



香港城市大學
City University of Hong Kong

專業 創新 胸懷全球
Professional · Creative
For The World

CityU Scholars

Shared Genetic Architecture of Blood Eosinophil Counts and Asthma in UK Biobank

Li, Boyang; Wang, Yuxuan; Wang, Zixiao; Li, Xinyue; Kay, Shannon; Chupp, Geoffrey L; Zhao, Hongyu; Gomez, Jose L

Published in:
ERJ Open Research

Published: 01/07/2023

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:
CC BY-NC

Publication record in CityU Scholars:
[Go to record](#)

Published version (DOI):
[10.1183/23120541.00291-2023](https://doi.org/10.1183/23120541.00291-2023)

Publication details:
Li, B., Wang, Y., Wang, Z., Li, X., Kay, S., Chupp, G. L., Zhao, H., & Gomez, J. L. (2023). Shared Genetic Architecture of Blood Eosinophil Counts and Asthma in UK Biobank. *ERJ Open Research*.
<https://doi.org/10.1183/23120541.00291-2023>

Citing this paper

Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights

Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission

Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy

Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.



Shared genetic architecture of blood eosinophil counts and asthma in UK Biobank

Boyang Li^{1,7}, Yuxuan Wang^{2,7}, Zixiao Wang¹, Xinyue Li³, Shannon Kay^{4,5}, Geoffrey L. Chupp⁴, Hongyu Zhao^{1,6,8} and Jose L. Gomez^{4,5,8}

¹Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA. ²Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. ³School of Data Science, City University of Hong Kong, Hong Kong SAR, China. ⁴Pulmonary, Critical Care and Sleep Medicine Section, Yale University, New Haven, CT, USA. ⁵Center for Precision Pulmonary Medicine (P2MED), Yale University, New Haven, CT, USA. ⁶Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA. ⁷These authors contributed equally to this work. ⁸These authors share senior authorship.

Corresponding author: Jose L. Gomez (jose.gomez-villalobos@yale.edu)



Shareable abstract (@ERSpublications)

Asthma and blood eosinophil counts are genetically correlated, and the STAT6 pathway is over-represented in this correlation <https://bit.ly/3X6qDFK>

Cite this article as: Li B, Wang Y, Wang Z, et al. Shared genetic architecture of blood eosinophil counts and asthma in UK Biobank. *ERJ Open Res* 2023; 9: 00291-2023 [DOI: 10.1183/23120541.00291-2023].

Copyright ©The authors 2023

This version is distributed under the terms of the Creative Commons Attribution Non-Commercial Licence 4.0. For commercial reproduction rights and permissions contact permissions@ersnet.org

Received: 4 May 2023
Accepted: 7 June 2023

Abstract

Rationale Asthma is a complex, heterogeneous disease strongly associated with type 2 inflammation, and blood eosinophil counts guide therapeutic interventions in moderate and severe asthma. Eosinophils are leukocytes involved in type 2 immune responses. Despite these critical associations between asthma and blood eosinophil counts, the shared genetic architecture of these two traits remains unknown. The objective of the present study was to characterise the genetic architecture of blood eosinophil counts and asthma in the UK Biobank.

Methods We performed genome-wide association studies (GWAS) of doctor-diagnosed asthma, blood eosinophil, neutrophil, lymphocyte and monocyte counts in the UK Biobank. Genetic correlation analysis was performed on GWAS results and validated in the Trans-National Asthma Genetic Consortium (TAGC) study of asthma.

Results GWAS of doctor-diagnosed asthma and blood eosinophil counts in the UK Biobank identified 585 and 3429 significant variants, respectively. *STAT6*, a transcription factor involved in interleukin-4 signalling, was a key shared pathway between asthma and blood eosinophil counts. Genetic correlation analysis demonstrated a positive correlation between doctor-diagnosed asthma and blood eosinophil counts ($r=0.38\pm 0.10$, $\text{correlation}\pm\text{SE}$; $p=4.7\times 10^{-11}$). As a validation of this association, we found a similar correlation between TAGC and blood eosinophil counts in the UK Biobank (0.37 ± 0.08 , $\text{correlation}\pm\text{SE}$; $p=1.2\times 10^{-6}$).

Conclusions These findings define the shared genetic architecture between blood eosinophil counts and asthma risk in subjects of European ancestry and point to a genetic link to the *STAT6* signalling pathway in these two traits.

Introduction

Eosinophils are leukocytes involved in type 2 (T2) immune responses against macroparasites, including helminths and ectoparasites [1]. T2 immune responses are activated in allergic diseases and some asthma endotypes [2]. In addition, eosinophilic airway inflammation in asthma is associated with reduced symptom control, airflow obstruction and exacerbations [3]. Therefore, blood eosinophil counts are an essential component of the clinical evaluation to identify T2 inflammation in patients with uncontrolled asthma. Approved therapeutic strategies targeting specific T2 immune response components include IgE, interleukin (IL)-5, and IL-4/IL-13 blockade [4–6].

Previous genome-wide association studies (GWAS) of asthma have identified links with allergy and T2 immune response traits, including IgE and blood eosinophil counts [7–10]. Thus, eosinophils are



prominent cells involved in disease pathogenesis. Nevertheless, we have an incomplete understanding of the shared genetic associations between eosinophil counts and asthma risk.

We hypothesised that blood eosinophil counts and asthma share a genetic architecture. To test this hypothesis, we performed a GWAS of doctor-diagnosed asthma and blood leukocytes in the UK Biobank. We then performed genetic correlation and colocalisation analyses to identify common loci associated with asthma and blood eosinophil counts.

Methods

Study population and design

This study used data from the UK Biobank. The study design is illustrated in figure 1. Details on the protocol and database have been described elsewhere [11, 12]. In brief, the UK Biobank is a prospective study of >500 000 UK residents aged 40–69 years, recruited during 2006–2010. Informed consent was obtained for all participants. This study focused on data from 361 194 individuals of European ancestry. We examined 11 717 doctor-diagnosed asthma cases (6715 female, 5002 male), 80 070 controls (43 838 female, 36 232 male) and 349 856 phenotypes for eosinophil counts (187 758 female, 162 098 male). Additional traits included the following: lymphocyte counts, monocyte counts and neutrophil counts. For more details, refer to supplementary table S1.

