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

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


This manuscript has recently been accepted for publication in the ERJ Open Research. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJOR online.

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

Boyang Li¹, Yuxuan Wang², Zixiao Wang¹, Xinyue Li³, Shannon Kay⁴,⁵, Geoffrey L Chupp⁴, Hongyu Zhao¹,⁶, a, Jose L Gomez⁴,⁵, a

Both authors contributed equally to this work

a These authors share senior authorship

Affiliations:

¹Department of Biostatistics, Yale School of Public Health, New Haven, CT 06510, USA

²Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA

³School of Data Science, City University of Hong Kong, Hong Kong SAR 999077

⁴Pulmonary, Critical Care and Sleep Medicine Section, Yale University, New Haven, CT 06520

⁵Center for Precision Pulmonary Medicine (P2MED), Yale University, New Haven, CT 06520

⁶Program of Computational Biology and Bioinformatics, Yale University, New Haven, CT 06510, USA

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

Address correspondence: jose.gomez-villalobos@yale.edu

300 Cedar Street (S419 TAC), New Haven, CT 06520-8057

Phone (203) 785-4163 / Fax (203) 785-3826

**Sources of support**

R01 HL153604, and R03 HL154275 to JLG.

R01 HL118346 to GLC.

NIH 1R01 GM134005 and NSF grant DMS 1902903 to HZ.

Total word count: 3323 words

**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.

**Objective:** To characterize 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 IL4 signaling, 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, $r = 0.38 \pm 0.10$ (correlation ± s.e., $P = 4.7 \times 10^{-11}$). As a validation of this association, we found a similar correlation between TAGC and blood eosinophil counts in UK Biobank ($0.37 \pm 0.08$ (correlation ± s.e., $P = 1.2 \times 10^{-6}$)

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

**Word Count:** 247  
**Key Words:** asthma; eosinophil; GWAS; genetic architecture; STAT6; IL4.
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, IL5, and IL4/IL13 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]. Eosinophils are thus prominent cells involved in disease pathogenesis. Nevertheless, we have an incomplete understanding of the shared genetic associations between eosinophil counts and asthma risk.

We hypothesized 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 colocalization 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 over 500,000 UK residents aged 40–69, 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 (female = 6,715; male = 5,002), 80,070 controls (female = 43,838; male = 36,232) and 349,856 phenotypes for eosinophil counts (female = 187,758; male = 162,098). Additional traits included the following: lymphocyte counts, monocyte counts, and neutrophil counts. For more details, see Supplementary Table 1.

Genome-wide association analysis

GWAS for all phenotypes was performed using a linear mixed model (LMM) 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 (IRNT) to the blood cell count phenotypes so that the standardized phenotypes approximated a normal distribution. Single nucleotide polymorphisms (SNPs) were excluded if the minor allele frequency was <1%, Hardy-Weinberg equilibrium p-values <1 x 10^{-5}, or an imputation score <0.8, with 5,067,514 remaining SNPs after imputation. QQ plots are presented in Supplementary Figures 1-5. P-values were adjusted by the genomic inflation factor (λ), specifically, test statistics for each SNP were divided by the genomic inflation factor (λ >1). Genome-wide significance was defined as P-value <5x10^{-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 2. Annotation of significant variants was performed with SNPeff.v5.0 [16]. Genetic loci were defined as the 5,000 base pair intervals, upstream and
downstream, from each variant. Previous GWAS results were obtained via the GWAS Catalog (https://www.ebi.ac.uk/gwas/docs/file-downloads, download: 2/8/2023). Linkage Disequilibrium (LD) clumping was performed using the Integrative Epidemiology Unit (IEU) Open GWAS portal (https://gwas.mrcieu.ac.uk/). The ‘ld_clump’ function of the ieugwasr 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 colocalization analyses. The Database of Immune Cell Expression (DICE) focused on the identification of SNPs affecting gene expression (eQTLs) in immune cells [18]. We used information for 15 cell types: B cell, classical monocyte, non-classical (M2) monocyte, natural killer (NK) cell, naive CD4 T cell, stimulated CD4 T cell, naive CD8 cell, stimulated CD8 cell, naive Treg, memory Treg, Th1 cell, Th2 cell, Th17 cell, Th1/17 cell, and T follicular helper cell (TFH). Using a False Discovery Rate (FDR) adjusted p value <0.05 and Transcripts Per Kilobase Million (TPM) >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 UCSC Genes track with the Ensembl release 109 (Feb 2023) and selected SIFT, PolyPhen-2, and functional role parameters. Upstream and downstream distance was set at 5,000 base pairs.
**Gene enrichment analysis.** Enrichment analyses were performed with MetaCore™ version 22.4.71100 (Clarivate Analytics, Philadelphia, PA). The single experiment workflow was implemented for enrichment analyses. We performed pathway maps, Gene Ontology (GO), and pathway network enrichment analysis on genes identified in the eQTL analyses of variants associated with Doctor-diagnosed asthma and blood eosinophil counts.

