Journal of Inflammation Research

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

# Screening and Validation of Potential Biomarkers of Immune Cells in Childhood Asthma Patients via Mendelian Randomization and Machine Learning

Yang Zhang[1–4], Yang Hai[1–3], Bangguo Song[4], Jing Xu[1–3], Liangjia Cao[5], Rukeye Yasen[5], Wenjuan Xu[5], Jiaxuan Zhang[5], Jihong Hu[6]

[1]Scientific Research and Experimental Center, Gansu University of Chinese Medicine, Lanzhou, People's Republic of China; [2]Key Laboratory of Dunhuang Medicine, Ministry of Education, Gansu University of Chinese Medicine, Lanzhou, People's Republic of China; [3]Research Center of Traditional Chinese Medicine, Gansu Province, Gansu University of Chinese Medicine, Lanzhou, People's Republic of China; [4]Clinical College of Chinese Medicine, Gansu University of Chinese Medicine, Lanzhou, People's Republic of China; [5]College of Public Health, Gansu University of Chinese Medicine, Lanzhou, People's Republic of China; [6]Laboratory and Simulation Training Center, Gansu University of Chinese Medicine, Lanzhou, People's Republic of China

Correspondence: Jihong Hu, Laboratory and Simulation Training Center, Gansu University of Chinese Medicine, Chengguan District, Lanzhou, Gansu, 73000, People's Republic of China, Email hujihonghappy@163.com



**Purpose:** Asthma is one of the most common chronic respiratory diseases affecting children, and there is currently no clear remedy. Immune cells play a key role in childhood asthma. Therefore, a deeper investigation of the correlation between immune cells and childhood asthma could lead to a better understanding of asthma's origin, the identification of potential treatment targets, and the development of personalized treatment strategies.

**Patients and Methods:** We used a two-sample Mendelian randomization (MR) analysis to investigate the possible causal relationship between childhood asthma and a total of 731 immune cells, including B cell (190), Maturation stages of T cell (79), Monocyte (43), Myeloid cell (64), TBNK (124), Treg (167), and CDC (64). LASSO logistic regression and SVM algorithms were used to identify key genes associated with childhood asthma. Specific signaling pathways associated with these key genes were further explored through gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA). Subsequently, the four key genes FCGR3A, TCTN3, ALOX5, and IL4R were verified in an established asthma mouse model using quantitative real-time PCR and Western blotting.

**Results:** MR analysis showed that 60 immune cells were associated with childhood asthma, of which 32 were associated with high risk and 28 were associated with low risk. LASSO logistic regression and SVM algorithm identified six key genes that affect childhood asthma as ATF4, FCGR3A, GAS5, MGAT3, TAB1, and TCTN3. In addition, four genes, FCGR3A, TCTN3, ALOX5, and IL4R, were verified through animal experiments.

**Conclusion:** Our findings confirmed that immune cells contribute to childhood asthma, highlighting the importance of key genes in the role of the immune microenvironment in this disease. These insights provide a new path for the exploration of the biological underpinnings of childhood asthma and the development of early intervention therapies.

**Keywords:** childhood asthma, immune cells, Mendelian randomization, biomarker, key genes

## Introduction

Asthma ranks among the most prevalent chronic respiratory diseases affecting children.[1] It is induced by chronic inflammation and airway cell hyperreaction and is clinically characterized by cough, shortness of breath, and chest tightness.[2] The global incidence of asthma is increasing, especially in low- and middle-income nations.[3] Additionally, asthma during infancy frequently remains undiagnosed and, consequently, untreated. According to the Global Burden of Disease (GBD) estimates, in 2021 there were approximately 95,721,185.84 cases of asthma among children aged 0–14 years globally, with a death rate of 0.41%.[4] Asthma significantly diminishes the quality of life and imposes a substantial

Received: 5 November 2024
Accepted: 7 February 2025
Published: 21 February 2025

economic burden, increasing the prevalence of asthma as a pressing global public health issue. To date, a definitive cure for *asthma* remains elusive.[3] Current therapeutic approaches, which are primarily focused on inhaled corticosteroids and bronchodilators, aim to alleviate symptoms without altering the disease trajectory.[5] Hence, elucidating the pathophysiological mechanisms underlying asthma is crucial for the development of new therapeutic strategies.

Recent investigations have elucidated the intricate relationship between asthma and factors such as immune regulation, genetic predisposition, viral infections, and allergen exposure, which collectively contribute to immune dysregulation, culminating in asthma.[6,7] Immune cells, notably dendritic cells (DCs) and regulatory T lymphocytes (Tregs) play pivotal roles in the onset of childhood asthma. For example, asthmatic children exhibit a marked reduction in the T regulatory cell population in their sputum, highlighting the compromised immune tolerance function of these patients.[8] Furthermore, individuals with asthma present elevated levels of peripheral blood dendritic cell subsets, underscoring the critical functions of DCs in antigen presentation, T-cell activation, and the genesis of asthma and allergic responses.[9] The Th2 cytokine pathway (mainly IL-4, IL-5, and IL-13) plays a central role in the pathogenesis of asthma. It severely affects the clinical manifestations and progress of asthma by driving airway inflammation, promoting eosinophil infiltration, and promoting airway remodeling.[10,11] However, the underlying mechanisms of asthma are unknown, casting doubts on the robustness of the inferred connection between immune cells and childhood asthma.

MR is a genetic epidemiological approach to investigating the causal effects of exposure factors on disease outcomes, free from the biases of confounding factors or reverse causality.[12,13] This study used dual-sample MR to investigate the possible causal relationship between childhood asthma and a total of 731 immune cells. Moreover, by leveraging machine learning techniques, we aspire to unravel potential molecular mechanisms and predict immune-centric therapeutic targets. Research has shown that using feature importance statistics can accurately predict asthma attacks through machine learning.[14] The application of machine learning technology in asthma research is steadily developing, especially in prediction, diagnosis, and treatment management.[15,16] Machine learning has become crucial for identifying biomarkers, improving the accuracy of clinical diagnosis, and determining optimal therapies in the biomedical field.[17–19]

# Materials and Methods

## Study Design

This study utilized MR analysis to assess the associations between a total of 731 immune cells that were categorized into 7 groups and childhood asthma through two-sample analysis. MR uses instrumental variables (IVs) to infer causality, adhering to three assumptions, including genetic variants that are associated with the exposures; not related to any confounders of the exposure-outcome relationship; and influencing the results solely through the exposure.[20]

The key genes that affect the onset of childhood asthma were identified via LASSO logistic regression and SVM algorithms. GSVA in combination with GSEA was subsequently used to further investigate genes involved in key signaling pathways to explore the molecular mechanisms through which key genes influence the progression of childhood asthma.

## Data Sources for Childhood Asthma Genome-Wide Association Studies (GWASs)

Childhood asthma outcome data were sourced primarily from individuals of European descent, utilizing summary data from the EBI database (EBI-A-GCST90018895), which includes publications, top associations, and comprehensive abstract statistics from the GWAS Catalog. Overall, 27,712 patients with childhood asthma and 411,131 control individuals were included.

## Immunity-Wide GWAS Data Sources

This study utilized 731 immune cell profiles from the GWAS Catalog database, maintained by EBI, covering different types of immune cells, including B cells (190), T cells (79), monocytes (43), myeloid cells (64), TBNKs (124), Tregs (167), and CDCs (64).

## Gene Expression Data Sources

Gene expression data were sourced from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), specifically the Series Matrix File of GSE65204, which includes 69 patients divided into control (33) and disease (36) groups.

## Mendelian Randomization Analysis

R program version 4.3.1 (http://www.Rproject.org) was used for data analysis. This study first used summary statistical data from hundreds of genome-wide association studies (GWAS) in the European Institute of Bioinformatics (EBI) database to extract causal relationships associated with 731 immune cells in the GWAS Catalog by screening outcome IDs. SNPs associated with a significance threshold for each immune cell across loci ($P<5\times10^{-8}$) were selected as potential instrumental variables (IVs). This strict P-value threshold is designed to ensure a strong correlation between the selected SNPs and exposure factors (immune cell characteristics), thereby reducing false positive results. Subsequently, the linkage disequilibrium (LD) between these SNPs was calculated, and the clustering method (clustering window size= 10,000 kb, $R^2<0.001$) was used to eliminate highly related SNPs to ensure the independence of each IVs and reduce bias caused by LD. To assess the causal impact of all cis and some cross-regional gene expression in whole blood on childhood asthma, this study used four Mendelian Randomization (MR) analysis methods: Inverse Variation Weighted (IVW),[21] MR Egger regression,[22] the weighted median method,[23] and the weighted mode method. The IVW method estimates the overall causal effect by weighted averaging the Wald ratio for each SNP, which is applicable if all IVs meet the null instrumental variable hypothesis; MR Egger allows for gene pleiotropy and detects and adjusts for potential directional pleiotropy by estimating the intercept term based on the assumption that instrument strength is independent of direct effects (InSIDE); The Weighted median method estimates causal effects by calculating a weighted median, allowing up to 50% of IVs to be invalid instrumental variables, ensuring that it can still provide robust estimates even when some IVs are invalid or pleiotropic; the Weighted mode method enhances the ability to detect causal effects and reduces bias and type I error rates by identifying and estimating the most common causal effect patterns. If there is only one SNP in a causal relationship as IVs, only the Wald ratio estimation method is used for causality assessment. To ensure the reliability of the screened causal relationships, this study performed a heterogeneity test (Cochran's IVWQ test) and a genetic diversity test to assess the heterogeneity among different IVs and the diversity of IVs selection and enhance the robustness of causal inference. There was no heterogeneity at $P>0.05$. In addition, sensitivity analysis was performed based on the results of multiple MR methods to assess the consistency and robustness of the causal relationship and eliminate potential confounding factors and pleiotropic effects. Through the above method steps, an overall causal effect estimate of the impact of all cis and some cross-regional gene expressions in whole blood on childhood asthma was finally obtained.

