University Medical Center Groningen

University of Groningen

Epigenome-Wide Association Study of DNA Methylation and Adult Asthma in the Agricultural Lung Health Study

Hoang, Thanh T; Sikdar, Sinjini; Xu, Cheng-Jian; Lee, Mi Kyeong; Cardwell, Jonathan; Forno, Erick; Imboden, Medea; Jeong, Ayoung; Madore, Anne-Marie; Qi, Cancan

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Hoang, T. T., Sikdar, S., Xu, C-J., Lee, M. K., Cardwell, J., Forno, E., Imboden, M., Jeong, A., Madore, A-M., Qi, C., Wang, T., Bennett, B. D., Ward, J. M., Parks, C. G., Beane-Freeman, L. E., King, D., Motsinger-Reif, A., Umbach, D. M., Wyss, A. B., ... London, S. J. (2020). Epigenome-Wide Association Study of DNA Methylation and Adult Asthma in the Agricultural Lung Health Study. *European Respiratory Journal*, *56*(3), [2000217]. https://doi.org/10.1183/13993003.00217-2020

Epigenome-wide association study of DNA methylation and adult asthma in the Agricultural Lung Health Study

Thanh T. Hoang [1,19], Sinjini Sikdar [1,2,19], Cheng-Jian Xu [3,4,19], Mi Kyeong Lee [1], Jonathan Cardwell [5], Erick Forno [6,7], Medea Imboden [8,9], Ayoung Jeong [8,9], Anne-Marie Madore [10], Cancan Qi [11], Tianyuan Wang [12], Brian D. Bennett [12], James M. Ward [12], Christine G. Parks [1], Laura E. Beane-Freeman [13], Debra King [14], Alison Motsinger-Reif [15], David M. Umbach [15], Annah B. Wyss [1], David A. Schwartz [5], Juan C. Celedón [6,7], Catherine Laprise [10,16,17], Carole Ober [18], Nicole Probst-Hensch [8,9], Ivana V. Yang [5], Gerard H. Koppelman [11] and Stephanie J. London [1]

ABSTRACT Epigenome-wide studies of methylation in children support a role for epigenetic mechanisms in asthma; however, studies in adults are rare and few have examined non-atopic asthma. We conducted the largest epigenome-wide association study (EWAS) of blood DNA methylation in adults in relation to non-atopic and atopic asthma.

We measured DNA methylation in blood using the Illumina MethylationEPIC array among 2286 participants in a case-control study of current adult asthma nested within a United States agricultural cohort. Atopy was defined by serum specific immunoglobulin E (IgE). Participants were categorised as atopy without asthma (n=185), non-atopic asthma (n=673), atopic asthma (n=271), or a reference group of neither atopy nor asthma (n=1157). Analyses were conducted using logistic regression.

No associations were observed with atopy without asthma. Numerous cytosine–phosphate–guanine (CpG) sites were differentially methylated in non-atopic asthma (eight at family-wise error rate (FWER) $p < 9 \times 10^{-8}$, 524 at false discovery rate (FDR) less than 0.05) and implicated 382 novel genes. More CpG sites were identified in atopic asthma (181 at FWER, 1086 at FDR) and implicated 569 novel genes. 104 FDR CpG sites overlapped. 35% of CpG sites in non-atopic asthma and 91% in atopic asthma replicated in studies of whole blood, eosinophils, airway epithelium, or nasal epithelium. Implicated genes were enriched in pathways related to the nervous system or inflammation.

We identified numerous, distinct differentially methylated CpG sites in non-atopic and atopic asthma. Many CpG sites from blood replicated in asthma-relevant tissues. These circulating biomarkers reflect risk and sequelae of disease, as well as implicate novel genes associated with non-atopic and atopic asthma.

Introduction

Asthma is a heterogeneous, chronic disease of the airways that affects over 300 million people worldwide [1]. Although asthma has a substantial familial component, only a small portion of the variation in disease risk and heritability is explained despite large genome-wide association studies (GWAS) [2, 3]. Epigenetic mechanisms may contribute to the unexplained variation and the best-studied epigenetic modification is DNA methylation.

Most epigenome-wide association studies (EWAS) of methylation in relation to asthma have been conducted in children [4]; however, asthma risk factors and pathogenetic mechanisms vary between children and adults [3]. Compared to childhood asthma, a higher proportion of adult asthma is non-atopic and pathogenesis of asthma may differ by atopy status. The few published EWAS in adults have fewer than 80 asthma cases, measured DNA methylation using the older Illumina Infinium 27K BeadChip or 450K BeadChip, or did not stratify by atopy status.

Using DNA from blood, we conducted the largest EWAS of DNA methylation and adult asthma to date, using the more comprehensive Illumina Infinium MethylationEPIC BeadChip, which assesses methylation at over 850 K cytosine–phosphate–guanine (CpG) sites. To better elucidate the pathogenesis of asthma, we stratified by atopy, defined objectively by specific immunoglobulin E (IgE). We assessed the potential functional impact of the differentially methylated CpG sites through enrichment of functional genomic features, pathway analyses, associations with gene expression and identification of druggable targets. We replicated our differentially methylated CpG sites in blood using existing studies with methylation assessed in whole blood, purified eosinophils, nasal epithelium, or bronchial airway epithelium.

Methods

Study population

Participants were enroled in the Agricultural Lung Health Study (ALHS), a case-control study of adult current asthma nested within the Agricultural Health Study (AHS). The AHS is a cohort of farmers and their spouses from Iowa and North Carolina. Details of the AHS and ALHS have been described previously [5, 6]. The ALHS enroled 3301 participants in the period 2009–2013.

Based on responses to an AHS questionnaire administered in the period 2005–2010 (data version P3REL201209.00), the ALHS enroled asthma cases using three definitions: 1) self-reported current diagnosed asthma without any self-reported diagnosis of chronic obstructive pulmonary disease (COPD) or emphysema (n=876); 2) potential undiagnosed asthma identified by self-report of current asthma symptoms or use of asthma medication without COPD or emphysema diagnosis among never smokers or light, former smokers (⩽10 pack-years) (n=309); or 3) current self-reported diagnosed asthma with COPD or emphysema diagnosis among never smokers or light, former smokers (⩽10 pack-years) (n=38). Controls (n=2078) were randomly selected from the AHS participants without the above criteria. A full description of the methods can be found in the supplementary material.

Affiliations: [1]Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Dept of Health and Human Services, Research Triangle Park, NC, USA. [2]Dept of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA. [3]Centre for Individualised Infection Medicine (CiiM), Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany. [4]Centre for Experimental and Clinical Infection Research (TWINCORE), Hannover Medical School and the Helmholtz Centre for Infection Research, Hannover, Germany. [5]Dept of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA. [6]Division of Pulmonary Medicine, UPMC Children's Hospital of Pittsburgh, Pittsburgh, PA, USA. [7]Dept of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [8]Chronic Disease Epidemiology Unit, Dept of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, Basel, Switzerland. [9]Dept of Public Health, University of Basel, Basel, Switzerland. [10]Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, Canada. [11]Dept of Pediatric Pulmonology and Pediatric Allergy, University Medical Center Groningen, University of Groningen, Beatrix Children's Hospital and GRIAC Research Institute, Groningen, The Netherlands. [12]Integrative Bioinformatics Support Group, National Institutes of Health, Dept of Health and Human Services, Research Triangle Park, NC, USA. [13]Occupational and Environmental Epidemiology Branch, National Cancer Institute, Bethesda, MD, USA. [14]Clinical Pathology Group, National Institute of Environmental Health Sciences, National Institutes of Health, Dept of Health and Human Services, Research Triangle Park, NC, USA. [15]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, Dept of Health and Human Services, Research Triangle Park, NC, USA. [16]Centre Intersectoriel en Santé Durable, Département des Sciences Fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, Canada. [17]Dept of Pediatrics, Centre Intégré Universitaire de Santé et de Services Sociaux du Saguenay–Lac-Saint-Jean, Saguenay, QC, Canada. [18]Dept of Human Genetics, University of Chicago, Chicago, IL, USA. [19]Joint first authors.

