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

Cheng-jian Xu, Cilla Soderhall, Mariona Bustamante, Nour Baz, Olena Gruzieva, Ulrike Gehring, Dan Mason, Leda Chatzi, Mikkel Basterrechea, Sabrina Llop, Maties Torrent, Francesco Forastiere, Maria Pia Fantini, Karin C Ladrup Carlsen, Tari Hahtela, Andréanne Morin, Marjan Kerkhof, Simon Kebede Merid, Bianca van Rijkm, Soesma A Jankipersadising, Marc Jan Bonder, Stephane Balleroux, Cornélis J Vermeulen, Raul Aguirre-Gamboa, Johan C de Jongste, Henriette A Smitt, Ashish Kumar, Goran Pershagen, Stefano Guerra, Judith García-Aynerich, Dario Greco, Lovisa Reinius, Rosemary R C McEachan, Raf Azad, Vegard Hovland, Petter Mowinckel, Harri Alenius, Nanna Fyhrquist, Nathanael Lemonnier, Johann Pellet, Charles Auffray, the BISOS Consortium, Pieter van der Vlies, Cleo C van Diemen, Yang Li, Ciska Wijmenga, Mihai G Netea, Miriam F Moffatt, William O C M Cookson, Josep M Anto, Jean Bousquet, Tania Lastikainen, Catherine Laprise, Kai-Håkan Carlsen, Davide Gori, Daniela Porta, Carmen Ihiguez, Jose Ramon Bilbao, Manolis Kogevinas, John Wright, Bert Brunekeef, Juha Kere, Martijn C Nawijn, Isabella Annesi-Maesano, Jordi Sunyer, Erik Melén,*, Gerard H Koppelman* | *Contributed equally

Summary
Background DNA methylation profiles associated with childhood asthma might provide novel insights into disease pathogenesis. We did an epigenome-wide association study to assess methylation profiles associated with childhood asthma.

Methods We did a large-scale epigenome-wide association study (EWAS) within the Mechanisms of the Development of ALLergy (MeDALL) project. We examined epigenome-wide methylation using Illumina Infinium Human Methylation450 BeadChips (450K) in whole blood in 207 children with asthma and 610 controls at age 4–5 years, and 185 children with asthma and 546 controls at age 8 years using a cross-sectional case-control design. After identification of differentially methylated CpG sites in the discovery analysis, we did a validation study in children (4–16 years; 185 cases and 2949 controls) from six additional European cohorts and meta-analysed the results. We next investigated whether replicated CpG sites in cord blood predict later asthma in 1316 children. We subsequently investigated cell-type specificity of the asthma association of the replicated CpG sites 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 with asthma, eight controls [age 2–56 years]) and compared this with whole-blood DNA samples of 74 individuals with asthma and 93 controls (age 1–79 years). Whole-blood transcriptional profiles associated with replicated CpG sites were annotated using RNA-seq data of subsets of peripheral blood mononuclear cells sorted by fluorescence-activated cell sorting.

Findings 27 methylated CpG sites were identified in the discovery analysis. 14 of these CpG sites were replicated and passed genome-wide significance (p=1.4E-10) after meta-analysis. Consistently lower methylation levels were observed at all associated loci across childhood from age 4 to 16 years in participants with asthma, but not in cord blood at birth. All 14 CpG sites were significantly associated with asthma in the second replication study using whole-blood DNA, and were strongly associated with asthma in purified eosinophils. Whole-blood transcriptional signatures associated with these CpG sites indicated increased activation of eosinophils, effector and memory CD8 T cells and natural killer cells, and reduced number of naïve T cells. Five of the 14 CpG sites were associated with asthma in respiratory epithelial cells, indicating cross-tissue epigenetic effects.

Interpretation Reduced whole-blood DNA methylation at 14 CpG sites acquired after birth was strongly associated with childhood asthma. These CpG sites and their associated transcriptional profiles indicate activation of eosinophils and cytotoxic T cells in childhood asthma. Our findings merit further investigations of the role of epigenetics in a clinical context.

Funding EU and the Seventh Framework Programme (the MeDALL project).