## LASSO Regression and SVM Algorithms for the Feature Selection Process

In this study, we used Lasso logistic regression and the support vector machine recursive feature elimination (SVM-RFE) algorithm to select features for SNP genes, improving the accuracy of disease diagnosis. Lasso logistic regression is implemented through the "glmnet" software package, and the regularization parameter (λ) is selected through cross-validation to optimize model performance and control overfitting. In addition, SVM-RFE, as a machine learning method based on support vector machines, identifies the best variables by iteratively deleting the least important features determined by SVM weights. It uses the "e1071" software package to build a support vector machine model, adopts a radial basis function kernel, and optimizes cost and gamma parameters through grid search. The Lasso and SVM-RFE methods were selected based on their complementary advantages in processing high-dimensional genetic data and enhancing the robustness of feature selection, thereby improving the transparency and reliability of identified SNP genes in disease diagnosis.

## GSVA and GSEA

Gene Set Variation Analysis (GSVA) is a non-parametric and unsupervised method to evaluate the enrichment of gene sets in the transcriptome. By comprehensively scoring predefined gene sets, expression changes at the gene level are transformed into changes at the pathway level, thereby judging differences in biological functions between samples. This study downloaded multiple functionally related gene sets from the Molecular Signatures Database (MSigDB) and used the GSVA algorithm to comprehensively score each gene set. A Z-score standardized method of gene expression is used during the scoring process to eliminate technical deviations between samples. In addition, the GSVA used an unsupervised approach with the significance level set to the adjusted *p*-value (FDR<0.05). The selection of these parameters ensures the robustness and reliability of the analysis results. The GSVA method was chosen based on its advantages in processing high-dimensional transcriptome data, eliminating the need to pre-specify sample categories, and being able to capture subtle changes in pathway activity, thereby improving the transparency and accuracy of the assessment of potential biological function changes in different samples.

GSEA was used to further analyze the differences in signaling pathways between high and low-expression groups. This study used the GSEA method, and the version 7.0 annotated gene set downloaded from the MsigDB database was used as the annotated gene set of subtype pathways. The significance threshold was set to an adjusted *p*-value (FDR) of less than 0.05 to control the false positive rate caused by multiple tests. During the analysis, the normalized enrichment score (NES) was used to rank significantly enriched gene sets, and signaling pathways that differed significantly between high and low-expression groups were identified based on the consistency score. The reason why the GSEA method was chosen is that it does not need to define a list of differentially expressed genes in advance, can comprehensively capture potential biological pathway changes in gene expression data, and has high sensitivity and specificity when processing complex transcriptome data, thereby effectively combining disease typing and biological significance to improve the transparency and reliability of analysis results.

## Establishment of a Mouse Asthma Model

The Laboratory Animal Ethics Committee of Gansu University of Traditional Chinese Medicine approved all procedures involving animals in this study. Twelve pathogen-free 3-week-old male BALB/c mice were provided by Changzhou Cavins Laboratory Animal Co., Ltd. (Changzhou, China) and randomly allocated into ova and vehicle groups. Ovalbumin (OVA) injection sensitization and nebulization stimulation were used to establish a young mouse model of asthma. The Vehicle group was given an identical amount of normal saline injection plus nebulization. At 28 days after model induction, all the mice were euthanized, and lung tissue samples were collected.

## RT–PCR

Total RNA was isolated from mouse lung tissue samples, and RNA concentrations were assessed using RNA purification kits. Next, the RNA samples were reverse transcribed into cDNA using a transcription kit. A real-time fluorescence quantification (PCR) protocol was subsequently applied to evaluate the results via the $2^{-\Delta\Delta Ct}$ method. Table 1 lists the sequences of primers used in this study (Table 1 See the end of the manuscript).

**Table 1** Sequences of the Primers Used in This Study

| Gene | Forward Primer | Reverse Primer |
|---|---|---|
| GAPDH | CCCTTAAGAGGGATGCTGCC | TACGGCCAAATCCGTTCACA |
| ALOX5 | CAAAATCTGGGTGCGTTCCA | TTGAAAAGGGGATGCACAGC |
| FCGR3A | AGGGAAAGCCTGAGGTCTTC | CTAAGCACGGAAAGGATCGC |
| IL4R | TGGCTGCTGACCTGGAATAA | TCAGCCTGGGTTCCTTGTAG |
| TCTN3 | CTGGGACGTATCTGGAGACC | GCGGTCTCAGAAAGCATCAG |

## Western Blotting

A sample of approximately 60 mg of mouse lung tissue was taken, thawed at room temperature, sterilized by high pressure with scissors, and cut into pieces. One milliliter of RIPA (Solarbio, R0010, China) lysis solution (containing PMSF, the final concentration of 1 mm) was added for lysis, the mixture was centrifuged, and the supernatant was collected. Protein quantification was carried out via the BCA method according to the manufacturer's instructions. Next, separation and concentration gels were prepared and assembled for SDS–PAGE. After electrophoresis, the proteins were transferred from the gel to a PVDF membrane. After blocking with 5% skim milk, primary antibodies against FCGR3A (dilution 1:1000, Bioss, China), TCTN3 (dilution 1:1000, Bioss, China), ALOX5 (dilution 1:1000, Bioss, China), IL4R (dilution 1:1000, Bioss, China), and GAPDH (dilution 1:1000, Bioss, China) were added dropwise and incubated overnight at 4°C. After the membrane was washed with TBST, it was incubated with a secondary antibody (1:4000) at 37°C for 90 min. Ultimately, it was developed with an enhanced chemiluminescence (ECL) luminescent solution, and the gray values of the protein bands were analyzed via ImageJ software to calculate the relative protein expression levels.

## Statistics

Statistical analysis was carried out via SPSS 21.0 software. Comparisons between two groups were performed with a *t* test, while multiple group comparisons were performed using analysis of variance (ANOVA), followed by an LSD test for pairwise group comparisons. When the variances were uneven, Dunnett's method was used for analysis. The data were drawn via Origin Pro 8.5 drawing software. $P<0.05$ indicated statistical significance.

# Results

## Examination of the Causal Link Between Childhood Asthma and Immune Cells

This investigation identified a correlation between features of 731 immune cells and childhood asthma through Mendelian randomization analysis, utilizing aggregated statistical data from 438,843 participants (comprising 27,712 asthma patients and 411,131 control individuals), as indicated by the outcome ID ebi-a-GCST90018895. Specifically, 60 immune cells were significantly associated with childhood asthma. For example, an elevated risk of developing asthma was linked to specific immune cells, including HLA DR on CD33− HLA DR+ (OR 1.114; 95% CI 1.057–1.174; $P$ <0.001) and CD4+ T regulatory cells (%CD4+ T cells) (OR 1.100; 95% CI 1.009–1.201; P = 0.031). Conversely, a decreased risk was noted for others, such as IgD+CD38−B-cell% lymphocytes[6] (OR 0.788; 95% CI 0.678–0.915; $P$ = 0.002) and CD28+CD28+CD45RA+CD8+ T cells (OR 0.858; 95% CI 0.744–0.990; $P$ = 0.036), illuminating the nuanced role of immune cell features in asthma (Figure 1A and B, Supplementary Files 1 and 2).

## Rigorous Evaluation of Causal Relationship Heterogeneity and Genetic Diversity

The integrity and reliability of these causal associations were further assessed through comprehensive heterogeneity and genetic diversity analyses, which demonstrated a lack of significant heterogeneity ($P > 0.05$), thereby strengthening the confidence in these findings (Supplementary Files 3 and 4). Sensitivity analyses conducted via the leave-one-out approach revealed a negligible impact on the overall results from the removal of any single SNP, underscoring the robustness of the identified causal relationships (Supplementary File 5). Finally, SNP-related genes in the GEO expression profiles (Supplementary File 6) were extracted via the get_variants function of the Gwasrapidd package, further enriching the study's foundational data.

## Identification and Screening of Asthma-Related Key Genes

To identify critical genes influencing childhood asthma, this study employed a combination of LASSO regression and SVM feature selection algorithms to meticulously screen 60 genes that exhibited causal relationships with positive outcomes in immune cells, leading to the identification of 20 feature genes. The utilization of the SVM-RFE algorithm allowed for the further refinement of six highly accurate feature genes, with a subsequent intersection revealing six pivotal genes: ATF4, FCGR3A, GAS5, MGAT3, TAB1, and TCTN3. These genes were earmarked for in-depth analysis in subsequent phases of the study (Figure 2A–D).