Correspondence: Stephanie J. London, National Institute of Environmental Health Sciences, PO Box 12233, MD A3-05, Research Triangle Park, NC 27709, USA. E-mail: london2@niehs.nih.gov

All participants provided informed consent and the Institutional Review Board at the National Institutes of Health approved this study.

Stratification by atopy

Atopy was determined by a positive blood IgE test (based on IgE $\geqslant 0.70$ IU·mL^{-1}) [7] to at least one of 10 common antigens: Bermuda grass, ragweed, Timothy grass, mountain cedar, *Alternaria*, dust mite, cat dander, milk, egg and wheat. IgE was measured at ImmuneTech (Foster City, CA, USA) using the Luminex platform (Luminex Corp., Austin, TX, USA). Asthma case-control status was stratified by atopy to categorise individuals into one of four mutually exclusive groups: atopy alone (individuals with atopy but without asthma), non-atopic asthma (individuals with asthma but no atopy), atopic asthma (individuals with asthma and atopy), or non-cases (individuals with neither asthma nor atopy).

DNA methylation and quality control

The DNA was bisulfite converted using the EZ-96 DNA Methylation kit (Zymo Research Corp., Irvine, CA, USA). Methylation was assessed for 2391 ALHS participants using the Infinium MethylationEPIC BeadChip, following Illumina's protocol.

Sample level quality control excluded 102 participants with either 1) >5% of CpG sites with detection p-values greater than 1.0×10^{-10} or 2) intensity values less than three standard deviations below the mean bisulfite control intensities. In addition, one participant was excluded due to a sex mismatch. CpG sites were removed if >5% of samples had detection p-values greater than 1.0×10^{-10} (n=31 533). Background correction and dye-bias correction were done using "Relic" in ENmix [8, 9]. Data were normalised using inter-array (quantile) normalisation [8]. Probe-type bias adjustment was conducted using the "Rcp" function in ENmix [10]. We corrected for batch effect (*i.e.* plate effect) using "ComBat" in sva [11]. We used "gaphunter" in minfi to trim extreme methylation outliers at individual CpG sites (threshold=0·3, outCutoff=0·0025) [12]. We analysed 817 235 autosomal CpG sites.

Among the 2286 participants with methylation data passing quality control and with complete information on smoking history, there were 185 with atopy alone, 673 with non-atopic asthma, 271 with atopic asthma and 1157 non-cases (supplementary figure E1).

Cell type proportions

Details of the peripheral blood smears have been described previously [13] and were available in 1894 participants. Monocytes, lymphocytes, neutrophils and eosinophils were counted in smears without platelet clumping and <20% smudged cells. Cell type proportions were calculated from counts for 1658 participants. For individuals without cell counts, cell type proportions were estimated from methylation data using the Houseman method [14] with the REINIUS *et al.* [15] reference panel.

Identification of differentially methylated CpG sites

Epigenome-wide analyses were conducted using logistic regression separately for three outcomes: atopy alone, non-atopic asthma and atopic asthma. Non-cases were the referent group. Untransformed β methylation values were used as the predictor. All analyses were adjusted for age (continuous), sex, body mass index (BMI) (continuous), smoking status (never, former, or current), pack-years of smoking (number of packs smoked per day•years of smoking, continuous), state of residence (North Carolina or Iowa) and four cell type proportions (monocyte, lymphocyte, neutrophil and eosinophil). As farmers were less likely to be enroled in the busy autumn harvest season, we also adjusted for autumn enrolment (yes/no). Significance was assessed using a family-wise error rate (FWER) of $p < 9 \times 10^{-8}$ [16], as well as a Benjamini–Hochberg false discovery rate (FDR) of less than 0.05 [17].

Several sensitivity analyses were conducted. To evaluate possible influences of asthma-related genotypes on the associations, we calculated a weighted polygenic risk score for asthma using the results from a large genome-wide meta-analysis of asthma [2]. Single nucleotide polymorphisms (SNPs) with $p < 5 \times 10^{-8}$ were clumped using PLINK ($r^2 = 0.5$) to identify independent loci within ±250 kb [18]. Methylation analyses were repeated adjusting for the polygenic risk score of asthma. We also repeated the analyses adjusting for current farming status (*i.e.* currently farming crops, working with farm animals, or working with pesticides). Analyses were conducted using the seven Houseman estimated cell types (monocyte, NK, B-cell, CD4$^+$, CD8$^+$, neutrophil and eosinophil) instead of the four measured cell types. As we used a stricter threshold to classify atopy than some previous studies, we conducted methylation analyses after re-stratifying asthma case-control status with an atopy cut-off of $\geqslant 0.35$ IU·mL^{-1}. Analyses were conducted in R version 3.4.0 (www.r-project.org).

Examination of polymorphic probes and additional filtering
We excluded cross-reactive probes [19], probes with SNPs at the extension base (minor allele frequency >1%) [20] and "ch" probes from downstream analyses (*i.e.* enrichment of functional genomic features, pathway analyses, expression quantitative trait methylation, druggable targets and replication look-ups). We visually inspected distributions of significant CpG sites for departure from unimodality.

Enrichment of functional genomic features
We localised CpG sites to four genomic features: CpG islands, CpG island shores, promoters and transcription factor binding sites. We evaluated enrichment or depletion for these features among FDR-significant CpG sites using a two-sided Fisher's exact test and applied a Bonferroni threshold ($p<0.05/4=0.0125$). Enrichment of transcription factor motifs was identified using eFORGE TF [21]. We used eFORGE version 2.0 to identify tissue-specific and cell type-specific enrichment in blood, lung and fetal lung for the following functionally relevant genomic features: DNase I hypersensitive sites, 15 chromatin states and five histone marks [21].

Pathway analyses
We conducted pathway analyses using the "gsameth" function in the missMethyl package in R [22], which accounts for the differing number of probes per gene. Pathways within the Kyoto Encyclopedia of Genes and Genomes (KEGG) gene sets from the Gene Set Enrichment Analysis Molecular Signatures Database (GSEA MSigDB) [23–25] were evaluated. We reported pathways with a nominal p-value of less than 0.05.

Expression quantitative trait methylation (eQTM) analyses
We evaluated association of significant CpG sites with expression of nearby genes (*cis*-eQTMs). We found no studies with both gene expression and blood methylation based on the EPIC array. Therefore, we evaluated eQTMs for the significant CpG sites present on the 450K array. Illumina 450K methylation and gene expression data from RNA-seq were available in blood samples from 3075 adults in the Biobank-based Integrative Omics Study (BIOS) consortium [26]. We evaluated *cis*-eQTMs using expression transcripts within ±250 kb of each significant CpG site. Significance was assessed at FDR less than 0·05.

Search for druggable targets among novel asthma genes
A previous systematic literature review [27] identified asthma-related genes from EWAS or a large asthma GWAS and we updated their search to incorporate more recent studies. Genes not previously associated with asthma were annotated to the ChEMBL database version 25 (released on January 02, 2019) to identify approved or in development drugs that target novel genes implicated by our significant CpG sites [28].

Replication of findings in blood, eosinophils, nasal and bronchial epithelium
We looked up our FDR significant CpG sites identified in atopic asthma in two studies with blood methylation: the Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults (SAPALDIA) and the Pregnancy and Childhood Epigenetics Consortium (PACE) meta-analysis [27]; in four studies with nasal methylation: the Epigenetic Variation and Childhood Asthma in Puerto Ricans (EVA-PR) study [29], the Inner-City Asthma Consortium (ICAC) study [30], the Prevention and Incidence of Asthma and Mite Allergy (PIAMA) study [29] and Project Viva [31]; in one study with eosinophil methylation: the Saguenay–Lac-Saint-Jean (SLSJ) study [32, 33]; and in one study with bronchial epithelium methylation [34]. FDR significant CpG sites identified in non-atopic asthma were looked up in SAPALDIA, the PACE meta-analysis, the EVA-PR study, Project Viva and the SLSJ study.

Results
Table 1 includes descriptive characteristics of the study population. The median age across the four outcome groups ranged from 60–63 years. Participants with atopic asthma were more likely to have adequately controlled asthma than participants with non-atopic asthma (58% *versus* 45%).