**Colocalization analysis**
To run the colocalization analysis, the genome must be divided into approximately independent windows. For each trait, we identified all genome-wide significant variants (p-value < \(5 \times 10^{-8}\)) 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 colocalization 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 utilized summary statistics-based Linkage
Disequilibrium-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 1,000 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 < 5x10^{-8} (Supplementary Table 3), with λ=1.047. (Figure 3A and Supplementary Table 3). A previous GWAS of 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 4). We found two independent variants in IL1RL1, ERMP1 and IL18R1, respectively.

To characterize the effects of variants associated with doctor-diagnosed asthma, two eQTL databases (DICE, and eQTLGen) were interrogated. This analysis identified 109 eQTLs (Supplementary Table 5). HLA 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 pathway 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 6). Only two variants, rs4193 and rs12722072 in HLA-DQA1, 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 IFN-gamma (FDR <0.05) and maturation and migration of dendritic cells in skin sensitization (FDR <0.05) (Supplementary Table 7).

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 3,429 genome-wide significant variants $p < 5 \times 10^{-8}$ (Figure 3B and Supplementary Table 8). Annotation with the two eQTL databases identified 1,081 unique eQTLs (Supplementary Table 9). The majority of these transcripts have been previously associated with eosinophil counts or percentages in the GWAS catalog, largely thanks to Astle et al.'s study [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 4). 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 10). VEP annotation identified a missense variant rs61731111 with large effects on blood eosinophil counts (0.13 ± 0.01, beta ± s.e.), 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 7). 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-gamma (FDR <0.05) and maturation and migration of dendritic cells in skin sensitization (FDR <0.05), were among the top enriched pathways. Furthermore, IFN-gamma and Th2 cytokines-induced inflammatory signaling 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 11).

**Genetic colocalization of blood eosinophil counts and asthma**

To identify shared variants between asthma and eosinophil counts, we performed a colocalization analysis of the SNPs identified in doctor-diagnosed asthma and blood eosinophil count GWAS (Supplementary Table 12). A posterior probability for H4 (PPH4) ≥ 0.80 was used to define colocalization. This analysis identified 26 colocalized variants between asthma and eosinophil counts; these variants were mapped to 36 eQTLs (Supplementary table 13). Beta coefficients for colocalized variants showed a strong correlation (r=0.92, p<0.01) (Figure 4A). Gene ontology analysis for these transcripts showed enrichment for interleukin-18-mediated signaling and positive regulation of leukocyte mediated immunity (Figure 4B; Supplementary Table 14). The only missense variant in the VEP of colocalization was rs16903574, in exon 8 of the OTU Deubiquitinase With Linear Linkage Specificity Like (OTULINL) (Supplementary Table 15). 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=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 ± 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 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 16). Blood eosinophil counts were higher in asthma compared to controls (223 vs. 163 cells/microliter, \( p < 0.001 \)), as well as neutrophil (4,243 vs. 4,064 cells/microliter, \( p < 0.001 \)) and monocyte (473 vs. 465, cells/microliter, \( p < 0.001 \)) counts (Figure 2).