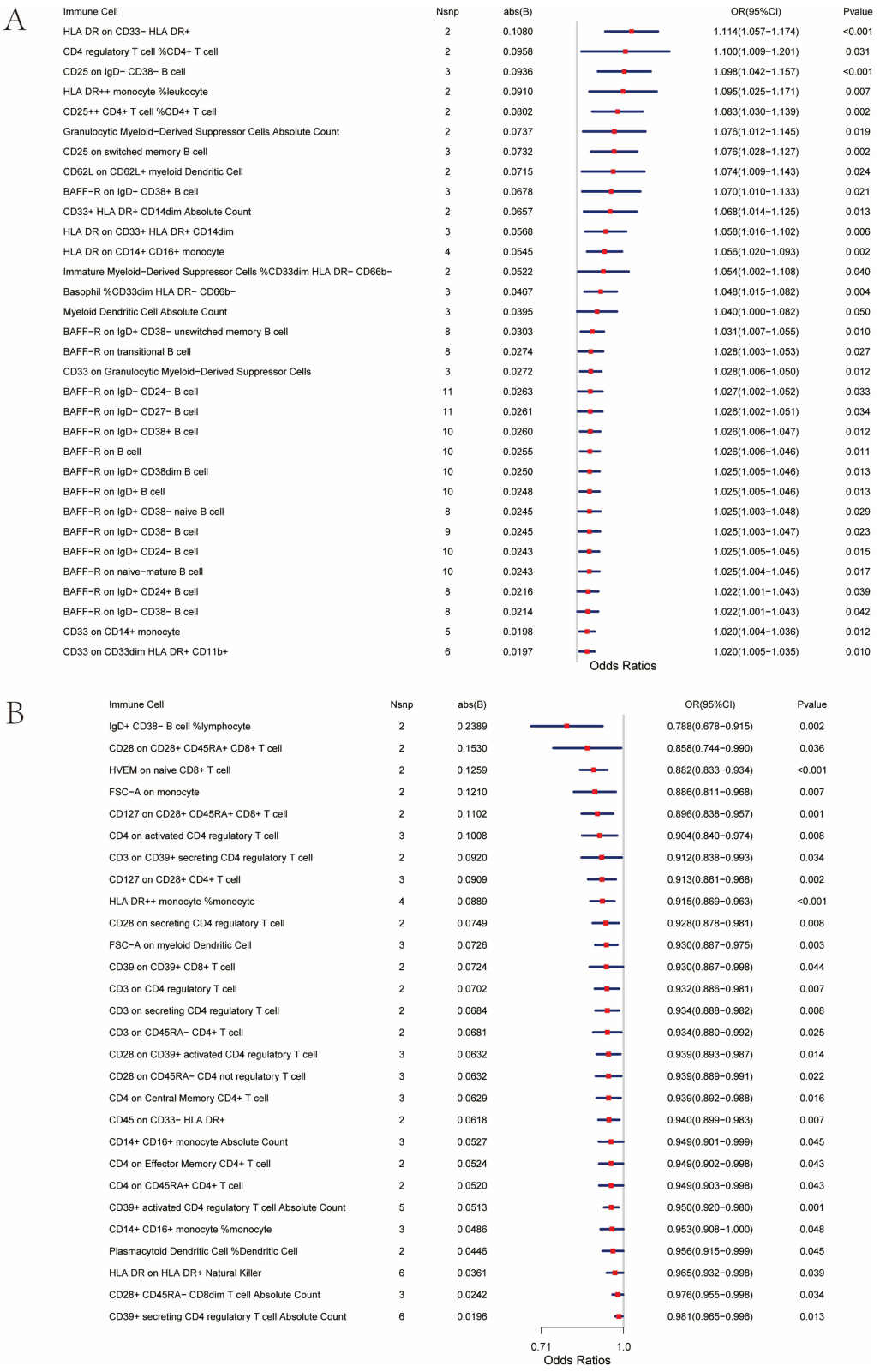


**Figure 1** (**A** and **B**) Forest plots of Mendelian randomization (MR) estimates between immune cells and childhood asthma according to *P* value correction. MR revealed that 60 immune cells were associated with childhood asthma; OR, odds ratio. (**A**) Displays immune cells associated with a higher risk of childhood asthma. (**B**) Displays immune cells associated with a lower risk of childhood asthma. Results with an Inverse-Variance Weighted (IVW) *P*-value<0.05 are considered statistically significant.

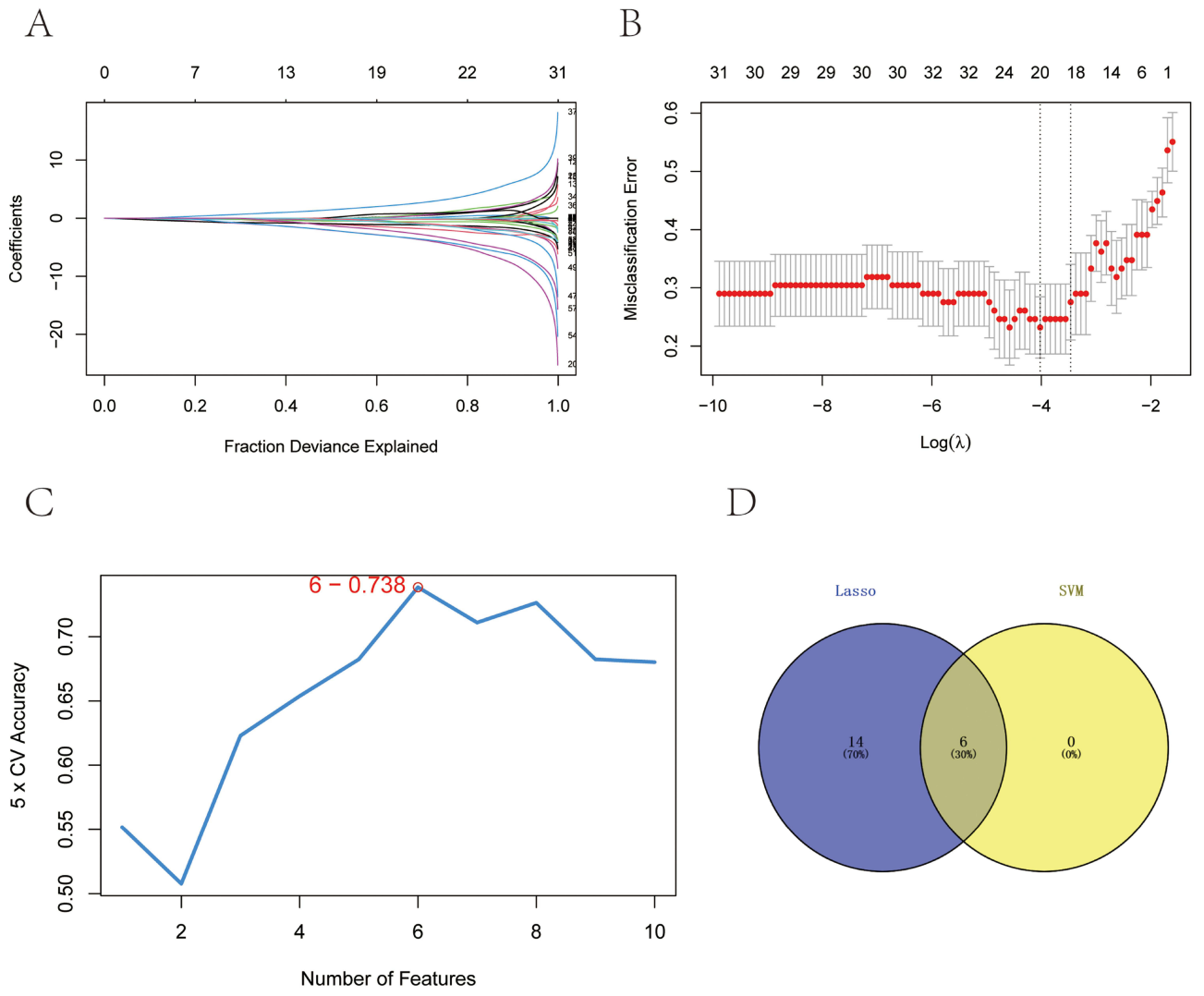


**Figure 2** (**A**–**D**) Screening of Diagnostic Markers and Identification of Key Genes. This study combined LASSO regression and an SVM feature selection algorithm to screen 60 genes with causal relationships to positive immune cell outcomes, identifying a total of 20 feature genes based on LASSO coefficients. (**A**) Distribution of LASSO coefficients for key genes. (**B**) Tenfold cross-validation of tuning parameter selection in LASSO models. (**C**) Accuracy curve based on 5-fold cross-validation via the SVM-RFE algorithm. (**D**) Six characteristic genes in childhood asthma patients were evaluated via the SVM-RFE algorithm and intersected with those screened via the LASSO regression algorithm, and a total of 6 intersecting genes were screened, diagram of six crossover genes as feature genes obtained via LASSO regression and the SVM-REF algorithm for feature selection.

## Exploration of the Associations Between Immune Factors and Significant Genes

Subsequent analyses delved into the relationships between these significant genes and various immune-related factors, including chemokines and cytokines, utilizing data from the TISIDB database. The results of this analysis reinforced the pivotal roles these key genes play within the immune microenvironment, revealing strong correlations with immune cell infiltration levels (Figure 3A–E).

## Enrichment of Key Genes in Specific Signaling Pathways

In this study, specific signaling pathways enriched with these key genes were profiled via GSVA and GSEA to identify possible molecular mechanisms through which these genes influence asthma progression in children. The analysis revealed that high expression of these genes was linked to important biological pathways.

GSVA revealed that high expression of ATF4 was connected to the IL2/STAT5 and PI3K/AKT/mTOR signaling pathways. High expression of FCGR3A was connected to the IL6/JAK/STAT3, PI3K/AKT/mTOR, and oxidative stress pathways. High expression of GAS5 was connected to fatty acid metabolism, myc target (v1), and allograft rejection.


