400 (Table S14).

The 14 replicated CpGs were significantly enriched for enhancer markers (13 out of 14) in whole blood Roadmap and Encode data<sup>28</sup> (Table S15-S16).

We next examined whether the replicated 14 CpGs were associated with whole blood gene expression in *cis* within 250 kb in two childhood cohorts. Four asthma associated CpGs were significantly associated with gene expression in *cis* (Table S17) in BAMSE, but this was not replicated in the smaller dataset INMA (Table S18). In the larger BIOS dataset<sup>32</sup>, we first analyzed the association of the 14 CpGs with gene expression *in cis* in 2,367 adults, and identified 10 of the 14 asthma associated CpGs to be associated with 22 gene transcripts (Table S21), with two CpG-gene transcript pairs (cg06483830-*SCPEP1* and cg14011077-*PPP1R26*) observed also in the BAMSE dataset (Table S17).

We then analyzed the association of asthma CpGs with whole genome gene expression. All CpGs showed significant associations with whole blood gene expression (mean number of associated transcripts 351, range 77 – 875). All associated transcripts are shown per CpG site in supplemental table S19). We identified three distinct gene expression patterns in whole blood by hierarchical clustering. The first cluster was inversely associated to methylation levels of seven CpGs, and encompassed genes such as *SLC29A1*, *SIGLEC8*, *IL5RA* and *ADORA3* (Figure 3a, Table S19). A second cluster was strongly associated with cg16592897 and six other CpGs, and included genes such as *CCR7*, *CHMP7* and *LEF1*. After correction for eosinophil blood counts, association of the first cluster was attenuated, confirming this as an eosinophilic gene signature, while the association of the second cluster became more prominent. In contrast, the third cluster, containing *TBX21*, *EOMES*, *CCL5* and *GZMB*, showed negative correlation with four asthma CpG methylation levels (Table S20, Figure 3b). An RNA-seq dataset of PBMCs FACS-sorted into highly defined leucocyte subsets<sup>29</sup> identified the second cluster as a naïve CD4 and CD8 T cell gene signature that was decreased, and the third cluster as an effector/memory CD8 T cell and NK cell profile that was increased with reduced CpG methylation (Figure S14).

## Discussion

This large consortium-based meta-analysis identified 14 CpGs in whole blood to be associated with childhood asthma. Consistently lower methylation of these CpGs was observed in asthma from preschool age until adolescence. Strongly reduced methylation at asthma-associated CpGs within isolated eosinophils clearly indicated that lower whole blood DNA methylation levels in asthma was not only due to altered eosinophil numbers in whole blood, but also to reduced methylation at those specific CpGs within this cell type. Finally, through clustering of the 14 asthma associated CpG sites with whole blood and immune cell specific gene expression signatures, we were able to annotate the asthma associated CpGs to (activated) eosinophils and CD8 positive T cells and NK cells.

DNA methylation may be strongly affected by ageing and environmental exposures<sup>8,9</sup>. We observed remarkably consistent patterns of DNA methylation associated with childhood asthma across different age groups and in participants from different countries. This indicates that our asthma associated CpGs are consistent findings in childhood asthma from the age of 4 years onwards but not at birth, indicating postnatal effects on methylation of these 14 CpGs that were identified in our study. Recently, cord blood methylation of a CpG in another asthma gene, *SMAD3*, was found to be associated with childhood asthma in children from asthmatic mothers<sup>11</sup>. Maternal smoking during pregnancy, a risk factor for childhood asthma, showed strong effects on methylation of 6,073 CpGs in DNA isolated from cord blood<sup>8</sup>. One of these CpGs was associated to asthma in our study (*LMAN2*), yet we did not identify a mediating effect of smoking on asthma through this CpG (data not shown). Thus, asthma associated DNA methylation patterns identified in this study are likely the result of postnatal environmental influences, and/or pathophysiological processes related to asthma. We performed detailed functional analysis of the asthma-associated CpGs to start identifying these processes.