Introduction Asthma is a heterogeneous chronic inflammatory airway disease, characterised by variable respiratory symptoms and reversible airflow limitation. Worldwide, more than 300 million people have asthma, with substantial morbidity, reduced quality of life (of patients and their families), and substantial health-care costs. Asthma originates predominantly in early childhood. According to twin studies, about 50–60% of asthma susceptibility is explained by genetic factors. Large-scale genome-wide association studies (GWAS) have uncovered several genetic variants related to asthma, some of which are specific to childhood-onset asthma. The dramatic increase in asthma prevalence in the past half-century
Research in context

Evidence before this study

We searched for articles in PubMed published in English up to Oct 17, 2017, with the search terms “epigenetics”, “asthma”, and “children”, and selected key candidate and epigenome-wide studies to be reviewed in more detail. We also searched for relevant references in review articles from experts in the field. We found evidence that epigenetic changes, including DNA methylation CpG sites, non-coding RNAs, and microRNAs had been associated with asthma in previous studies. However, it was largely unknown which DNA methylation sites were robustly associated with asthma across the genome and whether age-specific patterns can be observed in children from different countries with varying environmental exposures and lifestyles.

Added value of this study

To our knowledge, this is the largest epigenetics study on asthma to date, using data from more than 5000 children in ten cohorts. By applying stringent genome-wide significance criteria (p<1·14×10⁻⁸), our study identified reduced DNA methylation levels in 14 CpG sites to be associated with asthma across childhood from ages 4 to 16 years, but not at birth. Asthma associated whole blood DNA methylation profiles was strongly driven by lower methylation within eosinophils (average 18% difference in methylation levels in those with asthma vs those without), highlighting the importance of the eosinophil as an epigenetic determinant of asthma in childhood. Through clustering of the 14 asthma-associated CpG sites with whole blood and immune-cell-specific gene expression signatures, we were also able to annotate the asthma-associated CpG sites to activated immune cell subsets (CD8-positive T cells and natural killer cells).

Implications of all the available evidence

There is now accumulating evidence that methylation status in peripheral blood cells, as well as in disease-relevant tissues such as airway epithelial cells, is altered in patients with asthma and other respiratory diseases. Our study unequivocally shows that reduced whole-blood DNA methylation at specific CpG sites acquired after birth is strongly associated with childhood asthma. These CpG sites and their associated transcriptional profiles indicate activation of eosinophils and cytotoxic T cells. Although we were not able to assess causal links between methylation status and asthma development in our study, it is reasonable to believe that, by targeting methylation status of identified genes (ie, normalising levels), asthma-related disease activity may be regulated. These observations, together with evidence presented before our study, merit further investigation of the role of epigenetics in a clinical context.

Methods

Study design and participants

Whole-blood DNA from children from four MeDALL birth cohorts—BAMSE (Barn/Children, Allergy, Milieu, Stockholm, Epidemiology; Sweden), ECA (Étude des Déterminants pré et post natals du développement et de la santé de l’Enfant; France), INMA (Infancia y Medio Ambiente-Sabadel [SAB]; Spain), and PIAMA (Prevention and Incidence of Asthma and Mite Allergy; the Netherlands)—was investigated cell-type-specific methylation of the identified CpG sites in eosinophils and respiratory epithelial cells and their related gene-expression signatures.

All blood DNA samples from children not participating in the discovery phase were selected for the replication phase from seven MeDALL birth cohorts: BIB (UK), ECA (Norway), Karelia (Finland), PIAMA, RHEA (Greece), ROBBIC (Italy; subsets Bologna and Rome), and INMA (Spain; subsets Gipuzkoa [GIP], Menorca [MEN]) and Valencia [VAL]; appendix pp 18–19, 306–07). All of these cohorts, apart from Karelia, were included in the main meta-analysis.

We defined asthma as described previously in the MeDALL study, requiring two of three criteria present:
doctor diagnosis of asthma ever; use of asthma medication in the past 12 months; or wheezing or breathing difficulties in the past 12 months.

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

**Procedures**

Given the strong effects of age on whole-blood methylation patterns in childhood,\(^a\) we considered the possibility of age-specific methylation in asthma and performed a discovery analysis at age 4–5 years and age 8 years. We examined epigenome-wide DNA methylation using Illumina Infinium Human Methylation 450 BeadChips (450K) in 207 children with asthma and 610 controls at age 4–5 years, and 185 children with asthma and 546 controls at age 8 years using a cross-sectional case-control design (figure 1).