https://doi.org/10.2147/JIR.S498017

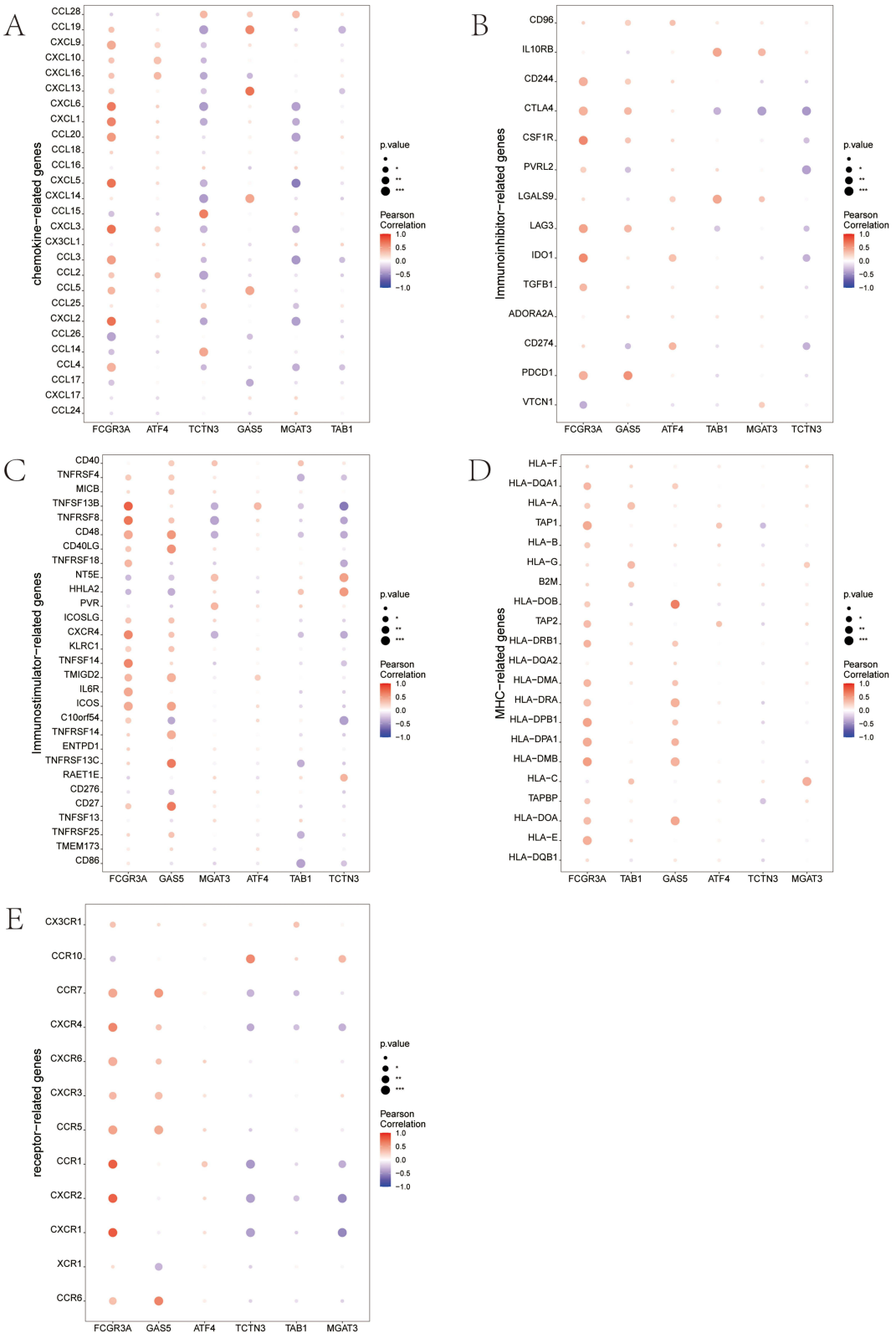


**Figure 3** (**A**–**E**) Correlations between key genes and immune cell infiltration. The relationships between these six key genes and different immune factors (including immunosuppressive factors, immunostimulatory factors, chemokines, MHC and receptors) were obtained from the TISIDB database. The relationships between immune factors and core genes related to asthma in children were determined. * represented *P*<0.05, ** represented *P*<0.01 and *** represented *P*< 0.001. The bigger the circle, the closer the *P*-value was to zero; The redder the color, the stronger the positive correlation; The deeper of the blue color, the stronger the negative correlation.

High MGAT3 expression is associated with TGF-beta, KRAS (DN), etc. High TAB1 expression is associated with mTOR signaling and other signaling pathways. High expression of TCTN3 was associated with bile acid metabolism, heme metabolism, the cell cycle, etc. (Figure 4A–F).

GSEA revealed that ATF4 was enriched in the chemokine signaling pathway, cytokine IL−17 signaling pathway, TNF alpha pathway, and other pathways. FCGR3A was enriched in pathways related to asthma, the MAPK pathway, the NF −kappa B pathway, and other pathways. GAS5 was enriched in asthma-related pathways, Th1 and Th2 cell differentiation pathways, and Th17 cell differentiation pathways (Figure 5A–F). MGAT3 was enriched in the B-cell receptor signaling pathway, the ECM−receptor interaction pathway, the JAK−STAT signaling pathway, and other pathways. TAB1 was enriched in the chemokine signaling pathway, propanoate metabolism pathway, retinol metabolism pathway, and other pathways. TCTN3 was enriched in lipoic acid metabolism, the PI3K−Akt signaling pathway, and other pathways (Figure 6A–F). These findings indicate that key genes might affect childhood asthma progression through these pathways.

## Disease Regulation Genes in Children with Asthma

This study used the GeneCards database (https://www.genecards.org/) to analyze childhood asthma disease gene regulation. The leading 20 genes according to the relevance rating were analyzed, and the expression levels of the 6 key genes were markedly associated with the disease regulatory genes. Among these key genes, FCGR3A and ALOX5 were significantly positively correlated ($r$=0.626), and TCTN3 was significantly negatively associated with IL4R ($r$= −0.463) (Figure 7).

## Validation of Gene and Protein Expression in Asthma Mouse Lung Tissue

In ovalbumin-sensitized/activated mouse lung tissue, FCGR3A, ALOX5, and IL4R mRNA levels were significantly increased, whereas TCTN3 mRNA levels were considerably decreased (Figure 8A–D).

At the protein level, FCGR3A, ALOX5, and IL4R are upregulated, whereas TCTN3 is downregulated. These findings suggest that these genes may play a practical role in the pathogenesis of asthma and that their expression is correlated with the disease state (Figure 9A–E).

# Discussion

Given that there is no cure for asthma and that treatment is limited to symptomatic improvement, the development of treatment methods for asthma is essential. Bioinformatics approaches to investigate disease-causing genes that are druggable have drawn the attention of scientists worldwide. An increasing number of investigations have suggested that immune mechanisms may play crucial roles in childhood asthma.

In recent years, MR has gained global recognition as a valuable resource for assessing potential causal connections between various diseases and risk factors, leveraging the wealth of publicly available genetic datasets. This study embarked on an exploratory journey to uncover causal associations between a group of 731 immune cell traits and childhood asthma, marking the first MR analysis to delve into the causal relationships between various immunophenotypes and childhood asthma. Among these immune cell types, 60 were recognized as significantly associated with childhood asthma. Notably, the presence of HLA DR on CD33- HLA DR+ and the percentage of CD4 regulatory T cells among CD4+ T cells were linked to an increased risk of developing childhood asthma. This investigation builds on the foundation laid by prior observational studies that have scrutinized the interplay between immune cells and childhood asthma. Allergic asthma, the most prevalent form of asthma in children, is characterized by a type 2 shift in the immune response, alongside diminished expression of innate immune genes.[24] The orchestration of the type 2 immune response involves both the innate and adaptive immune systems, engaging a myriad of immune cells, including T-helper (Th) 2 cells, group 2 innate lymphoid cells (ILCs), B cells, natural killer (NK) cells, NK T cells, basophils, eosinophils, and mast cells, alongside their principal cytokines.[25] The reduced expression of HLA-DR in monocytes during chronic inflammation suggests an anti-inflammatory function for these molecules.[26,27] The MHC class II cell surface receptor HLA-DR plays a pivotal role in lymphocyte activation and has been implicated in asthma pathogenesis.[28] Conversely, CD4+ T-cell subsets, distinguished by their cytokine secretion profiles, encompass various groups, such as Th1, Th2,

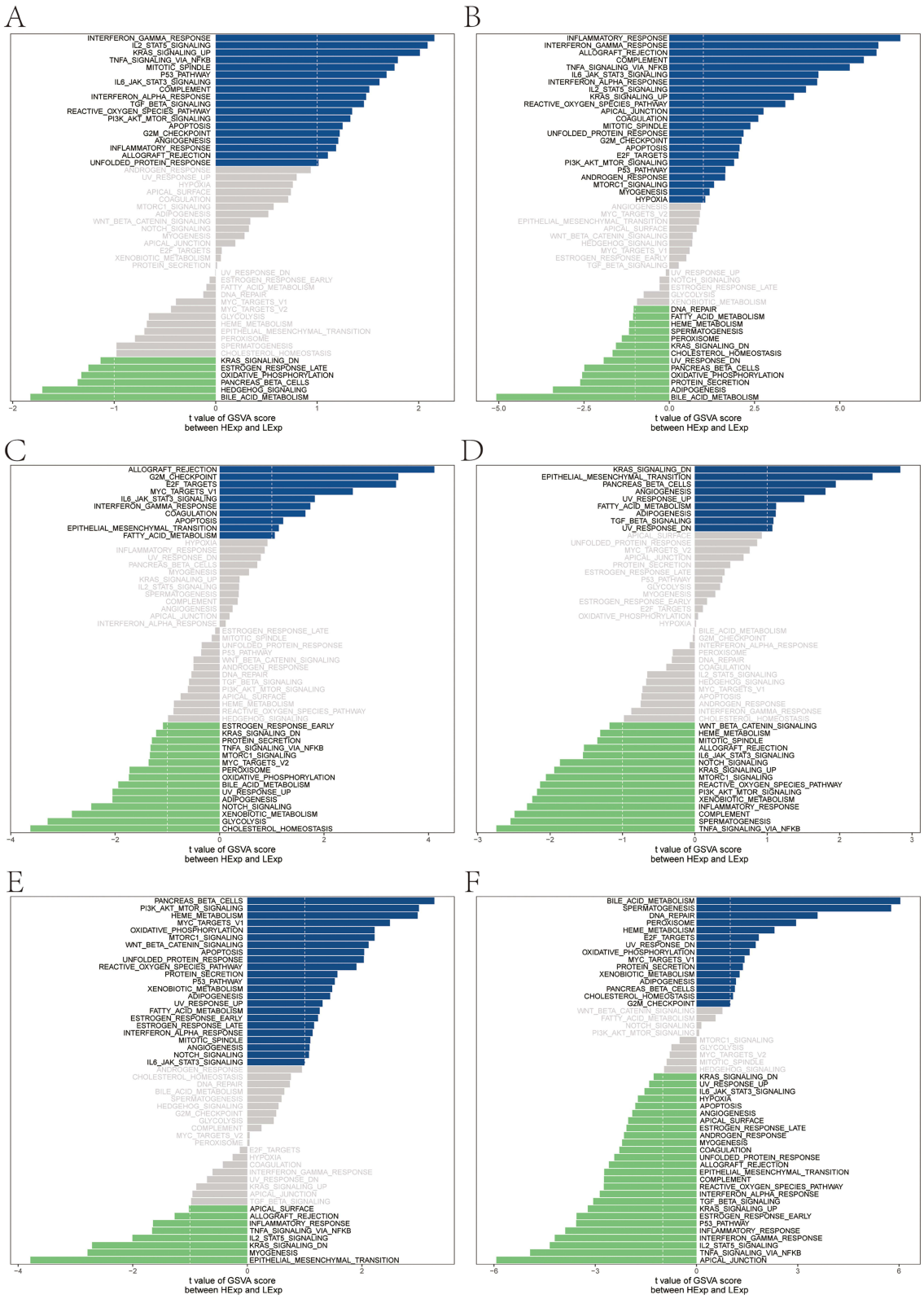


**Figure 4** (**A**–**F**) GSVA key gene enrichment analysis map. By studying the specific signaling pathways associated with the enrichment of the six key genes, we explored the potential molecular mechanism by which the key genes affect the progression of asthma in children. Signaling pathways involved in high levels of gene expression are shown in blue, green indicates signaling pathways involved in low levels of gene expression, and the background gene set is a hallmark.