Atopy alone
No CpG sites were differentially methylated among individuals with atopy without asthma (n=185) compared to non-cases (neither atopy nor asthma) (n=1157) at the FWER ($p<9\times10^{-8}$) or a FDR less than 0.05 ($\lambda=0.97$) (supplementary figure E2). Using a lower atopy cut-off ($\geqslant0.35$ IU·mL^{-1}), we compared 528 participants with atopy alone to 814 non-cases and no CpG sites were significant. No additional analyses were conducted.

Non-atopic asthma
For non-atopic asthma participants (n=673) compared to non-cases (n=1157), eight CpG sites were differentially methylated using the FWER ($p<9\times10^{-8}$) and 524 at a FDR less than 0.05 (figure 1). The top

TABLE 1 Characteristics of participants in the Agricultural Lung Health Study (ALHS) stratified by asthma and atopy status (n=2286)

Characteristic	Non-cases[#] (n=1157)	Atopy alone (n=185)	Non-atopic asthma (n=673)	Atopic asthma (n=271)
Sex				
Male	625 (54.0)	110 (59.5)	300 (44.6)	138 (50.9)
Female	532 (46.0)	75 (40.5)	373 (55.4)	133 (49.1)
Smoking status				
Never	754 (65.2)	121 (65.4)	456 (67.8)	192 (70.9)
Former	342 (29.6)	56 (30.3)	201 (29.9)	68 (25.1)
Current	61 (5.3)	8 (4.3)	16 (2.4)	11 (4.1)
State of residence				
Iowa	839 (72.5)	114 (61.6)	502 (74.6)	179 (66.1)
North Carolina	318 (27.5)	71 (38.4)	171 (25.4)	92 (33.4)
Autumn enrolment				
Yes	219 (18.9)	45 (24.3)	179 (26.6)	75 (27.7)
No	938 (81.1)	140 (75.7)	494 (73.4)	196 (72.3)
Currently farming				
Yes	797 (68.9)	130 (70.3)	440 (65.4)	180 (66.4)
No	360 (31.1)	55 (29.7)	233 (34.6)	91 (33.6)
Asthma control				
Adequately controlled	–	–	305 (45.3)	156 (57.6)
Inadequately controlled	–	–	131 (19.5)	44 (16.2)
Missing data	–	–	237 (35.2)	71 (26.2)
ICS use				
Yes	–	–	278 (41.3)	103 (38.0)
No	–	–	395 (58.7)	168 (62.0)
Age at home visit years	63 (55–72)	60 (55–68)	62 (55–71)	60 (52–68)
Pack-years of smoking[¶]	14.5 (2.3–35)	7.3 (0.9–25)	6 (1.5–19)	7.7 (1.5–21.3)
BMI kg·m^{-2}	28.8 (25.6–32.6)	29.0 (25.3–32.5)	30.5 (26.9–35.1)	29.6 (26.1–33.9)
Cell type				
Monocyte	0.08 (0.06–0.10)	0.08 (0.06–0.11)	0.07 (0.05–0.10)	0.08 (0.06–0.10)
Neutrophil	0.56 (0.48–0.63)	0.56 (0.47–0.63)	0.57 (0.48–0.65)	0.57 (0.50–0.65)
Eosinophil	0.01 (0.00–0.03)	0.02 (0.00–0.05)	0.02 (0.00–0.04)	0.02 (0.00–0.04)
Lymphocyte	0.34 (0.27–0.42)	0.34 (0.26–0.40)	0.33 (0.25–0.40)	0.32 (0.25–0.40)

Data are presented as n (%) or median (IQR). ICS: inhaled corticosteroids; BMI: body mass index; IQR: interquartile range. [#]: non-cases are participants with neither asthma nor atopy; [¶]: pack-years among former and current smokers.

30 CpG sites are shown in table 2, while all FDR CpG sites are given in supplementary table E1. In 516 of the 524 differentially methylated CpG sites, methylation was lower in non-atopic asthma participants compared to non-cases. Systematic inflation was minimal (λ=1.14) (supplementary figure E3). All differentially methylated CpG sites had a unimodal methylation distribution.

Results were not appreciably altered after adjusting for the polygenic risk score of asthma or current farming status (supplementary table E2). Repeating analyses using the seven estimated cell types, instead of the four measured and estimated cell types, resulted in greater inflation (λ=1.66; 127 CpG sites were significant at the FWER, while 2243 were for a FDR less than 0.05) (supplementary table E2). When individuals were recategorised using the less stringent atopy cut-off ($\geqslant$0.35 IU·mL^{-1}) (483 non-atopic asthma participants *versus* 814 non-cases), the number of associations decreased (nine CpG sites at a FDR less than 0.05) (supplementary table E2).

In case-only analyses (*i.e.* those restricted to non-atopic asthma participants), we calculated adjusted odds ratios (ORs) for the 524 differentially methylated CpG sites, comparing 278 inhaled corticosteroid (ICS) users to 395 non-users and found no significant associations (none had FDR less than 0.05). Adjusted ORs comparing 131 non-atopic asthma participants with inadequately controlled asthma to 305 with controlled asthma identified differential methylation (FDR less than 0.05) by asthma control at 17 of the 524 CpG sites (OR range 0.63–0.95) (supplementary table E3).
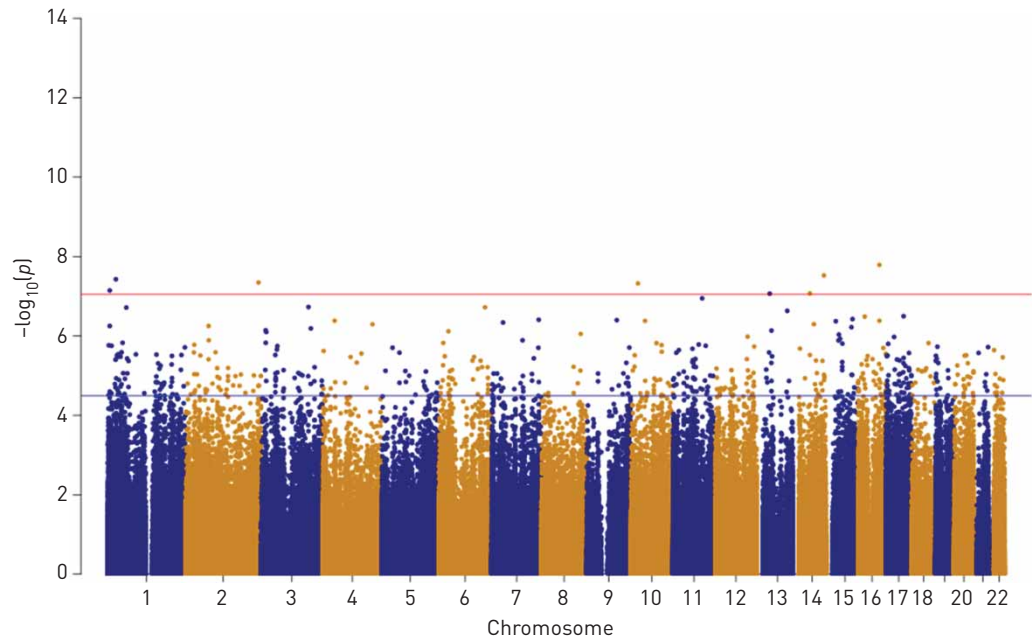
FIGURE 1 Manhattan plot of epigenome-wide analysis of non-atopic asthma. The red line represents the family-wise error rate (FWER) threshold and the blue line represents the false discovery rate (FDR) threshold. Analyses were adjusted for age, sex, body mass index (BMI), smoking status, pack-years of smoking, state of residence, autumn enrolment and four cell types (monocyte, lymphocyte, neutrophil and eosinophil).