We used an independent technology, iPlex (Agena Biosciences), to validate and replicate the selected CpG sites after quality control of the assays. The most significant CpG sites at both age 4–5 years and 8 years were selected for replication (appendix pp 20–22). Next we investigated whether replicated CpG sites in cord blood predict later asthma in 1316 children participating in BIB, EDEN, INMA-VAL, and ROBBIC (appendix p 23).

We studied cell-type specificity of the asthma association of the replicated CpG sites in 455 respiratory epithelial cell samples, collected by nasal brushing of 455 16-year-old children, as well as in DNA isolated from blood eosinophils (16 cases with asthma, eight controls [age 2–56 years]) and compared this with whole-blood DNA samples of 167 individuals (age 1–79 years) from families of the Saguenay–Lac-Saint-Jean (LSJ) region in Canada (appendix p 23).\(^b\)

The association of CpG methylation with single nucleotide polymorphisms (SNPs) was studied in 496 children, and with matched regional whole-blood gene-expression in 119 children (age 4 years, INMA), and 260 children (age 16 years, BAMSE) using Affymetrix Human Transcriptome Array 2.0 Genechips.\(^c\) Association between CpG methylation and whole-genome expression by RNA-seq was assessed in the BIOS consortium dataset.\(^d\) CpG sites were annotated by GREAT (Genomic Regions Enrichment of Annotations Tool) version 2.0.2.\(^e\) Functional enrichment analysis was done by overlapping our replicated CpG sites with histone marks and chromatin states of 27 blood cell types in the Roadmap epigenome project.\(^f\) Gene expression sets associated with replicated CpG sites were interrogated for cell-type specificity using the 500FG dataset.\(^g\)

**Statistical analyses**

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

We did inverse variance-weighted fixed-effects meta-analyses with METAL,\(^h\) or random-effects meta-analysis using Metafor\(^i\) package with statistical software environment R in case of a heterogeneity \(\chi^2\) test p value less than 0.05. Our replicated CpG sites 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 (\(p<1.14 \times 10^{-7}, 439\) 106 tests) after meta-analysis of all studies.

---

**Figure 1: Study design**

The EWAS study consisted of a discovery and replication phase, followed by a meta-analysis. The discovery phase was a case–control design with epigenome-wide DNA methylation in whole-blood DNA measured by Illumina 450K. The top significant CpG sites from the discovery phase were selected for replication using the iPlex design in asthma cases and controls from six additional European cohorts (\(n=3196\)). Finally, results from the discovery and replication phases were meta-analysed.
Role of the funding source

The funders of the study had no role in the design of the study, data gathering, analysis, interpretation, writing of the report, or in the decision to submit the report for publication. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

At age 4–5 years, we identified one genome-wide significant CpG site (cg07019303) annotated to ETTV6 and BCL2L14 (p=7.96×10⁻⁸; appendix pp 20, 308). Therefore, we used a looser threshold (false discovery rate [FDR]-adjusted p<0.10) to select 11 CpG sites for replication (appendix p 20) as we considered age-specific effects on asthma. We identified 26 CpG sites that passed the genome-wide significance threshold (p<1.14×10⁻⁷) at age 8 years (figure 2B, appendix pp 21–22, 308), with the strongest association (p=8.82×10⁻⁹) found for cg03695871 in AP5B1, a gene associated with eczema.35 Cg11456013 and cg1362844 were identified at both ages 4–5 years and 8 years (figure 2B, appendix pp 21–22, 308), with the strongest association (p=8.82×10⁻⁹) of cg03695871 in AP5B1, a gene associated with eczema.35 Cg11456013 and cg1362844 were identified at both ages 4–5 years and 8 years, and 35 CpG sites were selected for replication.