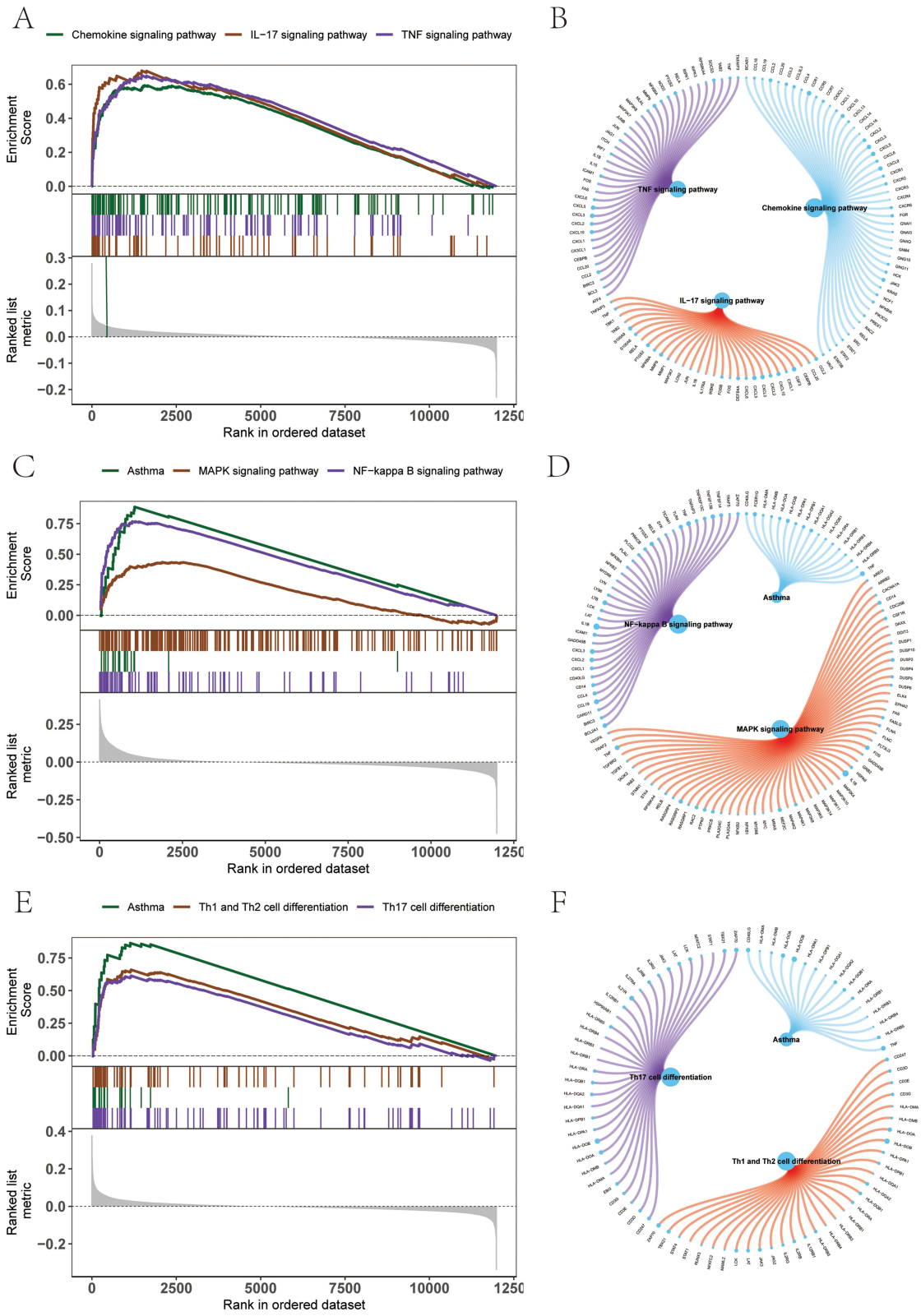


**Figure 5** (**A**–**F**) Molecular regulatory mechanism of core gene-related pathways. The enriched signaling pathways of key genes were analyzed by GSEA, which suggested that key genes may affect the progression of childhood asthma through these pathways. All pathways NOM *p*-val The condition<0.05 meets the statistical significance.


https://doi.org/10.2147/JIR.S498017

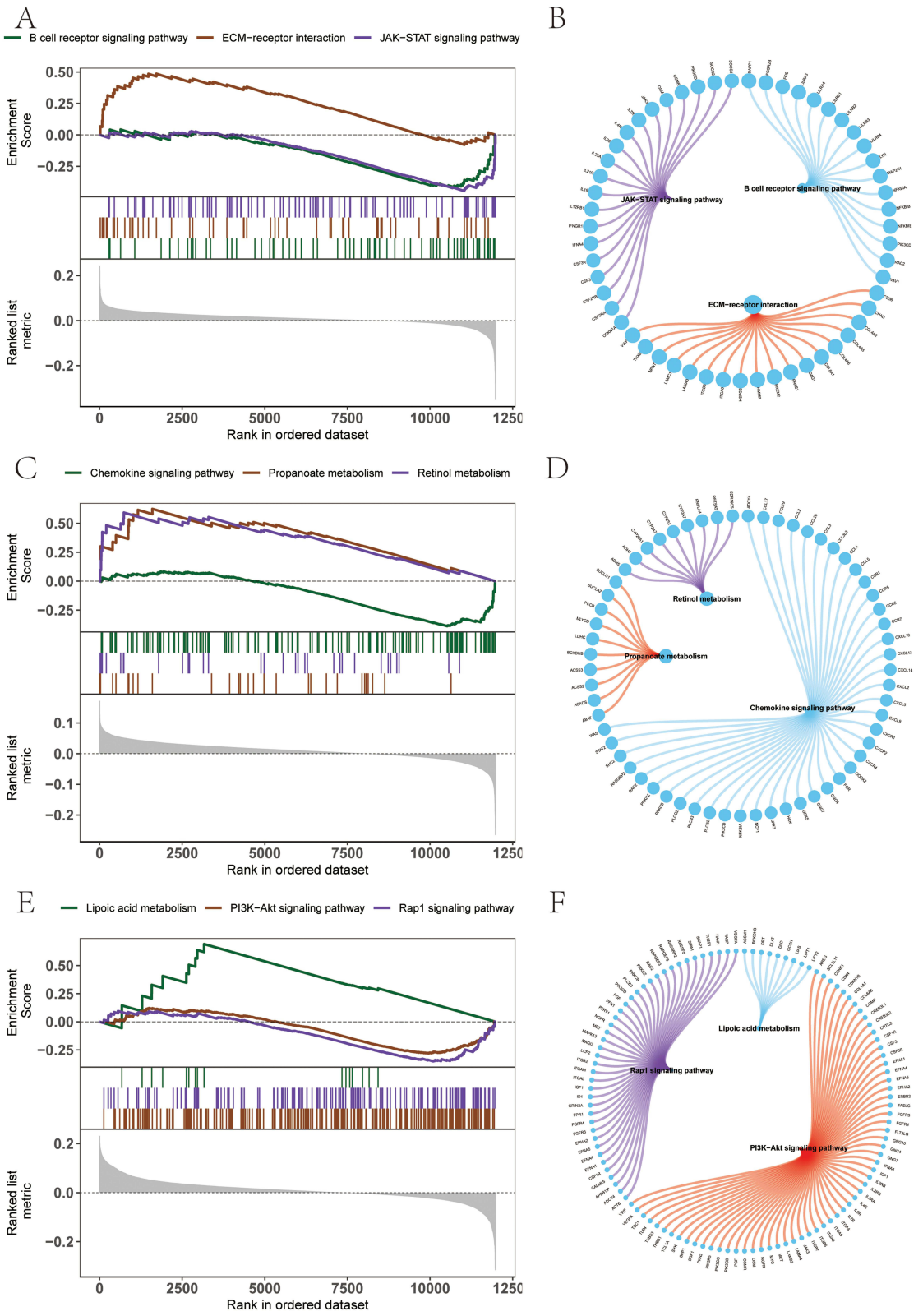


**Figure 6** (**A**–**F**) Molecular regulatory mechanism of core gene-related pathways. The enriched signaling pathways of key genes were analyzed by GSEA, which suggested that key genes may affect the progression of childhood asthma through these pathways. All pathways NOM *p*-val The condition <0.05 meets statistical significance.

