Journal of Translational Medicine

RESEARCH

Open Access

Genomic structural equation modeling elucidates the shared genetic architecture of allergic disorders

Jingsheng Ruan[1] and Xinglin Yi[2*]

Abstract

Background The intricate shared genetic architecture underlying allergic disorders—including allergic asthma, atopic dermatitis, contact dermatitis, allergic rhinitis, allergic conjunctivitis, allergic urticaria, anaphylaxis, and eosinophilic esophagitis—remains incompletely characterized.

Methods Our study employed genomic structural equation modeling (Genomic SEM) to define the common factor representing the shared genetic architecture of allergic disorders. Coupled with diverse post-GWAS analytical methods, we aimed to discover susceptible loci and investigate genetic associations with external traits. Furthermore, we explored enriched genetic pathways, cellular layers, and genomic elements, and investigated putative plasma protein biomarkers. Polygenic risk score (PRS) analyses, leveraging our integrated GWAS data, were conducted to assess chromosomal-level risk associations for allergic disorders.

Results A well-fitted genomic SEM integrated GWAS data, revealing the shared genetic architecture of allergic disorders. We identified a total of 2038 genome-wide significant SNP loci ($p < 5e\text{-}8$), including 31 previously unreported loci. Fine-mapping of variants and gene sets pinpointed 2 causal variants and 31 candidate susceptible genes. Genetic correlation analyses further illuminated the shared genetic architecture underlying multiple traits, notably psychiatric disorders. Preliminary findings identified four putative causal plasma protein biomarkers.

Conclusion Notably, this study presents the first comprehensive genetic characterization of allergic disorders through a GWAS analysis of an unmeasured composite phenotype, providing novel insights into shared etiological pathways across these conditions.

Keywords Genomic SEM, Allergic disorders, FAM114 A1, Rs145982144, Rs78017269

Introduction

Allergy exemplifies the dual nature of the immune system—it serves as both a protector and a potential threat. It is a systemic disorder caused by immune system dysregulation and encompasses conditions such as allergic asthma, atopic dermatitis, contact dermatitis, allergic rhinitis, allergic conjunctivitis, allergic urticaria, anaphylaxis, and eosinophilic esophagitis. Given the high prevalence of allergic diseases, the World Health Organization (WHO) has classified them among the top three diseases requiring priority prevention and treatment in the twenty-first century. Recent estimates from the Global Burden of Disease Study reveal substantial global allergic disease burdens, with approximately 260 million prevalent cases of asthma and 204 million cases of atopic dermatitis worldwide

*Correspondence:
Xinglin Yi
xinglinyi2024@163.com
[1] Department of Thoracic, Jinshan Hospital of Fudan University, Fudan University, Shanghai, China
[2] Department of Respiratory and Critical Care Medicine, Third Military Medical University, Chongqing, China

[1, 2]. Allergic rhinitis, meanwhile, is estimated to afflict approximately 20% of the global population [3]. Alarmingly, the prevalence of these conditions continues to escalate, reflecting a persistent upward trajectory in allergic disease morbidity that underscores urgent public health priorities [3].

Allergic disorders arise from complex gene-environment interactions and are closely associated with type 2 inflammation and epithelial barrier dysfunction. While these conditions are often viewed as distinct diseases, comorbidities and multimorbidities are common. For instance, cross-sectional studies indicate that individuals with other allergic disorders are more susceptible to developing asthma [4, 5]. A birth cohort study of 2,311 Canadian children revealed that infants diagnosed with atopic dermatitis at one year of age had an 11-fold and sevenfold increased risk of developing allergic asthma and allergic rhinitis, respectively, by the age of three, compared to those without atopic dermatitis [6]. This temporal pattern of disease progression, where different allergic disorders emerge sequentially over time, is known as the"atopic march".

The mechanisms underlying allergic disease comorbidities and multimorbidities are highly complex, with current research suggesting that genetic factors, gut microbiota, IgE-mediated and non-IgE-mediated pathways all contribute [7, 8]. With the advent of the Human Genome Project (HGP), genome-wide association studies (GWAS) and epigenome-wide association studies (EWAS) have underscored the role of genetic heritability in allergic disorders [9]. GWAS have identified key susceptibility loci, such as rs11740584 and rs2299007 in *KIF3 A* for food allergy, rs8111930 near *MRPL4* for allergic rhinitis, and rs2243250 in *IL4* and rs20541 in *IL13* for allergic asthma [10]. For instance, a recent study leveraging GWAS data and innovative analytical approaches identified multiple shared loci between eosinophilic disorders and allergic diseases (e.g., *C11orf30* and *SMAD3*) [11]. Using genomic structural equation modeling (Genomic-SEM), Gong et al. demonstrated substantial genetic correlation, represented by a common latent factor, among asthma, eczema, and allergic rhinitis [12]. These findings highlight the crucial role of genetic factors in allergic disease comorbidities and multimorbidities.

To date, numerous susceptibility single nucleotide polymorphisms (SNPs) and genes linked to specific allergic disorders have been identified. However, our understanding of the precise genetic and biological mechanisms underlying allergic disorder comorbidities remains limited, and existing GWAS findings are still underexplored in the context of combined allergic conditions.

To address these challenges, this study aims to integrate multiple genetic analysis methodologies and highly correlated association analysis tools to uncover potential molecular mechanisms and expand the understanding of disease interconnections. Specifically, we focus on genomic loci and chromosomal regions associated with allergic disorders to identify potential therapeutic targets. This research not only enhances our comprehension of allergic disorders but also provides a theoretical and practical foundation for prevention and intervention strategies aimed at reducing allergic disease burden and associated complications.

In this study, we constructed a Genomic-SEM utilizing publicly available GWAS summary statistics for specific allergic disorders. This method integrates genome-wide association data with structural equation modeling to elucidate shared genetic architectures and causal relationships between traits, while rigorously accounting for sample overlap and pleiotropic effects. The analytical framework has been validated in prior studies investigating comorbidities such as inflammatory bowel disease and asthma, as well as associations between body mass index and childhood asthma, among others [12–14]. By leveraging GWAS statistics, we estimated the associations of SNPs with latent phenotypes, effectively conducting a GWAS on an unmeasured allergic disorder construct. Furthermore, we employed integrative approaches from systems biology, defining unexplained genetic variation within allergic disorders as potential novel genetic markers and conducting extensive GWAS-related analyses. Although this approach does not fully capture the intricate interactions between allergic disorder pathways and multifactorial influences—given that allergic disorders are driven by a combination of genetic, environmental, and stochastic factors—our methodology minimizes confounding effects from known biomarkers, enabling analysis of previously challenging datasets. Additionally, we performed thousands of causal inference analyses based on GWAS summary data to help clinicians and biologists predict potential plasma biomarkers and clinical disease causal relationships, ultimately informing prevention strategies and potential interventions for patients.

Methods

A flowchart overview is presented in Fig. 1.

Genomic-SEM GWAS data source

The GWAS summary statistics used for Genomic-SEM analyses were derived from eight independent GWAS related to allergic disorders. These datasets were obtained from previously published GWAS studies, as well as the FinnGen R12 and UK Biobank (UKB) databases. The allergic phenotypes included in the analysis were allergic asthma, atopic dermatitis, contact

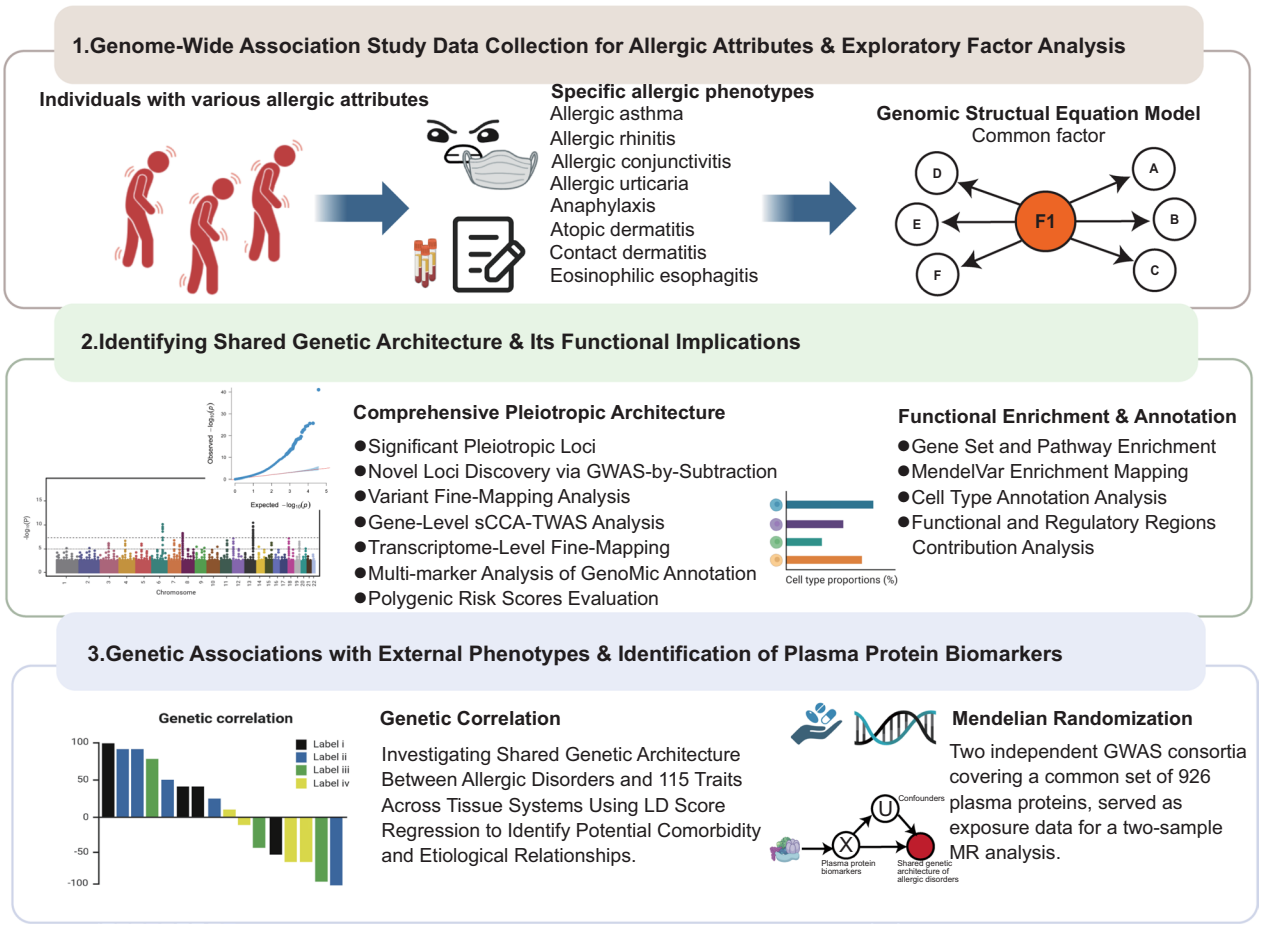

Fig. 1 Flowchart illustration

dermatitis, allergic rhinitis, allergic conjunctivitis, allergic urticaria, anaphylaxis, and eosinophilic esophagitis. All contributing GWAS studies received ethical approval from their respective institutional review boards, and all participants provided informed consent. The summary statistics underwent rigorous quality control procedures to ensure data integrity. A detailed list of the included GWAS datasets is available in Table S1.