Part of the reduced DNA methylation of the 14 CpGs in asthma was explained by eosinophil numbers, since the significance of the association reduced after correction for predicted eosinophil percentage in peripheral blood<sup>20</sup>. Importantly, strongly reduced methylation of the CpGs in purified eosinophils retained an association with asthma, highlighting that eosinophils are epigenetically altered in asthma. We observed a larger effect size in purified eosinophils (on average 18% lower in asthmatics) compared to whole blood (on average 1-2% lower in asthmatics) consistent with a functional effect on gene expression within eosinophils. As predicted eosinophil counts between asthma patients and controls were not significantly different in our discovery cohorts, our results reflect the presence of a different subset or activation state of eosinophils, rather than eosinophil counts per se. While eosinophils are linked to sub-phenotypes of asthma such as the Th2-type subset, our study provides unequivocal data supporting their involvement at the epigenetic level in mild-to-moderate asthma in children. This adds

to the genetic findings<sup>33</sup> that SNPs regulating blood eosinophil counts significantly overlap with asthma GWAS loci, such as *IL5*, *IL33*, *IL1RL1* and *TSLP*. Functional interpretation using Roadmap epigenomics data revealed that asthma associated CpGs were significantly enriched for enhancer markers again suggesting a cell-type specific regulatory role on gene expression<sup>34</sup>.

Correlation of our 14 CpGs with genome-wide gene expression revealed three clusters of associated genes. As all replicated 14 CpGs show consistently reduced DNA methylation in asthma patients, the cluster encompassing known eosinophil signature genes<sup>35</sup> displaying negative association with methylation levels of asthma-associated CpGs are therefore predicted to show increased expression in asthma. A second cluster with increased predicted expression in asthma is characterized by effector memory CD8 T cell and NK cell genes. A third cluster with reduced predicted expression in asthma was identified as naïve CD4 and CD8 T cells. These data indicate that whole blood DNA in childhood asthma carries CpG marks associated with reduced activity of naïve T cells and increased activity of effector/memory CD8 T cells and NK cells<sup>36</sup>. This asthma associated shift in cellular activity may be due to different environmental factors such as traffic related air pollution or infectious agents including intracellular bacteria or viruses<sup>37</sup>. Further studies are needed in well-characterised populations of children with asthma to address cell type specific methylation signatures together with RNA-seq data in effector memory CD8 positive T cells and NK cells to disentangle a cell proportion effect from differential cell activation in asthma. Thus, asthma associated whole blood methylation patterns represent an epigenetic fingerprint of (activated) immune cell subsets acquired in childhood.

Comparing the 14 CpGs with previously published DNA methylation study of asthma in US inner cities<sup>12</sup>, we did not identify overlapping CpG sites. However, three of the genome-wide significant CpGs (cg01770400, cg20503329, cg09447105) (Table S2-3) were previously reported in an EWAS for total IgE<sup>10</sup>, compatible with shared regulation of these CpGs in asthma and IgE-related traits. None of these annotated genes were previously implicated in (genetic) studies of asthma (Table S7 for an overview), yet several of these genes have known functions that may have relevance for asthma development. The most significant CpG site was annotated to *DICER1* (Figure S5). *DICER1* is a member of the ribonuclease III (RNaseIII) family, involved in the generation of microRNAs, which modulate gene expression at the posttranscriptional level. Next, *SYNTAXIN3* (*STX3*) is relevant for epithelial polarity and regulates exocytosis<sup>38</sup>. Epithelial polarity is important to maintain a functional epithelial barrier function, a process that is impaired in asthma<sup>39</sup>. *STX3* plays an important role in the release of the chemokines CXCL8, CCL2, CCL3, and CCL4 by human mast cells<sup>40</sup>. *LIPIN1* is a Mg(2+)-dependent phosphatidic acid phosphohydrolase, that was shown to negatively regulate mast cell degranulation and the anaphylactic response through inhibiting the PKC-SNAP-23 pathway<sup>41</sup>. *LIPIN1* also plays a role in macrophage

proinflammatory activation during TLR signaling. After TLR4 stimulation, Lipin-1-deficient macrophages showed a decreased production of diacylglycerol and activation of MAPKs and AP-1, accompanied by a reduced production of proinflammatory cytokines such as IL-6, IL-12, IL-23, or enzymes such as Inducible NO synthase<sup>42</sup>. Finally, Lectin, Mannose Binding 2, a Type I transmembrane lectin, shuttles between the endoplasmic reticulum, the Golgi apparatus and the plasma membrane, and was previously implicated in sphingolipid transport. Sphingolipid metabolism has been related to asthma<sup>43</sup> after the discovery of the role of *ORMDL3*, the first gene identified for childhood asthma through GWAS which may have a function in sphingolipid metabolism.