IPlex assays for 27 out of the 35 CpG sites selected were tested for association with childhood asthma. The red dotted horizontal line represents the Bonferroni-corrected threshold (p<1.14×10⁻⁷) of genome-wide significance. All 14 replicated CpG sites are marked in green dot and annotated with CpG site name. (A) Results for the EWAS analysis in children aged 4–5 years (n=817) from the four European cohorts BAMSE, EDEN, INMA and PIAMA. (B) Results for the EWAS meta-analysis in children aged 8 years (n=731) from BAMSE and PIAMA.
The expression and function of genes annotated to replicated CpG sites are described in the appendix (pp 30–35), as are the genomic annotations of all CpG sites in discovery cohorts (table 1, figure 2A, appendix pp 24–29). The most significant association (p=2.55×10⁻⁹) and regional co-methylation plots (appendix pp 30–35), as are the genomic annotations of all CpG sites in discovery cohorts (appendix p 336), and no bimodal potential SNP effects within the probe for the 14 replicated sites (p heterogeneity>0.05; SNP effect within the probe. Additionally, the residuals in the discovery cohorts (appendix p 332–35). To evaluate replication at age 4 years were consistently lower than those in controls for all participants with asthma (table 2), However, methylation of the 14 asthma-associated CpG sites in 1316 cord-blood DNA samples did not significantly predict asthma development at age 3–4 years in four different cohorts (appendix p 39). In our discovery samples, white blood cell composition was measured in the INMA cohort (appendix p 40) and estimated in all four cohorts (appendix p 40). Asthma was significantly associated with lower estimated CD4-positive and CD8-positive T-cell subsets in whole blood, but not with blood eosinophil counts. However, the association of CpG methylation with asthma became less significant after eosinophil and neutrophil correction (p values after correction ranged from 1.23×10⁻⁶ to 2.79×10⁻⁹ before correction; see http://bejerano.stanford.edu/great/). In the SLSJ cohort, we found all 14 CpG sites within eosinophils to be significantly associated with asthma in individuals with a very broad age range of 2–56 years (p=1.92×10⁻⁴, Fisher’s exact test; appendix pp 24–29). Using Reinius’ blood-cell-type specific methylation data, we found that the 11 of the 14 asthma-associated CpG sites had relatively lower methylation levels in eosinophils compared with other blood cell types (p=9.2×10⁻⁴, Fisher’s exact test; appendix pp 41–42).

In the SLSJ cohort, we found all 14 CpG sites within eosinophils to be significantly associated with asthma in individuals with a very broad age range of 2–56 years (appendix p 23). CpG methylation in eosinophils was on average 18% lower in participants with asthma (table 2, appendix pp 366–79). Furthermore, we found significant differential methylation in relation to asthma for five of 14 CpG sites in nasal respiratory epithelial cells from children aged 16 years (p corrected=3.8×10⁻⁴; table 2). Four of 14 CpG sites were associated with SNPs within 250 kb (cis-MeQTL; appendix p 43), including

discovery cohorts (table 1, figure 2A, appendix pp 24–29). The expression and function of genes annotated to replicated CpG sites are described in the appendix (pp 30–35), as are the genomic annotations of all CpG sites and regional co-methylation plots (appendix pp 309–30). The most significant association (p=2.55×10⁻⁹) was observed for cg01901579 in DICER1 (appendix p 331). Specific replication at age 4 years is provided in the appendix (p 36), but could not be done at age 8 years due to low power. Stratification of asthma cases at age 4 years by the presence of specific IgE levels to aeroallergens showed association of both specific IgE-positive and IgE-negative asthma with the top 11 CpG sites (from the 4-year discovery analyses), with, on average, stronger results for specific IgE-positive asthma (appendix pp 37–38). Methylation levels in patients with asthma were consistently lower than those in controls for all 14 replicated CpG sites, with small effects in whole blood in the discovery cohorts (appendix pp 332–35). To evaluate potential SNP effects within the probe for the 14 replicated CpG sites, the β value distributions were visually assessed in the discovery cohorts (appendix p 336), and no bimodal distribution was detected that could reveal an underlying SNP effect within the probe. Additionally, the residuals were normally distributed. Homogeneity testing showed consistent effects across childhood (4–16 years) for all replicated sites (p homogeneity>0.05; I² <25%; appendix pp 337–365). Strikingly, all 14 CpG sites were significantly associated with asthma in a second replication study using whole-blood DNA from Canadian families (age 1–79 years; table 2). However, methylation of the 14 asthma-associated CpG sites in 1316 cord-blood DNA samples did not significantly predict asthma development at age 3–4 years in four different cohorts (appendix p 39). In our discovery samples, white blood cell composition was measured in the INMA cohort (appendix p 40) and estimated in all four cohorts (appendix p 40). Asthma was significantly associated with lower estimated CD4-positive and CD8-positive T-cell subsets in whole blood, but not with blood eosinophil counts. However, the association of CpG methylation with asthma became less significant after eosinophil and neutrophil correction (p values after correction ranged from 1.23×10⁻⁶ to 2.79×10⁻⁹ before correction; see http://bejerano.stanford.edu/great/). In the SLSJ cohort, we found all 14 CpG sites within eosinophils to be significantly associated with asthma in individuals with a very broad age range of 2–56 years (appendix p 23). CpG methylation in eosinophils was on average 18% lower in participants with asthma (table 2, appendix pp 366–79). Furthermore, we found significant differential methylation in relation to asthma for five of 14 CpG sites in nasal respiratory epithelial cells from children aged 16 years (p corrected=3.8×10⁻⁴; table 2). Four of 14 CpG sites were associated with SNPs within 250 kb (cis-MeQTL; appendix p 43), including
Articles