Genome-wide association analysis

GWAS for all phenotypes was performed using a linear mixed model implemented in BOLT-LMM v2.3.4 [13–15]. The model was adjusted for age, sex and 20 genetic principal components as covariates. We performed two additional GWAS for asthma and blood eosinophil counts adding smoking as a covariate. We applied the rank-based inverse normal transformation to the blood cell count phenotypes so that the standardised phenotypes approximated a normal distribution. Single nucleotide polymorphisms (SNPs) were excluded if the minor allele frequency was <1%, Hardy–Weinberg equilibrium p-values < 1×10^{-5} or an imputation score <0.8, with 5 067 514 remaining SNPs after imputation. QQ plots are presented in supplementary figures S1–S5. p-values were adjusted by the genomic inflation factor (λ); specifically, test statistics for each SNP were divided by the genomic inflation factor ($\lambda > 1$). Genome-wide significance was defined as $p < 5 \times 10^{-8}$. SNPs were mapped to the *Homo sapiens* (human) genome assembly GRCh37 (hg19). Significant GWAS results for lymphocytes, monocytes and neutrophils are presented in supplementary table S2. Annotation of significant variants was performed using SNPeff.v5.0 [16]. Genetic loci were defined as the 5000 base pair intervals, upstream and downstream, from each variant. Previous GWAS results were obtained via the GWAS Catalog (www.ebi.ac.uk/gwas/docs/file-downloads, 8 February 2023). Linkage disequilibrium (LD) clumping was performed using the Integrative Epidemiology Unit

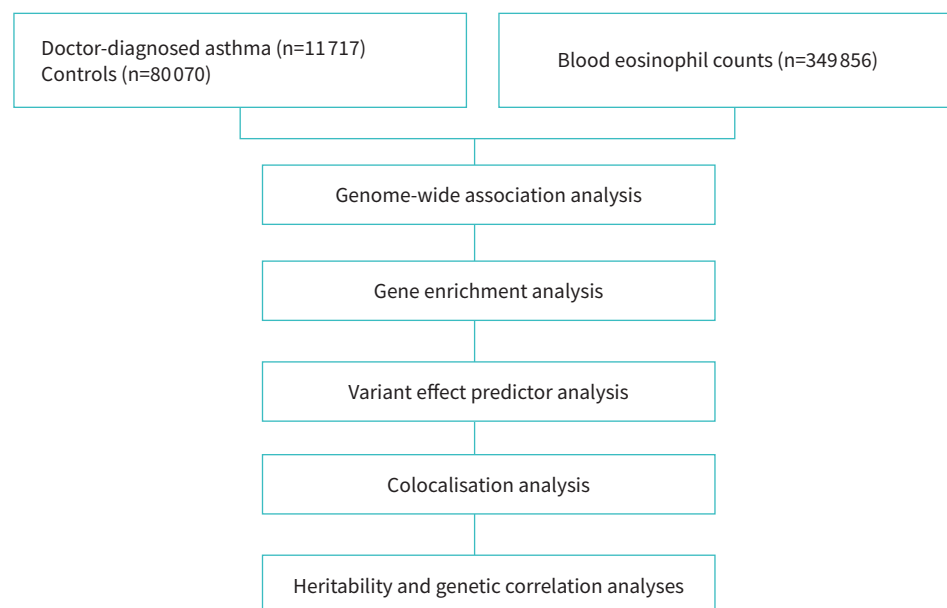


FIGURE 1 Study design.

open GWAS portal (<https://gwas.mrcieu.ac.uk/>). The “ld_clump” function of the ieu_gwas R package [17] was used to query the server with the European panel from the 1000 Genomes Project, with an r^2 threshold of 0.7 and a window size of 10 000 kb.

SNP annotation

Two databases were used to annotate variants identified in the GWAS of doctor-diagnosed asthma, eosinophil counts and colocalisation analyses. The Database of Immune Cell Expression (DICE) focused on the identification of SNPs affecting gene expression (expression quantitative trait loci (eQTLs)) in immune cells [18]. We used information for 15 cell types: B-cell, classical monocyte, nonclassical (M2) monocyte, natural killer cell, naive CD4 T-cell, stimulated CD4 T-cell, naive CD8 cell, stimulated CD8 cell, naive regulatory T-cell (Treg), memory Treg, type 1 T-helper (Th1) cell, Th2 cell, Th17 cell, Th1/17 cell and T follicular helper cell. We used a false discovery rate (FDR) adjusted p-value <0.05 and transcripts per kilobase million >1.0. In this database, if the expression of more than one gene is associated with a SNP, there will be separate entries for each polymorphism–gene combination. A second database, the eQTLGen Consortium, incorporates data from 37 cohorts [19]. We focused on statistically significant cis-eQTLs, FDR<0.05.

Variant effect predictor analysis

We performed a variant effect predictor (VEP) analysis to further annotate GWAS variants [20]. We used the University of California Santa Cruz Genes track with the Ensembl release 109 (February 2023) and selected SIFT, PolyPhen-2 and functional role parameters. Upstream and downstream distance was set at 5000 base pairs.

Gene enrichment analysis

Enrichment analyses were performed with MetaCore (version 22.4.71100; Clarivate Analytics, Philadelphia, PA, USA). The single experiment workflow was implemented for enrichment analyses. We performed pathway maps, Gene Ontology and pathway network enrichment analysis on genes identified in the eQTL analyses of variants associated with doctor-diagnosed asthma and blood eosinophil counts.

Colocalisation analysis

To run the colocalisation analysis, the genome must be divided into approximately independent windows. For each trait, we identified all genome-wide significant variants ($p < 5 \times 10^{-8}$) from the summary statistics as the index variants and sorted them by significance. We then defined a window of ± 100 kb on either side of each index variant. Within each of our defined windows for all the trait pairs, we ran colocalisation using the R coloc package with default parameters using SNPs present in both datasets [21].