There are some strengths and limitations of our work. To our knowledge, this is the first study that combined EWAS with an extensive validation and replication study of the CpGs using an independent laboratory method in over 5,000 children. One limitation of our work is that we used a questionnaire-based approach to define asthma at preschool age based on doctor's diagnosis, symptoms and use of asthma medication. Whereas we acknowledge that self-reported asthma at preschool age may be sensitive to misclassification of transient wheezing, our finding that 8 of the 11 CpGs selected from the age 4-5 years discovery sample were significantly associated with asthma across childhood strongly suggests that we found robust asthma CpGs. However, we cannot exclude that the challenges with diagnosing asthma in young children may partly explain why we identified only one genome-wide CpG in the initial EWAS discovery on 4-5-year olds. Since we a priori considered age specific effects on DNA methylation in asthma, two sets of CpG sites were selected for replication at age 4 and 8 years. Although different significance cut offs were used for this initial selection (FDR < 0,1 at age 4, Bonferroni significance at age 8 years), this did not affect the final outcome as all CpG passed robust criteria for replication (Bonferroni significance in the replication analysis only, as well as in in the final combined meta-analysis of discovery and replication studies). In addition, replication analyses of our 14 top hits in the SLSJ study allowed us to not only assess methylation profiles in purified eosinophils, but also to show that hypomethylation of these CpGs was observed in participants at a broader age range (age 1-79 years). A second limitation is that the CpG selection was limited to the 450K platform that by design only covers 1.6% of all methylation sites of the genome. We would expect to find many more asthma associated CpGs when using whole methylome sequencing in the future. Thirdly, we only can address CpG loci but not CpG regions in our study since CpG regions cannot not be replicated in the targeted iPLEX design. Fourthly, small differences in methylation levels between cases and controls were observed in the whole blood samples, whereas substantial differences were seen in purified eosinophils. This indicates dilution of the strong methylation effects (in eosinophils) if whole blood is used for analyses. Finally, we observed replication of 5 of the 14 CpG sites in respiratory nasal epithelial cells, yet cannot exclude the possibility that (part of) this signal is due to admixture of eosinophils in epithelial brushings.

In conclusion, our study identified consistent reduced DNA methylation levels in 14 CpGs to be associated with asthma across childhood from 4 to 16 years of age, but not at birth. Asthma associated whole blood DNA methylation profiles were strongly driven by lower methylation within eosinophils, highlighting the importance of the eosinophil as an epigenetic determinant of asthma. In addition, CpG methylation patterns identified an early-life shift from naïve T cell populations towards effector memory CD8 and NK cell subsets indicating a potentially crucial role for host-virus interactions in asthma inception. Our study reveals whole blood methylation signatures that represent an epigenetic fingerprint of (activated) eosinophil and CD8 positive T and NK cell subsets acquired in childhood.

### **Contributors**

GHK and EM conceived the project and coordinated the analyses. CJX analyzed epi-genomewide data for discovery and replication. GHK, EM, CJX and MCN drafted the manuscript. MCN contributed to the biological interpretation. CS, OG, SKM, AK, GP, DG, LR, JK, EM provided the biological materials and phenotype information of BAMSE. MBu, MBa, SL, MT, SG, JGA, CI, JRB and JS provided the biological materials and phenotype information of INMA. NB and IAM provide the biological materials and phenotype information of EDEN. UG, JCdJ, HAS and BB provide the biological materials and phenotype information of PIAMA. DM, RRCM, RA and JW provided the biological materials and phenotype information of BIB. LC and MK provided the biological materials and phenotype information of Rhea. FF, MPF, DG and DP provided the biological materials and phenotype information of Robbic. KCL, VH, PM and KHC provided the biological materials and phenotype information of ECA. TH, HA and NF and TL provided the biological materials and phenotype information of Karelia. OG implemented the shadow analysis of EWAS. AM, MFM, WOCMC, CL and implemented and coordinated replication in SLSJ. MK constructed the MedALL asthma phenotypes. SKM did eQTM analysis of BAMSE 16 years. CJX did eQTM analysis of BIOS data. MJB contributed Great and Roadmap data annotation. RAG, YL, CW and MGN implemented and coordinated RNA seq analysis in 500G. CJV did PIAMA RNA seq analysis. SB, NL, JP and CA implemented and coordinated eQTM analysis in INMA. BvR, SAJ, PvDV, CCvD produced the 450k and iplex data. JMA and JB coordinated the MedALL project. All authors were involved in data interpretation, read and approved the manuscript.

### **Declaration of Interests**



JP reports grants from European Commission, during the conduct of the study. JB reports personal fees from Almirall, AstraZeneca, Chiesi, GSK, Meda, Menarini, Merck, MSD, Novartis, Sanofi-Aventis, Takeda, Teva and Uriach, outside the submitted work. MCN reports grants from EU FP7 program - MeDALL project, during the conduct of the study, grants from GSK Ltd, personal fees from DC4U, outside the submitted work. CL reports a pending patent named “Refractory asthmatics with high levels of IgE and low levels of methylation”. GHK reports grants from European Union, grants from Lung Foundation of the Netherlands, during the conduct of the study, grants from TETRI foundation, grants from UBBO EMMIUS foundation, grants from TEVA the Netherlands, outside the submitted work. All other authors have no competing interests.