Quality control for genomic-SEM

We implemented a rigorous quality control pipeline based on recommended filtering criteria for all autosomal SNPs from the eight input GWAS datasets. To ensure consistency across studies, SNPs were filtered against the 1000 Genomes Phase 3 European (EUR) reference panel, excluding those with a minor allele frequency (MAF) <0.01, effect size estimates of zero, reference panel mismatches, or allele inconsistencies. Given that the GWAS datasets originated from different genomic repositories and study populations, we accounted for potential sample overlap to prevent inflation in test statistics and improve result robustness. Statistical methods were applied to estimate and adjust for any residual overlap, minimizing bias in effect size estimates and ensuring the reliability of downstream analyses.

Genomic-SEM construction

We applied Genomic-SEM using GenomicSEM R package (v0.0.5) to investigate the shared genetic architecture underlying allergic traits. Genomic-SEM is a recently developed multivariate method that enables the exploration of latent structures across multiple phenotypes by estimating a multivariable genetic model [15]. A key advantage of Genomic-SEM is that it is robust to sample overlap and differences in sample sizes, minimizing biases associated with these factors. Additionally, it facilitates the identification of genetic variants that influence only a subset of traits rather than all complex traits under investigation, thereby distinguishing trait-specific effects from broader cross-trait genetic susceptibility.

The Genomic-SEM analysis was conducted in two stages. In the first stage, we estimated the empirical genetic covariance matrix and the corresponding sampling covariance matrix. To achieve this, we compiled GWAS summary statistics for allergic disorders and applied a multivariate extension of Linkage Disequilibrium (LD) Score Regression (LDSC). LDSC is a robust statistical framework that quantifies polygenic architecture using GWAS summary statistics. By modeling the relationship between variant association statistics and their LD profiles, LDSC disentangles true polygenic signals from confounding biases such as cryptic relatedness and population stratification. In our study, we used a multivariate LDSC to generate an empirical genetic covariance matrix across eight allergic traits. This matrix served as the input for the common factor model in SEM, while SNP-based heritability estimates (h^2) for each trait are reported in Table S2.

In the second stage, we specified an SEM model to minimize the discrepancy between the hypothesized covariance structure and the empirical covariance matrix estimated in the first stage. Our primary objective was to identify the shared genetic factors underlying eight allergic disorders. To this end, we tested the fit of a single-factor model, evaluating model adequacy using Standardized Root Mean Square Residual (SRMR), model chi-square (χ^2), Akaike Information Criterion (AIC), and Comparative Fit Index (CFI) (see Table S3 and Table S4). By implementing an appropriate common factor SEM specification, we incorporated individual autosomal SNP associations into the genetic and sample covariance matrix, allowing for a genome-wide analysis of shared covariance across the eight allergic disorder GWAS datasets. Furthermore, to ensure that the SNPs identified by the novel genomic SEM exhibit consistent effect directions, we conducted a heterogeneity test for each SNP and excluded those with a Cochran's Q statistic p-value less than 0.05.

Multilevel evaluation of genomic structural equation model

In addition to the aforementioned methods for testing genomic-SEM fitting, we implemented another LDSC approach to evaluate the stability of the genomic-SEM fit. Specifically, we assessed the model using parameters such as the mean chi-square, lambda GC, maximum chi-square, h^2, intercept, and ratio (calculated as (LDSC intercept − 1)/(mean χ^2 − 1)). The detailed control of LDSC parameters included: retaining SNPs with missing values, retaining SNPs with an INFO score < 0.9, retaining SNPs with a MAF < 0.01, and excluding SNPs with p-values outside the valid range or with ambiguous strand orientation.

Identification of significant and novel genomic loci

We used FUMA (Functional Mapping and Annotation; https://fuma.ctglab.nl/) to identify genomic loci and to detect both independent significant and lead SNPs associated with the Genomic-SEM of allergic disorders [16]. These SNPs exhibited low LD with other variants ($r^2 < 0.6$ for independent SNPs and $r^2 < 0.1$ for lead SNPs) and surpassed the genome-wide significance threshold ($p < 5 \times 10^{-8}$). We then compared the identified lead SNPs and loci with those reported in single-trait GWAS to assess potential pleiotropy, cross-referencing significant published associations ($p < 5 \times 10^{-8}$) in the GWAS Catalog. Next, we performed a risk locus analysis of this novel model ($p < 5 \times 10^{-8}$) using FUMA. Subsequently, we conducted post-GWAS analysis with Multi-marker Analysis of GenoMic Annotation (MAGMA) to investigate gene-level and gene-set associations between genetic variants and phenotypes, applying a False Discovery Rate Correction (FDR)-corrected p-value threshold of < 0.05 for statistical significance. In addition, we applied the"GWAS-by-Subtraction"method, which involves comparing the lead loci identified through Genomic-SEM with those pinpointed using the genome-wide significance threshold ($p < 5 \times 10^{-8}$) in single-input GWAS. This approach facilitates the discovery of novel, high-utility loci by"subtracting"previously identified findings, thereby enhancing the efficiency and precision of genetic discoveries.

Finemap

To identify the most likely causal variants associated with our GWAS, we employed—fine-mapping methods of Bayesian Fine-Mapping Method (FINEMAP), that was implemented in the R package echolocatoR (v2.0.3). We set a posterior probability threshold of 0.95 to define a credible set of candidate causal variants. Specifically, we analyzed a 250-kb window around each lead SNP, computing each SNP's posterior probability of causality within these regions. Any variant with a posterior probability exceeding 0.95 was considered a putative causal variant. FINEMAP is a sophisticated Bayesian fine-mapping tool designed to pinpoint the most probable causal variants driving observed associations. FINEMAP leverages a Bayesian statistical approach to compute the posterior probability of causality for each SNP within a defined genomic region, adjusting for LD.

Transcriptome-wide association study

Because proximal SNP-level fine-mapping can be overly simplistic, we performed a cross-tissue sCCA-TWAS (sparse canonical correlation analysis-based Transcriptome-Wide Association Study) after identifying putative causal variants. This approach aims to pinpoint genes

most relevant to the Genomic-SEM of allergic disorders by leveraging 37,920 precomputed eQTL features from the GTEx (v8) dataset [17]. We then selected genes with FDR-corrected TWAS p-values < 0.05 for further analysis using fine-mapping of gene sets (FOCUS), which evaluates the likelihood of a causal gene–phenotype relationship based on the FOCUS posterior inclusion probability. FOCUS applies a Bayesian framework to probabilistically attribute causality to genes associated with GWAS signals. By integrating multi-omics data, it addresses pleiotropy and LD confounding, calculating posterior inclusion probabilities (PIPs) to quantify the likelihood of a gene's causal role. A PIP threshold of > 0.8 is commonly adopted to identify genes with strong evidence of causality.

Gene set and pathway enrichment analysis

We performed pathway enrichment analyses using the Molecular Signatures Database (MsigDB) [18] with Gene Set Enrichment Analysis (GSEA) to investigate potential relationships between allergic disorders and related pathways for the genes identified by MAGMA through FUMA gene-to-function. Additionally, we conducted further gene enrichment analyses using MendelVar (https://mendelvar.mrcieu.ac.uk/submit/).

Cell type annotation analysis and regional contribution analysis

To identify etiologically relevant cell types associated with genomic-SEM–based allergic disorders, we used the CELLECT pipeline, which integrates cell-type expression specificity from single-cell RNA sequencing (scRNA-seq) data [19]. We employed the Tabula Muris dataset, comprising transcriptomic data from Mus musculus across multiple organs and tissues [20]. Preprocessing and normalization were performed in CELLEX to compute expression specificity likelihood scores for each region [21]. We then carried out cell-type–specific analyses with Stratified LD Score Regression (S-LDSC), classified the cell types, and applied an FDR-adjusted p-value threshold of < 0.05 to determine significance.

To evaluate each region's contribution to phenotypic variance, we also employed S-LDSC to estimate partitioned heritability across genomic annotations—such as genes, enhancers, and repressors. Specifically, S-LDSC uses a weighted LD matrix, allele frequency data, and GWAS summary statistics to apportion heritability among distinct genomic segments [22].

Genetic correlations with external traits

We estimated the genetic correlation (rg) between the newly constructed genomic-SEM GWAS dataset and each of the 115 GWAS datasets from the FinnR12 database for common diseases in respiratory, digestive, circulatory, and other categories, as listed in Table S5.

Identification of plasma protein biomarkers

Next, we also leveraged plasma protein GWAS data from the deCODE (Iceland) [23] and UKB-PPP databases [24] as exposures, with allergic disorders as the outcome, to identify putative plasma biomarkers. We performed Mendelian randomization (MR) analyses using the MendelianRandomization R package (v0.7.0), with random-effect inverse-variance weighting (IVW) as the primary method for assessing associations. A Bonferroni correction was applied to account for multiple comparisons, with a significance threshold set at $p < 0.05$. The deCODE cohort acted as the discovery dataset, and the UKB-PPP cohort served as the validation dataset. Instrumental variables were selected based on the following criteria: $p < 5 \times 10^{-8}$ for association with plasma protein levels, not in LD with any other selected SNP ($r^2 < 0.001$), and F-statistic > 10.