Trans-National Asthma Genetic Consortium

Summary statistics from the Trans-National Asthma Genetic Consortium (TAGC) were used for validation. The TAGC study included 23 948 asthma cases and 118 538 controls of diverse genetic ancestries [22]. Specifically, we focused on GWAS associations identified for European ancestry (19 954 asthma cases, 107 715 controls) for comparability with the UK Biobank discovery cohort.

Heritability and genetic correlation analyses

We estimated the SNP-based heritability and genetic correlation among traits in the UK Biobank by applying BOLT-REML to the same set of individuals. BOLT-REML constructed a Bayesian mixture model that used individual-level genotype and phenotype information to estimate variance–covariance components *via* a Monte Carlo algorithm. We applied Wald tests to evaluate the statistical significance of genetic correlations estimated with BOLT-REML. We utilised summary statistics-based LD-score regression [23] to estimate genetic correlations involving the asthma phenotype in the TAGC cohort where individual-level genotype and phenotype data were not available. Specifically, we used the European-ancestry meta results from the random-effects model in the TAGC cohort to estimate genetic correlations with two other traits in the UK Biobank: doctor-diagnosed asthma and blood eosinophil counts. LD scores were estimated using European samples from the 1000 Genomes Project (phase 3) as a reference panel [24, 25]. When computing LD, SNPs with <5% minor allele frequency in the reference panel were filtered.

Results

GWAS of doctor-diagnosed asthma

To identify loci associated with doctor-diagnosed asthma, we used a linear mixed model implemented in BOLT-LMM. This GWAS analysis identified 585 genome-wide significant variants using a $p < 5 \times 10^{-8}$ (supplementary table S3), with $\lambda = 1.047$ (figure 2a and supplementary table S3). A previous GWAS of

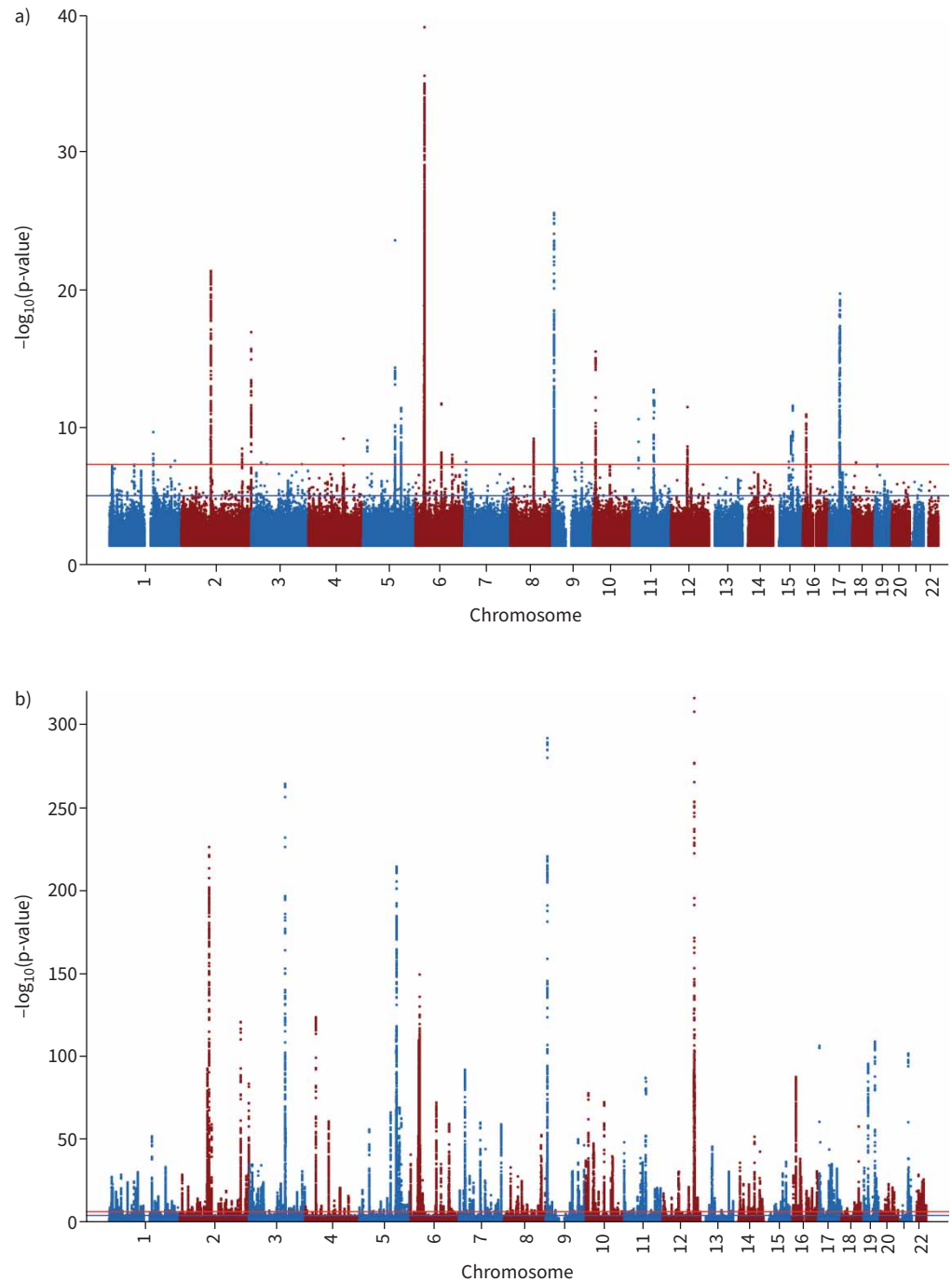


FIGURE 2 a) Manhattan plot of doctor-diagnosed asthma genome-wide association study in UK Biobank ($\lambda=1.047$); b) Manhattan plot of blood eosinophil count genome-wide association study in UK Biobank ($\lambda=1.311$).

asthma in the UK Biobank yielded similar results [26, 27]. To identify the presence of multiple independent signals in a genetic location, we performed a conditional and joint analysis of multiple SNPs associated with doctor-diagnosed asthma (supplementary table S4). We found two independent variants in *IL1RL1*, *ERMP1* and *IL18R1*.