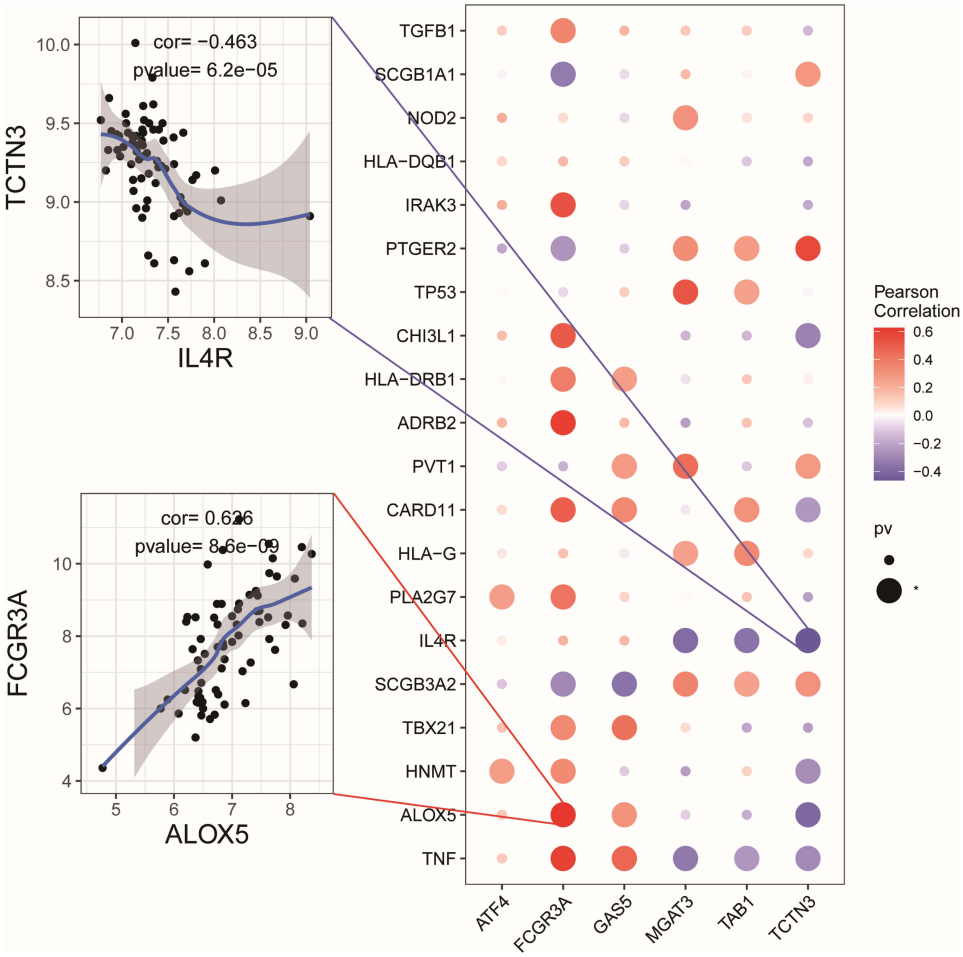


**Figure 7** Disease gene expression levels. There was a significant correlation between the 6 key genes' expression levels and disease-regulatory genes' expression levels. * represented *P*<0.05, Richer shades of red indicate stronger positive correlations, while deeper shades of blue signify stronger negative correlations.

Th17, follicular helper T (Tfh), and regulatory T cells, each of which plays distinct roles in asthma onset and progression.[29,30] Increased levels of Th2-associated cytokines (IL-4, IL-5, and IL-13) have been documented in the bronchoalveolar lavage (BAL) fluid and blood of patients with allergic asthma.[31,32] In animal models, the attenuation of asthma hallmark features following the depletion of OVA-specific Th2 cells further underscores the critical role of these cells in the disease.[33] Furthermore, interventions such as Ma Huang Tang have been shown to ameliorate asthma by modulating Th1/Th2 cytokines and inhibiting Th17 cells in sensitized mice.[34] Tregs, known for their role in suppressing inflammatory responses and maintaining immune tolerance, exhibit altered abundance in asthma, with a shift away from Tregs toward an increased presence of Th17 cells. The balance of Tregs in childhood asthma has been linked to clinical measures such as FEV1 and the levels of specific microRNAs (miR-146a-5p and miR-210-3p).[35] Additionally, B regulatory cells (Bregs), a subset of IL-10-secreting B cells, have been identified as key players in modulating allergic asthma in both murine models and human studies, controlling airway hyperresponsiveness and remodeling.[36,37] The above studies demonstrated a causal relationship between immune cells and childhood asthma.

In addition, we utilize both LASSO regression and SVM feature selection algorithms to pinpoint six principal genes that significantly influence childhood asthma: ATF4, FCGR3A, GAS5, MGAT3, TAB1, and TCTN3. Our comprehensive analysis revealed a strong association between genes and immune-related factors, highlighting their pivotal roles within the immune microenvironment. We delved deeper into the signaling pathways characterized by these enriched genes via GSVA and GSEA methodologies. Our findings indicate their involvement in critical pathways, such as the IL-2/STAT5 pathway, the PI3K/Akt/mTOR pathway, and the TNF pathway. Analysis of the expression levels of disease-related genes


https://doi.org/10.2147/JIR.S498017

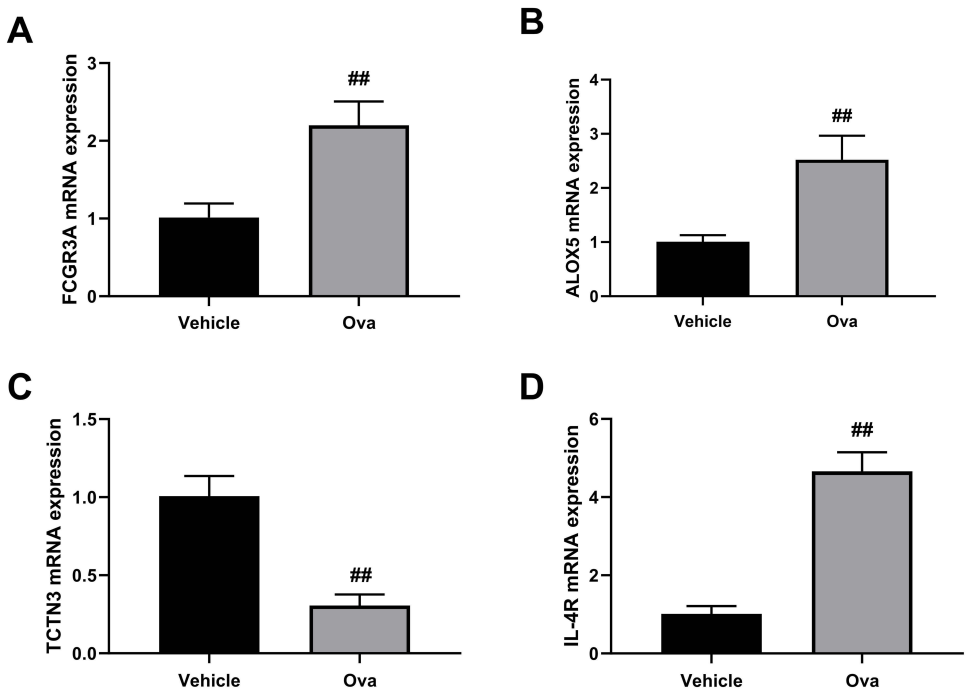


**Figure 8** (**A**–**D**) Gene expression of FCGR3A, TCTN3, ALOX5, and IL4R in ovalbumin-sensitized/activated mouse lung tissue ($^{\#\#}p<0.01$).

in childhood asthma patients revealed that the expression levels of six key genes were significantly correlated with the expression levels of disease-related genes. Notably, FCGR3A and ALOX5 were significantly positively correlated ($r$=0.626), whereas TCTN3 and IL4R were significantly negatively correlated ($r$=−0.463). These correlations revealed the molecular mechanisms through which these key genes may influence the progression of childhood asthma. We validated the gene and protein expression levels of FCGR3A, TCTN3, ALOX5, and IL4R in a juvenile asthma mouse model, and the results were consistent with our bioinformatics analysis, which revealed significant predictive effects.

The Fc fragment of IgG receptor IIIa (FCGR3A, also known as CD16a), a cell membrane Fc receptor from the Fcγ receptor family that is primarily found on the surfaces of immune cells such as NK cells and macrophages, is essential for the immune response and antibody-mediated immune effects.[38] In our recent research, we discovered that FCGR3A has a crucial effect on the progression of childhood asthma by influencing various signaling pathways. These include not only direct asthma-related pathways but also the MAPK signaling pathway, the NF-kappa B, the IL6/JAK/STAT3, and the PI3K/AKT/mTOR signaling pathway, among other pathways, such as the reactive oxygen species pathway. In a distinct study, Theodorou et al[39] highlighted the importance of the MAPK pathway and its inhibitor, dual-specificity phosphatase-β 1 (DUSP 1), in the onset of childhood asthma and the protective effects mediated by the environment. Furthermore, research by Kim et al[40] revealed that essential oils from Mentha species could mitigate asthma symptoms when exposed to PM10 through the IL-6/JAK2/STAT3 signaling pathway. These findings highlight the multifaceted mechanisms through which various signaling pathways and genetic factors contribute to the complex pathogenesis of asthma, suggesting new targets for therapeutic intervention and research directions. Additionally, we examined TCTN3, a protein within the TCTN family found in the transition zone of cilia, located on human chromosome 10.[41] While previous studies on TCTN3 have concentrated on its involvement in ciliopathies, its connection to asthma has not been investigated.[42] Our research suggests that TCTN3 may regulate biological processes related to childhood asthma through pathways such as lipoic acid metabolism, the PI3K–Akt signaling pathway, and the Rap1 signaling pathway. For example, the PI3K-Akt pathway is involved in immune cell function and inflammation,[43] whereas the Rap1 pathway is related to cell adhesion and migration.[44] Abnormalities in these pathways may promote inflammation and allergic reactions in asthma, thereby affecting disease progression. This opens a new avenue for future research, highlighting the potential significance of TCTN3 in understanding and treating asthma. Therefore, the findings of this study provide new