The Netherlands Twin register, we analysed the association of SLC29A1, levels of seven CpG sites, and included genes such as the first cluster was inversely associated with methylation patterns in whole blood by hierarchical clustering. The (pp 49–238). We identified three distinct gene expression transcripts are shown per CpG site in the appendix. All associated transcripts 351 [SD 228]). All associated transcripts are shown per CpG site in the appendix (pp 49–238). We identified three distinct gene expression patterns in whole blood by hierarchical clustering. The first cluster was inversely associated with methylation levels of seven CpG sites, and included genes such as SLC29A1, SIGLEC8, IL5RA, and ADORA3 (figure 3A, appendix pp 49–238). A second cluster was strongly associated with cg16592897 and six other CpG sites, 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, whereas 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 (figure 3B, appendix pp 239–305). An RNA-seq dataset of peripheral blood mononuclear cells fluorescence-activated cell-sorted into highly defined leucocyte subsets identified the second cluster as a naive CD4 and CD8 T cell gene signature that was decreased, and the third cluster as an effector and memory CD8 T cell and natural killer cell profile that was increased with reduced CpG methylation (appendix pp 380–82).

Discussion

This large consortium-based meta-analysis identified 14 CpG sites in whole blood to be associated with childhood asthma. Consistently lower methylation of these CpG sites was observed in individuals with asthma from age 4–5 years until adolescence. Strongly reduced methylation at asthma-associated CpG sites within isolated eosinophils clearly indicated that lower whole-blood DNA methylation levels in asthma were due not only to altered eosinophil numbers in whole blood, but also to reduced methylation at those specific CpG sites in 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

| Gene name       | Whole blood coefficient* | Whole blood p value† | Eosinophils coefficient‡ | Eosinophils p value§ | Nasal epithelial coefficient¶ | Nasal epithelial p value|||
|-----------------|--------------------------|----------------------|--------------------------|----------------------|-------------------------------|--------------------------|
| cg01501579      | -0.0188                  | 1.44×10⁻⁴            | -0.180                   | 0.0030               | -0.0122                        | 0.0008                   |
| cg12628444      | -0.0152                  | 5.08×10⁻⁴            | -0.225                   | 0.0009               | -0.0159                        | 0.0037                   |
| cg19764973      | -0.0072                  | 0.0001               | -0.264                   | 0.0228               | -0.0113                        | 0.0052                   |
| cg01445399      | -0.0153                  | 0.0005               | -0.212                   | 0.0191               | -0.0124                        | 0.0148                   |
| cg13835688      | -0.0159                  | 7.54×10⁻⁴            | -0.216                   | 0.0004               | -0.0120                        | 0.0190                   |
| cg11456013      | -0.0147                  | 0.0004               | -0.189                   | 0.0023               | -0.0069                        | 0.2074                   |
| cg15144640      | -0.0089                  | 0.0009               | -0.236                   | 0.0038               | -0.0035                        | 0.3413                   |
| cg01770400      | -0.0050                  | 0.0111               | -0.140                   | 0.0017               | -0.0032                        | 0.3610                   |
| cg06481820      | -0.0107                  | 0.0060               | -0.185                   | 0.0064               | 0.0046                         | 0.4942                   |
| cg0131767       | -0.0106                  | 0.0117               | -0.215                   | 0.0037               | -0.0045                        | 0.6359                   |
| cg01042874      | -0.0201                  | 1.16×10⁻⁴            | -0.191                   | 0.0006               | -0.0019                        | 0.6745                   |
| cg08085999      | -0.0148                  | 0.0014               | -0.229                   | 0.0092               | 0.0027                         | 0.7121                   |
| cg16592897      | -0.0066                  | 0.0223               | -0.070                   | 0.0021               | -0.0004                        | 0.9072                   |
| cg14011077      | -0.0013                  | 0.0205               | -0.108                   | 0.0050               | -0.0100                        | 0.9099                   |