To characterise the effects of variants associated with doctor-diagnosed asthma, two eQTL databases (DICE and eQTLGen) were interrogated. This analysis identified 109 eQTLs (supplementary table S5). Human

leukocyte antigen genes were the most commonly identified ($n=17$), followed by *IL1R1*, *IL1RL1*, *IL18R1*, *GSDMA*, *GSDMB*, *ORMDL3* and *IL18RAP*. All these genes have been found to be associated with asthma [22]. We found novel associations including multiple genes in the complement pathways *C4A* and *C4B*.

We performed VEP analysis to predict the functional effect of these variants and their overlap with regulatory elements. Most variants (70%, $n=409$) had multiple predicted effects (supplementary table S6). Only two variants, rs4193 and rs12722072 in *HLA-DQAI*, were predicted to be stop-gain mutations [28].

To identify biological processes that were enriched for the genes identified in the eQTL analysis, we performed pathway analysis. The top two pathways in this analysis were induction of the antigen presentation machinery by interferon (IFN)- γ (FDR<0.05) and maturation and migration of dendritic cells in skin sensitisation (FDR<0.05) (supplementary table S7).

GWAS of blood eosinophil counts

To identify loci associated with blood eosinophil count, we used a linear mixed model implemented in BOLT-LMM. This analysis identified 3429 genome-wide significant variants $p < 5 \times 10^{-8}$ (figure 2b and supplementary table S8). Annotation with the two eQTL databases identified 1081 unique eQTLs (supplementary table S9). The majority of these transcripts have been previously associated with eosinophil counts or percentages in the GWAS Catalog, largely thanks to the study by ASTLE *et al.* [29]. As compared to the GWAS Catalog, our analysis identified 414 new eQTLs. Similar to asthma eQTLs, we identified associations with complement genes *C4A* and *C4B*. In addition, we observed an association with *IL4*, a key gene involved in T2 inflammation and a drug target for dupilumab [30]. The conditional and joint analysis of multiple SNPs for variants associated with blood eosinophil counts identified 236 unique loci and eQTLs (supplementary table S4). Multiple cytokine receptors *IL1RL1*, *IL2RA*, *IL4R*, *IL5RA*, *CSF2RB*, *IL18R1* and *ACKR2* were identified, in addition to complement genes *C4A* and *C4B*.

We performed a VEP analysis in variants associated with blood eosinophil counts to determine their effect on protein function (supplementary table S10). VEP annotation identified a missense variant, rs61731111, with large effects on blood eosinophil counts ($\beta \pm SE$ 0.13 \pm 0.01), mapped to *S1PR4* (sphingosine-1-phosphate receptor 4). A different missense variant in *S1PR4*, rs3746072, has been previously associated with lower total white blood cell and neutrophil counts [31]. Additional variants with similar effects on eosinophil counts were located in *BCL2*, *RPN1*, *RAB7A*, *GATA2*, *RPN1*, *IL1RL1* and *RUNX1* (supplementary table S7). Together, these analyses demonstrate the clustering of variants with large effects on eosinophil counts in multiple genes involved in T2 inflammation.

Similar to the enrichment for doctor-diagnosed asthma, induction of the antigen presentation machinery by IFN- γ (FDR<0.05) and maturation and migration of dendritic cells in skin sensitisation (FDR<0.05), were among the top enriched pathways. Furthermore, IFN- γ and Th2 cytokine-induced inflammatory signalling in normal and asthmatic airway epithelium (FDR<0.05) and role of type 2 innate lymphoid cells in airway allergic inflammation and tissue repair (FDR<0.05) were among the top 10 enriched pathways (supplementary table S11).

Genetic colocalisation of blood eosinophil counts and asthma

To identify shared variants between asthma and eosinophil counts, we performed a colocalisation analysis of the SNPs identified in doctor-diagnosed asthma and blood eosinophil count GWAS (supplementary table S12). A posterior probability for $H_4 \geq 0.80$ was used to define colocalisation. This analysis identified 26 colocalised variants between asthma and eosinophil counts; these variants were mapped to 36 eQTLs (supplementary table S13). β -coefficients for colocalised variants showed a strong correlation ($r=0.92$, $p < 0.01$) (figure 3a). Gene ontology analysis for these transcripts showed enrichment for IL-18-mediated signalling and positive regulation of leukocyte-mediated immunity (figure 3b, supplementary table S14). The only missense variant in the VEP of colocalisation was rs16903574, in exon 8 of the OTU deubiquitinase with linear linkage specificity like (*OTULINL*) (supplementary table S15). The functional effects of this mutation are unknown.

Heritability and genetic correlation of asthma and blood leukocytes

To determine the SNP-based heritability and genetic correlation between doctor-diagnosed asthma and blood leukocyte counts, we implemented the variance components method BOLT-REML. The estimated SNP-based heritability was 0.16 ($SE \pm 0.01$) for blood eosinophil counts and 0.05 for doctor-diagnosed asthma (table 1). We identified significant genetic correlation between doctor-diagnosed asthma and blood eosinophil counts ($r=0.38 \pm 0.10$). Because smoking is a risk factor for asthma and may affect the eosinophil count in the blood, the genetic correlation may be confounded by smoking. To further