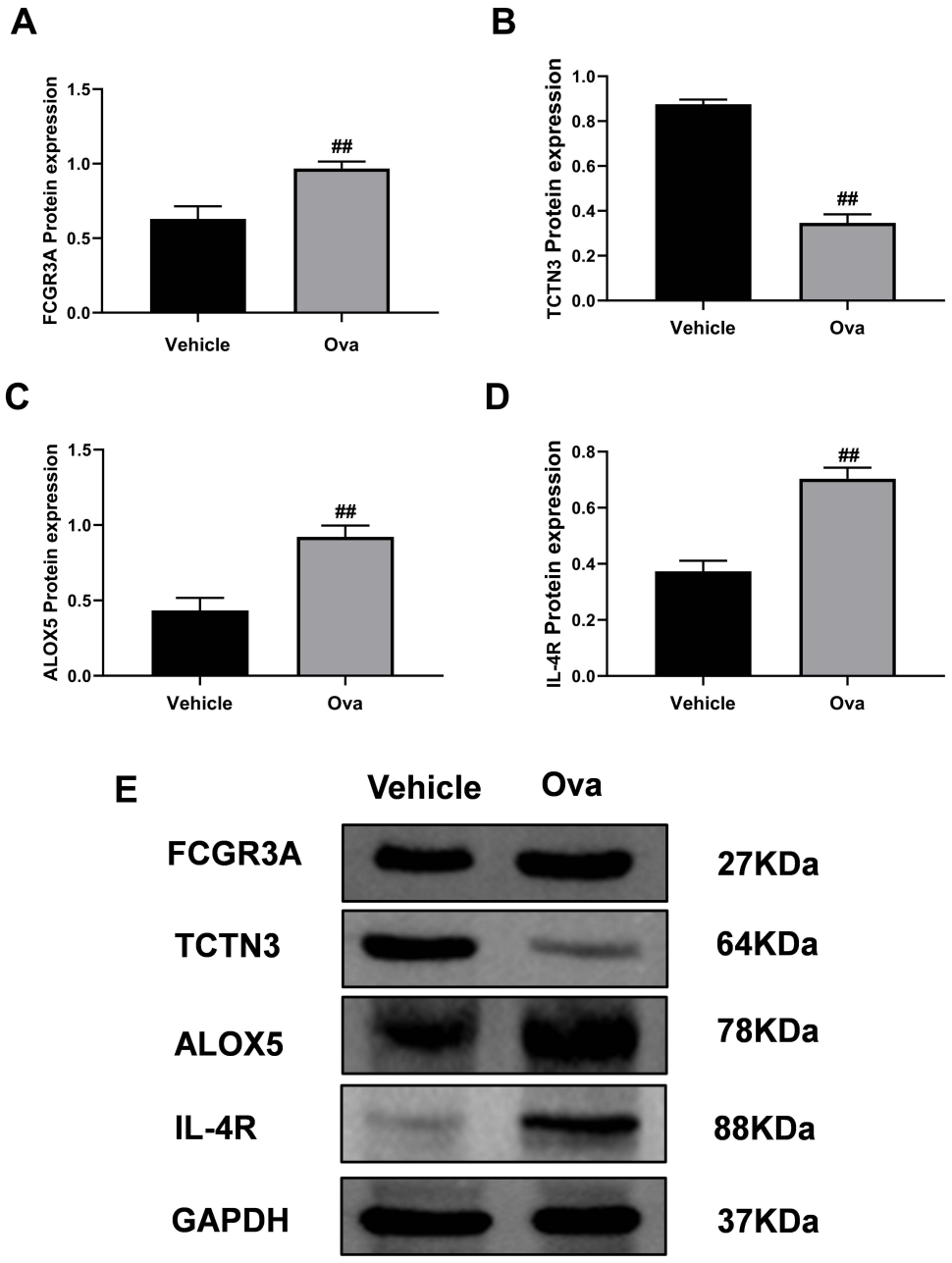


**Figure 9** (**A**–**E**) Protein expression of FCGR3A, TCTN3, ALOX5, and IL4R in ovalbumin-sensitized/activated mouse lung tissue ($^{\#\#}p<0.01$).

guidance for our future research. ALOX5 and IL4R are disease-regulating genes associated with childhood asthma. ALOX5 (arachidonate 5-lipoxygenase) plays a crucial role in the synthesis of inflammatory mediators, such as leukotrienes, which are produced by leukocytes and are closely related to asthma inflammation and allergic reactions. Studies suggest that the expression levels or genetic variations of ALOX5 may impact the risk and severity of asthma.[45] IL4R (Interleukin-4 Receptor) is closely associated with childhood asthma. Research has shown that genetic variations in IL4R are linked to the susceptibility to and severity of asthma. For example, some studies have shown that specific single nucleotide polymorphisms (SNPs) in IL4R are related to an increased risk of asthma in children.[46] Our research revealed a notable positive association between FCGR3A and ALOX5 and a significant negative correlation between TCTN3 and IL4R. However, a review of the literature revealed that there are no reports on the correlations between FCGR3A and ALOX5, TCTN3, or IL4R in childhood asthma, which opens a new direction for future research.


https://doi.org/10.2147/JIR.S498017

Certain limitations of this study must be recognized. This study only verified the gene and protein expression levels of two key genes and disease regulatory genes in asthma disease models through animal experiments. In the future, we need to conduct larger-scale clinical cohort studies and experiments to confirm the relationship between the expression levels of two key genes and the expression levels of disease-regulatory genes and clarify their regulatory role in biological processes related to childhood asthma.

## Conclusion

In our research, we successfully established a causal relationship between immune mechanisms and the development of pediatric asthma via MR analysis. This method significantly reduces the impact of unavoidable confounding factors. Additionally, we used machine learning techniques to identify six key genes associated with pediatric asthma. Through enrichment analysis, we were able to identify specific signaling pathways through which these key genes may influence disease progression. Moreover, by validating the protein and gene expression levels of key genes through animal experiments, we revealed their roles in the development or regulation of asthma. These findings help in understanding disease mechanisms and identifying potential therapeutic targets. This study not only deepens our understanding of the molecular basis of asthma but also paves the way for more effective prevention and treatment of this common childhood disease.

## Abbreviations

ATF4, Activating transcription factor 4; ALOX5, Arachidonate 5-Lipoxygenase; Bregs, Regulatory B cells; BAL, Bronchoalveolar lavage; CARAS, Combined allergic rhinitis and asthma syndrome; DCs, Dendritic cells; DUSP 1, Dual-specificity phosphatase-β 1; FCGR3A, Fc fragment of IgG receptor IIIa; GAS5, Growth arrest-specific transcript 5; GBD, Global Burden of Disease; GEO, Gene Expression Omnibus; GSEA, Gene set enrichment analysis; GSVA, Gene set variation analysis; IVs, Instrumental variables; IVW, Inverse variance-weighted; IL4R, Interleukin-4 Receptor; LCs, Innate lymphoid cells; MAPK, Mitogen-activated protein kinase; MEO, Essential oil derived from Mentha species; MGAT3, Mannose acetylglucosaminyltransferase 3; MR, Mendelian randomization; NCBI, National Center for Biotechnology Information; NK, Natural killer; OVA, Ovalbumin; STAT5, Signal transducer and activator of transcription 5; SVM, Support vector machine; TAB1, Transforming growth factor β-activated kinase 1 binding protein 1; Tfh, Follicular helper T; TP53, Tumor protein p53; Treg, regulatory T; Tregs, Regulatory T lymphocytes.

## Data Sharing Statement

The datasets supporting the conclusions of this study are available in the IEU open GWAS project repository (https://gwas.mrcieu.ac.uk/).

## Ethics Approval and Informed Consent

The Laboratory Animal Ethics Committee of Gansu University of Traditional Chinese Medicine approved all procedures involving animals in this study (ethical lot number: SY2024-247), and the guideline followed is "GB/T35892-2018 Guidelines for Ethical Review of Laboratory Animal Welfare". Our population data uses publicly available GWAS summary statistics. According to national legislative guidelines, this research can be exempted from ethical review. For example, item 1 of Article 32 of the "Measures for the Ethical Review of Life Sciences and Medical Research Involving Human Beings" issued by China on February 18, 2023: Research uses legally obtained public data or data generated through observation without interfering with public behavior; item 2: Research uses anonymous information data. All population data in this study came from public databases and met the above two requirements.

## Consent for Publication

Written informed consent for publication of this study was obtained from all participants.

## Acknowledgments

The authors thank the GWAS Catalog database for publishing GWAS summary statistics of immune cells, the NCBI GEO public database for publishing gene expression profile data, and the GeneCards database for publishing disease regulatory genes of childhood asthma.

## Funding

This study was supported by grants from the National Natural Science Foundation of China (grant number 81960614); Open Subjects of the Key Laboratory of Dunhuang Medicine, Ministry of Education (DHYX20-17); Open Subjects of the Research Center of Traditional Chinese Medicine, Gansu Province (ZYZX-2023-05); Gansu Provincial Administration of Traditional Chinese Medicine Scientific Research Project (GZKP-2023-42); Scientific Research and Innovation Fund Project of Gansu University of Chinese Medicine (2023KCYB-7); General project of Gansu Provincial Joint Scientific Research Fund (No. 24JRRA880).