Functional enrichment

Removing 15 potentially cross-reactive CpG sites [19] left 509 for downstream analysis. These CpG sites were depleted for CpG islands ($p=5.4\times10^{-27}$), CpG island shores ($p=0.0004$) and promoters ($p=1.7\times10^{-18}$) (supplementary table E4), enriched for transcription factor binding sites ($p=1.4\times10^{-13}$) and enriched (FDR less than 0.05) for two transcription factor motifs (V_FOXO3A_Q1 and V_AIRE_01) (supplementary table E5). Enrichment of DNase I hypersensitivity was observed in blood, fetal lung and lung (supplementary figure E4). Of the 15 chromatin states, we observed enrichment for active transcription start sites in blood, as well as enrichment for enhancers, weak transcription and strong transcription in blood and lung (supplementary figure E4). Among histone marks, we found enrichment for H3K4me1 in blood, fetal lung and lung, for H3K4me3 in blood and fetal lung, and for H3K36me3 in blood (supplementary figure E4).

Pathway analyses

Nine pathways showed enrichment ($p<0.05$), including Alzheimer's disease, amyotrophic lateral sclerosis, long-term potentiation, vascular smooth muscle contraction and calcium signalling (supplementary table E6 and supplementary figures E5 and E6). Two asthma relevant pathways had p-values between 0.05 and 0.10: asthma ($p=0.08$) and sphingolipid metabolism ($p=0.055$).

cis-eQTM

Of the 509 CpG sites associated with non-atopic asthma, 169 CpG sites were on the Illumina 450K array and 168 were available in BIOS [26]. 120 CpG sites (71.4%) were significantly associated (FDR less than 0.05) with 377 gene expressions in *cis* (supplementary table E7).

Druggable targets

The 509 differentially methylated CpG sites annotated to 479 genes. After excluding genes identified from previous asthma EWAS and GWAS (supplementary table E8), 382 genes were novel (including *AZU1* and *PDE4B*). In ChEMBL [28], *PDE4B* is a target of several approved drugs related to asthma or other respiratory diseases (*e.g.* Roflumilast, Dyphylline and Theophylline) (supplementary table E9).

Replication

Five replication studies assessed methylation in blood, nasal epithelium, or eosinophils. Replication studies generally had smaller sample sizes and not all stratified by atopy (table 3). Of the 509 differentially methylated CpG sites identified in non-atopic asthma, 169 were on the 450K array and 84 of these (50%)

TABLE 2 Top 30 significant cytosine–phosphate–guanine (CpG) sites in non-atopic asthma

Chromosome	Position [#]	CpG	Gene name [¶]	n	OR [+]	95% CI	p-value	FDR	Coefficient [§]
16	68804850	cg06085527	–	1829	0.93	0.91–0.95	1.62E-08	7.76E-03	−0.0093
14	100610407	cg14084609	DEGS2	1830	0.82	0.76–0.88	2.99E-08	7.76E-03	−0.0033
1	26091858	cg08028384	SELENON	1829	0.93	0.90–0.95	3.73E-08	7.76E-03	−0.0099
2	234608559	cg10180919	UGT1A6	1829	0.89	0.85–0.93	4.52E-08	7.76E-03	−0.006
10	22743835	cg15737719	LOC100499489	1829	0.76	0.68–0.84	4.75E-08	7.76E-03	−0.0023
1	6341287	cg09249800	GPR153	1828	0.93	0.90–0.95	7.16E-08	8.81E-03	−0.0085
14	55603874	cg23575099	LGALS3	1830	0.81	0.76–0.88	8.46E-08	8.81E-03	−0.0029
13	41631052	cg07908654	WBP4	1830	0.86	0.81–0.91	8.62E-08	8.81E-03	−0.0036
11	93456334	cg23338316	SCARNA9	1828	0.87	0.82–0.91	1.13E-07	1.03E-02	−0.0048
3	152215571	cg07952576	TMEM14EP	1828	0.97	0.96–0.98	1.87E-07	1.31E-02	−0.026
6	149450585	cg09639771	TAB2	1830	0.95	0.93–0.97	1.90E-07	1.31E-02	−0.0128
1	59543930	cg18581916	LINC01358	1830	0.95	0.93–0.97	1.93E-07	1.31E-02	−0.0126
13	97846783	cg23933458	LINC00456	1829	0.96	0.94–0.97	2.33E-07	1.46E-02	−0.0158
17	56269767	cg01955639	EPX	1830	0.88	0.84–0.92	3.18E-07	1.58E-02	−0.0038
16	21831372	cg07611887	RRN3P1	1829	0.88	0.84–0.92	3.26E-07	1.58E-02	−0.0053
15	85186517	cg17194668	WDR73	1830	0.86	0.81–0.91	3.76E-07	1.58E-02	−0.0037
7	150773709	cg06807926	FASTK	1828	0.96	0.94–0.98	3.91E-07	1.58E-02	−0.0141
9	95800911	cg03234093	SUSD3	1829	0.95	0.94–0.97	3.99E-07	1.58E-02	−0.013
4	38110810	cg07456972	TBC1D1	1825	0.81	0.75–0.88	4.13E-07	1.58E-02	−0.0025
16	68804845	cg21468244	–	1827	0.95	0.93–0.97	4.14E-07	1.58E-02	−0.0133
10	45495435	cg01614759	ZNF22	1829	0.87	0.82–0.92	4.17E-07	1.58E-02	−0.0037
15	31505024	cg08884974	LINC02352	1829	0.88	0.83–0.92	4.26E-07	1.58E-02	−0.0043
7	36077237	cg24598141	LOC101928618	1827	0.97	0.95–0.98	4.59E-07	1.63E-02	−0.0175
14	68713369	cg01171954	LOC100996664	1828	0.94	0.92–0.96	5.04E-07	1.66E-02	−0.0102
4	159969044	cg15834151	C4orf45	1830	0.93	0.91–0.96	5.08E-07	1.66E-02	−0.0076
2	74612222	cg17988187	DCTN1-AS1	1830	0.89	0.85–0.93	5.62E-07	1.70E-02	−0.0051
1	6341230	cg21220721	GPR153	1828	0.94	0.92–0.97	5.62E-07	1.70E-02	−0.009
15	81623226	cg25466522	TMC3-AS1	1828	0.82	0.76–0.89	6.02E-07	1.76E-02	−0.0024
3	159847104	cg17991030	IL12A-AS1	1829	0.97	0.95–0.98	6.46E-07	1.82E-02	−0.019
3	14708144	cg15106081	C3orf20	1827	0.76	0.68–0.85	7.20E-07	1.93E-02	−0.0017

OR: odds ratio; CI: confidence interval; FDR: false discovery rate. [#]: genome build GRCh37/hg19; [¶]: annotated using HOMER version 4.10.3 and an associated human genome database (hg19, 5v.10) [35]; [+]: logistic regression model adjusted for age, sex, body mass index (BMI), smoking status, pack-years of smoking, state of residence, autumn enrolment and four cell types (monocyte, lymphocyte, neutrophil and eosinophil). ORs indicate the multiplicative change in the odds of having non-atopic asthma in relation to an increase in DNA methylation; [§]: adjusted mean methylation difference based on non-atopic asthma status. The coefficient is from a linear regression model with robust standard error estimation, where methylation is the outcome and non-atopic asthma status is the predictor, adjusting for the same covariates as were used in the logistic regression. A coefficient of −0.01 means that the mean methylation value is 0.01 lower in participants with non-atopic asthma compared to non-cases. Possible methylation values range from zero to one.

were replicated (p<0.05, with the same direction of association) in at least one study (supplementary table E10). Project Viva used the EPIC array, which had an additional 314 CpG sites, of which 56 were replicated (supplementary table E10).

Atopic asthma

Comparing atopic asthma participants (n=271) to non-cases (n=1157), we identified 181 differentially methylated CpG sites using the FWER (p<9×10^{-8}) and 1086 with a FDR less than 0.05 (figure 2). The top 30 CpG sites are shown in table 4 and all FDR CpG sites are given in supplementary table E11). In 99.5% of the 1086 differentially methylated CpG sites, methylation was lower in atopic asthma participants than in non-cases. All differentially methylated CpG sites had a unimodal methylation distribution. Systematic inflation was not observed (λ=0·98) (supplementary figure E7). Of the 1086 CpG sites, 104 (9.6%) were FDR-significant in non-atopic asthma (supplementary table E12).