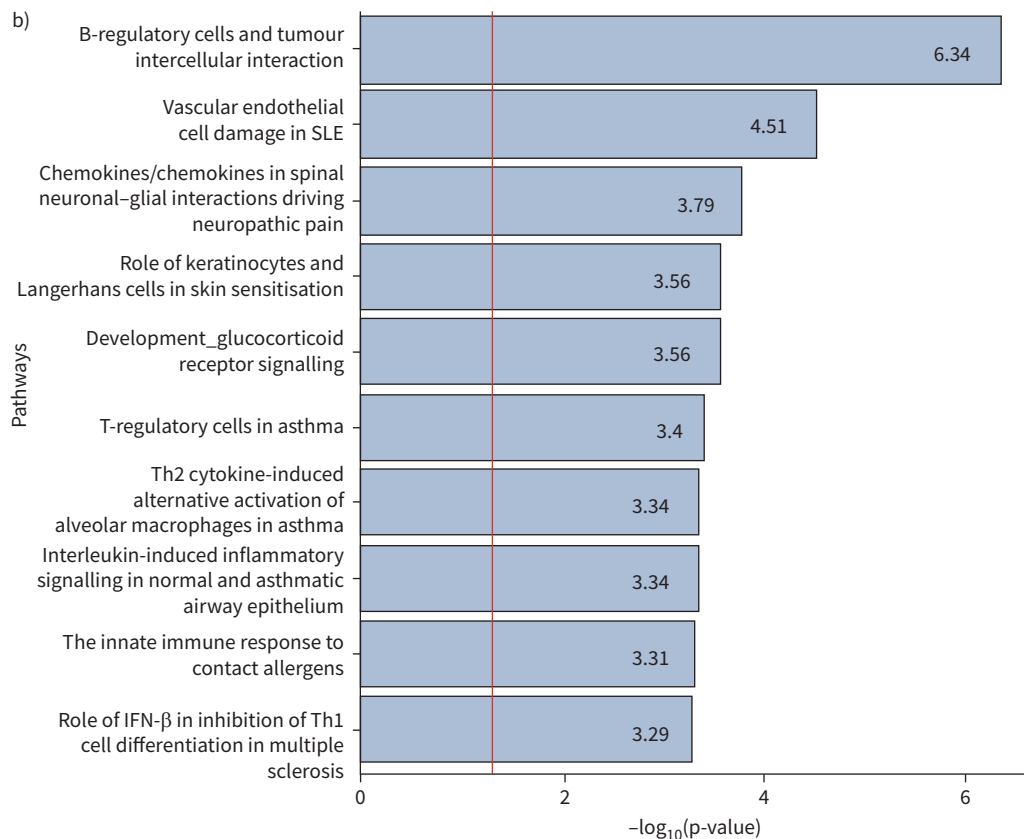
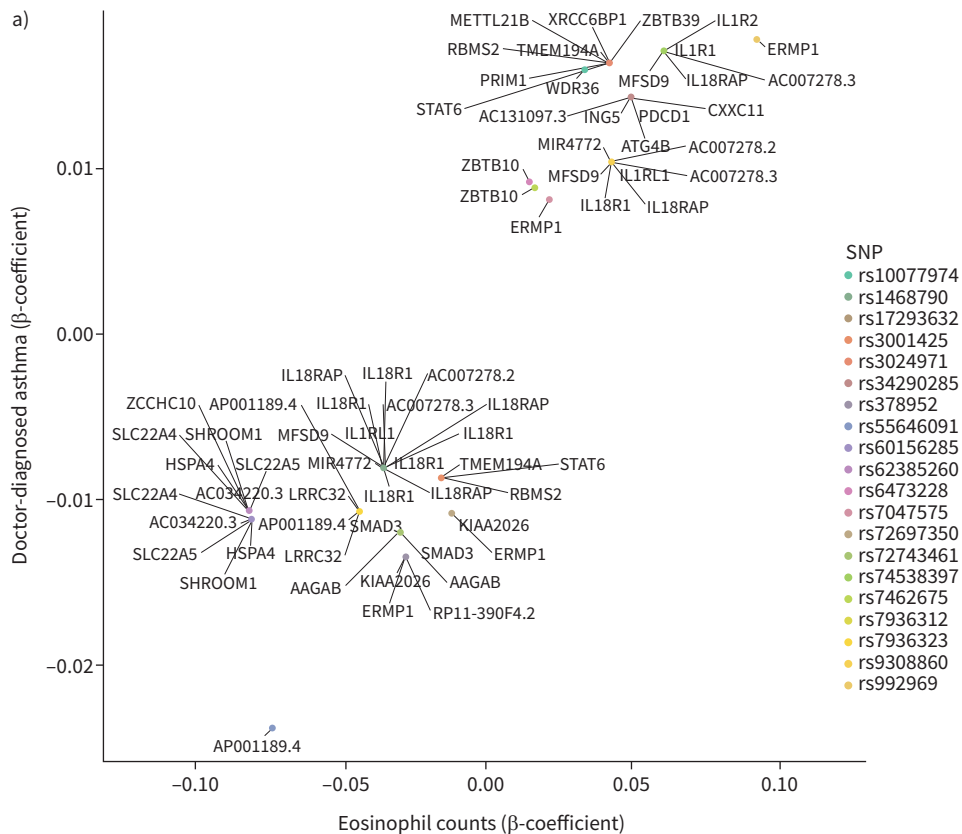


FIGURE 3 a) Colocalisation of single nucleotide polymorphisms (SNPs) associated with asthma and blood eosinophil counts, probability for H4 ≥ 0.80 ; b) colocalisation pathway enrichment. SLE: systemic lupus erythematosus; Th: T-helper cell; IFN: interferon.

understand the genetic correlation between doctor-diagnosed asthma and blood eosinophil counts conditional on smoking, we performed additional genetic correlation estimation where we adjusted smoking as a covariate in the GWAS of both traits. The smoking-adjusted genetic correlation analysis confirmed the positive correlation between doctor-diagnosed asthma and blood eosinophil counts ($r=0.44 \pm 0.03$). No significant genetic correlation was identified between doctor-diagnosed asthma and other blood leukocyte counts (lymphocytes, monocytes and neutrophils) (supplementary table S16). Blood eosinophil counts were higher in asthma compared to controls (223 versus 163 cells· μL^{-1} , $p<0.001$), as were neutrophil (4243 versus 4064 cells· μL^{-1} , $p<0.001$) and monocyte (473 versus 465 cells· μL^{-1} , $p<0.001$) counts (figure 4).