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Asher MI, García-Marcos L, Pearce NE, Strachan DP. Trends in worldwide asthma prevalence. *Eur Respir J*. 2020;56(6):2002094. doi:10.1183/13993003.02094-2020
2. Wu JHJ. Highlights of bronchial asthma in children guideline (2016 edition) for its diagnosis and treatment. *World Clin Drugs*. 2018;39(8):512.
3. Pawankar R. Allergic diseases and asthma: a global public health concern and a call to action. *World Allergy Organ J*. 2014;7(1):12. doi:10.1186/1939-4551-7-12
4. Global Burden of Disease (GBD) [homepage on the Internet]. Seattle: Institute for Health Metrics and Evaluation (IHME); 2024. Available from: http://www.healthdata.org/gbd. Accessed September 17, 2024.
5. El-Husseini ZW, Gosens R, Dekker F, Koppelman GH. The genetics of asthma and the promise of genomics-guided drug target discovery. *Lancet Respir Med*. 2020;8(10):1045–1056. doi:10.1016/S2213-2600(20)30363-5
6. Laubhahn K, Schaub B. From preschool wheezing to asthma: immunological determinants. *Pediatr Allergy Immunol*. 2023;34(10):e14038. doi:10.1111/pai.14038
7. Deckers J, Madeira FB, Hammad H. Innate immune cells in asthma. *Trends Immunol*. 2013;34(11):540–547. doi:10.1016/j.it.2013.08.004
8. Hartl D, Koller B, Mehlhorn AT, et al. Quantitative and functional impairment of pulmonary CD4+CD25hi regulatory T cells in pediatric asthma. *J Allergy Clin Immunol*. 2007;119(5):1258–1266. doi:10.1016/j.jaci.2007.02.023
9. Spears M, McSharry C, Donnelly I, et al. Peripheral blood dendritic cell subtypes are significantly elevated in subjects with asthma. *Clin Exp Allergy*. 2011;41(5):665–672. doi:10.1111/j.1365-2222.2010.03692.x
10. Habib N, Pasha MA, Tang DD. Current understanding of asthma pathogenesis and biomarkers. *Cells*. 2022;11(17):2764. doi:10.3390/cells11172764
11. Zhu J. T helper 2 (Th2) cell differentiation, type 2 innate lymphoid cell (ILC2) development and regulation of interleukin-4 (IL-4) and IL-13 production. *Cytokine*. 2015;75(1):14–24. doi:10.1016/j.cyto.2015.05.010
12. Smith GD, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22. doi:10.1093/ije/dyg070
13. Larsson SC, Burgess S. Causal role of high body mass index in multiple chronic diseases: a systematic review and meta-analysis of Mendelian randomization studies. *BMC Med*. 2021;19(1):320. doi:10.1186/s12916-021-02188-x
14. Huang AA, Huang SY. Use of feature importance statistics to accurately predict asthma attacks using machine learning: a cross-sectional cohort study of the US population. *PLoS One*. 2023;18(11):e0288903. doi:10.1371/journal.pone.0288903
15. Kaplan A, Cao H, FitzGerald JM, et al. Artificial intelligence/machine learning in respiratory medicine and potential role in asthma and COPD diagnosis. *J Allergy Clin Immunol Pract*. 2021;9(6):2255–2261. doi:10.1016/j.jaip.2021.02.014
16. Khanam U, Ulfat A, Adamko, D, et al. A scoping review of asthma and machine learning. *J Asthma*. 2023;60(2):213–226. doi:10.1080/02770903.2022.2043364
17. Wang X, Meng L, Zhang J, et al. Identification of angiogenesis-related genes in diabetic foot ulcer using machine learning algorithms. *Heliyon*. 2023;9(12):e23003. doi:10.1016/j.heliyon.2023.e23003
18. Wang Q, Liu D, Gao T, et al. Experimental verification and identifying biomarkers related to insomnia. *Front Neurol*. 2023;14:1189076. doi:10.3389/fneur.2023.1189076
19. Xu B, Li H, Chen H, et al. Identification and prediction of molecular subtypes of atherosclerosis based on m6A immune cell infiltration. *Biochim Biophys Acta Gen Subj*. 2024;1868(2):130537. doi:10.1016/j.bbagen.2023.130537
20. Sanderson E, Glymour MM, Holmes MV, et al. Mendelian randomization. *Nat Rev Meth Primers*. 2022;2(1):6. doi:10.1038/s43586-021-00092-5
21. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res*. 2017;26(5):2333–2355. doi:10.1177/0962280215597579
22. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-egger method. *Eur J Epidemiol*. 2017;32(5):377–389. doi:10.1007/s10654-017-0255-x

23. Bowden J, Smith GD, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304–314. doi:10.1002/gepi.21965
24. Krusche J, Basse S, Schaub B. Role of early life immune regulation in asthma development. *Semin Immunopathol*. 2020;42(1):29–42. doi:10.1007/s00281-019-00774-z
25. Akdis M, Aab A, Altunbulakli C, et al. Interleukins (from IL-1 to IL-38), interferons, transforming growth factor β, and TNF-α: receptors, functions, and roles in diseases. *J Allergy Clin Immunol*. 2016;138(4):984–1010. doi:10.1016/j.jaci.2016.06.033
26. Tumer G, Simpson B, Roberts TK. *Genetics, Human Major Histocompatibility Complex (MHC)*. Florida, US: StatPearls Publishing; 2023.
27. Wang C, Zhu D, Zhang D, et al. Causal role of immune cells in schizophrenia: Mendelian randomization (MR) study. *BMC Psychiatry*. 2023;23(1):590. doi:10.1186/s12888-023-05081-4
28. Li X, Howard TD, Zheng SL, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. *J Allergy Clin Immunol*. 2010;125(2):328–335.e11. doi:10.1016/j.jaci.2009.11.018
29. Zhao Y, Wang C. Review of CD4+ T cell subsets in asthma phenotypes: molecular mechanisms and biologic treatment options. *J Shanghai Jiao Tong Univ*. 2023;43(8):1064–1070.
30. Luo W, Hu J, Xu W, Dong J. Distinct spatial and temporal roles for Th1, Th2, and Th17 cells in asthma. *Front Immunol*. 2022;13:974066. doi:10.3389/fimmu.2022.974066
31. Coquet JM, Schuijs MJ, Smyth MJ, et al. Interleukin-21-producing CD4(+) T cells promote type 2 immunity to house dust mites. *Immunity*. 2015;43(2):318–330. doi:10.1016/j.immuni.2015.07.015
32. Tibbitt CA, Stark JM, Martens L, et al. Single-Cell RNA sequencing of the T helper cell response to house dust mites defines a distinct gene expression signature in airway Th2 cells. *Immunity*. 2019;51(1):169–184.e5. doi:10.1016/j.immuni.2019.05.014
33. Cohn L, Homer RJ, Marinov A, Rankin J, Bottomly K. Induction of airway mucus production by T helper 2 (Th2) cells: a critical role for interleukin 4 in cell recruitment but not mucus production. *J Exp Med*. 1997;186(10):1737–1747. doi:10.1084/jem.186.10.1737
34. Ma CH, Ma ZQ, Fu Q, Ma SP. Ma Huang Tang ameliorates asthma through modulation of Th1/Th2 cytokines and inhibition of Th17 cells in ovalbumin-sensitized mice. *Chin J Nat Med*. 2014;12(5):361–366. doi:10.1016/S1875-5364(14)60044-3
35. Eldosoky MA, Hammad R, Rushdi A, et al. MicroRNA-146a-5p and microRNA-210-3p correlate with T regulatory cells frequency and predict asthma severity in Egyptian pediatric population. *J Asthma Allergy*. 2023;16:107–121. doi:10.2147/JAA.S398494
36. Kliem CV, Schaub B. The role of regulatory B cells in immune regulation and childhood allergic asthma. *mol Cell Pediatr*. 2024;11(1):1. doi:10.1186/s40348-023-00174-2
37. Habener A, Happle C, Grychtol R, et al. Regulatory B cells control airway hyperreactivity and lung remodeling in a murine asthma model. *J Allergy Clin Immunol*. 2021;147(6):2281–2294.e7. doi:10.1016/j.jaci.2020.09.041
38. Patel KR, Roberts JT, Subedi GP, Barb AW. Restricted processing of CD16a/Fc γ receptor IIIa N-glycans from primary human NK cells impacts structure and function. *J Biol Chem*. 2018;293(10):3477–3489. doi:10.1074/jbc.RA117.001207
39. Theodorou J, Nowak E, Böck A, et al. Mitogen-activated protein kinase signaling in childhood asthma development and environment-mediated protection. *Pediatr Allergy Immunol*. 2022;33(1):e13657. doi:10.1111/pai.13657
40. Kim MH, Park SJ, Yang WM. Inhalation of essential oil from Mentha piperita ameliorates PM10-exposed asthma by targeting IL-6/JAK2/STAT3 pathway based on a network pharmacological analysis. *Pharmaceuticals*. 2020;14(1):2. doi:10.3390/ph14010002
41. Gong S, Ji F, Wang B, Zhang Y, Xu X, Sun M. Tectonic proteins are important players in non-motile ciliopathies. *Cell Physiol Biochem*. 2018;50(1):398–409. doi:10.1159/000494017
42. Huang LX, Lu XG, Liu JX, et al. Case report and a brief review: analysis and challenges of prenatal imaging phenotypes and genotypes in Joubert syndrome. *Front Genet*. 2022;13:1038274. doi:10.3389/fgene.2022.1038274
43. Wang S, Guo L, Gu F, et al. Quercetin restores respiratory mucosal barrier dysfunction in Mycoplasma gallisepticum-infected chicks by enhancing Th2 immune response. *Phytomedicine*. 2024;133:155953. doi:10.1016/j.phymed.2024.155953
44. Johnson AK, Lorimer EL, Szabo A, et al. Rap1A, Rap1B, and β-Adrenergic signaling in autologous HCT: a randomized controlled trial of propranolol. *Yale J Biol Med*. 2022;95(1):45–56. doi:10.1016/j.bbmt.2020.03.002
45. Kalayci O, Birben E, Sackesen C, et al. ALOX5 promoter genotype, asthma severity and LTC production by eosinophils. *Allergy*. 2006;61(1):97–103. doi:10.1111/j.1398-9995.2006.00979.x
46. Li L, Li Y, Zeng XC, Li J, Du XY. Role of interleukin-4 genetic polymorphisms and environmental factors in the risk of asthma in children. *Genet Mol Res*. 2016;15(4)::1–8. doi:10.4238/gmr15048873