## Acknowledgements

We are grateful to all children and families that participated in this study. We especially thank Professor Dirkje Postma for helpful comments and support during the course of this study. The Mechanisms of the Development of ALLergy (MeDALL) EU project was supported by the seventh Framework programme (grant agreement number 261357). The Biobank-Based Integrative Omics Studies (BIOS) Consortium is funded by BBMRI-NL, a research infrastructure financed by the Dutch government (NWO 184.021.007).

## References

- 1 Eder W, Ege MJ, von Mutius E. The Asthma Epidemic. *N Engl J Med* 2006; **355**: 2226–35.
- 2 Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and Wheezing in the First Six Years of Life. *N Engl J Med* 1995; **332**: 133–8.
- 3 McGeachie MJ, Stahl EA, Himes BE, *et al.* Polygenic heritability estimates in pharmacogenetics: focus on asthma and related phenotypes. *Pharmacogenet Genomics* 2013; **23**: 324–8.
- 4 Moffatt MF, Gut IG, Demenais F, *et al.* A Large-Scale, Consortium-Based Genomewide Association Study of Asthma. *N Engl J Med* 2010; **363**: 1211–21.
- 5 Ober C, Nicolae DL. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011; **43**: 887–92.
- 6 Stein MM, Hrusch CL, Gozdz J, *et al.* Innate Immunity and Asthma Risk in Amish and Hutterite Farm Children. *N Engl J Med* 2016; **375**: 411–21.

- 7 Ege MJ, Mayer M, Normand A-C, *et al.* Exposure to Environmental Microorganisms and Childhood Asthma. *N Engl J Med* 2011; **364**: 701–9.
- 8 Joubert BR, Felix JF, Yousefi P, *et al.* DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis. *Am J Hum Genet* 2016; **98**: 680–96.
- 9 Xu C-J, Bonder MJ, Söderhäll C, *et al.* The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics* 2017; **18**: 25.
- 10 Liang L, Willis-Owen SAG, Laprise C, *et al.* An epigenome-wide association study of total serum immunoglobulin E concentration. *Nature* 2015; **520**: 670–4.
- 11 DeVries A, Wlasiuk G, Miller SJ, *et al.* Epigenome-wide analysis links SMAD3 methylation at birth to asthma in children of asthmatic mothers. *J Allergy Clin Immunol* DOI:10.1016/j.jaci.2016.10.041.
- 12 Yang IV, Pedersen BS, Liu A, *et al.* DNA methylation and childhood asthma in the inner city. *J Allergy Clin Immunol*; **136**: 69–80.
- 13 Yang IV, Richards A, Davidson EJ, *et al.* The Nasal Methylome: A Key to Understanding Allergic Asthma. *Am J Respir Crit Care Med* 2017; **195**: 829–31.
- 14 Nicodemus-Johnson J, Myers RA, Sakabe NJ, *et al.* DNA methylation in lung cells is associated with asthma endotypes and genetic risk. *JCI Insight* 2016; **1**. DOI:10.1172/jci.insight.90151.
- 15 Nicodemus-Johnson J, Naughton KA, Sudi J, *et al.* Genome-Wide Methylation Study Identifies an IL-13–induced Epigenetic Signature in Asthmatic Airways. *Am J Respir Crit Care Med* 2016; **193**: 376–85.
- 16 Austin PJ, Tsitsiou E, Boardman C, *et al.* Transcriptional profiling identifies the long noncoding RNA plasmacytoma variant translocation (PVT1) as a novel regulator of the asthmatic phenotype in human airway smooth muscle. *J Allergy Clin Immunol* 2017; **139**: 780–9.
- 17 Maes T, Cobos FA, Schleich F, *et al.* Asthma inflammatory phenotypes show differential microRNA expression in sputum. *J Allergy Clin Immunol* 2016; **137**: 1433–46.
- 18 Nadeau K, McDonald-Hyman C, Noth EM, *et al.* Ambient air pollution impairs regulatory T-cell function in asthma. *J Allergy Clin Immunol* 2010; **126**: 845–852.e10.
- 19 Arathimos R, Suderman M, Sharp GC, *et al.* Epigenome-wide association study of asthma and wheeze in childhood and adolescence. *Clin Epigenetics* 2017; **9**: 112.
- 20 Chen W, Wang T, Pino-Yanes M, *et al.* An epigenome-wide association study of total serum IgE in Hispanic children. *J Allergy Clin Immunol* DOI:10.1016/j.jaci.2016.11.030.
- 21 Bousquet J, Anto J, Auffray C, *et al.* MeDALL (Mechanisms of the Development of ALLergy): an integrated approach from phenotypes to systems medicine. *Allergy* 2011; **66**: 596–604.
- 22 Pinart M, Benet M, Annesi-Maesano I, *et al.* Comorbidity of eczema, rhinitis, and asthma in IgE-sensitized and non-IgE-sensitized children in MeDALL: a population-based cohort study. *Lancet Respir Med* 2014; **2**: 131–40.
- 23 Laprise C. The Saguenay-Lac-Saint-Jean asthma familial collection: the genetics of asthma in a young founder population. *Genes Immun* 2014; **15**: 247–55.