Validation in TAGC

To validate the findings from the UK Biobank, we performed a separate genetic correlation analysis with the TAGC cohort [22]. We identified the genetic correlation between asthma phenotype in the TAGC and UK Biobank doctor-diagnosed asthma to be 0.987 ± 0.08 (correlation \pm SE, $p=1.1 \times 10^{-33}$), which suggests that genetic effects on asthma are highly consistent between these two cohorts. In addition, the genetic correlation between Asthma in the TAGC and blood eosinophil counts in the UK Biobank was estimated to be $r=0.373 \pm 0.08$ (correlation \pm SE, $p=1.2 \times 10^{-6}$), similar to the correlation between doctor-diagnosed asthma and blood eosinophil counts in the UK Biobank: $r=0.457 \pm 0.07$ (correlation \pm SE, $p=4.7 \times 10^{-11}$).

Discussion

This study shows that asthma and blood eosinophil counts in the UK Biobank share a genetic architecture. A key strength of this study is its combination of a large cohort with blood leukocyte measurements and a robust asthma definition. Using this approach, we found a positive genetic correlation between blood eosinophil counts and doctor-diagnosed asthma. Furthermore, no other blood leukocytes showed a genetic correlation with asthma in this study. Colocalisation analysis showed enrichment for multiple pathways involved in asthma including Treg cells in asthma and Th2 cytokine-induced alternative activation of alveolar macrophages in asthma. Additionally, a strong correlation was found between β -coefficients of the colocalised variants, indicating a consistent direction of association for both traits.

We used a stringent definition of asthma and identified eQTLs in specific immune cells and tissues. In addition to concordant associations with previous GWAS in the UK Biobank [26, 27], our analysis also revealed novel eQTLs. A new association was found between *C4A* and *C4B*, two genes involved in the complement pathway. As a result of these novel observations, the list of genes associated with the complement pathway has grown following the previous identification of *C2* in TAGC [32]. The identification of these novel associations may be partly explained by the annotation methodology, which relies on eQTLs rather than computational predictions. Nevertheless, we also examined a particularly strong disease phenotype in a large cohort, suggesting that phenotype selection may have played a role. Thus, our results extend the list of genetic associations with asthma.

Similar to the GWAS of asthma, our GWAS of blood eosinophil counts builds upon previous studies of blood traits [29]. Using eQTL-based annotation, we identified 414 new eQTLs in our GWAS, expanding the number of pathways previously associated with blood eosinophil counts. Remarkably, we found an association with *IL4*, a gene encoding a key T2 cytokine. In addition to these results, colocalisation results demonstrating IL4 signalling via *STAT6* highlight the importance of this pathway in asthma and blood

TABLE 1 Estimated proportions of variance in trait heritability (h²g) explained by single nucleotide polymorphisms (SNPs)

	Asthma [#] h ² g \pm se	Eosinophil count [¶] h ² g \pm se
Estimated SNP heritability	0.05 \pm 0.01	0.16 \pm 0.01
[#] : cases n=11 717, controls n=80 070; [¶] : n=349 856.		