Adjustment for the polygenic risk score of asthma or current farming status did not materially alter the results (supplementary table E2). When we adjusted for the seven estimated cell types, instead of the four cell types, the λ value increased to 1.21 (407 CpG sites were significant using the FWER and 2947 were significant with a FDR less than 0·05, including 1076 of the original 1086 CpG sites) (supplementary table E11). When we classified atopy using the cut-off ⩾0.35 IU·mL^{-1} (461 atopic asthma participants *versus* 814 non-cases), the

TABLE 3 Summary of look-up replication

Study	Tissue	Study population	Methylation array	Stratified by atopy?	Asthma cases	Controls	Overlap[#]	Replicated[¶]	Replicated and overlap in ALHS asthma	
									Non-atopic	Atopic
Non-atopic asthma										
SAPALDIA	Blood	Adults	450K	Non-atopic	70	134	169 (100)	0 (0)		0
PACE meta-analysis	Blood	Children	450K	Did not stratify	631	2862	168 (99.4)	68 (40.5)		41
EVA-PR	Nasal	Children	450K	Non-atopic	66	104	146 (86.4)	8 (5.5)		5
Project Viva	Nasal	Children	EPIC	Did not stratify	65	398	471 (92.5)	143 (30.4)		86
SLSJ	Eosinophil	Children and adults	450K	Did not stratify	16	8	168 (99.4)	49 (29.2)		39
Atopic Asthma										
SAPALDIA	Blood	Adults	450K	Atopic	91	134	349 (100)	0 (0)	0	
PACE meta-analysis	Blood	Children	450K	Did not stratify	631	2862	347 (99.4)	288 (83.0)	41	
EVA-PR	Nasal	Children	450K	Atopic	169	104	299 (83.3)	235 (78.6)	28	
ICAC	Nasal	Children	450K	Atopic	36	36	349 (100)	222 (63.6)	26	
PIAMA	Nasal	Children	450K	Atopic	27	219	349 (100)	63 (18.1)	16	
Project Viva	Nasal	Children	EPIC	Atopic	36	265	1008 (94.2)	875 (86.8)	86	
SLSJ	Eosinophil	Children and adults	450K	Did not stratify	16	8	338 (96.8)	274 (81.1)	39	
Bronchial epithelium	Airway epithelial	Adults	450K	Did not stratify	74	41	254 (72.8)	49 (19.3)	4	

Data are presented as n or n (%). ALHS: Agricultural Lung Health Study; SAPALDIA: Swiss Cohort Study on Air Pollution and Lung and Heart Diseases in Adults; PACE: Pregnancy and Childhood Epigenetics Consortium; EVA-PR: Epigenetic Variation and Childhood Asthma in Puerto Ricans; ICAC: Inner-City Asthma Consortium; PIAMA: Prevention and Incidence of Asthma and Mite Allergy; SLSJ: Saguenay–Lac-Saint-Jean. [#]: overlap with ALHS false discovery rate (FDR) results. For 450K, percentages were based on 169 cytosine–phosphate–guanine (CpG) sites for non-atopic asthma and 349 CpG sites for atopic asthma. For EPIC, percentages were based on 509 CpG sites for non-atopic asthma and 1070 CpG sites for atopic asthma; [¶]: CpG sites in the replication study with p<0.05 and the same direction of association as observed in the ALHS. Percentages were based on the number of CpG sites that overlap with the replication study.

number of associations decreased (124 CpG sites were significant using the FWER and 847 were significant with a FDR less than 0.05, including 669 of the original 1086 CpG sites) (supplementary table E2).

Conducting adjusted logistic regression analyses restricted to participants with atopic asthma, none of the 1086 significant CpG sites were differentially methylated (all FDR greater than 0.05) in relation to ICS use (103 users *versus* 168 non-users) or asthma control (44 with inadequately controlled asthma *versus* 156 with adequately controlled asthma).

Functional enrichment

The removal of 16 potentially cross-reactive probes [19] left 1070 differentially methylated CpG sites for downstream analysis. We found depletion for CpG islands ($p=4.3\times10^{-66}$), CpG island shores ($p=1.3\times10^{-6}$) and promoters ($p=1.0\times10^{-46}$), enrichment for transcription factor binding sites ($p=0.0057$) (supplementary table E4) and enrichment for eight transcription factor motifs (including GATA3_primary and MA0029.1-Evi1) (supplementary table E5). Among the 15 chromatin states, we identified enrichment of enhancers, genic enhancers, weak transcription and strong transcription in blood and lung (supplementary figure E8). The FDR significant CpG sites were enriched for H3K4me1 and H3K36me3 in blood, fetal lung and lung (supplementary figure E8).

Pathway analyses

Ten KEGG pathways had p<0.05, including insulin signalling, type II diabetes, starch and sucrose metabolism, and valine, leucine, and isoleucine degradation (supplementary table E6 and supplementary figures E5 and E6). The asthma pathway was marginally significant (p=0.055). Figure 3 shows a network plot of nine disease and biological pathways implicated by genes identified in our non-atopic and atopic asthma analyses.
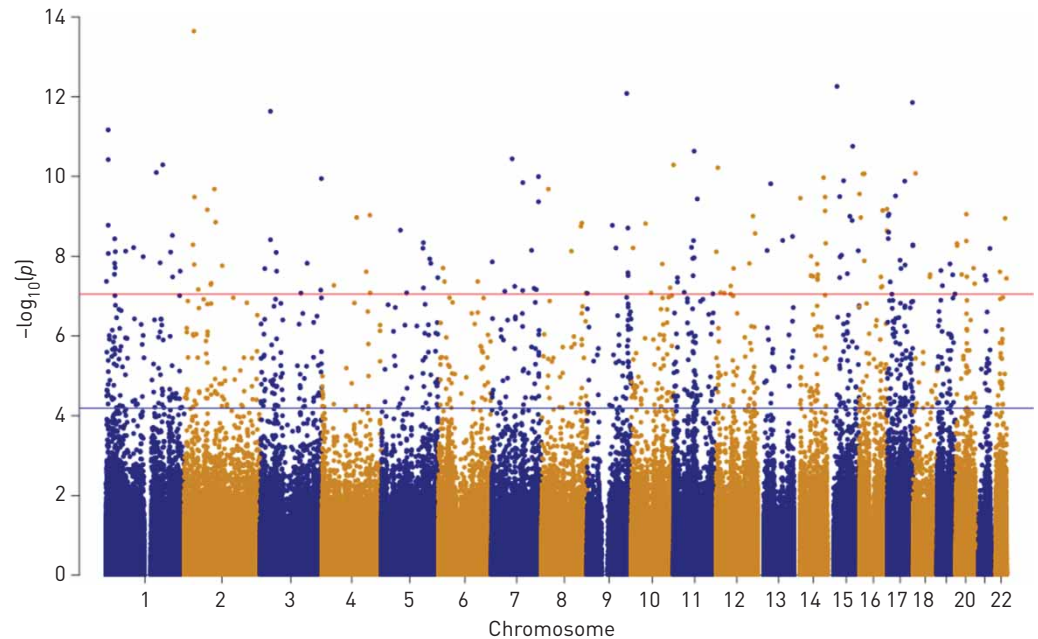
FIGURE 2 Manhattan plot of epigenome-wide analysis of atopic asthma. The red line represents the family-wise error rate (FWER) threshold and the blue line represents the false discovery rate (FDR) threshold. Analyses were adjusted for age, sex, body mass index (BMI), smoking status, pack-years of smoking, state of residence, autumn enrolment and four cell types (monocyte, lymphocyte, neutrophil and eosinophil).

cis-eQTM

Of the 1070 CpG sites associated with atopic asthma, 349 were present in the 450K array and 346 were available in BIOS. 242 CpG sites (69.9%) were significantly associated with 671 gene expressions in *cis* (supplementary table E13).

Druggable targets

The 1070 CpG sites annotated to 885 genes, of which 569 have not been reported in previous asthma EWAS or GWAS (including *PDE4B* and *PPARG*). In ChEMBL [28], *PPARG* is a target of a candidate asthma drug (Rosiglitazone) [36, 37] (supplementary table E14).