Polygenic risk scores evaluation

We constructed PRS from genome-wide summary-level data to assess the genetic contributions of distinct chromosomal regions to disease susceptibility. Specifically, we applied PRS with Continuous Shrinkage (PRS-CS) algorithm, which estimates posterior SNP effect sizes by incorporating an external LD reference panel. This Bayesian regression framework integrates summary-level GWAS data with LD information to produce shrinkage-based effect estimates and, ultimately, the PRS.

Results

Structural equation model fitting

Based on LD-Score regression analysis, the Z-scores of genetic heritability contributions for the 8 GWAS that constitute the allergic disorders genomic-SEM are almost greater than 3, indicating strong statistical significance and robust heritable components (Table S2). The genetic covariance values between each pair of traits are presented in Table S6 and Fig. 2. The factor model of the genetic covariance matrix and the empirical covariance matrix in the Genomic-SEM fits well (CFI = 0.964, SRMR = 0.115 (Supplementary Table S7)). Estimates of the latent variable loadings and residual covariances are provided in Table S8. These results provide evidence for the presence of shared genetic factors within the Genomic-SEM. After filtering SNPs with significant heterogeneity (Q p-value < 0.05), 511,970 SNPs were removed. The final genomic-SEM analysis generated an indirect measurement-based

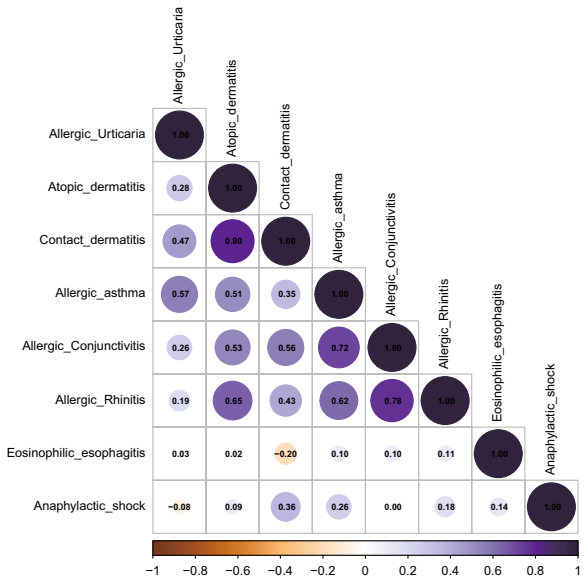

Fig. 2 Genetic Correlation Matrix of Allergic Disorders. The color intensity and circle size represent correlation strength, ranging from − 1 to + 1

GWAS comprising 6,918,772 SNPs to investigate the genetic architecture of allergic disorder multimorbidity.

Assessment of genomic structural equation model stability via LDSC for genomic-SEM

Using the LDSC method, we applied parameter controls and removed a total of 6,296,401 SNPs. The average chi-square value for the retained SNPs for genomic-SEM of allergic disorders was Mean chi-square = 1.514, the Lambda GC was 1.323, the Max chi-square was 121.56, the was h^2 = 0.0019 (SE = 0.0001), and the intercept was 1.0915 (SE = 0.0133) in the regression model, indicating a relatively good fit, with minor signs of population stratification or residual confounding. Additionally, the attenuation ratio of 0.1817 a moderate influence of confounding factors, with a stronger contribution from genetic effects.

Risk genetic loci

In the novel genomic-SEM, a total of 6,918,772 SNPs were analyzed, with 2,038 variants having a p-value < 5e-8 (Figure S1). Employing the GWAS-by-Subtraction methodology, we identified 31 novel significant variants that were not previously detected in the input trait GWAS statistics or documented in the GWAS Catalog databases (Fig. 3, Table S9). Functional annotation via the FUMA tool revealed that the majority of these loci were situated in intergenic (46.6%) and intronic (36.4%) regions, with a smaller fraction of 49 SNPs (1.0%) located in exonic regions. Further analysis identified 135 independent SNPs (r^2 < 0.6) and 83 lead SNPs (r^2 < 0.1) (Table S10). Among the lead SNPs, three novel loci—chr1: rs114695117

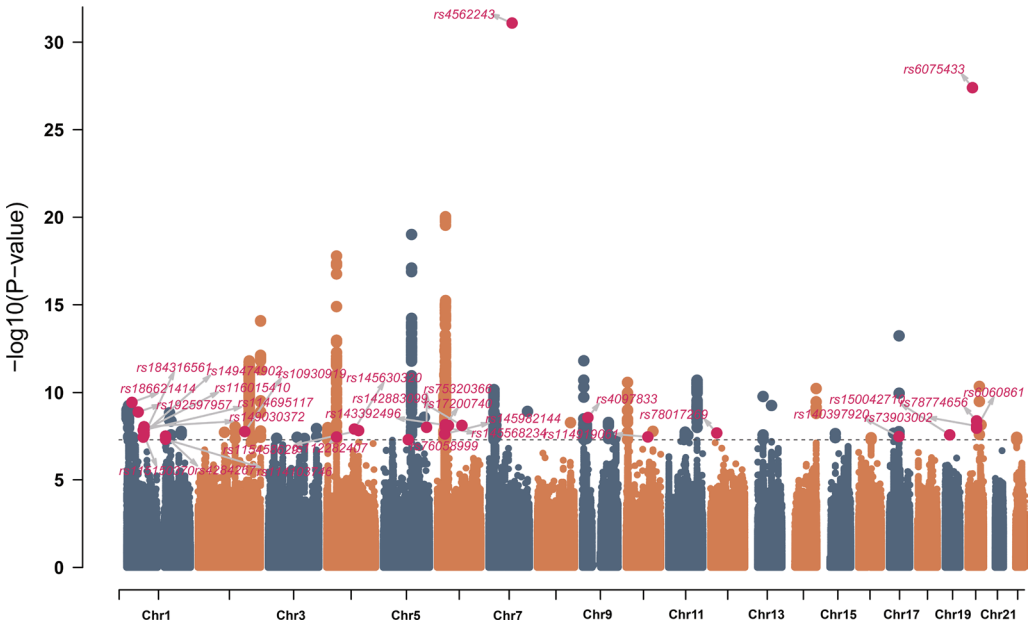

Fig. 3 Manhattan Plot of Novel Genomic-SEM Results for Allergic Disorders. The x-axis denotes chromosomal positions, while the y-axis represents the negative logarithm of the P-value (– log10(P)). The dashed line indicates the genome-wide significance threshold at $P = 5 \times 10^{-8}$. Labeled points correspond to novel SNPs identified through the novel GWAS approach

(near RAVER2), chr6: rs145982144 (near UBDP1), and chr17: rs140397920 (near MED24)—were identified, representing genomic regions previously unreported in association with allergic disorders. These lead SNPs contributed to a total of 66 distinct risk loci (Table S11). Gene-based analysis using MAGMA highlighted 62 potential allergic disorder-related genes, with the most significant enrichment observed on chromosome 2 and chromosome 6. Functional annotation analysis revealed that 59 out of 62 candidate genes (95.2%) encompassed genomic regions harboring at least one variant with high predicted pathogenicity (CADD Scaled C-Score ≥ 10), with 9 genes (14.5%) containing extremely deleterious variants (CADD ≥ 20) (Table S12, Figure S2).

Finemap

Fine-mapping analyses revealed strong associations at two genomic loci exceeding a posterior probability (PP) threshold of 0.95 and achieving genome-wide significance (GWAS $P < 5\times 10^{-8}$). These loci include chr6: rs145982144 (near *UBDP1*) and chr12: rs78017269 (near *SOX5*), with fine-mapping t-statistics of −5.61 and −5.59, and corresponding GWAS P-values of 2.03×10^{-8} and 2.33×10^{-8}, respectively (Fig. 4).

Gene-level identification of susceptibility

We performed sCCA-based TWAS, identifying 1350 genes that surpassed the FDR significance threshold. Among these, 35 genes had a TWAS P-value <5e-8 (Fig. 5A), while 60 genes exhibited a TWAS Z-value >6 (Fig. 5B). Subsequently, we conducted fine-mapping analysis using the FOCUS method on the genomic structural equation data, identifying 127 genes with a posterior inclusion probability (PIP) greater than 0.8, which suggests they may represent potential pathogenic signals. To further confirm the"high-confidence"gene-level associations, we performed an intersection test to include 31 unique genes (Table 1). The TWAS identified FAM114 A1 as the most strongly associated locus, demonstrating the highest positive association statistic ($Z = 11.13$) and most stringent significance threshold (FDR-P $= 1\times10^{-24}$). This was followed by RFTN2 ($Z = 6.93$, FDR-P $= 7\times10^{-9}$) and PLCL1 ($Z = 6.76$, FDR-P $= 1.6\times10^{-8}$), collectively indicating that elevated predicted expression of these genes is robustly correlated with allergic disorder susceptibility. Conversely, RP5-1115 A15.1 exhibited the strongest inverse association ($Z = -6.81$, FDR-P $= 1\times10^{-8}$), with RERE ($Z = -6.62$, FDR-P $= 3.9\times10^{-8}$) ranking second, suggesting that reduced expression of these loci may confer increased disease risk.