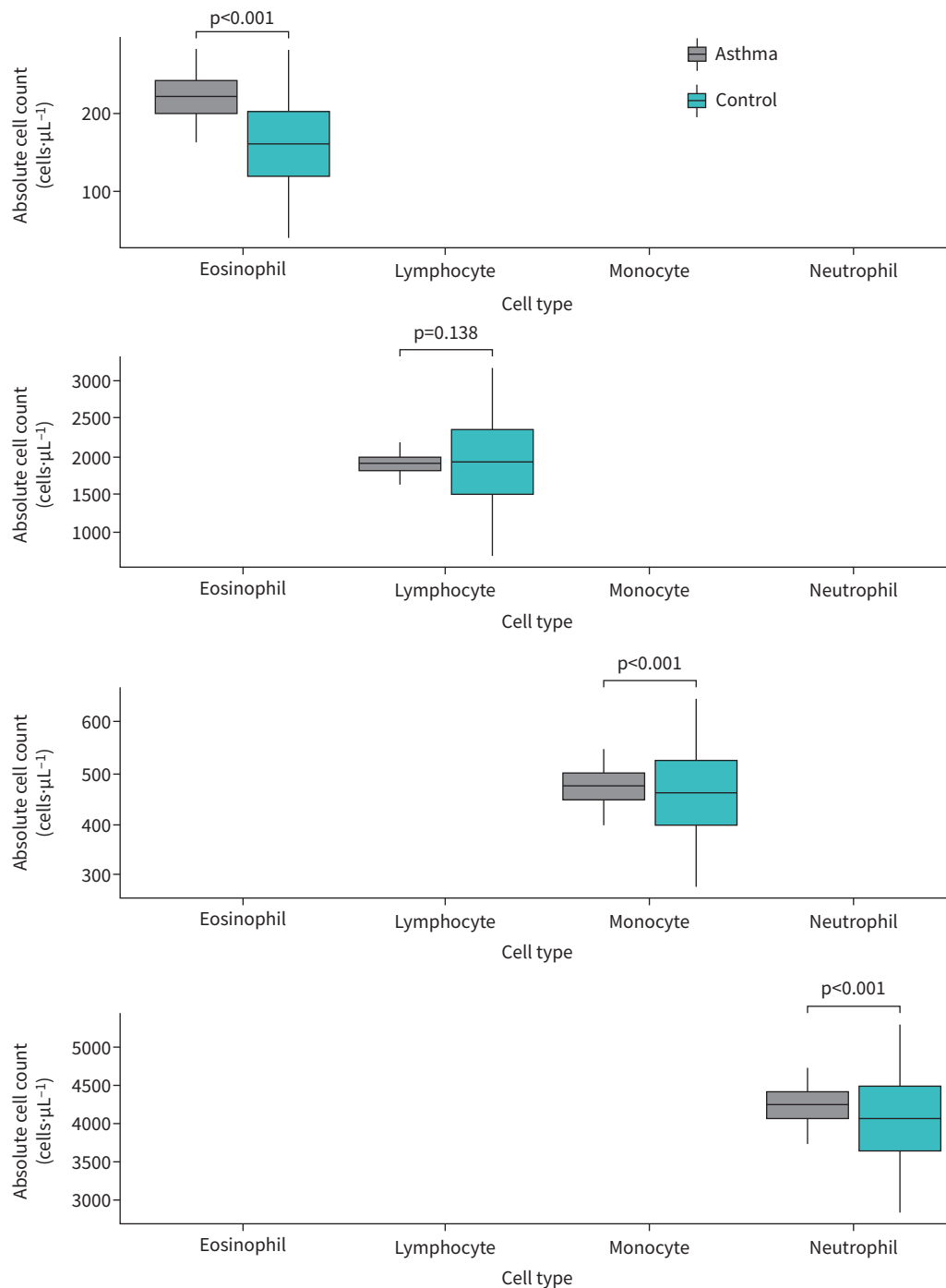


FIGURE 4 Leukocyte blood counts in UK Biobank in asthma (n=14 810) and controls (n=101 327). a) Eosinophil; b) lymphocyte; c) monocyte; d) neutrophil.

eosinophil counts. The heritability of blood eosinophil counts was similar to other blood leukocytes; however, eosinophils were the only leukocytes correlated with doctor-diagnosed asthma. Enrichment analyses of candidate genes associated with blood eosinophil counts revealed that many of the top pathways share genes with airway inflammation in asthma including tumour necrosis factor- α , IFN- γ and Th2 cytokines. These observations underscore the importance of eosinophils in asthma, as further evidenced by current therapies targeting these pathways and leading to eosinophil depletion in the circulation and airway [4].

Previous studies have demonstrated an association between blood eosinophil counts and asthma [8, 29]. Our study sought to characterise the genomic overlap between asthma and blood eosinophil counts using multiple complementary approaches. First, we found that in the UK Biobank, blood eosinophils were the only leukocytes with a positive genetic correlation with doctor-diagnosed asthma. Second, our colocalisation analysis found enrichment for multiple genes involved in the regulation of lymphocyte mediated immunity including *IL18R1*, *IL1R1*, *IL18RAP*, *STAT6* and *HSP70*. The role of *STAT6*, involved in IL-4 signalling, has been extensively studied in asthma [33], while gene expression of *IL18R1* and *IL1R1*, members of the IL-1 receptor family, is associated with severe asthma and a high Th2 signature in the sputum. In addition, there was a positive correlation between *IL18R1* and eosinophil counts [34]. Furthermore, six loci associated with moderate-to-severe asthma were also present in our colocalisation analysis including *IL1RL1*, *WDR36*, *STAT6*, *SMAD3*, *D2HGDH* and *C11orf30* [35]. Although we found a genetic correlation between asthma and blood eosinophil counts in the UK Biobank from individuals of European descent, a study of Japanese individuals identified a similarly strong genetic correlation ($r=0.35$; $p=3.76\times 10^{-5}$) [36]. Lastly, we determined the genetic correlation between the UK Biobank traits and TAGC, a large consortium of asthma GWAS [22]. This analysis confirmed that doctor-diagnosed asthma and blood eosinophil counts in UK Biobank are correlated with asthma in TAGC. These analyses identify a robust genetic correlation between asthma and blood eosinophil counts in the UK Biobank and the existing GWAS of asthma.

Blood eosinophil counts can be used as a quantitative marker of T2 immune responses and are used in the clinic to identify subjects experiencing moderate to severe poorly controlled asthma who are amenable to treatment with IL-5 and IL-4/IL-13 blockades [4, 30]. Therefore, understanding the genetic architecture of both traits can be relevant to the characterisation of the pathogenesis of T2 responses. In the colocalisation analysis of the UK Biobank, variants with large concordant effects in doctor-diagnosed asthma and blood eosinophil counts are concentrated in two pathways, *STAT6* and IL-1 receptor family. A possible mechanism for genetic associations between asthma and blood eosinophil counts is IL-4 signalling through *STAT6*.

We are aware of the limitations associated with European ancestry as the primary genetic background in our analyses. However, our confidence in the approach is reinforced by the presence of genetic correlations between doctor-diagnosed asthma, eosinophilic asthma and asthma in TAGC, the colocalisation of variants in *STAT6* between GWAS of asthma and our GWAS of blood eosinophil counts, and a correlation in the Japanese Biobank [36]. A GWAS in African Americans did not identify any significant associations with blood eosinophil counts [37]; moreover, the top associations in their study did not overlap with blood eosinophil associated loci from the UK Biobank, reported here or previously [29]. It is possible that this may be explained by variation in these associations across different genetic backgrounds or by differences in the study design. Moreover, we do not know how variation in blood eosinophil counts precedes asthma onset and whether any potential relationship varies across a lifespan. Nevertheless, these limitations should not affect the interpretation of our observations or their validity based on the consistent signal across phenotypes.

In conclusion, we have defined the shared genetic architecture between asthma and blood eosinophil counts using a large cohort in the UK Biobank and the GWAS Catalog. The most notable overlap between doctor-diagnosed asthma and blood eosinophil counts was the *STAT6* signalling pathway and the IL-1 receptor family. Together, these data support a link between blood eosinophil counts and asthma-related loci mediated through overlapping variants associated with *STAT6* signalling.

Provenance: Submitted article, peer reviewed.

Acknowledgements: This research has been conducted using the UK Biobank Resource. Access to data was obtained under application number 29900. We want to thank the participants and researchers from UK Biobank who significantly contributed or collected data. In addition, we thank the TAGC consortium for providing GWAS summary statistic data.

Author contributions: Conception and design: B. Li, Y. Wang, X. Li, H. Zhao and J.L. Gomez. Data acquisition and analysis: B. Li, Y. Wang, Z. Wang, X. Li and J.L. Gomez. Article drafting/revision: all authors. Final approval: all authors.