Replication

Eight replication studies assessed methylation in blood, nasal epithelium, bronchial epithelium, or eosinophils (table 3). The 1076 CpG sites identified in atopic asthma included 349 on the 450K array and 341 of these (98%) were replicated in at least one study (supplementary table E15). An additional 684 CpG sites were available in Project Viva, which assessed nasal methylation using the EPIC array and 595 of these (87%) were replicated (supplementary table E15).

Discussion

Several hundred CpG sites in blood were differentially methylated in blood DNA from adults with non-atopic or atopic asthma compared to adults with neither asthma nor atopy. Many more CpG sites were differentially methylated in relation to atopic rather than non-atopic asthma and about 10% of those for atopic asthma overlapped with non-atopic asthma. Our findings highlight the importance of distinguishing between non-atopic and atopic types to improve our understanding of asthma and show that differential methylation in atopic asthma is not driven by atopy *per se*.

We performed several analyses to further evaluate our findings. Results were similar after adjustment for the polygenic risk score of asthma, suggesting that our findings are not influenced by known genetic variations of asthma. In most participants, we directly measured differential blood counts for monocytes, lymphocytes, neutrophils and eosinophils. Adjustment for all seven estimated cell types instead, which disaggregates lymphocytes into NK, B-cell, CD4[+] and CD8[+] types, resulted in some inflation; however, results were similar. Eosinophils are higher in individuals with asthma than the general population [38]. Estimation of eosinophils from a small reference panel of healthy participants may not be highly accurate, as suggested in our data (Spearman correlation between measured and estimated=0·09). Using a less stringent atopy cut-off

TABLE 4 Top 30 significant cytosine–phosphate–guanine (CpG) sites in atopic asthma

Chromosome	Position [#]	CpG	Gene name [¶]	n	OR [+]	95% CI	p-value	FDR	Coefficient [§]
2	31302644	cg26382374	GALNT14	1425	0.76	0.71–0.82	2.27E-14	1.86E-08	−0.0101
15	31248701	cg16606719	MTMR10	1428	0.72	0.66–0.79	5.51E-13	2.25E-07	−0.0066
9	126497597	cg01745810	DENND1A	1428	0.76	0.71–0.82	8.29E-13	2.26E-07	−0.0082
17	79851485	cg12380988	ANAPC11	1428	0.74	0.68–0.80	1.40E-12	2.86E-07	−0.0075
3	33112660	cg26396322	TMPPE	1428	0.72	0.66–0.79	2.32E-12	3.79E-07	−0.006
1	6341287	cg09249800	GPR153	1428	0.88	0.85–0.91	6.80E-12	9.26E-07	−0.0161
15	81623226	cg25466522	TMC3-AS1	1426	0.70	0.63–0.78	1.75E-11	2.04E-06	−0.0051
11	65546210	cg05300717	AP5B1	1421	0.87	0.84–0.91	2.32E-11	2.37E-06	−0.0121
7	65439512	cg04290133	GUSB	1422	0.81	0.75–0.86	3.61E-11	3.09E-06	−0.008
1	6341327	cg11699125	GPR153	1426	0.90	0.87–0.93	3.78E-11	3.09E-06	−0.0191
1	180940378	cg11649969	STX6	1428	0.60	0.51–0.70	5.10E-11	3.49E-06	−0.0032
10	135061670	cg12227660	MIR202HG	1419	0.88	0.84–0.91	5.13E-11	3.49E-06	−0.0158
12	6342778	cg20560376	CD9	1427	0.71	0.64–0.79	6.03E-11	3.79E-06	−0.0056
1	160309220	cg09332506	NCSTN	1428	0.53	0.44–0.64	7.96E-11	4.15E-06	−0.0027
18	8720395	cg27550672	MTCL1	1428	0.68	0.61–0.76	8.37E-11	4.15E-06	−0.0038
16	16143033	cg24304533	ABCC1	1426	0.73	0.67–0.81	8.52E-11	4.15E-06	−0.0057
16	11709855	cg24517604	LITAF	1425	0.77	0.71–0.83	8.64E-11	4.15E-06	−0.0068
7	149543136	cg05184016	ZNF862	1427	0.68	0.61–0.76	1.01E-10	4.59E-06	−0.0048
14	95615731	cg01901579	DICER1	1428	0.78	0.72–0.84	1.07E-10	4.60E-06	−0.0074
3	195974300	cg02803925	SLC51A	1424	0.77	0.71–0.83	1.13E-10	4.62E-06	−0.0051
15	52427175	cg05853552	BCL2L10	1428	0.74	0.67–0.81	1.28E-10	4.87E-06	−0.0063
17	55190679	cg13947225	AKAP1	1424	0.81	0.76–0.87	1.31E-10	4.87E-06	−0.0083
7	99507261	cg17717565	TRIM4	1421	0.84	0.79–0.88	1.43E-10	5.08E-06	−0.0103
13	41631052	cg07908654	WBP4	1428	0.78	0.73–0.84	1.53E-10	5.21E-06	−0.0068
2	96965099	cg01923915	SNRNP200	1428	0.73	0.67–0.81	2.07E-10	6.54E-06	−0.005
8	22169041	cg16427256	PIWIL2	1428	0.76	0.69–0.82	2.08E-10	6.54E-06	−0.0052
16	616212	cg04497992	NHLRC4	1428	0.87	0.83–0.91	2.75E-10	8.32E-06	−0.0127
17	25897247	cg06127160	LGALS9	1428	0.87	0.83–0.91	3.08E-10	8.65E-06	−0.0114
15	40093898	cg18852698	LOC105370941	1427	0.82	0.77–0.87	3.20E-10	8.65E-06	−0.007
2	32946576	cg18399629	MIR4765	1428	0.69	0.62–0.78	3.28E-10	8.65E-06	−0.004

OR: odds ratio; CI: confidence interval; FDR: false discovery rate. [#]: genome build GRCh37/hg19; [¶]: annotated using HOMER version 4.10.3 and an associated human genome database (hg19, 5v.10) [35]; [+]: logistic regression model adjusted for age, sex, body mass index (BMI), smoking status, pack-years of smoking, state of residence, autumn enrolment and four cell types (monocyte, lymphocyte, neutrophil and eosinophil). ORs indicate the multiplicative change in the odds of having atopic asthma in relation to an increase in DNA methylation; [§]: adjusted mean methylation difference based on atopic asthma status. The coefficient is from a linear regression model with robust standard error estimation, where methylation is the outcome and atopic asthma status is the predictor, adjusting for the same covariates as were used in the logistic regression. A coefficient of −0.01 means that the mean methylation value is 0.01 lower in participants with atopic asthma compared to non-cases. Possible methylation values range from zero to one.

($\geqslant 0.35$ IU·mL^{-1}) yielded fewer significant findings, consistent with previous evidence that the lower threshold is less discriminatory [7]. Notably, results for atopic asthma were more robust in response to the threshold change (more CpG sites remained significant) than the results for non-atopic asthma.

A substantial proportion of our differentially methylated CpG sites were replicated in at least one published study. A higher proportion were replicated in atopic asthma than in non-atopic asthma, probably reflecting the much larger quantity of studies available for atopic asthma and thus the higher power. Most CpG sites that were replicated in non-atopic asthma overlapped with our findings in atopic asthma (table 3), suggesting that these are asthma-related CpG sites independent of atopy status. Our results in atopic asthma were highly consistent in a small cohort of participants with methylation measured in isolated eosinophils, a critical cell type for atopic asthma. Nasal epithelium is a good surrogate for lower respiratory epithelium [39] and 89% of our findings in atopic asthma were replicated in nasal tissue, suggesting that methylation in blood is a good surrogate for nasal tissue in methylation studies of atopic asthma. This is an important observation, given the wide availability of stored blood for methylation analyses.