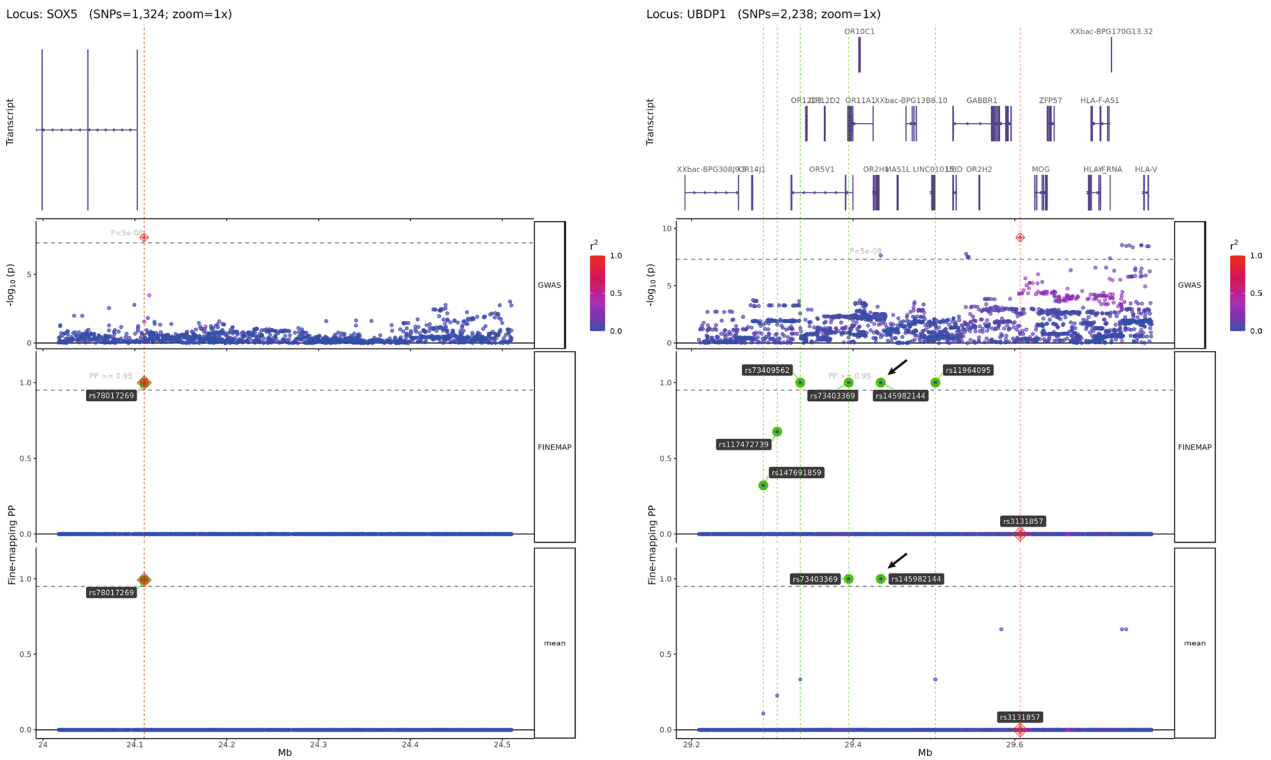

Fig. 4 Fine-mapping Results of Genomic Loci with Strong Associations (PP > 0.95) Identified by FINEMAP

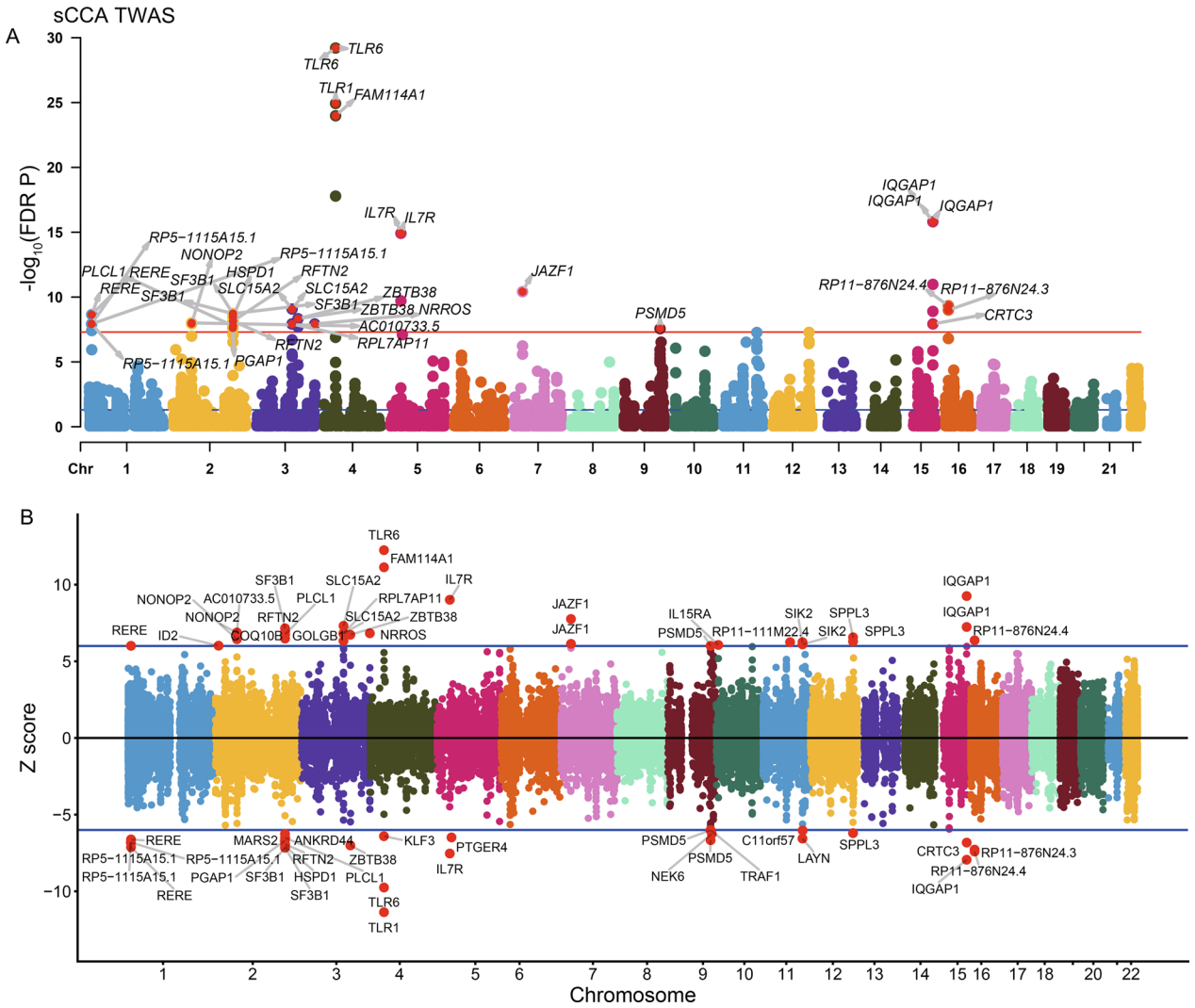

Fig. 5 Manhattan Plot of Results from sCCA TWAS Analysis for Allergic Disorders. Manhattan plot of the TWAS P-values for allergic disorders. The x-axis represents chromosomes, and the y-axis shows the negative logarithm of the P-value (– log10(p)). The horizontal red line corresponds to a threshold of TWAS P = 1e-12 (**A**). Manhattan plot of the TWAS Z-scores for allergic disorders. The x-axis represents chromosomes, and the y-axis displays the Z-scores. The horizontal blue lines mark the absolute Z-score value of 7, which represents the threshold for significance (**B**)

Pathway and cell type enrichment analysis

GSEA of MAGMA-derived risk loci highlighted REACTOME_ADAPTIVE_IMMUNE_ SYSTEM as the most significantly enriched pathway within the Curated Gene Sets collection. Delving into Canonical Pathways, we observed robust associations within three prominent immune-related modules including REACTOME_ADAPTIVE_IMMUNE_SYSTEM, REACTOME_ DISEASES_OF_IMMUNE_SYSTEM and REACTOME_COSTIMULATION_BY_THE_ CD28_ FAMILY. Notably, these enriched pathways encompassed genes critically involved in Th2 polarization, antigen presentation, and T-cell receptor signaling (Table S13). This mechanistic convergence strongly links these pathways to established allergic phenotypes such as asthma and atopic dermatitis, as well as immune-mediated comorbidities like type 1 diabetes and systemic lupus erythematosus. Furthermore, cataloged associations for these pathways extend beyond allergic disorders, encompassing conditions like asthma, type 1 diabetes, and systemic lupus erythematosus, and also extending to neuropsychiatric conditions such as autism spectrum disorder and schizophrenia (Figure S3). Additionally, through MendelVar enrichment mapping, we observed significant gene overlap with diseases

Table 1 Genetic associations with allergic disorders in sCCA and FOCUS analysis

Gene	CHR	Srart position	End position	Heritability Squared	TWAS Z	TWAS FDR P	FOCUS pip
RP5-1115 A15.1	1	8,424,644	8,424,645	0.1611	− 6.81478	1.2E-08	0.923
RERE	1	8,817,642	8,817,643	0.3321	− 6.61544	3.9E-08	0.997
RFTN2	2	197,676,044	197,676,045	0.1435	6.93279	7.2E-09	0.969
PLCL1	2	197,804,701	197,804,702	0.1789	6.757755	1.6E-08	0.837
IL1R2	2	101,991,843	101,991,844	0.4395	4.325578	1.8E-03	0.95
FARP2	2	241,356,242	241,356,243	0.3243	3.11431	5.0E-02	0.932
ATP1B3	3	141,876,123	141,876,124	0.0905	− 4.92	2.0E-04	0.999
FAM114 A1	4	38,867,676	38,867,677	0.3223	11.1337	1.0E-24	1
KLF3	4	38,680,586	38,680,587	0.0735	− 6.40443	1.3E-07	0.999
FAM114 A1	4	38,867,676	38,867,677	0.5296	5.56709	1.1E-05	1
FAM114 A1	4	38,867,676	38,867,677	0.3542	5.0457	1.2E-04	0.997
FAM114 A1	4	38,867,676	38,867,677	0.3542	5.0457	1.2E-04	0.971
RFC1	4	39,366,374	39,366,375	0.1071	4.52361	9.2E-04	0.997
RFC1	4	39,366,374	39,366,375	0.1071	4.52361	9.2E-04	1
SPATA5	4	122,923,073	122,923,074	0.2112	3.83259	8.5E-03	0.98
U91328.19	6	25,992,661	25,992,662	0.7309	5.824	3.0E-06	0.964
U91328.19	6	25,992,661	25,992,662	0.475	− 5.64151	7.9E-06	0.998
ABT1	6	26,596,951	26,596,952	0.2157	− 3.6894	1.3E-02	0.902
PVT1	8	127,794,532	127,794,533	0.2175	5.58441	1.0E-05	0.883
RANBP6	9	6,015,624	6,015,625	0.1052	4.31736	1.8E-03	0.956
KIAA2026	9	6,007,824	6,007,825	0.1628	4.26985	2.1E-03	0.984
TGFBR1	9	99,104,037	99,104,038	0.1129	− 3.76475	1.0E-02	0.985
IL15RA	10	5,978,186	5,978,187	0.3351	6.06753	8.8E-07	1
SUFU	10	102,503,986	102,503,987	0.109	5.95708	1.5E-06	0.916
RBM17	10	6,088,986	6,088,987	0.2453	5.18371	6.9E-05	0.994
CUEDC2	10	102,432,660	102,432,661	0.0362	− 4.91882	2.0E-04	0.874
RP11-563 J2.2	10	6,277,686	6,277,687	0.2516	− 3.89475	7.0E-03	0.993
PRKCQ-AS1	10	6,580,418	6,580,419	0.1954	3.78693	9.8E-03	0.999
SIK2	11	111,602,390	111,602,391	0.1017	6.103014	7.2E-07	0.998
C11orf1	11	111,878,934	111,878,935	0.1407	− 4.76819	3.7E-04	0.933
ELF1	13	41,061,272	41,061,273	0.0806	− 5.17324	7.1E-05	0.983
CHP1	15	41,230,838	41,230,839	0.2169	5.69736	6.0E-06	0.95
TOM1L2	17	17,972,421	17,972,422	0.0527	− 4.43104	1.2E-03	0.936
KRT15	17	41,522,528	41,522,529	0.3557	− 4.138814	3.2E-03	0.834
CCDC116	22	21,632,715	21,632,716	0.4066	− 4.66051	5.7E-04	0.945
LINC01637	22	20,957,091	20,957,092	0.2046	− 3.639734	1.5E-02	0.823