Support statement: This study was supported by US National Heart, Lung, and Blood Institute grants R01 HL153604 and R03 HL154275 to J.L. Gomez, and R01 HL118346 to G.L. Chupp; and US National Institutes of Health

1R01 GM134005 and US National Science Foundation grant DMS 1902903 to H. Zhao. Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: B. Li has nothing to disclose. Y. Wang has nothing to disclose. Z. Wang has nothing to disclose. X. Li has nothing to disclose. S. Kay has nothing to disclose. G.L. Chupp reports grants and personal fees from GSK, AstraZeneca, Genentech and Teva, personal fees from BI, and grants and personal fees from Sanofi and Regeneron, outside the submitted work. H. Zhao has nothing to disclose. J.L. Gomez has nothing to disclose.

References

- 1 Palm NW, Rosenstein RK, Medzhitov R. Allergic host defences. *Nature* 2012; 484: 465–472.
- 2 Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. *Nat Med* 2012; 18: 716–725.
- 3 Newby C, Agbetile J, Hargadon B, et al. Lung function decline and variable airway inflammatory pattern: longitudinal analysis of severe asthma. *J Allergy Clin Immunol* 2014; 134: 287–294.
- 4 Nair P, Pizzichini MM, Kjarsgaard M, et al. Mepolizumab for prednisone-dependent asthma with sputum eosinophilia. *N Engl J Med* 2009; 360: 985–993.
- 5 Hanania NA, Alpan O, Hamilos DL, et al. Omalizumab in severe allergic asthma inadequately controlled with standard therapy: a randomized trial. *Ann Intern Med* 2011; 154: 573–582.
- 6 Rabe KF, Nair P, Brusselle G, et al. Efficacy and safety of dupilumab in glucocorticoid-dependent severe asthma. *N Engl J Med* 2018; 378: 2475–2485.
- 7 Weidinger S, Gieger C, Rodriguez E, et al. Genome-wide scan on total serum IgE levels identifies FCER1A as novel susceptibility locus. *PLoS Genet* 2008; 4: e1000166.
- 8 Gudbjartsson DF, Bjornsdottir US, Halapi E, et al. Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. *Nat Genet* 2009; 41: 342–347.
- 9 Moffatt MF, Gut IG, Demenais F, et al. A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010; 363: 1211–1221.
- 10 Olafsdottir TA, Theodors F, Bjarnadottir K, et al. Eighty-eight variants highlight the role of T cell regulation and airway remodeling in asthma pathogenesis. *Nat Commun* 2020; 11: 393.
- 11 Collins R. What makes UK Biobank special? *Lancet* 2012; 379: 1173–1174.
- 12 Allen NE, Sudlow C, Peakman T, et al. UK biobank data: come and get it. *Sci Transl Med* 2014; 6: 224ed4.
- 13 Loh P-R, Tucker G, Bulik-Sullivan BK, et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* 2015; 47: 284–290.
- 14 Loh P-R, Bhatia G, Gusev A, et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat Genet* 2015; 47: 1385–1392.
- 15 Loh P-R, Kichaev G, Gazal S, et al. Mixed-model association for biobank-scale datasets. *Nat Genet* 2018; 50: 906–908.
- 16 Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 2012; 6: 80–92.
- 17 Elsworth B, Lyon M, Alexander T, et al. The MRC IEU OpenGWAS data infrastructure. *bioRxiv* 2020; preprint [https://doi.org/10.1101/2020.08.10.244293].
- 18 Schmiedel BJ, Singh D, Madrigal A, et al. Impact of genetic polymorphisms on human immune cell gene expression. *Cell* 2018; 1701–1715.
- 19 Vösa U, Claringbould A, Westra H-J, et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat Genet* 2021; 53: 1300–1310.
- 20 McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol* 2016; 17: 122.
- 21 Wallace C. Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genet* 2020; 16: e1008720.
- 22 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.
- 23 Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet* 2015; 47: 1236–1241.
- 24 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature* 2015; 526: 68–74.
- 25 Sudmant PH, Rausch T, Gardner EJ, et al. An integrated map of structural variation in 2,504 human genomes. *Nature* 2015; 526: 75–81.
- 26 Ferreira MA, Vonk JM, Baurecht H, et al. Shared genetic origin of asthma, hay fever and eczema elucidates allergic disease biology. *Nat Genet* 2017; 49: 1752–1757.
- 27 Zhu Z, Lee PH, Chaffin MD, et al. A genome-wide cross-trait analysis from UK Biobank highlights the shared genetic architecture of asthma and allergic diseases. *Nat Genet* 2018; 50: 857–864.

- 28 Neville MDC, Choi J, Lieberman J, *et al.* Identification of deleterious and regulatory genomic variations in known asthma loci. *Respir Res* 2018; 19: 248.
- 29 Astle WJ, Elding H, Jiang T, *et al.* The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 2016; 167: 1415–1429.
- 30 Castro M, Corren J, Pavord ID, *et al.* Dupilumab efficacy and safety in moderate-to-severe uncontrolled asthma. *N Engl J Med* 2018; 378: 2486–2496.
- 31 CHARGE Consortium Hematology Working Group. Meta-analysis of rare and common exome chip variants identifies S1PR4 and other loci influencing blood cell traits. *Nat Genet* 2016; 48: 867–876.
- 32 Kim S, Forno E, Yan Q, *et al.* SNPs identified by GWAS affect asthma risk through DNA methylation and expression of *cis*-genes in airway epithelium. *Eur Respir J* 2020; 55: 1902079.
- 33 Foster PS, Maltby S, Rosenberg HF, *et al.* Modeling T_H 2 responses and airway inflammation to understand fundamental mechanisms regulating the pathogenesis of asthma. *Immunol Rev* 2017; 278: 20–40.
- 34 Rossios C, Pavlidis S, Hoda U, *et al.* Sputum transcriptomics reveal upregulation of IL-1 receptor family members in patients with severe asthma. *J Allergy Clin Immunol* 2018; 141: 560–570.
- 35 Shrine N, Portelli MA, John C, *et al.* Moderate-to-severe asthma in individuals of European ancestry: a genome-wide association study. *Lancet Respir Med* 2019; 7: 20–34.
- 36 Kanai M, Akiyama M, Takahashi A, *et al.* Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet* 2018; 50: 390–400.
- 37 Reiner AP, Lettre G, Nalls MA, *et al.* Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). *PLoS Genet* 2011; 7: e1002108.