- 24 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**: 2190–1.
- 25 Viechtbauer W. Conducting Meta-Analyses in R with the **metafor** Package. *J Stat Softw* 2010; **36**. DOI:10.18637/jss.v036.i03.
- 26 Gref A, Merid SK, Gruzieva O, *et al.* Genome-Wide Interaction Analysis of Air Pollution Exposure and Childhood Asthma with Functional Follow-up. *Am J Respir Crit Care Med* 2017; **195**: 1373–83.
- 27 McLean CY, Bristor D, Hiller M, *et al.* GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010; **28**: 495–501.
- 28 Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A. Integrative analysis of 111 reference human epigenomes. *Nature* 2015; **518**. DOI:10.1038/nature14248.
- 29 Aguirre-Gamboa R, Joosten I, Urbano PCM, *et al.* Differential Effects of Environmental and Genetic Factors on T and B Cell Immune Traits. *Cell Rep* 2016; **17**: 2474–87.
- 30 Paternoster L, Standl M, Chen C-M, *et al.* Meta-analysis of genome-wide association studies identifies three new risk loci for atopic dermatitis. *Nat Genet* 2012; **44**: 187–92.
- 31 Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlen SE, Greco D. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One* 2012; **7**. DOI:10.1371/journal.pone.0041361.
- 32 Bonder MJ, Luijk R, Zhernakova DV, *et al.* Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* 2017; **49**: 131–8.
- 33 Astle WJ, Elding H, Jiang T, *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell*; **167**: 1415–1429.e19.
- 34 Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nat Rev Mol Cell Biol* 2015; **16**: 144–54.
- 35 Choy DF, Jia G, Abbas AR, *et al.* Peripheral blood gene expression predicts clinical benefit from anti-IL-13 in asthma. *J Allergy Clin Immunol*; **138**: 1230–1233.e8.
- 36 Abdelsamed HA, Moustaki A, Fan Y, *et al.* Human memory CD8 T cell effector potential is epigenetically preserved during in vivo homeostasis. *J Exp Med* 2017; **214**: 1593.
- 37 Youngblood B, Hale JS, Ahmed R. Memory CD8 T cell transcriptional plasticity. *F1000Prime Rep* 2015; **7**: 38.
- 38 Vogel GF, Klee KMC, Janecke AR, Müller T, Hess MW, Huber LA. Cargo-selective apical exocytosis in epithelial cells is conducted by Myo5B, Slp4a, Vamp7, and Syntaxin 3. *J Cell Biol* 2015; **211**: 587.
- 39 Davies DE. Epithelial Barrier Function and Immunity in Asthma. *Ann Am Thorac Soc* 2014; **11**: S244–51.
- 40 Frank SPC, Thon K-P, Bischoff SC, Lorentz A. SNAP-23 and syntaxin-3 are required for chemokine release by mature human mast cells. *Mol Immunol* 2011; **49**: 353–8.

- 41 Shin J, Zhang P, Wang S, Wu J, Guan Z, Zhong X-P. Negative control of mast cell degranulation and the anaphylactic response by the phosphatase lipin1: Innate immunity. *Eur J Immunol* 2013; **43**: 240–8.
- 42 Meana C, Peña L, Lordén G, *et al.* Lipin-1 Integrates Lipid Synthesis with Proinflammatory Responses during TLR Activation in Macrophages. *J Immunol* 2014; **193**: 4614–22.
- 43 Levy BD. Sphingolipids and Susceptibility to Asthma. *N Engl J Med* 2013; **369**: 976–8.