such as *Mayer-Rokitansky-Kuster-Hauser syndrome*, *osteochondrodysplasia*, and *immune system diseases* like *primary immunodeficiency*. Diseases related to glucose metabolism and liver conditions, including *liver cirrhosis*, were also enriched (Figure S4). The most significant cell types enriched in the allergic disorder GWAS were primarily immune-related, with a strong representation of T cells, natural killer (NK) cells, and B cells. These cell types, including Fat T cells, Marrow B cells, Limb Muscle T cell, Lung T cells, and Marrow regulatory T cell, Trachea blood cell, Spleen T cell, play crucial roles in immune regulation and allergic responses. Additionally, myeloid cells and leukocytes from various tissues, such as the Lung myeloid cells and Kidney leukocytes, were also significantly enriched (Table S14).

Heritability enrichment across genomic functional and regulatory regions

Based on the genetic enrichment analysis of genomic functional regions using S-LDSC (Table S15), we observed a significant enrichment pattern of heritability in epigenetic regulatory elements and core functional regions of genes. Specifically, evolutionarily enhancer regions, conserved elements, and DNase I hypersensitive sites (DGF) showed the strongest positive genetic contribution, suggesting that these regulatory elements play a key role in influencing complex traits through cis-regulatory networks. These findings underscore the critical role of cis-regulatory networks, orchestrated through these genomic elements, in influencing the complex genetic architecture of allergic disorders, highlighting that variations within these regulatory regions are likely key drivers of disease susceptibility and multimorbidity.

Genetic correlations with external traits and plasma biomarkers identification

Comprehensive genetic correlation analyses using LDSC revealed significant pleiotropy between allergic disorders and 12 out of 115 common diseases. Notably, acute upper respiratory infections exhibited the strongest positive genetic correlation (rg = 0.143 ± 0.046, P = 0.002), followed by bipolar affective disorders (rg = 0.136 ± 0.05, P = 0.007) and Crohn's disease (rg = 0.148 ± 0.075, P = 0.048). Conversely, inverse associations were observed for disorders of lipoprotein metabolism (rg = − 0.101 ± 0.043, P = 0.02) and heart failure (rg = − 0.088 ± 0.043, P = 0.04), suggesting potential protective mechanisms (Fig. 6). Two-sample MR analysis revealed seven plasma proteins with genetically predicted levels significantly associated with allergic disorders after Bonferroni correction in the deCODE discovery cohort: *MANF*, *GLB1*, *HEXIM1*, *KYNU*, *ICAM4*, *KRT5*, and *LAYN*. Of these, four proteins (*MANF*, *GLB1*, *ICAM4*, and *LAYN*) were consistently replicated in the UKB-PPP cohort using inverse-variance weighted models, demonstrating bidirectional concordance (Fig. 7).

Construction of polygenic risk scores from summary data

Our analysis demonstrates that the variants in our PRS are strongly associated with disease onset risk, and that the genetic contribution to disease susceptibility varies significantly across chromosomal regions (Table S16). Specifically, when considering all SNPs, chromosomes 14 and 20 exhibit the highest positive genetic contribution. However, among the top 1000 most contributing SNPs, chromosomes 6 (86 SNPs) and 5 (85 SNPs) are most prominent, while within the top 10,000, chromosomes 2 (821 SNPs) and 1 (805 SNPs) are most frequent. This shift in prominent chromosomes suggests that different sets of genes and regulatory elements, located in these regions, influence disease susceptibility at different levels of genetic contribution.

Discussion

Allergic disorders, encompassing conditions such as allergic asthma, atopic dermatitis, allergic rhinitis, and others, represent a significant global health burden, characterized by not only highly prevalent but also exhibit complex comorbidities and multimorbidities, often progressing through a well-documented sequence known as the"atopic march". The genetic basis of allergic multimorbidity remains underexplored, as traditional single-trait GWAS approaches face inherent limitations in dissecting the intricate genetic and phenotypic interplay underlying co-occurring allergic disorders. Previously, genomic SEM has demonstrated utility in identifying shared genetic architectures, such as the comorbidity between inflammatory bowel disease and asthma, as well as associations between allergic disorders and inflammatory diseases, body mass index, and gastroesophageal reflux disease [12–14]. We applied this novel approach to integrate genome-wide association data by first constructing a genetic covariance matrix via multivariable LDSC, followed by specifying a latent factor model representing eight allergic traits within the SEM framework. The application of Genomic-SEM to metabolic syndrome research recently successfully identified 159 novel SNPs not previously cataloged in the GWAS Catalog, along with 82 additional SNPs exhibiting significant pleiotropic effects independent of the constituent metabolic syndrome traits [25]. Similarly, Andrew et al. used this method to analyze gene expression profiles across 13 major psychiatric disorders. This analysis successfully identified genes associated with both shared and disorder-specific genetic risks, leading to the proposal of repurposing existing drugs as potential therapeutic agents [26].

Through the novel established model and a suite of integrative post-GWAS analyses, we identified 31 novel, pleiotropic variants and prioritized 31 key genes significantly associated with allergic disorder multimorbidity. Among the lead SNPs, three—chr1:rs114695117 (near *RAVER2*), chr6:rs145982144 (near *UBDP1*), and chr17:rs140397920 (near *MED24*)—represent previously unreported associations with allergic disorders. The variant rs114695117 is located near *RAVER2*, which encodes a ribonucleoprotein involved in the regulation of alternative splicing. Previous studies have implicated *RAVER2* in splicing events within the thymic epithelium [27] and identified it as a susceptibility gene for liver injury in rheumatoid arthritis patients treated with immunomodulatory drugs [28]. Consequently, *RAVER2* may contribute to allergic inflammation by modulating the splicing patterns