**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 ± s.e., \( P = 1.1 \times 10^{-33} \)), 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 ± s.e., \( P = 1.2 \times 10^{-6} \)), similar to the correlation between doctor-diagnosed asthma and blood eosinophil counts in the UK Biobank \( r = 0.457 \pm 0.07 \) (correlation ± s.e., \( P = 4.7 \times 10^{-11} \)).
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. Colocalization analysis showed enrichment for multiple pathways involved in asthma including T regulatory cells in asthma and Th2 cytokine-induced alternative activation of alveolar macrophages in asthma. Additionally, a strong correlation was found between beta coefficients of the colocalized 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 CFB and C4A, 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, colocalization results demonstrating IL4 signaling via STAT6 highlight the
importance of this pathway in asthma and blood 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 TNF-alpha, IFN-gamma, 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 characterize 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 colocalization 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 IL4 signaling, has been extensively studied in asthma [33], while gene expression of IL18R1 and IL1R1, members of the IL1 receptor family, is associated with severe asthma and a high Th2 signature in the sputum. There was also a positive correlation between IL18R1 and eosinophil counts [34]. Furthermore, six loci associated with moderate-to-severe asthma were also present in our colocalization 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 characterization of the pathogenesis of T2 responses. In the colocalization 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 IL1 receptor family. A possible mechanism for genetic associations between asthma and blood eosinophil counts is IL4 signaling 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 colocalization 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. We also 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 \textit{STAT6} signaling pathway and the IL1 receptor family. Together, these data support a link between blood eosinophil counts and asthma-related loci mediated through overlapping variants associated with \textit{STAT6} signaling.
Acknowledgments

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.
REFERENCES


**Figure 1.** Study Design

**Figure 2.** Leukocyte blood counts in UK Biobank in asthma (n=14,810) and controls (n=101,327).

**Figure 3.** A. Manhattan plot of doctor-diagnosed asthma genome-wide association study in UK Biobank. B. Manhattan plot of blood eosinophil count genome-wide association study in UK Biobank

**Figure 4.** A. Colocalization of SNPs associated with asthma and blood eosinophil counts, probability for $H_4$ (PPH4) $\geq 0.80$. B. Colocalization pathway enrichment.
Table 1. Estimated proportions of variance in trait heritability explained by SNPs.

<table>
<thead>
<tr>
<th></th>
<th>Asthma $h^2_g$ (s.e.)</th>
<th>Eosinophil Counts $h^2_g$ (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases (n=11,717)</td>
<td>Controls (n=80,070)</td>
</tr>
<tr>
<td></td>
<td>0.05 (0.01)</td>
<td>0.16 (0.01)</td>
</tr>
<tr>
<td></td>
<td>(n=349,856)</td>
<td></td>
</tr>
</tbody>
</table>
Conflict of Interest Disclosure
Dr. Li has nothing to disclose.
Mr. Wang has nothing to disclose.
Dr. Li has nothing to disclose.
Dr. Kay has nothing to disclose.
Dr. Chupp reports grants and personal fees from GSK, grants and personal fees from Astra Zeneca, grants and personal fees from Genentech, grants and personal fees from Teva, personal fees from BI, grants and personal fees from Sanofi, grants and personal fees from Regeneron, outside the submitted work.
Dr. Zhao has nothing to disclose.
Dr. Gomez has nothing to disclose.
Figure 1. Study Design

Doctor-diagnosed asthma (n=11,717)
Controls (n=80,070)  

Blood eosinophil counts (n=349,856)

- Genome-wide association analysis
- Gene enrichment analysis
- Variant effect predictor analysis
- Colocalization analysis
- Heritability and genetic correlation analyses
Figure 2. Leukocyte blood counts in UK Biobank in asthma (n=14,810) and controls (n=101,327).
Figure 3A. Doctor-Diagnosed Asthma, $\lambda=1.047$
Figure 3B. Eosinophil count, lambda=1.311
Figure 4. A. Colocalization of SNPs associated with asthma and blood eosinophil counts, probability for $H4$ (PHF4) $\geq 0.80$. 
Figure 4. B. Colocalization pathway enrichment.
Supplementary Figure 1. Q-Q plot of Doctor-Diagnosed GWAS UK Biobank (unadjusted for genomic inflation).

\( \lambda = 1.10 \)
Supplementary Figure 2. Q-Q plot of blood eosinophil count GWAS UK Biobank (unadjusted for genomic inflation).
Supplementary Figure 3. Q-Q plot of blood lymphocyte count GWAS UK Biobank (unadjusted for genomic inflation)

\[ \lambda = 1.37 \]
**Supplementary Figure 4.** Q-Q plot of blood neutrophil count GWAS UK Biobank (unadjusted for genomic inflation)

\[ \lambda = 1.31 \]
Supplementary Figure 5. Q-Q plot of blood monocyte count GWAS UK Biobank (unadjusted for genomic inflation)

\[ \lambda = 1.31 \]