In our study and in most of the replication studies, methylation at significant CpG sites tended to be lower in participants with asthma than in individuals without asthma. We hypothesise that lower methylation in individuals with asthma reflects the combination of a cell-type composition which displays lower methylation at these CpG sites [40] and the activation of cell types related to disease processes. For example, there might be a shift from naïve T-cells to effector memory CD8T cells and NK cells [40], or activation of
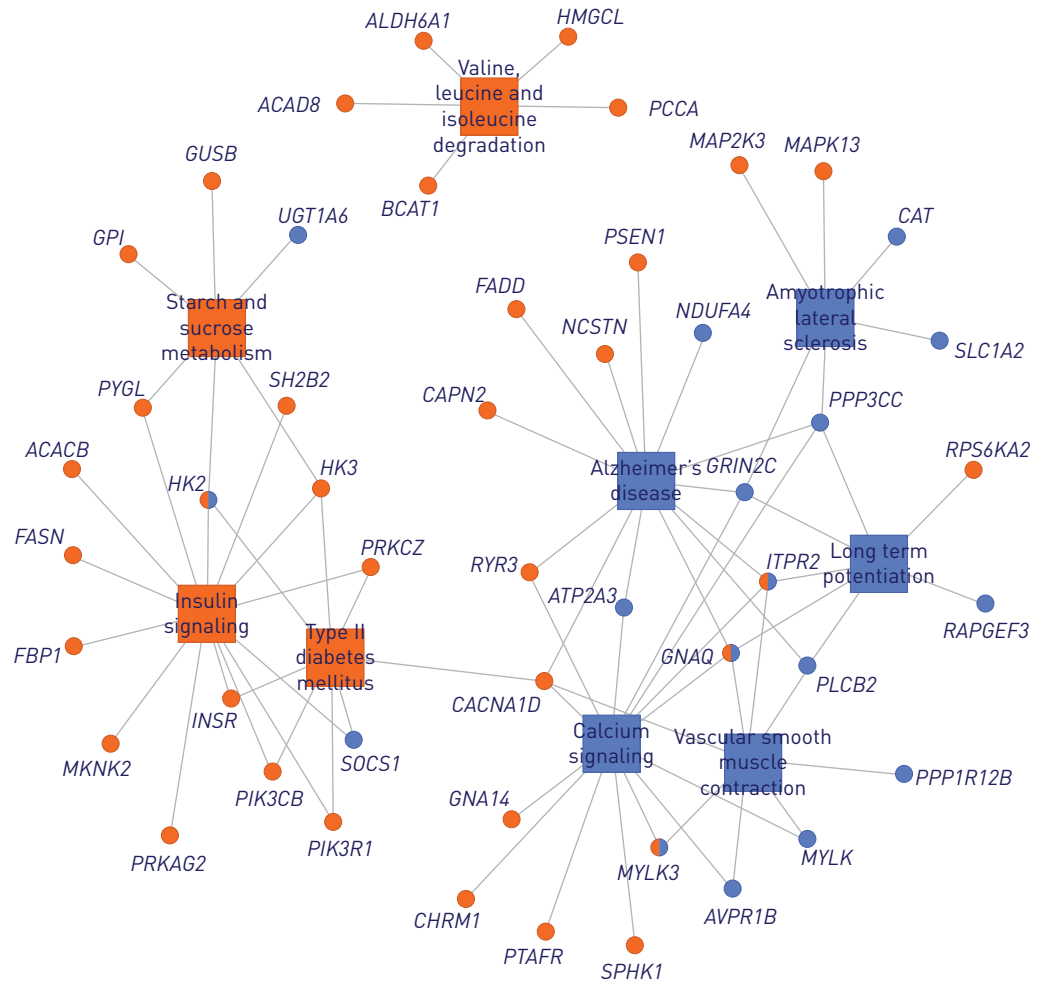
FIGURE 3 Network plot of nine enriched pathways (squares) with connecting implicated genes (circles) from non-atopic and/or atopic asthma. Blue represents genes or pathways implicated in non-atopic asthma. Orange represents genes or pathways implicated in atopic asthma.

basal cells [41]. Mechanistic studies are needed to investigate why methylation is generally lower among individuals with asthma than those without asthma at differentially methylated CpG sites across the studied tissues.

This study provides substantial replication of prior studies. Our analysis of non-atopic asthma participants implicated 479 genes, including 96 identified in previous asthma GWAS or EWAS (supplementary table E16). Of 885 genes implicated in our analysis of atopic asthma, 315 were identified in other GWAS or EWAS of asthma (supplementary table E17). We also identified many novel asthma-related genes not previously implicated in GWAS or EWAS. For example, high expression of *AZU1*, an inflammatory mediator that regulates neutrophils, is correlated with poor asthma control [42]. Additionally, *DEGS2* is involved in sphingolipid metabolism, which has been implicated in respiratory outcomes [43] including asthma [40].

Enrichment of functional genomic features and pathway analyses provide biological insight into our differentially methylated CpG sites. Our results highlighted certain genomic features, histone marks in blood, lung and fetal lung, and transcription factor motifs that might be biologically relevant to asthma. For example, transcription factors FOXO3A and GATA3 help regulate type-2 helper T-cells [44, 45], an important cell type in asthma and EVI-1 interacts with *SMAD3* [46], a known asthma gene. The differentially methylated CpG sites in non-atopic asthma were enriched in signalling pathways related to the nervous system (*e.g.* sphingolipid metabolism, calcium signalling and vascular smooth muscle contraction). Sphingolipids and calcium are signalling molecules involved in vascular smooth muscle contraction [47–50], suggesting that mechanisms related to innervation of the airways may play a greater role in non-atopic asthma *versus* atopic asthma. Atopic asthma CpG sites were enriched in pathways involved in inflammatory response (*i.e.* insulin signalling) [51] or characterised by chronic inflammation (*i.e.* diabetes) [52], consistent with the inflammatory nature of atopic asthma.

Previous studies have identified CpG sites differentially methylated in relation to total serum IgE [33, 53–55], atopy [29, 56, 57] and allergies [56, 58] but, in our study, no CpG sites were differentially methylated in relation to atopy alone. As previous studies have been conducted in children, or a combination of children and adults, and our study is comprised of older adults (median age 62 years), it is possible that differential methylation related to atopy in children may not persist into adulthood.

This study has some limitations. Our population was rural and may differ in environmental factors compared to urban populations (*e.g.* air pollution). However, we found high levels of replication across geographically distinct populations from published studies. Furthermore, heterogeneity in disease phenotypes likely remains even though we stratified asthma by atopy status and, as the study is cross-sectional, we cannot determine if findings reflect the pathogenesis or consequences of asthma.

This study also has several strengths. It is the largest EWAS of adult asthma that has been conducted and the substantial replication in nasal epithelium and purified eosinophils suggests that methylation in blood is a good proxy for asthma-relevant tissues in discovering novel differential methylation in atopic asthma. Pathway analyses support the biological plausibility of our findings and the CpG sites implicated in blood are enriched for localisation to functional genomic elements in blood and lung. Many CpG sites were associated with gene expression in both non-atopic and atopic asthma, further supporting the potential functional impact of the differential methylation we identified at these loci. Different findings for non-atopic and atopic asthma confirm the importance of separating forms of asthma to discover novel associations, which may better inform disease aetiology and lead to improved treatment strategies.

References

1   GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet Respir Med* 2017; 5: 691–706.

2   Demenais F, Margaritte-Jeannin P, Barnes KC, *et al*. Multiancestry association study identifies new asthma risk loci that colocalize with immune-cell enhancer marks. *Nat Genet* 2018; 50: 42–53.

3   Pividori M, Schoettler N, Nicolae DL, *et al*. Shared and distinct genetic risk factors for childhood-onset and adult-onset asthma: genome-wide and transcriptome-wide studies. *Lancet Respir Med* 2019; 7: 509–522.

4   Edris A, den Dekker HT, Melen E, *et al*. Epigenome-wide association studies in asthma: a systematic review. *Clin Exp Allergy* 2019; 49: 953–968.

5   Alavanja MC, Sandler DP, McMaster SB, *et al*. The Agricultural Health Study. *Environ Health Perspect* 1996; 104: 362–369.

6   House JS, Wyss AB, Hoppin JA, *et al*. Early-life farm exposures and adult asthma and atopy in the Agricultural Lung Health Study. *J Allergy Clin Immunol* 2017; 140: 249–256.

7   Van Hoeyveld E, Nickmans S, Ceuppens JL, *et al*. Defining thresholds of specific IgE levels to grass pollen and birch pollen allergens improves clinical interpretation. *Clin Chim Acta* 2015; 450: 46–50.