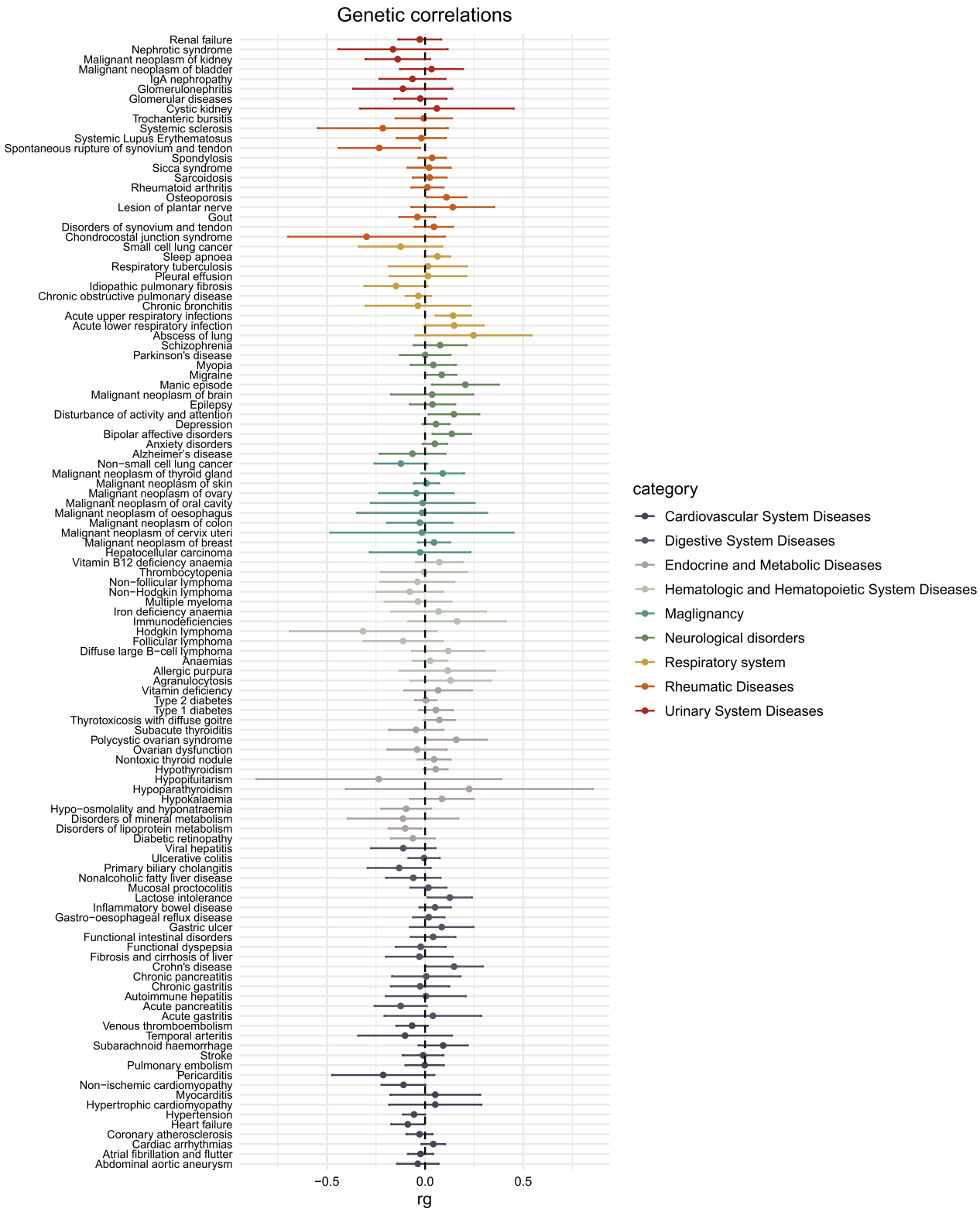

Fig. 6 Genetic associations between allergic disorders and external traits

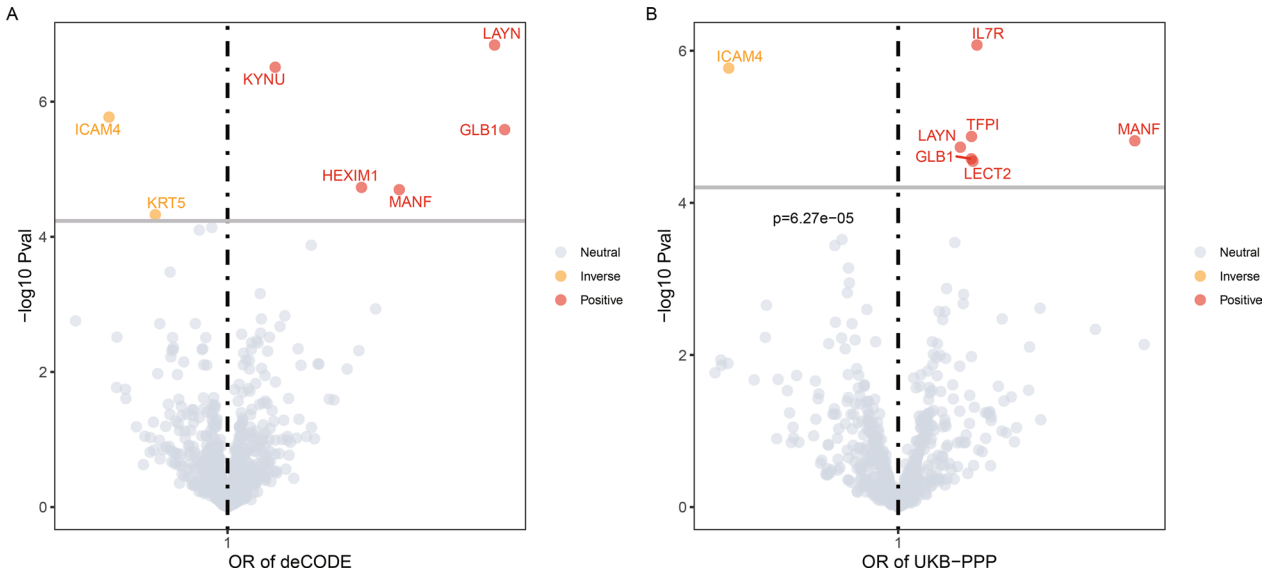

Fig. 7 Volcano Plot of Plasma Protein Exposures and Associations with Allergic Disorders in Mendelian Randomization Analysis. Volcano plots of Mendelian randomization results for plasma protein exposures and allergic disorders in the deCODE discovery cohort (**A**) and UKB-PPP cohort (**B**)

of immune-related genes. The variant rs145982144 lies near the pseudogene *UBDP1*, a component of the ubiquitin system. While the precise role of *UBDP1* in immune regulation remains to be elucidated, ubiquitination is known to play a critical role in protein degradation, DNA repair, and immune signaling, and its dysregulation has been implicated in autoimmunity via modulation of NF-κB and STAT signaling pathways [29]. Therefore, rs145982144 might influence allergic disease susceptibility by altering *UBDP1* function, potentially affecting the activity of these pathways. Finally, rs140397920 is located near *MED24*, a subunit of the Mediator complex, which bridges transcription factors and RNA polymerase II, thereby regulating the expression of specific genes. Within the immune system, the Mediator complex is involved in the transcriptional regulation of T cell programs [30, 31]. Notably, *MED24* has previously been identified within a region containing asthma risk alleles, supporting its potential role in allergic disease susceptibility [32].

Furthermore, fine-mapping analyses identified two putative causal SNPs associated with allergic disorders—rs145982144 (near *UBDP1*) and rs78017269 (near *SOX5*). Although these SNPs represent novel discoveries in allergic disease GWAS, their positional genes have prior indirect evidence supporting immunological relevance. Specifically, *SOX5* regulates T-cell receptor signaling to maintain immune tolerance—a mechanism critically implicated in suppressing type 2 inflammation [33]. This aligns with existing GWAS evidence linking *SOX5* variants to asthma susceptibility, further supporting its pathological significance in allergic pathogenesis [34, 35].

Our sCCA and subsequent FOCUS fine-mapping identified 31 genes with potential causal roles in allergic disorders. While some of these genes, such as *IL1R2*, *IL15RA*, *TGFBR1*, and *PVT1*, have established associations with allergic traits [36, 37], others represent novel candidates with plausible links to allergic pathways. Notably, *FAM114 A1* ($Z = 11.13$, FDR-P $= 1 \times 10{-}24$), *RFTN2* ($Z = 6.93$, FDR-P $= 7 \times 10{-}9$), *PLCL1* ($Z = 6.76$, FDR-P $= 1.6 \times 10{-}8$), *RP5-1115 A15.1* ($Z = -6.81$, FDR-P $= 1.2 \times 10{-}8$), and *RERE* ($Z = -6.62$, FDR-P $= 3.9 \times 10{-}8$) emerged as high-confidence loci, underscoring their potential critical roles in allergic pathogenesis. *FAM114 A1*, a relatively uncharacterized gene, has been implicated in immune responses and apoptosis [38]. Furthermore, its role in myocardial fibrosis via regulation of angiotensin II has been demonstrated [39], and angiotensin II itself has been linked to increased asthma risk and airway inflammation [40, 41]. This suggests a potential mechanism whereby *FAM114 A1* might influence allergic processes, possibly modulated by Angiotensin-(1–7). However, the precise function of *FAM114 A1* remains largely unexplored and warrants further investigation. *RFTN2* is involved in, or upstream of, dsRNA transport and is predicted to correlate with type I interferon expression, a pathway implicated in ulcerative colitis susceptibility [42]. Recent reports also highlight *RFTN2*'s function in neuroglia [43, 44], adding further complexity to its potential roles. *PLCL1*, located on chromosome 2,

encodes a phospholipase C-like enzyme involved in signal transduction. It generates second messengers, such as IP3 and DAG, activating calcium signaling and protein kinase C, thereby influencing B and T cell function. Because allergic diseases involve IgE production by B cells and mast cell degranulation, increased *PLCL1* expression could potentially enhance these immune cell responses, elevating allergic risk. While direct studies linking *PLCL1* to allergy are limited, research on rheumatoid arthritis demonstrates its involvement in regulating key inflammatory cytokines, including IL-6 and IL-1β [45], suggesting a possible parallel mechanism in allergic diseases. Intriguingly, several other identified genes, while not directly implicated in allergic diseases in previous reports, have demonstrated roles in the pathogenesis of glioma, including *SUFU* [46], *RBM17* [47], *FAM114 A1* [48], *CHP1* [49], *CUEDC2* [50], *RFC1* [51], *RANBP6* [52], *PRKCQ-AS1* [53], and *FARP2* [54]. Given the emerging concept of glioma as a marker AllergoOncology disease, these findings may shed light on the complex genetic underpinnings of AllergoOncology. The enrichment of genes associated with Th2 polarization, T-cell receptor signaling, and other immune regulatory processes in this analysis aligns seamlessly with previous research. Han et al. identified genetic variants in these immune pathways as significant contributors to asthma susceptibility, demonstrating that asthma-associated variants are enriched in open chromatin regions of immune cells, particularly $CD4^+$ and $CD8^+$ T lymphocytes. Their findings also revealed an overrepresentation of T-cell receptor signaling pathways among asthma risk loci [55]. Recent studies have further highlighted the role of key genes, such as *CD52*—encoding a membrane glycoprotein expressed on various leukocytes—in linking T-cell-related pathways to asthma phenotypes [55–57]. Furthermore, the PRS analysis underscores the variability and intricate genetic architecture underlying allergic disorders. This observation highlights the distinct roles of specific chromosomes in the pathogenesis of these conditions.

Our novel SEM analysis explored the shared genetic architecture and its correlation with external traits, identifying strong comorbidities, primarily neuropsychiatric disorders. These included bipolar disorder, manic episode, migraine, and attention deficit hyperactivity disorder (ADHD). This finding aligns with prior clinical observations and post-GWAS research on allergic traits like asthma, atopic dermatitis, allergic rhinitis, and mental disorders [58–62]. It suggests that the complex comorbidity of allergic traits is driven by a shared genetic and environmental etiology, with our research indicating that the link between allergic disorders and neuropsychiatric conditions may stem from this shared genetic architecture. This further underscores the importance of investigating causal loci.

Our MR findings highlight four plasma proteins—MANF, GLB1, ICAM4, and LAYN—exhibiting robust associations indicative of a potential causal role in allergic disorders. This finding addresses limitations of prior retrospective studies [63–65], providing genetic evidence of causal effects. These proteins, collectively implicated in regulating immune homeostasis and inflammation, offer novel mechanistic targets in allergy. The replicated associations and suggestive causal evidence underscore their translational value as potential therapeutic interventions or diagnostic biomarkers, urgently necessitating functional studies to elucidate their precise regulatory networks in allergic inflammation.