rs9425436, which was associated with CpG methylation of cg01770400 in SERPINC1 (p=5·6×10⁻⁴). rs9425436 was in strong linkage disequilibrium (r²=0·95) with rs4652298 in ZBTB37, which has previously been shown to be associated with asthma (p=2·7×10⁻³) in the GABRIEL GWAS. Mediation analysis revealed that 14·4% of the SNP rs9425436 effect on asthma was due to CpG methylation of cg01770400 (appendix p 43). 13 of 14 replicated CpG sites were significantly enriched for enhancer markers in whole-blood Roadmap and Encode data (appendix pp 44–45).

Four asthma-associated CpG sites were significantly associated with gene expression in cis in BAMSE (appendix p 46), but this was not replicated in the smaller INMA dataset (appendix p 47). In the larger BIOS dataset, which contains participants from the Dutch LifeLines Deep study, and the Netherlands Twin register, we analysed the association of the 14 CpG sites with gene expression in cis in 2367 adults, and identified ten of the 14 asthma-associated CpG sites to be associated with 22 gene transcripts (appendix p 48), with two CpG-gene transcript pairs (cg06483830-SCPEP1 and cg14011077-PPP1R26) also observed in the BAMSE dataset (appendix p 46).

All CpG sites showed significant associations with whole-blood gene expression (mean number of associated transcripts 351 [SD 228]). All associated transcripts are shown per CpG site in the appendix (pp 49–238). We identified three distinct gene expression patterns in whole blood by hierarchical clustering. The first cluster was inversely associated with methylation levels of seven CpG sites, and included genes such as SLC29A1, SIGLEC8, IL5RA, and ADORA3 (figure 3A, appendix pp 49–238). A second cluster was strongly associated with cg16592897 and six other CpG sites, 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, whereas 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 (figure 3B, appendix pp 239–305). An RNA-seq dataset of peripheral blood mononuclear cells fluorescence-activated cell-sorted into highly defined leucocyte subsets identified the second cluster as a naive CD4 and CD8 T cell gene signature that was decreased, and the third cluster as an effector and memory CD8 T cell and natural killer cell profile that was increased with reduced CpG methylation (appendix pp 380–82).

Discussion

This large consortium-based meta-analysis identified 14 CpG sites in whole blood to be associated with childhood asthma. Consistently lower methylation of these CpG sites was observed in individuals with asthma from age 4–5 years until adolescence. Strongly reduced methylation at asthma-associated CpG sites within isolated eosinophils clearly indicated that lower whole-blood DNA methylation levels in asthma were due not only to altered eosinophil numbers in whole blood, but also to reduced methylation at those specific CpG sites in 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
CpG sites to (activated) eosinophils and CD8-positive T cells and natural killer cells.

DNA methylation can be strongly affected by ageing and environmental exposures.8,9 We observed remarkably consistent patterns of DNA methylation associated with childhood asthma across different age groups and in participants from different countries. This suggests that our asthma-associated CpG sites are consistent findings in childhood asthma from the age of 4 years onwards, but we did not find this association at birth, indicating postnatal effects on methylation of the 14 CpG sites identified in our study. Recently, cord-blood methylation of a CpG site in another asthma gene, \( SMAD3 \), was found to be associated with childhood asthma in children of mothers with asthma.10 Maternal smoking during pregnancy, a risk factor for childhood asthma, showed strong effects on methylation of 6073 CpG sites in DNA isolated from cord blood. One of these CpG sites was associated with asthma in our study (LMAN2), but 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 to be the result of postnatal environmental influences, pathophysiological processes related to asthma, or both. We performed detailed functional analysis of the asthma-associated CpG sites to start identifying these processes.

Part of the reduced DNA methylation of the 14 CpG sites in asthma was explained by eosinophil numbers.