8   Xu Z, Niu L, Li L, *et al*. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucleic Acids Res* 2016; 44: e20.

9   Xu Z, Langie SA, De Boever P, *et al*. RELIC: a novel dye-bias correction method for Illumina Methylation BeadChip. *BMC Genomics* 2017; 18: 4.

10 Niu L, Xu Z, Taylor JA. RCP: a novel probe design bias correction method for Illumina Methylation BeadChip. *Bioinformatics* 2016; 32: 2659–2663.

11 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; 8: 118–127.

12 Andrews SV, Ladd-Acosta C, Feinberg AP, *et al*. "Gap hunting" to characterize clustered probe signals in Illumina methylation array data. *Epigenetics Chromatin* 2016; 9: 56.

13 Fessler MB, Carnes MU, Salo PM, *et al*. House dust endotoxin and peripheral leukocyte counts: results from two large epidemiologic studies. *Environ Health Perspect* 2017; 125: 057010.

14 Houseman EA, Accomando WP, Koestler DC, *et al*. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; 13: 86.

15 Reinius LE, Acevedo N, Joerink M, *et al*. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One* 2012; 7: e41361.

16 Mansell G, Gorrie-Stone TJ, Bao Y, *et al*. Guidance for DNA methylation studies: statistical insights from the Illumina EPIC array. *BMC Genomics* 2019; 20: 366.

17 Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc B* 1995; 57: 289–300.

18 Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559–575.

19 Pidsley R, Zotenko E, Peters TJ, *et al*. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol* 2016; 17: 208.

20 Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res* 2017; 45: e22.

21 Breeze CE, Reynolds AP, van Dongen J, *et al*. eFORGE v2.0: updated analysis of cell type-specific signal in epigenomic data. *Bioinformatics* 2019; 35: 4767–4769.

22 Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* 2016; 32: 286–288.

23 Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000; 28: 27–30.

24 Kanehisa M, Sato Y, Furumichi M, *et al*. New approach for understanding genome variations in KEGG. *Nucleic Acids Res* 2019; 47: D590–D595.

25 Subramanian A, Tamayo P, Mootha VK, *et al*. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005; 102: 15545–15550.

26 Bonder MJ, Luijk R, Zhernakova DV, *et al*. Disease variants alter transcription factor levels and methylation of their binding sites. *Nat Genet* 2017; 49: 131–138.

27 Reese SE, Xu CJ, den Dekker HT, *et al*. Epigenome-wide meta-analysis of DNA methylation and childhood asthma. *J Allergy Clin Immunol* 2019; 143: 2062–2074.

28   Gaulton A, Hersey A, Nowotka M, *et al.* The ChEMBL database in 2017. *Nucleic Acids Res* 2017; 45: D945–D954.

29   Forno E, Wang T, Qi C, *et al.* DNA methylation in nasal epithelium, atopy, and atopic asthma in children: a genome-wide study. *Lancet Respir Med* 2019; 7: 336–346.

30   Yang IV, Pedersen BS, Liu AH, *et al.* The nasal methylome and childhood atopic asthma. *J Allergy Clin Immunol* 2017; 139: 1478–1488.

31   Cardenas A, Sordillo JE, Rifas-Shiman SL, *et al.* The nasal methylome as a biomarker of asthma and airway inflammation in children. *Nat Commun* 2019; 10: 3095.

32   Laprise C. The Saguenay–Lac-Saint-Jean asthma familial collection: the genetics of asthma in a young founder population. *Genes Immun* 2014; 15: 247–255.

33   Liang L, Willis-Owen SAG, Laprise C, *et al.* An epigenome-wide association study of total serum immunoglobulin E concentration. *Nature* 2015; 520: 670–674.

34   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: e90151.

35   Heinz S, Benner C, Spann N, *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; 38: 576–589.

36   Richards DB, Bareille P, Lindo EL, *et al.* Treatment with a peroxisomal proliferator activated receptor gamma agonist has a modest effect in the allergen challenge model in asthma: a randomised controlled trial. *Respir Med* 2010; 104: 668–674.

37   Sandhu MS, Dimov V, Sandhu AK, *et al.* The use of the peroxisome proliferator-activated receptors γ agonist rosiglitazone to treat airway hyperreactivity. *Ann Allergy Asthma Immunol* 2012; 109: 75–77.

38   Jacinto T, Malinovschi A, Janson C, *et al.* Differential effect of cigarette smoke exposure on exhaled nitric oxide and blood eosinophils in healthy and asthmatic individuals. *J Breath Res* 2017; 11: 036006.

39   Brugha R, Lowe R, Henderson AJ, *et al.* DNA methylation profiles between airway epithelium and proxy tissues in children. *Acta Paediatr* 2017; 106: 2011–2016.

40   Xu CJ, Soderhall C, Bustamante M, *et al.* DNA methylation in childhood asthma: an epigenome-wide meta-analysis. *Lancet Respir Med* 2018; 6: 379–388.

41   Vermeulen CJ, Xu CJ, Vonk JM, *et al.* Differential DNA methylation in bronchial biopsies between persistent asthma and asthma in remission. *Eur Respir J* 2020; 55: 1901280.

42   Kasaian MT, Lee J, Brennan A, *et al.* Proteomic analysis of serum and sputum analytes distinguishes controlled and poorly controlled asthmatics. *Clin Exp Allergy* 2018; 48: 814–824.

43   Chakinala RC, Khatri A, Gupta K, *et al.* Sphingolipids in COPD. *Eur Respir Rev* 2019; 28: 190047.

44   Chen T, Hou X, Ni Y, *et al.* The Imbalance of FOXP3/GATA3 in Regulatory T Cells from the Peripheral Blood of Asthmatic Patients. *J Immunol Res* 2018; 2018: 3096183.

45   Ohkura N, Sakaguchi S. Foxo1 and Foxo3 help Foxp3. *Immunity* 2010; 33: 835–837.

46   Kurokawa M, Mitani K, Irie K, *et al.* The oncoprotein Evi-1 represses TGF-β signalling by inhibiting Smad3. *Nature* 1998; 394: 92–96.

47   Grimmer B, Kuebler WM. The endothelium in hypoxic pulmonary vasoconstriction. *J Appl Physiol* 2017; 123: 1635–1646.

48   Proia RL, Hla T. Emerging biology of sphingosine-1-phosphate: its role in pathogenesis and therapy. *J Clin Invest* 2015; 125: 1379–1387.

49   Sturgill JL. Sphingolipids and their enigmatic role in asthma. *Adv Biol Regul* 2018; 70: 74–81.

50   Erle DJ, Sheppard D. The cell biology of asthma. *J Cell Biol* 2014; 205: 621–631.

51   van Niekerk G, Christowitz C, Conradie D, *et al.* Insulin as an immunomodulatory hormone. *Cytokine Growth Factor Rev* 2020; 52: 34–44.

52   Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. *Nat Rev Immunol* 2011; 11: 98–107.

53   Peng C, Cardenas A, Rifas-Shiman SL, *et al.* Epigenome-wide association study of total serum immunoglobulin E in children: a life course approach. *Clin Epigenetics* 2018; 10: 55.

54   Ek WE, Ahsan M, Rask-Andersen M, *et al.* Epigenome-wide DNA methylation study of IgE concentration in relation to self-reported allergies. *Epigenomics* 2017; 9: 407–418.

55   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* 2017; 140: 571–577.

56   Peng C, Van Meel ER, Cardenas A, *et al.* Epigenome-wide association study reveals methylation pathways associated with childhood allergic sensitization. *Epigenetics* 2019; 14: 445–466.

57   Everson TM, Lyons G, Zhang H, *et al.* DNA methylation loci associated with atopy and high serum IgE: a genome-wide application of recursive Random Forest feature selection. *Genome Med* 2015; 7: 89.

58   Martino D, Joo JE, Sexton-Oates A, *et al.* Epigenome-wide association study reveals longitudinally stable DNA methylation differences in CD4+ T cells from children with IgE-mediated food allergy. *Epigenetics* 2014; 9: 998–1006.