Our study has several limitations. First, the shared genetic architecture of allergic disorders was investigated primarily in individuals of European ancestry. Since genomic-SEM relies on LDSC—a method sensitive to ancestry-specific LD patterns—our analysis was confined to a genetically homogeneous population. Consequently, validation of this model across diverse ethnic populations is imperative to establish its generalizability in future studies. Second, the susceptibility loci identified in this study require experimental validation, particularly functional studies at the protein level, to elucidate their translational relevance and mechanistic roles in allergic pathophysiology. Third, while Cochran's Q-test was applied to assess heterogeneity (Q_{SNP}), this approach prioritizes the detection of SNPs with pleiotropic effects across multiple traits rather than those exhibiting heterogeneous effects on specific allergic phenotypes. This methodological focus may have obscured SNPs with trait-specific associations critical to individual allergic disorders. Finally, the scarcity of GWAS data for allergic disorders in East Asian populations precluded validation of our European ancestry-derived genetic model in other ancestries. Addressing this gap through expanded sample sizes in underrepresented populations remains a critical priority to enhance the robustness and global applicability of our findings. Targeting pleiotropic loci identified in future research holds significant promise for drug development, as these loci could enable simultaneous therapeutic intervention across multiple allergic diseases. However, given the intricate relationship between allergy and immunity, it is equally crucial to actively investigate the role of these genes in immune regulation. In this study, we identified a genetic correlation between allergic disorders and other traits, notably psychiatric disorders. However, integrating biological experiments such as functional validation of these loci through gene

expression analysis or CRISPR-based gene editing can substantiate the biological significance of these genetic findings and elucidate their underlying mechanisms [66, 67]. Moreover, given the recent successes of clustering and non-clustering machine learning algorithms, as well as artificial intelligence, developing additional computational methods represents a viable and promising strategy to further elucidate the mechanisms underlying allergic disorders [68–71].

Conclusion

Leveraging genomic SEM, our novel GWAS elucidated the shared genetic architecture of allergic traits. Employing a suite of post-GWAS methodologies, we robustly identified 2038 genome-wide significant SNP loci, including 31 previously unreported loci. Furthermore, integrating sCCA with FOCUS, we precisely pinpointed 31 candidate causal genes. Genetic correlation analyses further illuminated the shared genetic architecture underlying multiple traits, notably psychiatric disorders. Moreover, we identified four putative causal plasma protein biomarkers: MANF, GLB1, ICAM4, and LAYN.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06465-8.

Supplementary material 1. Additional file 1. Figure S1. QQ Plot of GWAS Summary Statistics for Allergic Disorders.Figure S2. Manhattan Plot of GWAS Results for Allergic Disorders from MAGMA Analysis.Figure S3. Pathway Enrichment Analysis Using Catalog Gene Enrichment from the MSigDB Database. Figure S4. Disease Enrichment Analysis from the MendelVar Database.

Supplementary material 2.Additional file 2. Table S1. GWAS summary source.Table S2. SNP heritability of Genomic-SEM phenotypes. Table S3. Fit Indices for Genomic-SEM Model. Table S4. Genomic-SEM Factor Loadings and Variance Estimates for Allergic Traits. Table S5. Common Diseases or Disorders in the FinnGen R12 Dataset. Table S6. Genetic Correlations Between Allergic Disorders. Table S7. Model Fit Indices for the Genomic Structural Equation Model. Table S8. Factor Loadings and Residual Covariances in Genomic Structural Equation Model of Allergic Disorders. Table S9. Novel SNP Variants Identified by Genomic-SEM. Table S10. Lead SNP Identified by Genomic-SEM Table S11. Risk locus Identified by Genomic-SEM. Table S12. MAGMA Risk Gene Annotation Using Genomic-SEM. Table S13. Enriched Pathways by MsigDB. Table S14. Enriched Cell Types in GWAS for Allergic Disorders. Table S15. Heritability Enrichment Across Genomic Functional and Regulatory Regions. Table S16. Polygenic Risk Score and Genetic Contribution Across Chromosomal Regions.

Author contributions

X.L. Y. conceived and conducted the analysis, R.J.S. completed the visualization, and they jointly wrote the manuscript.

Funding

Not available.

Data availability

Data generated during this study are available in the Supplementary Files.

Declarations

Ethics approval and consent to participate

This study utilized publicly available data and therefore did not require ethical approval.

Competing interests

Not available.

Received: 6 March 2025 Accepted: 7 April 2025
Published online: 15 April 2025

References

1. Shin YH, Hwang J, Kwon R, Lee SW, Kim MS, GBD, Allergic Disorders Collaborators, et al. Global, regional, and national burden of allergic disorders and their risk factors in 204 countries and territories from 1990 to 2019: a systematic analysis for the Global Burden of Disease Study. Allergy. 2019;78(8):2232–54.
2. Tian J, Zhang D, Yang Y, Huang Y, Wang L, Yao X, et al. Global epidemiology of atopic dermatitis: a comprehensive systematic analysis and modelling study. Br J Dermatol. 2024;190:55–61.
3. Santra G. A comprehensive review on allergic disorders, their epidemiological trend and barriers in management. J Assoc Physicians India. 2024;72:73–80.
4. Rogliani P, Laitano R, Ora J, Beasley R, Calzetta L. Strength of association between comorbidities and asthma: a meta-analysis. Eur Respir Rev. 2023;32:220202.
5. Torabizadeh M, Aghaei M, Saki N, Vahid MA, Bitaraf S, Bandar B. The association of nasal and blood eosinophils with serum IgE level in allergic rhinitis and asthma: a case-control study. Health Sci Rep. 2024;7:e70191.
6. Tran MM, Lefebvre DL, Dharma C, Dai D, Lou WYW, Subbarao P, et al. Predicting the atopic march: Results from the Canadian Healthy Infant Longitudinal Development Study. J Allergy Clin Immunol. 2018;141:601-7.e8.
7. Zysk W, Mesjasz A, Trzeciak M, Horvath A, Plata-Nazar K. Gastrointestinal comorbidities associated with atopic dermatitis—a narrative review. Int J Mol Sci. 2024;25:1194.
8. Kallio S, Jian C, Korpela K, Kukkonen AK, Salonen A, Savilahti E, et al. Early-life gut microbiota associates with allergic rhinitis during 13-year follow-up in a Finnish probiotic intervention cohort. Microbiol Spectr. 2024;12:e0413523.
9. Demirdag Y, Bahna S. The role of genetics in food allergy. Expert Rev Clin Immunol. 2022;18:401–11.
10. Wang J, Zhou Y, Zhang H, Hu L, Liu J, Wang L, et al. Pathogenesis of allergic diseases and implications for therapeutic interventions. Signal Transduct Target Ther. 2023;8:138.
11. Li X, Li J, Xue S, Gao Y, Wan L, Wang C, et al. Shared genetic investigation of asthma and blood eosinophils in relation to chronic rhinosinusitis. Allergy Asthma Clin Immunol. 2025;21:11.
12. Gong T, Kuja-Halkola R, Harder A, Lundholm C, Smew AI, Lehto K, et al. Shared genetic architecture between gastro-esophageal reflux disease, asthma, and allergic diseases. Commun Biol. 2024;7:1–11.
13. Han X, Zhu Z, Xiao Q, Li J, Hong X, Wang X, et al. Obesity-related biomarkers underlie a shared genetic architecture between childhood body mass index and childhood asthma. Commun Biol. 2022;5:1–10.
14. Gong T, Brew BK, Lundholm C, Smew AI, Harder A, Kuja-Halkola R, et al. Comorbidity between inflammatory bowel disease and asthma and allergic diseases: a genetically informed study. Inflamm Bowel Dis. 2024;30:1556–65.
15. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nat Hum Behav. 2019;3:513–25.
16. Hur HJ, Yang HJ, Kim MJ, Lee K-H, Kim M-S, Park S. Association of polygenic variants with type 2 diabetes risk and their interaction with lifestyles in Asians. Nutrients. 2022;14:3222.

17. Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, Wucher V, Gewirtz ADH, Cotter DJ, et al. The impact of sex on gene expression across human tissues. Science. 2020. https://doi.org/10.1530/ey.18.14.14.
18. Castanza AS, Recla JM, Eby D, Thorvaldsdóttir H, Bult CJ, Mesirov JP. Extending support for mouse data in the molecular signatures database (MSigDB). Nat Methods. 2023;20:1619–20.
19. Zhang L, Ge Q, Sun Z, Zhang R, Li X, Luo X, et al. Association and shared biological bases between birth weight and cortical structure. Transl Psychiatry. 2025;15:1–9.
20. Tabula Muris Consortium, Overall coordination, Logistical coordination, Organ collection and processing, Library preparation and sequencing, Computational data analysis, et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature. 2018. https://doi.org/10.1038/s41586-018-0590-4.
21. Timshel PN, Thompson JJ, Pers TH. Genetic mapping of etiologic brain cell types for obesity. Elife. 2020;9:e55851.
22. Ostrom QT, Edelson J, Byun J, Han Y, Kinnersley B, Melin B, et al. Partitioned glioma heritability shows subtype-specific enrichment in immune cells. Neuro Oncol. 2021;23:1304–14.
23. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet. 2021;53:1712–21.
24. Sun BB, Chiou J, Traylor M, Benner C, Hsu Y-H, Richardson TG, et al. Plasma proteomic associations with genetics and health in the UK Biobank. Nature. 2023;622:329–38.
25. Park S, Kim S, Kim B, Kim DS, Kim J, Ahn Y, et al. Multivariate genomic analysis of 5 million people elucidates the genetic architecture of shared components of the metabolic syndrome. Nat Genet. 2024;56:2380–91.
26. Grotzinger AD, Singh K, Miller-Fleming TW, Lam M, Mallard TT, Chen Y, et al. Transcriptome-wide structural equation modeling of 13 major psychiatric disorders for cross-disorder risk and drug repurposing. JAMA Psychiat. 2023;80:811–21.
27. Padonou F, Gonzalez V, Provin N, Yayilkan S, Jmari N, Maslovskaja J, et al. Aire-dependent transcripts escape Raver2-induced splice-event inclusion in the thymic epithelium. EMBO Rep. 2022;23:e53576.
28. Eektimmerman F, Swen JJ, den Broeder AA, Hazes JMW, Kurreeman FS, Verstappen SMM, et al. Genome-wide association study of methotrexate-induced liver injury in rheumatoid arthritis patients. Clin Pharmacol Ther. 2023;113:916–23.
29. Khanra S, Singh S, Singh TG. Mechanistic exploration of ubiquitination-mediated pathways in cerebral ischemic injury. Mol Biol Rep. 2024;52:22.
30. Chaudhuri SM, Weinberg SE, Wang D, Yalom LK, Montauti E, Iyer R, et al. Mediator complex subunit 1 architects a tumorigenic Treg cell program independent of inflammation. Cell Rep Med. 2024;5: 101441.
31. Li Z, Sun M, Yang R, Wang Z, Zhu Q, Zhang Y, et al. Mediator complex subunit 1 promotes oral squamous cell carcinoma progression by activating MMP9 transcription and suppressing CD8+ T cell antitumor immunity. J Exp Clin Cancer Res. 2024;43:270.
32. Ilchuk LA, Kubekina MV, Okulova YD, Silaeva YY, Tatarskiy VV, Filatov MA, et al. Genetically Engineered Mice Unveil In Vivo Roles of the Mediator Complex. Int J Mol Sci. 2023;24:9330.
33. Kanno T, Konno R, Miyako K, Nakajima T, Yokoyama S, Sasamoto S, et al. Characterization of proteogenomic signatures of differentiation of CD4+ T cell subsets. DNA Res. 2023. https://doi.org/10.1093/dnares/dsac054.
34. Chen Z, Shang Y, Zhang X, Duan W, Li J, Zhu L, et al. METTL3 mediates SOX5 m6A methylation in bronchial epithelial cells to attenuate Th2 cell differentiation in T2 asthma. Heliyon [Internet]. 2024 [cited 2025 Mar 27];10. Available from: https://www.cell.com/heliyon/abstract/S2405-8440(24)04915-6
35. Chen Z, Shang Y, Zhang X, Duan W, Li J, Zhu L, et al. METTL3 mediates SOX5 m6A methylation in bronchial epithelial cells to attenuate Th2 cell differentiation in T2 asthma. Heliyon. 2024;10:e28884.
36. Ma L, Zhang Q, Hao J, Wang J, Wang C. LncRNA PVT1 exacerbates the inflammation and cell-barrier injury during asthma by regulating miR-149. J Biochem Mol Toxicol. 2020;34:e22563.
37. Olafsdottir TA, Theodors F, Bjarnadottir K, Bjornsdottir US, Agustsdottir AB, Stefansson OA, et al. Eighty-eight variants highlight the role of T cell regulation and airway remodeling in asthma pathogenesis. Nat Commun. 2020;11:393.
38. Zhang H, Zeng Y, Ye C, Cai J, Hu X. Inhibition of FAM114A1 suppresses hepatocellular carcinoma by targeting AKT1 signaling. Ann Clin Lab Sci. 2024;54:378–87.
39. Subbaiah KCV, Wu J, Tang WHW, Yao P. FAM114A1 influences cardiac pathological remodeling by regulating angiotensin II signaling. JCI Insight. 2022;7:e152783.
40. Liu G, Chen Y, Wang Y, Deng X, Xiao Q, Zhang L, et al. Angiotensin II enhances group 2 innate lymphoid cell responses via AT1a during airway inflammation. J Exp Med. 2022;219:e20211001.
41. Patel M, Kurade M, Rajalingam S, Bhavsar R, Mustafa SJ, Ponnoth DS. Role of angiotensin II type 1 (AT1) and type 2 (AT2) receptors in airway reactivity and inflammation in an allergic mouse model of asthma. Immunopharmacol Immunotoxicol. 2019;41:428–37.
42. Gao H, Liu Z. The latest breakthrough on genetic characteristics of inflammatory bowel disease in Chinese and other East Asian ancestries. Precis Clin Med. 2023. https://doi.org/10.1093/pcmedi/pbad017.
43. Wei W-H, Ma S, Fu B, Song R, Guo H. Human-specific insights into candidate genes and boosted discoveries of novel loci illuminate roles of neuroglia in reading disorders. Genes Brain Behav. 2024;23:e12899.
44. Xiong Z, Guan H, Pei S, Wang C. Identification of metabolism-related subtypes and feature genes of pre-eclampsia. Sci Rep. 2025;15:4986.
45. Luo S, Li X-F, Yang Y-L, Song B, Wu S, Niu X-N, et al. PLCL1 regulates fibroblast-like synoviocytes inflammation via NLRP3 inflammasomes in rheumatoid arthritis. Adv Rheumatol. 2022;62:25.
46. Li Q, Liu Y, Wu J, Zhu Z, Fan J, Zhai L, et al. P4HA2 hydroxylates SUFU to regulate the paracrine Hedgehog signaling and promote B-cell lymphoma progression. Leukemia. 2024;38:1751–63.
47. Lu J, Li Q, Cai L, Zhu Z, Guan J, Wang C, et al. RBM17 controls apoptosis and proliferation to promote Glioma progression. Biochem Biophys Res Commun. 2018;505:20–8.
48. Liu B, Su Q, Xiao B, Zheng G, Zhang L, Yin J, et al. RAB42 promotes glioma pathogenesis via the VEGF signaling pathway. Front Oncol. 2021;11:657029.
49. Zhao B, Xuan R, Yang G, Hu T, Chen Y, Cai L, et al. A novel golgi related genes based correlation prognostic index can better predict the prognosis of glioma and responses to immunotherapy. Discov Oncol. 2025;16:212.
50. Liu X-F, Tang C-X, Zhang L, Tong S-Y, Wang Y, Abdulrahman AA, et al. Down-regulated CUEDC2 increases GDNF expression by stabilizing CREB through reducing its ubiquitination in glioma. Neurochem Res. 2020;45:2915–25.
51. Almeida J, Costa J, Coelho P, Cea V, Galesio M, Noronha JP, et al. Adipocyte proteome and secretome influence inflammatory and hormone pathways in glioma. Metab Brain Dis. 2019;34:141–52.
52. Oldrini B, Hsieh W-Y, Erdjument-Bromage H, Codega P, Carro MS, Curiel-García A, et al. EGFR feedback-inhibition by Ran-binding protein 6 is disrupted in cancer. Nat Commun. 2017;8:2035.
53. Maimaiti A, Tuerhong M, Wang Y, Aisha M, Jiang L, Wang X, et al. An innovative prognostic model based on autophagy-related long noncoding RNA signature for low-grade glioma. Mol Cell Biochem. 2022;477:1417–38.
54. Toledano S, Sabag AD, Ilan N, Liburkin-Dan T, Kessler O, Neufeld G. Plexin-A2 enables the proliferation and the development of tumors from glioblastoma derived cells. Cell Death Dis. 2023;14:41.
55. Han Y, Jia Q, Jahani PS, Hurrell BP, Pan C, Huang P, et al. Genome-wide analysis highlights contribution of immune system pathways to the genetic architecture of asthma. Nat Commun. 2020;11:1776.
56. Shafiei-Jahani P, Helou DG, Hurrell BP, Galle-Treger L, Howard E, Quach C, et al. CD52-targeted depletion by Alemtuzumab ameliorates allergic airway hyperreactivity and lung inflammation. Mucosal Immunol. 2021;14:899–911.
57. Stikker BS, Hendriks RW, Stadhouders R. Decoding the genetic and epigenetic basis of asthma. Allergy. 2023;78:940–56.
58. Cao S, Zhang Z, Liu L, Li Y, Li W, Li Y, et al. Causal relationships between atopic dermatitis and psychiatric disorders: a bidirectional two-sample Mendelian randomization study. BMC Psychiatry. 2024;24:16.
59. Brown ES. The complex relationship between asthma and psychiatric illnesses. J Allergy Clin Immunol Pract. 2023;11:809–10.
60. Wang Y, Wang S, Wu J, Liu X, Zhang L. Causal association between allergic diseases and dementia: evidence from multivariate mendelian randomization study. J Alzheimers Dis. 2024;98:505–17.

61. Oh H, Stickley A, Lincoln KD, Koyanagi A. Allergies, infections, and psychiatric disorders among Black Americans: findings from the National Survey of American Life. Ethn Health. 2022;27:74–82.
62. Conway AE, Verdi M, Kartha N, Maddukuri C, Anagnostou A, Abrams EM, et al. Allergic Diseases and Mental Health. J Allergy Clin Immunol Pract. 2024;12:2298–309.
63. Bahreiny SS, Ahangarpour A, Aghaei M, Mohammadpour Fard R, Jalali Far MA, Sakhavarz T. A closer look at Galectin-3: its association with gestational diabetes mellitus revealed by systematic review and meta-analysis. J Diabetes Metab Disord. 2024;23:1621–33.
64. Aghaei M, Khademi R, Bahreiny SS, Saki N. The need to establish and recognize the field of clinical laboratory science (CLS) as an essential field in advancing clinical goals. Health Sci Rep. 2024;7:e70008.
65. Torfi E, Bahreiny SS, Saki N, Khademi R, Sarbazjoda E, Nezhad IA, et al. Evaluation of Pro-BNP biomarker in heart failure patients and its relationship with complete blood count parameters: a case-control study. Health Sci Rep. 2024;7:e70083.
66. Eftekhar Z, Aghaei M, Saki N. DNA damage repair in megakaryopoiesis: molecular and clinical aspects. Expert Rev Hematol. 2024;17:705–12.
67. Saki N, Haybar H, Aghaei M. Subject: Motivation can be suppressed, but scientific ability cannot and should not be ignored. J Transl Med. 2023;21:520.
68. Liu Z, Qiu H, Letchmunan S, Deveci M, Abualigah L. Multi-view evidential *c*-means clustering with view-weight and feature-weight learning. Fuzzy Sets Syst. 2025;498:109135.
69. Maiti B, Biswas S, Ezugwu AE-S, Bera UK, Alzahrani AI, Alblehai F, et al. Enhanced crayfish optimization algorithm with differential evolution's mutation and crossover strategies for global optimization and engineering applications. Artif Intell Rev. 2025. https://doi.org/10.1093/pcmedi/pbad017.
70. Biswas S, Singh G, Maiti B, Ezugwu AE-S, Saleem K, Smerat A, et al. Integrating Differential Evolution into Gazelle Optimization for advanced global optimization and engineering applications. Comp Meth Appl Mech Eng. 2025. https://doi.org/10.1016/j.cma.2024.117588.
71. Rateb R, Hadi AA, Tamanampudi VM, Abualigah L, Ezugwu AE, Alzahrani AI, et al. An optimal workflow scheduling in IoT-fog-cloud system for minimizing time and energy. Sci Rep. 2025;15:3607.

Publisher's Note