Figure 3: Heatmap and hierarchical clustering of \( z \)-scores of the genes most significantly associated with methylation levels of the 14 CpG sites

(A) Selection of the top 99 associated genes using BIOS data from 2367 individuals. The heatmap was plotted on \( z \)-scores from the meta-analysis of three cohorts (Lifelines Deep, Leiden Longevity Study, and Netherlands Twin Register). (B) Selection of the top 99 genes after correction for eosinophil counts in the model. The heatmap was plotted on \( z \)-scores from meta-analysis of the same three cohorts.
since the significance of the association was reduced after correction for predicted eosinophil percentage in peripheral blood. Importantly, strongly reduced methylation of the CpG sites 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 individuals with asthma than in those without asthma) compared with whole blood (on average 1–2% lower in individuals with asthma than in those without asthma), consistent with a functional effect on gene expression in eosinophils. As predicted, eosinophil counts were not significantly different between patients with asthma and controls in our discovery cohorts, suggesting that our results reflect the presence of a different subset or activation state of eosinophils, rather than a change in eosinophil counts per se. Although eosinophils are linked to sub-phenotypes of asthma such as the T helper 2-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 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 CpG sites were significantly enriched for enhancer markers, again suggesting a cell-type-specific regulatory role in gene expression.

Correlation of our 14 CpG sites with genome-wide gene expression revealed three clusters of associated genes. Because all replicated 14 CpG sites show consistently reduced DNA methylation in patients with asthma, the cluster encompassing known eosinophil signature genes that are negatively associated with methylation levels of asthma-associated CpG sites is therefore predicted to show increased expression in asthma. A second cluster with increased predicted expression in asthma is characterised by effector and memory CD8 T cell and natural killer cell genes. A third cluster with reduced predicted expression in asthma was identified as naive CD4 and CD8 T cells. These data indicate that whole-blood DNA from children with asthma carries CpG methylation marks associated with reduced activity of naive T cells and increased activity of effector and memory CD8 T cells and natural killer cells. This asthma-associated shift in cellular activity might be due to different environmental factors such as traffic-related air pollution or infectious agents, including intracellular bacteria or viruses. 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 natural killer 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. Given that large methylation differences were seen in eosinophils from individuals with asthma compared with those without asthma, we believe that our findings could have a direct clinical impact in the future by aiding diagnostics, guiding drug discovery programmes, or both. For example, rapid tests for eosinophil methylation status available for clinical use could be a valuable diagnostic tool to identify children with asthma, and we encourage such new research efforts. Although we are not able to assess causal links between methylation status and asthma development in our study, it is reasonable to believe that, by targeting the methylation status of identified genes (ie, normalising levels), asthma-related disease activity might be regulated.

Comparing the 14 CpG sites with a DNA methylation study of asthma in US inner cities, we did not identify overlapping CpG sites. However, three of the genome-wide significant CpG sites (cg01770400, cg0503329, cg0947105; appendix pp 20–22) were previously reported in an EWAS for total IgE, compatible with shared regulation of these CpG sites in asthma and IgE-related traits. None of these annotated genes was previously implicated in (genetic) studies of asthma (see appendix, pp 30–35, for an overview); however, several of these genes (DICER1, STX3, LIPIN1) have known functions that might be relevant to asthma development. The most significant CpG site was annotated to DICER1 (appendix p 331). DICER1 is a member of the ribonuclease III family, involved in the generation of microRNAs, which modulate gene expression at the post-transcriptional level. SYNTAXIN3 (STX3) is relevant for epithelial polarity and regulates exocytosis. Epithelial polarity is important to maintain epithelial barrier function, a process that is impaired in asthma. STX3 plays an important part in the release of chemokines CXC8, CCL2, CCL3, and CCL4 by human mast cells. LIPIN1 is a Mg²+-dependent phosphatidic acid phosphohydrolase, which has been shown to negatively regulate mast cell degranulation and the anaphylactic response through inhibiting the PKC–SNAP-23 pathway. LIPIN1 also has a role in macrophage proinflammatory activation during Toll-like receptor (TLR) signalling. After TLR4 stimulation, Lipin-1-deficient macrophages showed a decreased production of diacylglycerol and activation of MAP kinases and AP-1, accompanied by a reduced production of proinflammatory cytokines such as IL-6, IL-12, and IL-23, or enzymes such as inducible nitric oxide synthase.

Finally, LMAN2 (lectin, mannose binding 2) encodes a type 1 transmembrane lectin that shuttles between the endoplasmic reticulum, the Golgi apparatus, and the plasma membrane, and has previously been implicated in sphingolipid transport. Our study has some strengths and some limitations. To our knowledge, this is the first study to combine EWAS with an extensive validation and replication study of CpG
sites using an independent laboratory method in more than 5000 children. One limitation of our work is that we used a questionnaire-based approach to define asthma at age 4–5 years based on doctor’s diagnosis, symptoms, and use of asthma medication. Although we acknowledge that self-reported asthma at age 4–5 years could be sensitive to misclassification of transient wheezing, our finding that eight of the 11 CpG sites selected from the age 4–5 years discovery sample were significantly associated with asthma throughout childhood strongly suggests that we found robust asthma CpG sites. However, we cannot ignore that the challenges of diagnosing asthma in young children might partly explain why we identified only one genome-wide CpG site 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 4–5 years and 8 years. Although different significance cutoffs were used for this initial selection (FDR <0.1 at age 4–5 years, Bonferroni significance at age 8 years), this did not affect the final outcome because all CpG sites passed robust criteria for replication (Bonferroni significance in the replication analysis, as well as in the final combined meta-analysis of discovery and replication studies).

Additionally, replication analyses of our 14 top hits in the SLSJ cohort allowed us not only to assess methylation profiles in purified eosinophils, but also to show that hypomethylation of these CpG sites was observed in participants with asthma 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 CpG sites using whole-methylome sequencing in the future. Third, we can address CpG loci but not CpG regions in our study since CpG regions cannot be replicated in the targeted iPlex design. Fourth, 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 five of the 14 CpG sites in respiratory nasal epithelial cells, but cannot exclude the possibility that (part of) this signal is due to admixture of eosinophils in epithelial brushings.

In conclusion, our study identified consistently reduced DNA methylation levels in 14 CpG sites to be associated with asthma across childhood from ages 4 to 16 years, but not at birth. Asthma-associated whole-blood DNA methylation profiles were strongly driven by lower methylation within eosinophils, highlighting the importance of eosinophils as an epigenetic determinant of asthma. Additionally, CpG methylation patterns identified an early-life shift from naïve T cell populations towards effector and memory CD8 and natural killer 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 cell and natural killer cell subsets acquired in childhood.

Contributors
GHK and EM conceived the project and coordinated the analyses. CJX analysed epigenome-wide data for discovery and replication. GHK, EM, CJX, CS and MCN drafted the manuscript. MCN contributed to the biological interpretation. CS, OG, SKM, AK, GP, DG, LR, JK, and 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 provided the biological materials and phenotype information of EDE. UG, JCG, HAS, and BB provided 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, MFP, DG, and DP provided the biological materials and phenotype information of ROBBIC. KCL, VH, PM, and RKC provided the biological materials and phenotype information of ECA. TH, HA, NF, and TI provided the biological materials and phenotype information of Karelia. OG implemented the shadow analysis of EWAS. CL designed, recruited, evaluated, and managed the SLSJ study. AM, MFM, WOCMC, and CL 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, SAI, PrDv, CCoD, and CS designed and produced the 450K and iPLEX data. JMA and JB coordinated the MeDALL project. All authors were involved in data interpretation, and read and approved the manuscript.

Declaration of interests
JP reports grants from the European Commission, during the conduct of the study; JB reports personal fees from Almirall, AstraZenecca, Chiesi, GSK, Meda, Menarini, Merck, MSD, Novartis, Sanofi-Aventis, Takeda, Teva, and Uitch, outside of the submitted work. MCN reports grants from EU FP7 program - MeDALL project, during the conduct of the study; and grants from GSK Ltd and personal fees from DC4U, outside of the submitted work. CI reports a pending patent named “Refractory asthmatics with high levels of IgE and low levels of methylation”. EM reports grants from the EU (ERC) during conduct of the study, and lecture fees from Thermo Fisher Scientific and Meda outside of the submitted work. GHK reports grants from the EU and grants from Lung Foundation of the Netherlands, during the conduct of the study; and grants from TETRI foundation, grants from UBBO EMMIUS foundation, and grants from TEVA the Netherlands, outside of the submitted work. All other authors declare no competing interests.

Acknowledgments
We are grateful to all children and families who 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). ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya.