


Comprehensive Characterization of Th2/Th17 Cells-Related Gene in Systemic Juvenile Rheumatoid Arthritis: Evidence from Mendelian Randomization and Transcriptome Data Using Multiple Machine Learning Approaches

Mei Wang^{1,*}, Jing Wang^{2,*}, Fei Lv³, Aifeng Song¹, Wurihan Bao¹, Huiyun Li¹, Yongsheng Xu³ 

¹Department of Rheumatology and Immunology, Inner Mongolia Autonomous Region People's Hospital, Hohhot, Inner Mongolia, 010017, People's Republic of China; ²Department of Rheumatology and Immunology, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot, Inner Mongolia, 010050, People's Republic of China; ³Orthopedic Center, Inner Mongolia Autonomous Region People's Hospital, Hohhot, Inner Mongolia, 010017, People's Republic of China

*These authors contributed equally to this work

Correspondence: Huiyun Li, Department of Rheumatology and Immunology, Inner Mongolia Autonomous Region People's Hospital, Hohhot, Inner Mongolia, 010017, People's Republic of China, Email Lihuiyun2572@163.com; Yongsheng Xu, Orthopedic Center, Inner Mongolia Autonomous Region People's Hospital, Hohhot, Inner Mongolia, 010017, People's Republic of China, Email myfs202406@163.com

Background: Growing research has demonstrated that alterations in Th2 and Th17 cell composition were linked to systemic juvenile rheumatoid arthritis (sJRA). Nevertheless, whether these associations indicate a causal link remains unclear, and the potential effects of Th2/Th17-related molecules have not been clarified.

Methods: Mendelian randomization (MR) alongside transcriptome examination was implemented to ascertain the links between the Th2/Th17 cells and sJRA. Subsequently, we established an innovative machine learning (ML) framework encompassing 12 ML approaches and their 111 permutations to generate a unified Th2/Th17 classifier, which underwent verification across three separate cohorts. The hub Th2/Th17-related genes' level in the sJRA patients was substantiated via qRT-PCR. Lastly, the SHapley Additive exPlanations (SHAP) in conjunction with the XGBoost algorithm to pinpoint ideal Th2/Th17-linked genes.

Results: Based on MR analyses of two sJRA GWAS, 2 immunophenotypes (lymphocyte and IgD+ B cell) were causally linked to sJRA. Based on IOBR algorithms, we revealed that lymphocyte Th2/Th17 proportion was markedly changed in sJRA from seven cohorts. WGCNA and differential analysis in two merged GEO cohorts identified 64 Th2/Th17-related genes. Based on the average AUC (0.844) and model stability in four cohorts, we converted 12 ML techniques into 111 combinations, from which we chose the optimal algorithm to generate an ML-derived diagnostic signature (Th2/Th17 classifier). qRT-PCR verified results. Moreover, immune cell infiltration and functional enrichment analysis suggested hub Th2/Th17-related gene potentially mediated sJRA onset. XGBoost algorithm and SHAP detected HRH2 as crucial genetic markers, which may be an important target for sJRA.

Conclusion: A diagnostic model (Th2/Th17 classifier) via 111 ML algorithm combinations in six independent cohorts was generated and validated, which stands as an effective instrument for sJRA detection. The identification of essential immune components and molecular cascades, along with HRH2, could emerge as vital therapeutic targets for sJRA intervention, providing an enhanced understanding of its fundamental processes.

Keywords: systemic juvenile rheumatoid arthritis, Th2/Th17 cells, machine learning approaches, Mendelian randomization, transcriptome

Introduction

Juvenile rheumatoid/idiopathic arthritis (JRA/JIA) represents a diverse collection of arthritis with an unclear etiology that affects children under 16 and lasts for at least six weeks.¹ Systemic JRA (sJRA) is a distinctive subtype of JRA and accounts for approximately 10–20% of JRA.² sJRA is the most serious subtype due to multisystem involvement, and complications like macrophage activation syndrome, pericarditis, pleuritis, and so on, resulting in high rates of mortality and disability.³ While fever, joint involvement and fever with rash are classic clinical indicators of sJIA, early clinical manifestations might vary and certain symptoms may be minor or atypical, making detection and diagnosis challenging, primarily by exclusion.⁴ In order to expedite the diagnosis process and enable the implementation of efficient treat-to-target strategies, updated categorization criteria for sJIA have been proposed, which now need at least 6 weeks of arthritis.⁵ Additionally, there are not many options for treating this childhood illness. Treatment challenges can be avoided or reduced with effective identification and intervention. Thus, identifying emerging biomarkers remains crucial for sJIA detection, patient-specific risk evaluation, therapy target selection, and better comprehension of its underlying pathology.

The pathogenesis of sJIA is highly intricate and is acknowledged as an autoimmune inflammatory disease. The dysregulated immune system, including immune cells (ICs), serves a crucial function in this disease.⁶ Additionally, environmental elements will trigger its onset in individuals with genetic predisposition. For example, A 2012 study first described a markedly higher proportion of Th1 and Th17 helper T-cell subsets in the peripheral blood of individuals with sJIA.⁷ In acute sJIA, circulating Tregs exhibit a Th17-associated transcriptional profile, accompanied by IL-17 secretion, while retaining their inhibitory functions. In chronic JIA instead, the Th17-specific gene pattern predominantly manifests in circulating effector cells, indicating a shift in T cell phenotype during disease progression.⁸ These suggest that the pathophysiology and prognosis of sJIA may be directly impacted by the diversity of ICs. A knowledge gap exists regarding the comprehensive immune cellular landscape affecting sJIA and its function in disease development. Given the scarcity of randomized controlled trials, Mendelian randomization (MR) addresses biases stemming from reverse causation and confounding factors. In the past decade, MR has developed into a valuable approach for evaluating stronger causal relationships between exposure and clinical outcomes through genetic markers serving as instrumental variables (IVs).⁹

Moreover, the emergence of high-throughput RNA sequencing datasets has created unprecedented opportunities for the identification of novel biomarkers. As bioinformatics continues to evolve, numerous machine-learning (ML) techniques have been established and have become standard tools for identifying significant variables and developing predictive models. Currently, LASSO-Cox remains the predominant algorithm for developing comprehensive predictive signatures.^{10,11} However, the distinctiveness and limitations of certain modeling approaches have led to the development of models with notable deficiencies, which hampers their applicability in clinical settings. The exploration of emerging biomarkers through the integration of transcriptome data and sophisticated ML techniques for sJIA remains constrained. Furthermore, the potential of an integrated approach that combines diverse ML algorithms to generate a consensus model for sJIA diagnosis has yet to be fully realized.

Initially, MR analyses were executed to establish causal links between 731 ICs and sJIA, based on the comprehensive GWAS conducted thus far in Asian and European populations. The expression profiles of seven sJRA datasets (GSE7753-GPL570, GSE80060-GPL570, GSE8650-GPL96, GSE11907-GPL96, GSE13501, GSE17590, and GSE112057) were retrieved via exploration of the GEO repository. IOBR tools (CIBERSORT, TIMER, xCell, MCPcounter, ESTIMATE, EPIC, Quantiseq) along with ssGSEA were employed to assess the levels of immune cell infiltration (ICI). Weighted gene coexpression network analysis (WGCNA) was then executed to identify notable Th2/Th17-related modules within two merged GEO cohorts. Subsequently, an innovative ML structure was developed, incorporating 12 ML approaches and 111 permutations, to generate a unified Th2/Th17 classifier and confirm it across three independent cohorts. The hub Th2/Th17-related genes level in sJRA patients was confirmed using qRT-PCR. The investigation examined ICI and molecular mechanisms underlying sJRA onset, exploring their connections to central Th2/Th17-associated genes. Finally, the SHapley Additive exPlanations (SHAP) method and the XGBoost algorithm were employed to ascertain the most significant Th2/Th17-related genes, with the Mantel test being conducted to explore

the link between the selected genes and essential molecular cascades. It is hypothesized that the developed Th2/Th17 classifier may prove effective in diagnosing sJRA. Furthermore, the identification of hub Th2/Th17-related genes, fundamental molecular mechanisms, and immune components provides crucial understanding regarding sJRA pathogenesis, potentially highlighting targetable elements for therapeutic intervention.

Methodologies and Materials

GWAS Data Sources for sJRA and IC Traits

Two large-scale GWAS data on sJRA (GWAS ID: ebi-a-GCST90018653, ebi-a-GCST90018873) were procured from the IEU OpenGWAS project (<https://gwas.mrcieu.ac.uk/>). Two sJRA GWAS datasets come from different ethnic groups, namely East Asian and European populations. Table 1 depicts basic data procured from sJRA GWAS datasets.

The summary statistics of immunological traits from GWAS are publicly accessible through the GWAS Catalog (accession numbers GCST90001391 to GCST90002121).¹² This dataset encompasses 731 unique immune phenotypes, including morphological parameters (n=32), absolute cell counts (n=118), relative cell counts (n=192), and median fluorescence intensities indicative of surface antigen expression levels (n=389). Table 1 depicts the fundamental data procured through GWAS on ICs.

Primary MR Analysis

In MR, genetic variants serve as IVs that exhibit robust associations with the exposure and remain independent of alternative elements potentially impacting the outcome, thereby effectively mitigating the impact of confounders. To ensure the validity of an MR study, the IVs must satisfy three fundamental criteria: (1) Relevance Assumption; (2) Independence Assumption; (3) Exclusivity Assumption.

SNPs associated with immunological characteristics were identified using a threshold of $P < 1 \times 10^{-5}$, consistent with earlier MR investigations.¹³ Additionally, the analysis of genetic correlation (LD) between SNPs was based on European ancestry reference data obtained from the 1000 Genomes Project. We selected SNPs employing an LD threshold ($r^2 < 0.001$) and positioned beyond 10Mb, to establish SNPs demonstrating independent genetic effects.

Furthermore, when shared SNPs were absent between exposure and outcome, substitute SNPs in linkage disequilibrium ($r^2 \geq 0.8$) were incorporated. To minimize weak IV bias, we selected SNPs with F-statistic values exceeding 10 (an indicator of IV's strength).

MR analysis utilizing the TwoSampleMR package (version 0.5.6) was employed to determine the causal links between immunological traits and sJRA in two distinct sJRA GWAS cohorts. The MR estimates are reported as odds ratios (OR) with 95% CI for dichotomous data, or as beta values with standard errors (SE) for continuous variables. To enhance the robustness and reliability of the causal association, the intersection of outcomes from the two sJRA GWAS cohorts was considered.

MR Sensitivity and Heterogeneity Analysis

To assess the robustness of MR impact estimates against potentially invalid genetic variants, sensitivity analyses were conducted using the weighted median (WM), MR-Egger regression, simple mode, and weighted mode. In order to account for potential heterogeneity, the Cochran Q statistic was applied, calculated through MR-Egger regression and IVW methodologies. Visualizations were generated utilizing funnel plots, and the MR-Egger intercept was employed to investigate horizontal pleiotropy, with a threshold of $P < 0.05$ applied to both tests. Additionally, MR Steiger filtering

Table 1 Genome-Wide Association Studies Summary Statistics Datasets Used for Genetic Analyses

| GWAS ID | Year | Consortium | Population | Ncase/Ncontrol | Number of SNPs | Phenotypes |
|--|------|------------|------------|-------------------|-------------------|--------------|
| ebi-a-GCST90018653 | 2021 | NA | East Asian | 110/173,268 | 12,453,962 | JRA |
| ebi-a-GCST90018873 | 2021 | NA | European | 216/409,001 | 24,190,478 | JRA |
| ebi-a-GCST90001391 to ebi-a-GCST90002121 | 2020 | NA | European | Sample size: 3757 | Nearly 22,000,000 | Immune cells |

Abbreviations: GWAS, genome-wide association studies; SNP, single-nucleotide polymorphism; JRA, juvenile rheumatoid arthritis; NA, not available.

was performed to determine the causal direction between exposure and outcome. Lastly, a leave-one-out (LOO) evaluation was implemented to recalculate the aggregate effect magnitude and examine if individual SNPs could affect the relationship by systematically omitting each exposure-linked SNP.

Transcriptome Dataset Collection, and Preprocessing

The GEO database (<http://www.ncbi.nlm.nih.gov/geo>) underwent comprehensive exploration to identify sJRA-associated transcriptomic datasets spanning from May 2003 to February 2024. The transcriptome expression datasets were required to fulfill these specifications: a) organism: *Homo sapiens*; b) expression profiling via high-throughput sequencing or array technologies. In the end, seven GEO datasets were selected for both quantitative and qualitative analyses. Considering that GSE8650 and GSE11907 (GPL96), GSE7753 and GSE80060 (GPL570) have the same annotation platform and the similar baseline expression level of the same genes, GSE8650 and GSE11907, GSE7753 and GSE80060 were merged as GEO-GPL96, and GEO-GPL570, respectively, with batch correction. GEO-GPL96 were set as discovery cohorts, and GEO-GPL570, GSE13501, GSE17590, and GSE112057 were used as the external validation cohorts. For GEO datasets, a comprehensive multi-array average examination was executed, encompassing quantile standardization, background adjustment, and data consolidation.¹⁴

Assessment of ICI

Eight distinct algorithms were employed to assess ICI in sJRA patients across seven GEO cohorts. These algorithms, encompassing Quantiseq, MCPcounter, CIBERSORT, TIMER, xCell, ESTIMATE, and EPIC, were implemented through the R package “IOBR” (version 0.99.8),¹⁵ whereas ssGSEA was utilized to ascertain the infiltrating ICs.¹⁶ Furthermore, comparisons of the relative abundance of Th1, Th2, Th17, Th2/Th17 cells, and B cells were made between healthy/control individuals and the sJRA cohort.

Wgcna

A weighted gene co-expression network was constructed in two merged GEO cohorts (GEO-GPL96 and GEO-GPL570, both characterized by large sample sizes) utilizing the R package “WGCNA” (version 1.71). This approach enabled the identification of operational genetic clusters that exhibited strong correlations with Th2/Th17 cell percentages, making them suitable for further screening.¹⁷ The identified modules were then associated with four traits (Th1, Th2, Th17, and Th2/Th17 cell proportions). Genes from the module showing notable correlations with Th2/Th17 cells were subsequently selected for additional analysis.

Detection of Th2/Th17-Related Genes and Functional Enrichment Analysis (FEA)

Standardized data, comprising the GEO-GPL96 and GEO-GPL570 datasets from both healthy and sJRA specimens, underwent variance examination via the NetworkAnalyst online Gene Expression Table (<https://www.networkanalyst.ca/>). Subsequently, the overlap of DEGs and notable Th2/Th17 module genes from the two sJRA cohorts was identified, with the intersecting genes being classified as Th2/Th17-related. These genes were then visualized using the UpSetR package. GO and KEGG enrichment analyses were performed on the Th2/Th17-related genes with the clusterProfiler R package (version 4.8.3).

Development and Establishment of Th2/Th17 Discriminator Utilizing Diverse ML Approaches

The subsequent steps were implemented to develop an agreed-upon diagnostic model for sJRA: (1) Twelve classical algorithms were initially combined, including StepAIC, gradient boosting machine (GBM), eXtreme Gradient Boosting (XGBoost), Linear Discriminant Analysis (LDA), NaiveBayes, support vector machine (SVM), random forest (RF), glmBoost, LASSO, plsRglm, ridge regression, and elastic network (Enet). Among these, LASSO, RF, StepAIC, and glmBoost were equipped with feature selection capabilities. Furthermore, 111 distinct algorithm combinations were generated as predictive models employing the leave-one-out cross-validation (LOOCV) framework (2). In the subsequent

phase, the GEO-GPL96 dataset for sJRA was utilized to train the model, applying the 111 combinations to independently develop classifiers based on Th2/Th17-related genes. (3) Finally, in the testing cohorts (GEO-GPL570, GSE13501, and GSE17590), the Th2/Th17 score for each cohort was calculated by applying the model derived from the training set. The optimal consensus diagnostic model for sJRA was finally identified by averaging the area under the curve (AUC) across the four cohorts, alongside considerations of model simplicity and generalizability.

The above process can avoid the potential overfitting problems of ML. For example, 1. The model exhibits excessive complexity: Achieving precise fitting outcomes leads to an overly intricate system, yielding impractical parameters, poor performance, and predictive power. 2. Noise in data due to random errors or irrelevant information in the data.

Diagnostic Value, and Clinical Usefulness of Th2/Th17 Classifier

Receiver operating characteristic (ROC) examination was performed to assess the diagnostic performance of the Th2/Th17 classifier across five sJRA datasets. Principal component analysis (PCA),¹⁸ employing key Th2/Th17-related gene expression profiles, was conducted on the same five datasets. Finally, the clinical relevance of the Th2/Th17 classifier was evaluated through decision curve analysis (DCA).

Clinical Specimen Collection and qRT-PCR

Twenty-five patients in our hospital, ten of them were diagnosed with systemic juvenile arthritis (sJRA) based on the International League of Associations for Rheumatology's (ILAR) classification criteria¹ and 15 control patients (general surgical patients), were recruited. The Inner Mongolia Autonomous Region People's Hospital's institutional review board (ethics committee) sanctioned all procedures involving human subjects. At admission, peripheral blood specimens and related clinical information were gathered. Within four hours of blood collection, peripheral blood mononuclear cells (PBMCs) were separated. The family members of qualified participants provided written consent before specimen collection.

Total RNA was procured from PBMCs per the specified protocol utilizing Trizol (Invitrogen). RNA reverse transcription was performed with the RevertAid RT Reverse Transcription Kit (Thermo Scientific). Quantitative PCR was carried out with PowerUp™ SYBR™ Green Master Mix, employing Gapdh as the internal control. Quantitative reverse transcription-PCR was performed on the ABI 7500 real-time PCR system (Applied Biosystems). The fold change in gene expression was subsequently assessed utilizing the $2^{-\Delta\Delta C_t}$ technique. Gene-specific PCR primers are depicted in [Supplementary Material 1](#).

Expression Levels, and Correlation Pattern of Hub Th2/Th17-Related Genes

The mRNA levels of hub Th2/Th17-related genes in both the sJRA and control cohorts were examined and validated across five GEO datasets and a hospital cohort. Furthermore, the hub Th2/Th17-related genes within these datasets underwent correlation analysis.

Evaluation of ICI in sJRA and Its Connection with Hub Th2/Th17-Related Genes

Building upon the outcomes derived from eight distinct algorithms utilized to quantify ICI, the scores for immune, stromal, and estimate components were compared across two merged sJRA cohorts (GEO-GPL96 and GEO-GPL570). Furthermore, Spearman correlation analyses were executed to explore the link between hub Th2/Th17-related genes and IC types.

Exploration of Possible Pathways in sJRA and Its Connection with Hub Th2/Th17-Related Genes

ssGSEA analysis was conducted on diverse characteristic gene sets associated with immune, inflammatory, and cell death-related pathways¹⁶ across two sJRA cohorts. Subsequently, gene set variation analysis (GSVA)¹⁹ was employed to identify signaling cascades based on the KEGG (c2.cp.kegg.v7.4.symbols.gmt), GO (c5.all.v2023.2.Hs.symbols.gmt), and HALLMARK (h.all.v2023.2.Hs.symbols.gmt) gene sets procured from the Molecular Signatures Database. In the

pathway analysis, differential enrichment scores between the sJRA and control cohorts were assessed. Finally, correlation analyses were performed to explore the association between Th2/Th17-related genes and essential biological mechanisms.

XGBoost and SHAP

The XGBoost algorithm has recently gained widespread recognition for its superior performance in ML competitions, particularly when compared with algorithms like k-nearest neighbors, logistic regression, Lasso, RF, and SVM.²⁰ In this investigation, the XGBoost algorithm was utilized to develop a diagnostic model for sJRA, focusing on Th2/Th17-related genes. The significance of each feature was subsequently assessed utilizing SHAP values, which were visually represented through summary and bar plots that ranked the features by importance. Lastly, the Mantel test was employed to examine the relationship between the selected Th2/Th17-related genes and key biological pathways.

Statistical Analysis

Statistical analyses were executed employing RevMan version 5.3 (The Cochrane Collaboration, Copenhagen, Denmark), GraphPad Prism 9, SPSS 22, and R software (R version 4.3.1). Continuous data were evaluated utilizing Student's t-tests, Mann–Whitney *U*-tests, or Wilcoxon signed-rank tests, with two-tailed assessments depending on variable distributions. Statistical significance was established at $P < 0.05$, unless particular *P*-values were specified.

Results

Genetic Instruments for ICs

The significance threshold ($P < 1 \times 10^{-5}$), along with harmonization, LD clumping, and F-statistics > 10 , was employed to identify genetic IVs for 731 ICs. As a result, negligible instrument bias was demonstrated in the present study.

MR Analysis

While investigating the causative influence of immune characteristics on sJRA, we identified 9284 and 18189 SNPs as IVs from 731 ICs for sJRA ebi-a-GCST90018653 and ebi-a-GCST90018873, respectively ([Supplementary material 2](#) and [3](#)). A sum of 32 immunophenotypes (ebi-a-GCST90018653) ([Supplementary material 4](#) and [Figure S1](#)) and 38 immunophenotypes (ebi-a-GCST90018873) ([Supplementary material 5](#) and [Figure S2](#)) demonstrated causative influences on sJRA utilizing the IVW approach ($P < 0.05$) ([Table 2](#) and [Figure S1](#)). Through sensitivity evaluation, 24 immune phenotypes (ebi-a-GCST90018653) ([Supplementary material 4](#) and [Figure S3](#)) and 34 immune phenotypes (ebi-a-GCST90018873) ([Supplementary material 5](#) and [Figure S4](#)) utilizing the IVW approach aligned with outcomes from MR-Egger regression, WM, simple model, and weight model. To enhance the robustness and significance of the causal associations, we analyzed the overlap between two sJRA GWAS (ebi-a-GCST90018653 and ebi-a-GCST90018873) while considering immune trait directionality. This analysis identified two immune phenotypes (Lymphocyte %leukocyte, IgD+ B cell absolute count) exhibiting causal links to sJRA ([Figure 1A](#)). In the sJRA ebi-a-GCST90018653 cohort from East Asian population, Host-genetic-driven increase in Lymphocyte %leukocyte (OR= 1.27, 95% CI=1.03–1.55, $P_{IVW} = 0.022$), IgD+ B cell absolute count (OR= 1.29, 95% CI=1.02–1.63, $P_{IVW} = 0.035$) markedly increases sJRA risk ([Table 2](#) and [Figure 1B](#)). Comparable findings emerged in the sJRA ebi-a-GCST90018873 cohort from the European population ([Table 2](#) and [Figure 1C](#)).

Furthermore, no notable heterogeneity was noted in the Cochran Q test employing MR-Egger regression and IVW (all $P > 0.05$, [Table 2](#), [Supplementary materials 6](#) and [7](#), [Figures S5](#) and [S6](#)). The lack of deviation from zero in the MR-Egger regression intercepts confirms the absence of horizontal pleiotropy ($P > 0.05$ for all intercepts, [Table 2](#), [Supplementary materials 8](#) and [9](#)). Furthermore, the MR Steiger test results validated the SNP selection, thereby supporting the proposed causal direction of immunophenotypes' influence on sJRA (all $P > 0.05$). Finally, the LOO analyses demonstrated that none of the identified causal relationships were influenced by any individual IV ([Figures S7](#) and [S8](#)).

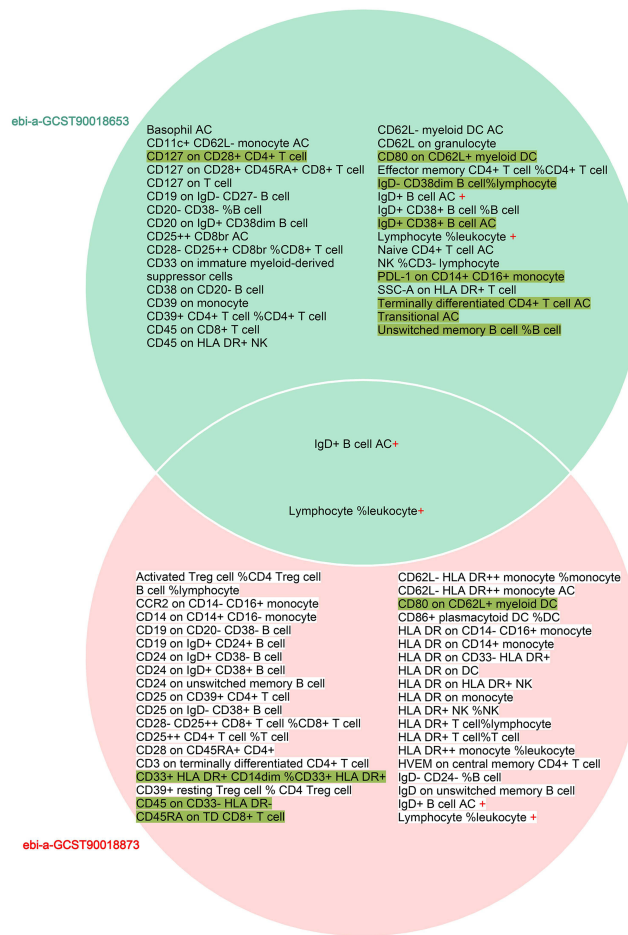
Table 2 Mendelian Randomization Estimating the Effect of Immune Cells Exposure on the Risk of Juvenile Rheumatoid Arthritis in Two GWAS Datasets

| GWAS ID | Exposure | Method | Number of SNP | Odds ratio (95% CI) | Beta \pm SE | P | Q-statistics | P_h | Egger intercept | Pintercept |
|--------------------|----------------------------|----------|---------------|---------------------|-------------------|-------|--------------|-------|-----------------|------------|
| ebi-a-GCST90018653 | Lymphocyte %leukocyte | IVW | 11 | 1.27 (1.03–1.55) | 0.236 \pm 0.103 | 0.022 | 6.78 | 0.77 | 0.017 | 0.78 |
| | | MR Egger | 11 | 1.25 (0.99–1.57) | 0.221 \pm 0.117 | 0.092 | 6.70 | 0.67 | | |
| | | WM | 11 | 1.25 (0.94–1.67) | 0.226 \pm 0.146 | 0.12 | | | | |
| | IgD+ B cell Absolute Count | IVW | 12 | 1.29 (1.02–1.63) | 0.254 \pm 0.121 | 0.035 | 6.96 | 0.80 | | |
| | | MR Egger | 12 | 1.14 (0.86–1.51) | 0.129 \pm 0.144 | 0.39 | 4.35 | 0.93 | 0.096 | 0.14 |
| | | WM | 12 | 1.21 (0.90–1.61) | 0.189 \pm 0.148 | 0.20 | | | | |
| ebi-a-GCST90018873 | Lymphocyte %leukocyte | IVW | 22 | 1.20 (1.03–1.40) | 0.183 \pm 0.080 | 0.021 | 25.80 | 0.21 | -0.009 | 0.82 |
| | | MR Egger | 22 | 1.22 (1.01–1.47) | 0.195 \pm 0.097 | 0.057 | 25.73 | 0.17 | | |
| | | WM | 22 | 1.23 (0.89–1.69) | 0.205 \pm 0.164 | 0.21 | | | | |
| | IgD+ B cell Absolute Count | IVW | 17 | 1.19 (1.06–1.35) | 0.177 \pm 0.061 | 0.004 | 15.04 | 0.52 | | |
| | | MR Egger | 17 | 1.11 (0.96–1.29) | 0.108 \pm 0.074 | 0.16 | 12.18 | 0.67 | 0.064 | 0.11 |
| | | WM | 17 | 1.14 (0.97–1.34) | 0.133 \pm 0.081 | 0.10 | | | | |

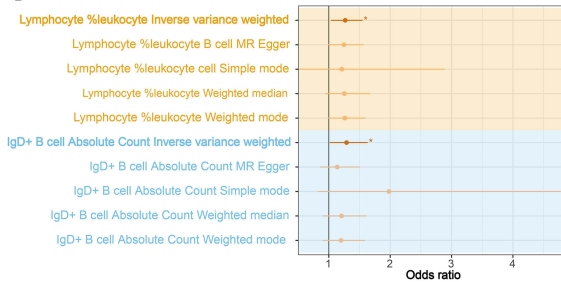
Notes: Odds ratios, 95% CI, and *P*-values were obtained from Mendelian randomization analysis. The heterogeneity test in the IVW and MR-Egger method was performed using Cochran's *Q* statistic.

Abbreviations: GWAS, genome-wide association studies; SNP, single nucleotide polymorphism; CI, confidence interval; P_h , *P*-value for heterogeneity; Pintercept, *P*-value for the intercept of the MR-Egger regression; IVW, inverse-variance-weighted; MR, Mendelian randomization; WM, Weight median.

A



B



C

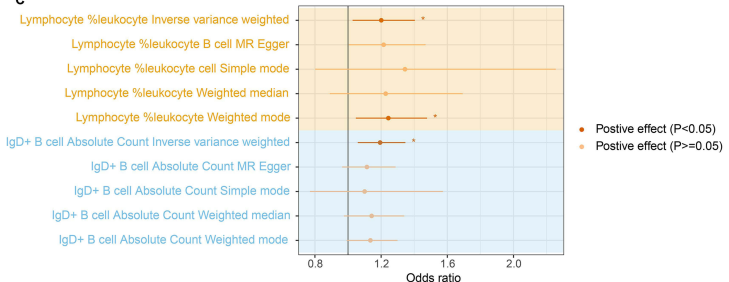


Figure 1 Mendelian randomization (MR) estimating the effect of immune cells exposure on the risk of juvenile rheumatoid arthritis (JRA) in two Genome-wide association studies (GWAS) datasets. **(A)** Venn diagram displays the shared immune cells subtypes in two JRA GWAS datasets (GWAS ID: ebi-a-GCST90018653, and ebi-a-GCST90018873). + represent positive correlation, - represent negative correlation, text with color shading represent the result direction of inverse variance weighted (IVW) method exists inconsistent with those of sensitivity analysis (MR-Egger regression, weighted median, simple model and weight model). **(B and C)** MR evaluated the effects of 2 immunophenotypes (lymphocyte %leukocyte and IgD+ B cell absolute count) on JRA. **(B)** Forrest plot showing causal effect of genetically predicted immunological traits on sJRA (GWAS ID: ebi-a-GCST90018653). **(C)** Forrest plot showing causal effect of genetically predicted immunological traits on sjRA (GWAS ID: ebi-a-GCST90018873). * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Analysis of the Infiltration Level of Th2/Th17 and B Cells in sJRA Transcriptome Data

The results of MR from two sJRA GWAS cohorts uncovered causal relationships between sJRA with lymphocyte and B cells. Lymphocytes are known to consist of T cells, B cells, and NK cells. A thorough examination of ICI inferred using eight algorithms in seven GEO datasets found that lymphocyte Th17 cells (proinflammatory) were generally more abundant in sJRA, while lymphocyte Th2 cells (anti-inflammatory) levels decreased. Interestingly, the lymphocyte Th2/Th17 cell ratio is markedly decreased in sJRA patients than in healthy control patients in seven GEO datasets (Figure 2A-G). According to ssGSEA, CIBERSORT, xCell, EPIC, TIMER, and Quantiseq algorithms, there was almost no significant difference in B cell infiltrating

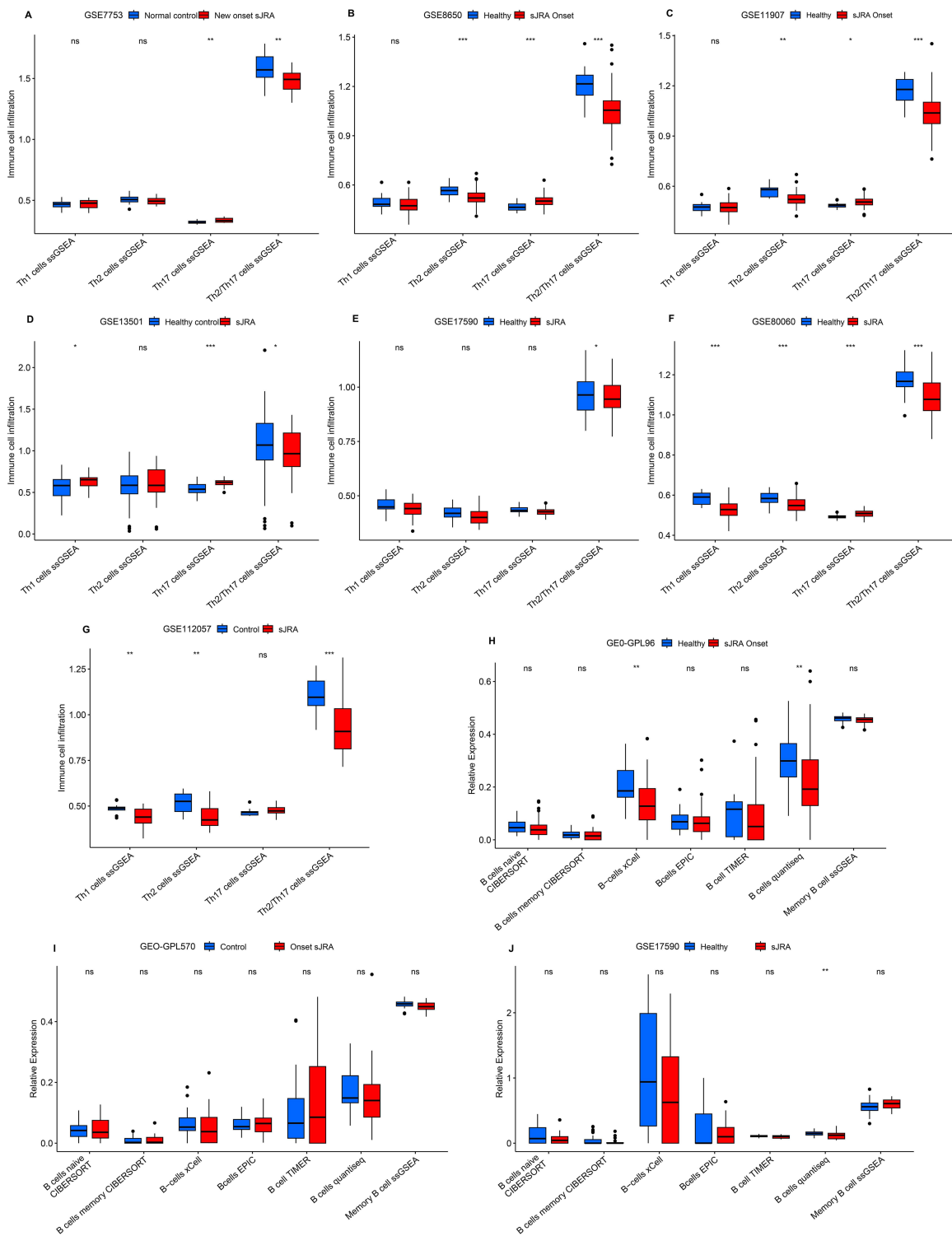


Figure 2 Comparison of the relative abundance of Th1, Th2, Th17, Th2/Th17 and B cells in seven systemic JRA (sJRA) transcriptome data using six immune cell infiltration algorithms (CIBERSORT, ssGSEA, xCell, TIMER, EPIC, and Quantiseq). **(A)** Th1, Th2, Th17, Th2/Th17 in GSE7753 dataset using ssGSEA. **(B)** Th1, Th2, Th17, Th2/Th17 in GSE8650 dataset using ssGSEA. **(C)** Th1, Th2, Th17, Th2/Th17 in GSE11907 dataset using ssGSEA. **(D)** Th1, Th2, Th17, Th2/Th17 in GSE13501 dataset using ssGSEA. **(E)** Th1, Th2, Th17, Th2/Th17 in GSE17590 dataset using ssGSEA. **(F)** Th1, Th2, Th17, Th2/Th17 in GSE80060 dataset using ssGSEA. **(G)** Th1, Th2, Th17, Th2/Th17 in GSE112057 dataset using ssGSEA. **(H)** B cells in GEO-GPL96 dataset using CIBERSORT, ssGSEA, xCell, TIMER, EPIC, and Quantiseq. **(I)** B cells in GEO-GPL570 dataset using CIBERSORT, ssGSEA, xCell, TIMER, EPIC, and Quantiseq. **(J)** B cells in GSE17590 dataset using CIBERSORT, ssGSEA, xCell, TIMER, EPIC, and Quantiseq. The asterisks indicate a significant statistical *p* value calculated using the Wilcoxon test (**p* < 0.05; ***p* < 0.01; ****p* < 0.001).

from three GEO cohorts with large sample sizes (Figure 2H-J). These findings emphasized the crucial significance of lymphocyte Th2/Th17 ratio dysregulation in sJRA development.

Wgcna

WGCNA was performed on two merged sJRA datasets to identify key Th2/Th17 module genes. Initially, pairwise gene correlations were calculated to construct a similarity matrix. A scale-free network was subsequently designed after testing soft threshold power (β) values ranging from 1 to 20. In the GEO-GPL96 and GEO-GPL570 cohorts, the optimal β values were determined to be 6 and 4, respectively, and applied to transform the similarity matrix into an adjacency matrix (Figures 3A and S9A). Moreover, a negative link ($R^2 = 0.71$, slope = -1.99 in GSE119600; $R^2 = 0.92$, slope = -1.71 in GSE159676) between $\log_{10}(k)$ and $\log_{10}(p(k))$ is illustrated in Figures 3B and S9B, demonstrating the proximity of the transformed adjacency matrix to a scale-free network for subsequent analysis. The adjacency matrix was then used to generate the TOM for dynamic tree-clustering of genes (Figures 3C and S9C), allowing for the detection of gene modules associated with Th2/Th17. Four characteristics—Th1 cells, Th2 cells, Th17 cells, and Th2/Th17 cells—were found to correlate with the discovered gene modules. As depicted in Figures 3D and S9D, the ME red, ME salmon, ME blue, and ME light cyan modules exhibited significant correlations with Th2/Th17 in the GEO-GPL96 and GEO-GPL570 cohorts. Furthermore, the Th2/Th17 modules were re-identified by computing the mean absolute GS value of the Th2/Th17 gene in each module in order to verify the validity of the identification results. The mean absolute GS values for Th2/Th17 were largest for the ME red module and ME salmon module in GEO-GPL96 cohort (Figure 3E), ME blue module and ME light cyan module in GEO-GPL570 cohort (Figure S9E). Finally, the genetic components within the ME red module ($n = 1873$) and ME salmon module ($n = 183$) from the GEO-GPL96 group, along with the ME light cyan module ($n = 1522$) and ME blue module ($n = 572$) from GEO-GPL570 group were identified as essential Th2/Th17 module genes for subsequent investigation.

Detection of Th2/Th17-Associated Genes and FEA

DEGs were identified employing the NetworkAnalyst online Gene Expression Table, with 3441 DEGs in the GEO-GPL96 dataset and 12,218 in the GEO-GPL570 dataset. The intersection of DEGs and notable Th2/Th17 module genes was then determined across the two datasets, yielding a sum of 64 genes identified as Th2/Th17-related in sJRA (Figure 4A). Subsequently, GO term enrichment analysis of these genes revealed the top five clusters with notable enrichment across cellular components (CCs), biological processes (BPs), and molecular functions (MFs) (Figure 4B). These clusters included the cytoplasmic side of the plasma membrane, presynaptic membrane, synaptic membrane, and postsynaptic membrane. For the KEGG pathways, the genes were predominantly associated with the IL-17 signaling pathway, cell adhesion molecules, inositol phosphate metabolism, and the phosphatidylinositol signaling system, among others (Figure 4C).

Development and Establishment of Th2/Th17 Discriminator Utilizing Diverse ML Methodologies

To establish a robust diagnostic model for sJRA, 64 hubs Th2/Th17-related genes were integrated into the LOOCV framework. Prediction models were constructed using 10-fold cross-validation and 111 algorithm combinations within the GEO-GPL96 training cohort. For the testing cohorts, the average AUC value was computed for each algorithm. The ultimate model was identified as the integration of glmBoost and RF, which demonstrated the highest average AUC (0.844), the lowest variability, and optimal stability across the four cohorts, as depicted in Figure 4D. Subsequently, a diagnostic model, referred to as the Th2/Th17 classifier, was developed based on six hub Th2/Th17-related genes (HRH2, ISL1, MRPL22, FTCD, CD2, and E2F1). Furthermore, Th2/Th17 scores for each sample across all four cohorts were calculated using the expression profiles of the six genes incorporated into the Th2/Th17 classifier.

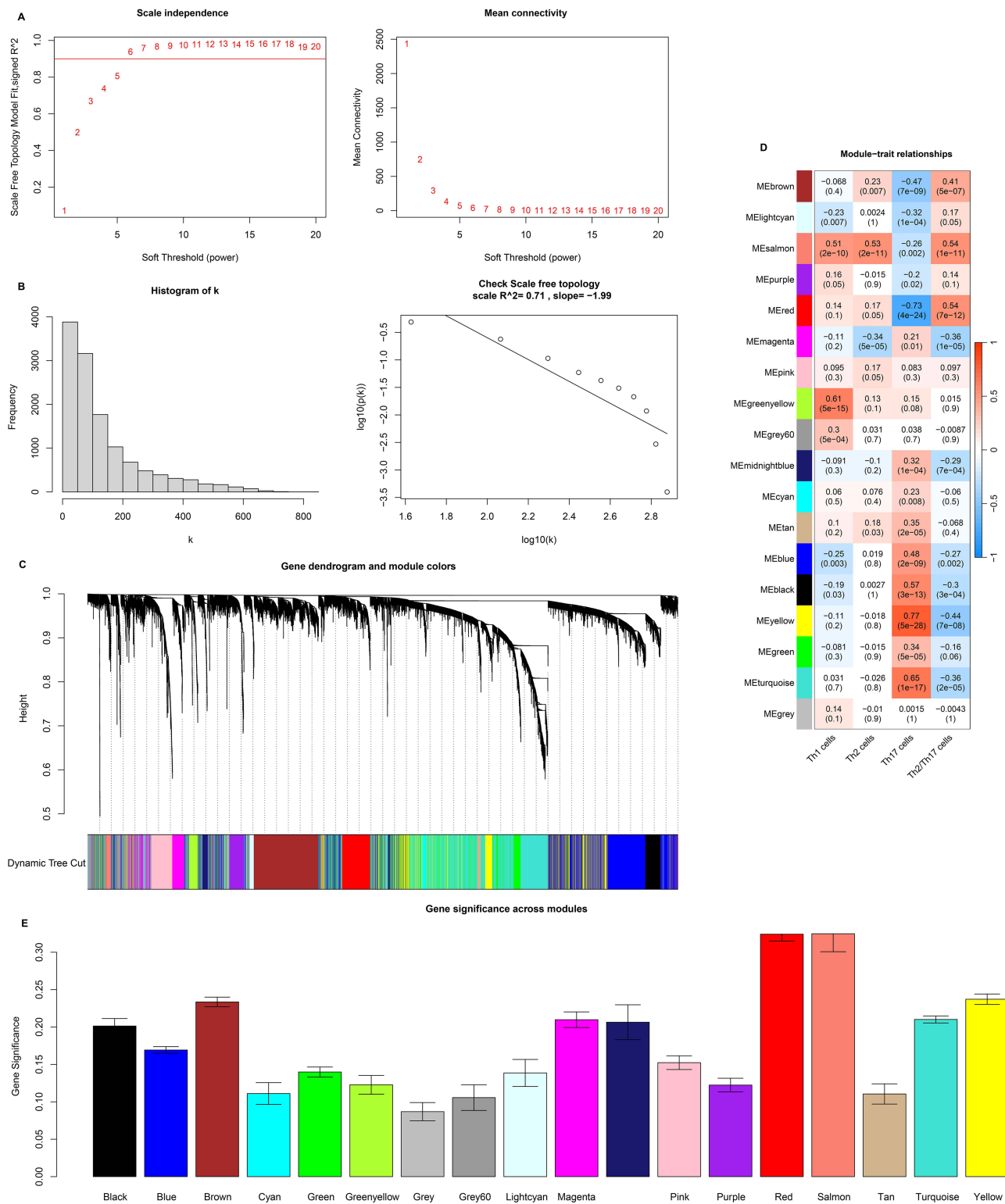


Figure 3 Weighted gene coexpression network analysis (WGCNA) for screening functional gene modules highly associated with Th2/Th17 proportion in GEO-GPL96 dataset. **(A)** Analysis of network topology for various soft-thresholding powers (weighting coefficient, β). The x-axis represents different soft-thresholding powers. The y-axis represents the correlation coefficient between $\log(k)$ and $\log[P(k)]$. The red line indicates a correlation coefficient of 0.9. Average network connectivity under different weighting coefficients. **(B)** Distribution of nodes with the degree of connection, k , and Correlation of $\log(k)$ and $\log[P(k)]$. **(C)** Clustering dendrograms of all genes, with dissimilarity based on topological overlap, together with assigned module colors. Altogether, 18 coexpression modules were constructed and displayed in different color. **(D)** The heatmap showed the relationship between five WGCNA-identified modules and two traits (Th1, Th2, Th17 and Th2/Th17 cells). Each cell includes the correlation coefficient and p value. **(E)** Bar plots of mean gene significance (GS) across modules. The higher the mean GS in a module is, the more significantly associated with Th2/Th17 cells the module will be.

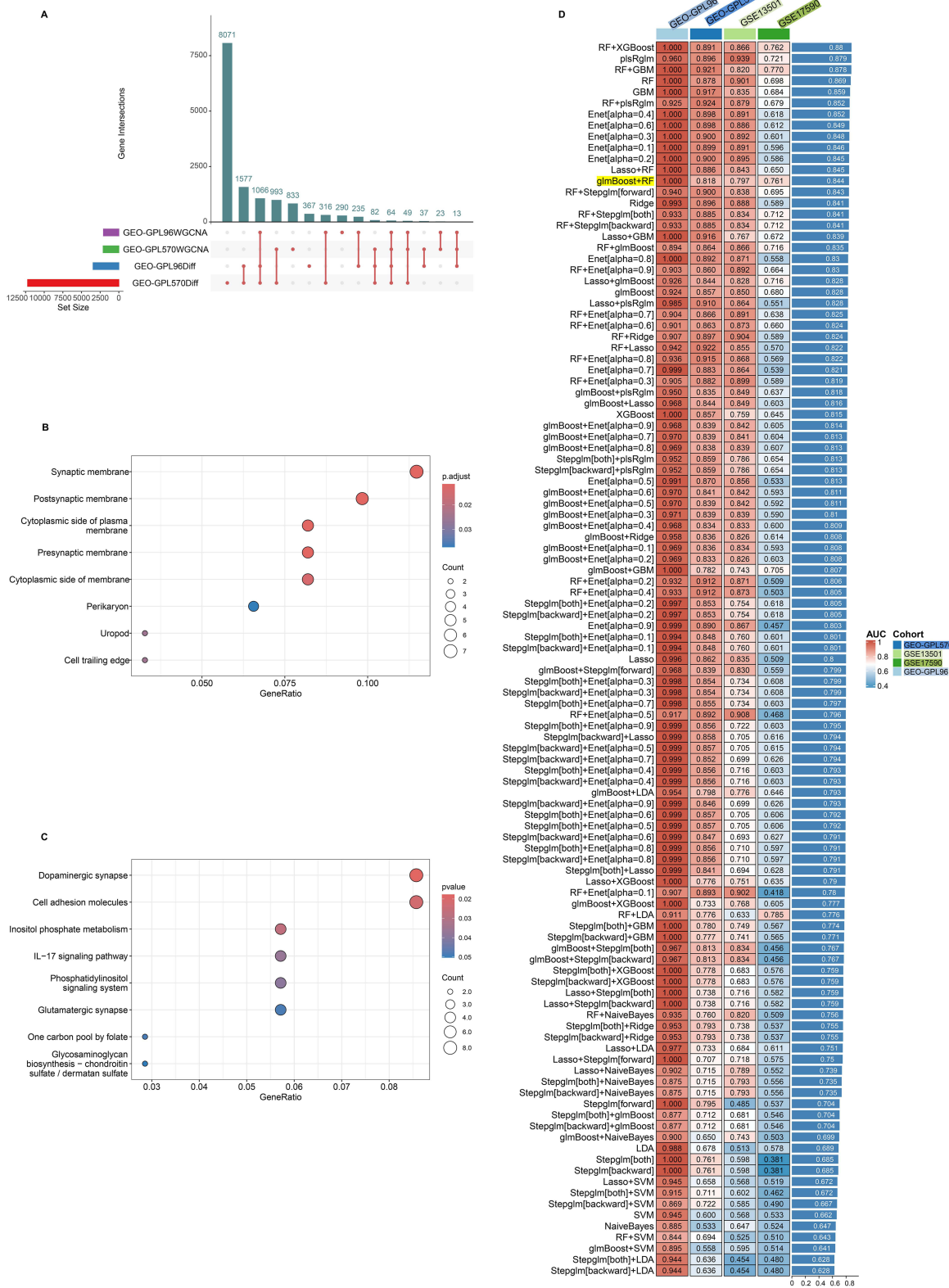


Figure 4 Identification of Th2/Th17-related genes, functional enrichment analysis and construction and testing of the artificial intelligence-derived diagnostic signature (Th2/Th17 classifier) for sJRA. (A) UpSet plot presents the intersection of the results of differentially expressed genes (DEGs) and significant Th2/Th17 module genes in two merged sJRA datasets (GEO-GPL96, and GEO-GPL570 with large sample size). (B and C) Functional enrichment analysis of Th2/Th17-related genes in sJRA. (B) Gene ontology (GO) analysis on biological process (BP), cellular component (CC), and molecular function (MF). (C) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. (D) The area under curve (AUC) of 111 machine-learning algorithm combinations in the training (GEO-GPL96) and testing (GEO-GPL570, GSE13501, and GSE17590) cohorts.

Diagnostic Value, and Clinical Usefulness of Th2/Th17 Classifier

ROC curves indicated that the Th2/Th17 classifier demonstrated exceptional diagnostic performance, achieving AUC values of 1.000 for the GEO-GPL96 cohort (Figure 5A), 0.818 in the GEO-GPL570 cohort (Figure 5B), 0.794 in the GSE13501 cohort (Figure 5C), 0.761 in the GSE17590 cohort (Figure 5D), 0.840 in the GSE112057 cohort (Figure 5E). PCA suggested that six hub Th2/Th17-related genes level successfully distinguished sJRA patients from healthy individuals in GEO-GPL96 (Figure 5F), GEO-GPL570 (Figure 5G), GSE13501 (Figure 5H), GSE17590 (Figure 5I), and GSE112057 cohorts (Figure 5J). Notably, the DCA chart suggested that across threshold probabilities ranging from 0 to 1, utilizing the Th2/Th17 classifier (prediction model) yielded superior net benefits relative to universal intervention or non-intervention strategies in GEO-GPL96 (Figure 5K), GEO-GPL570 (Figure 5L), GSE13501 (Figure 5M), GSE17590 (Figure 5N), and GSE112057 cohorts (Figure 5O).

Abundance Patterns and Interrelationships Among Hub Th2/Th17-Related Genes

In GEO-GPL96 (Figure S10A), GEO-GPL570 (Figure S10B), GSE13501 (Figure S10C), GSE17590 (Figure S10D), and GSE112057 cohorts (Figure S10E), distinct expression patterns emerged among six essential Th2/Th17-associated genes. qRT-PCR analysis confirmed that, in comparison to controls, the mRNA expression levels of HRH2, FTCD, and E2F1 were notably up-regulated in sJRA patients. In contrast, CD2 and MRPL22 levels were notably reduced in the sJRA group relative to the controls (Figure S10K-S10P). Furthermore, as shown in Figure S10F, HRH2 displayed a significant positive correlation with E2F1 ($R = 0.45$, $P < 0.05$), and FTCD ($R = 0.54$, $P < 0.05$); while demonstrating negative link to CD2 ($R = -0.46$, $P < 0.05$) and MRPL22 ($R = -0.46$, $P < 0.05$). Similarly, the other four sJRA cohorts validated the robust coexpression relationships among six hubs Th2/Th17-related genes (Figure S10G-S10J).

Examination of ICI in sJRA and Its Connection to Hub Th2/Th17-Related Genes

The immune profile composition variation between individuals with sJRA and healthy controls was assessed. The ESTIMATE algorithms revealed that immune scores, ESTIMATE scores, and stromal scores were markedly increased in the sJRA cohort relative to the healthy control cohort within the GEO-GPL96 dataset (Figure 6A-C). A similar trend was observed in the GEO-GPL570 dataset (Figure 6D-F). Furthermore, an inverse correlation was found between the expression levels of hub Th2/Th17-related genes and the frequency of Th2/Th17 cells in the two merged sJRA cohorts (Figure 6G-I), and vice versa. Lastly, we further explore relationships between hub Th2/Th17-related genes and ICI inferred through seven methods and revealed that elevated levels of hub Th2/Th17-related gene typically exhibited positive correlations with neutrophils, fibroblasts, erythrocytes, megakaryocytes, and basophils; however, they exhibited negative linked to M2 macrophages, osteoblast in two merged sJRA cohorts (Figures S11A and S11B), and vice versa.

Exploration of Possible Pathways in sJRA and Its Connection to Hub Th2/Th17-Related Genes

To investigate the potential biological pathways involved in sJRA, ssGSEA and GSVA were performed to evaluate the enrichment scores of pre-defined biological pathways, along with GO, KEGG, and HALLMARK gene sets. In relation to the pre-defined biological pathways across the two sJRA cohorts (Figure 7A and B), it was found that the sJRA group displayed distinct activation of pathways such as pyroptosis, ferroptosis, autophagy, complement and coagulation cascades, the Toll-Like Receptor pathway, NF-kappa B signaling, CCR, parainflammation, T cell co-inhibition, APC co-inhibition, HLA, cytolytic activity, checkpoint regulation, and IL6 JAK-STAT3 signaling, in contrast to the control group. Moreover, correlation analyses identified significant positive associations between the elevated expression of hub Th2/Th17-related genes and pathways, including IL6 JAK-STAT3 signaling, complement and coagulation cascades, pyroptosis, ferroptosis, parainflammation, and CCR. In contrast, these genes were negatively correlated with HLA expression and cytolytic activity (Figure 7C and D).

The differential analysis of GO/KEGG/HALLMARK gene sets identified the top 60 GO pathways, 25 KEGG pathways, and 15 hALLMARK pathways by intersecting data from the two sJRA cohorts (Figure 8 and S12). In contrast to the control cohort, several GO pathways were predominantly enriched in the sJRA group, including platelet activation,

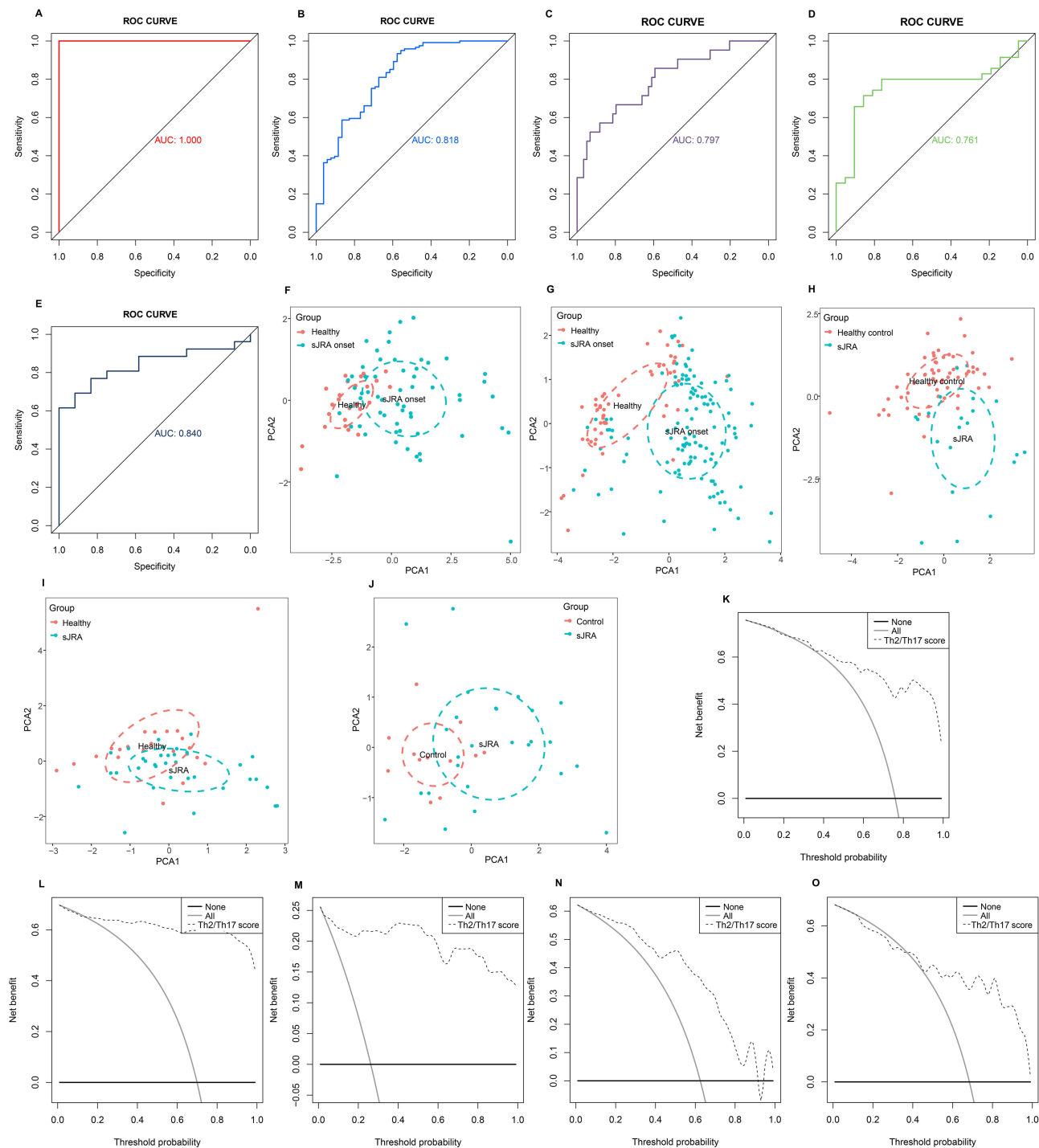


Figure 5 Evaluation of diagnostic value and clinical usefulness of Th2/Th17 classifier. (A–E) Receiver operating characteristic (ROC) curves with AUC values to evaluate predictive efficacy of Th2/Th17 classifier in GEO-GPL96 (A), GEO-GPL570 (B), GSE13501 (C), GSE17590 (D), GSE112057 (E) datasets. (F–J) Principal component analysis for the expression profiles of six hub Th2/Th17-related genes to distinguish sJRA patients from healthy control patients in GEO-GPL96 (F), GEO-GPL570 (G), GSE13501 (H), GSE17590 (I), GSE112057 (J) datasets. (K–O) Decision curve analysis was applied to evaluate the clinical usefulness of Th2/Th17 classifier in GEO-GPL96 (K), GEO-GPL570 (L), GSE13501 (M), GSE17590 (N), GSE112057 (O) datasets. The Y-axis represents the net benefit. The black line represents the hypothesis that no patients treatment. The X-axis represents the threshold probability. The threshold probability is where the expected benefit of treatment is equal to the expected benefit of avoiding treatment.

regulation of coagulation, macrophage activation, monocyte activation, suppression of extrinsic apoptotic signaling cascade, receptor-mediated endocytosis, response to Fe^{2+} , acute phase response, erythrocyte homeostasis, and neutrophil-mediated immunity. Similarly, KEGG pathways encompassing the TGF-beta signaling cascade, chemokine signaling

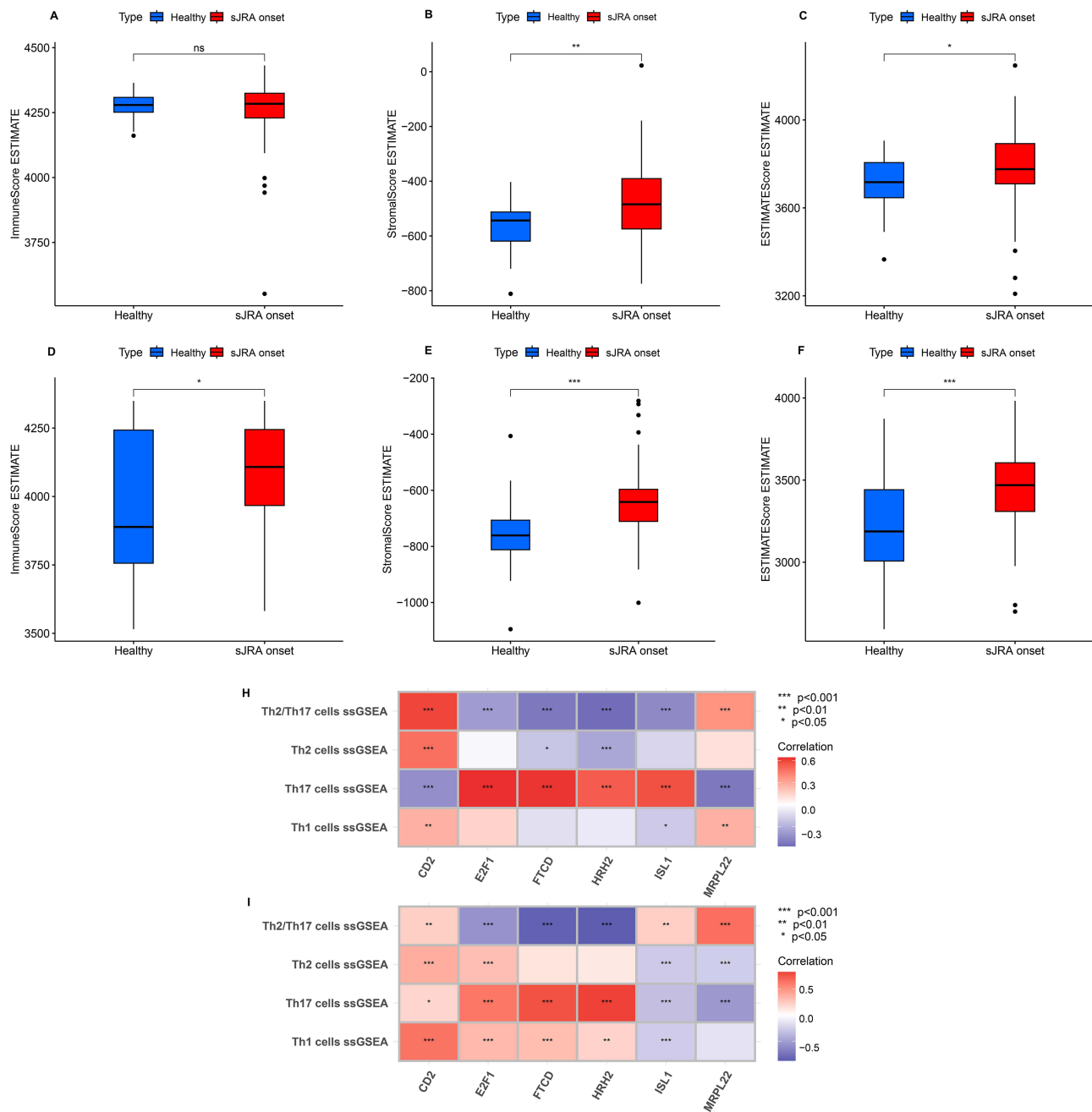


Figure 6 Analyses of immune cell infiltration in sjRA and its relationship with hub Th2/Th17-related genes. **(A-F)** Analyzing the microenvironment (equal to immune + stromal) between sjRA and healthy control groups in two merged sjRA datasets (GEO-GPL96, and GEO-GPL570 with large sample size) using ESITMATE algorithms. **(A)** immune score in GEO-GPL96 dataset. **(B)** stromal score in GEO-GPL96 dataset. **(C)** ESITMATE score in GEO-GPL96 dataset. **(D)** immune score in GEO-GPL570 dataset. **(E)** stromal score in GEO-GPL570 dataset. **(F)** ESITMATE score in GEO-GPL570 dataset. The asterisks indicate a significant statistical *p* value calculated using the Wilcoxon test (**p* < 0.05; ***p* < 0.01; ****p* < 0.001). **(H and I)** Correlation of six hub Th2/Th17-related genes with Th1, Th2, Th17 and Th2/Th17 cells infiltration using ssGSEA in GEO-GPL96 **(H)**, and GEO-GPL570 **(I)** datasets. Correlation coefficient and *p* value were calculated by Spearman correlation analysis. The asterisks indicate a statistically significant *p* value calculated using the Spearman correlation analysis (**p* < 0.05; ***p* < 0.01; ****p* < 0.001).

cascade, complement and coagulation cascades, Fc gamma R-mediated phagocytosis, Nod-like receptor signaling cascade, MAPK signaling cascade, and ErbB signaling cascade were markedly upregulated in sjRA. In contrast, HALLMARK pathways associated with complement, heme metabolism, IL2-STAT5 signaling, TNF- α /NF-kappa B signaling, interferon gamma response, and PI3K-AKT-mTOR signaling were primarily involved in sjRA. On the other hand, the control group exhibited enrichment in antigen processing and presentation, upregulation of T cell cytotoxic responses, butanoate metabolism, oxidative phosphorylation, and MYC targets v1. Furthermore, statistical

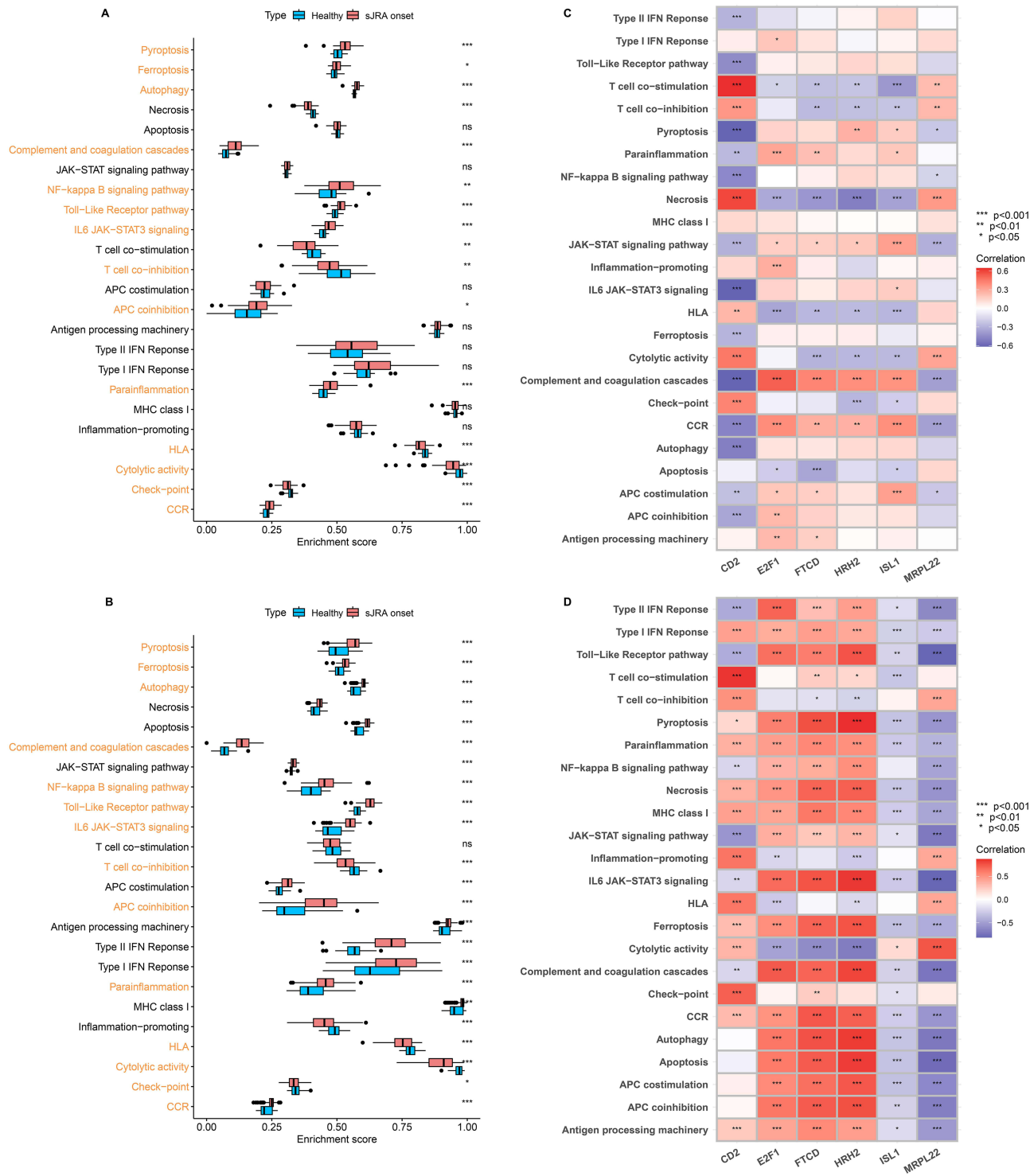


Figure 7 Analyses of several representative gene sets (immune/inflammatory/cell death-related pathways) in sJRA and its relationship with hub Th2/Th17-related genes. **(A and B)** Comparison of pre-defined biological processes between sJRA and control groups in GEO-GPL96 **(A)** and GEO-GPL570 **(B)** datasets based on ssGSEA algorithm. *p* value was calculated using the Wilcoxon test. **(C and D)** The heatmap plot depicted correlation between seven hub Th2/Th17-related genes and pre-defined biological biological pathways in GEO-GPL96 **(C)** and GEO-GPL570 **(D)** datasets. Correlation coefficient and *p* value were calculated by Spearman correlation analysis. The asterisks indicate a statistically significant *p* value calculated using the Spearman correlation analysis (**p* < 0.05; ***p* < 0.01; ****p* < 0.001).

correlation assessment demonstrated that abundantly expressed central Th2/Th17-associated genes exhibited predominantly positive associations with enhanced signaling cascades in sJRA, while displaying inverse relationships with suppressed pathways, and vice versa **(Figure S13)**.

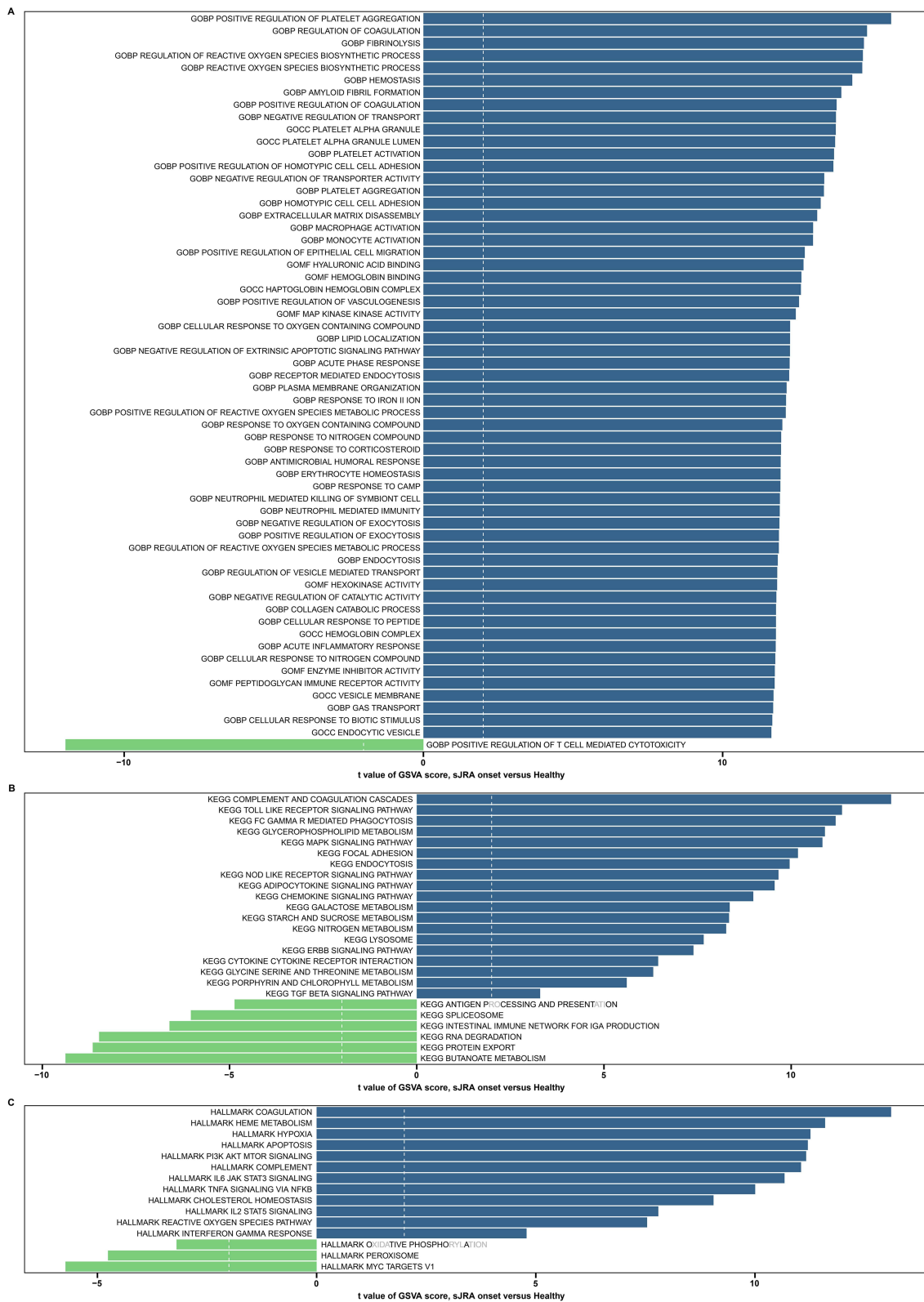


Figure 8 Comparison of GO/KEGG/HALLMARK gene sets from Molecular Signatures Database between sJRA and control groups in GEO-GPL96 cohort based on gene set variation analysis (GSVA). **(A)** GO gene sets. **(B)** KEGG gene sets. **(C)** HALLMARK gene sets. t value and p value was calculated using the limma R package.

XGBoost and SHAP

In the effort to identify genes with diagnostic significance for sJRA, SHAP and XGBoost were employed to assess the importance of features. The summary plot (Figure 9A and C) ranks these features based on their impact, with HRH2 emerging as a key determinant in both the GEO-GPL96 and GEO-GPL570 cohorts. The influence of each feature on the model's outputs is illustrated in Figure 9B and D. Notably, HRH2 emerged as the primary indicator for sJRA identification. Moreover, the Mantel analysis revealed that HRH2 exhibited positive associations with coagulation, heme metabolism, acute phase response, erythrocyte homeostasis, neutrophil mediated immunity, platelet aggregation, MAP kinase activity, and response to iron II ion and so on in two sJRA cohorts (Figure 9E and S14).

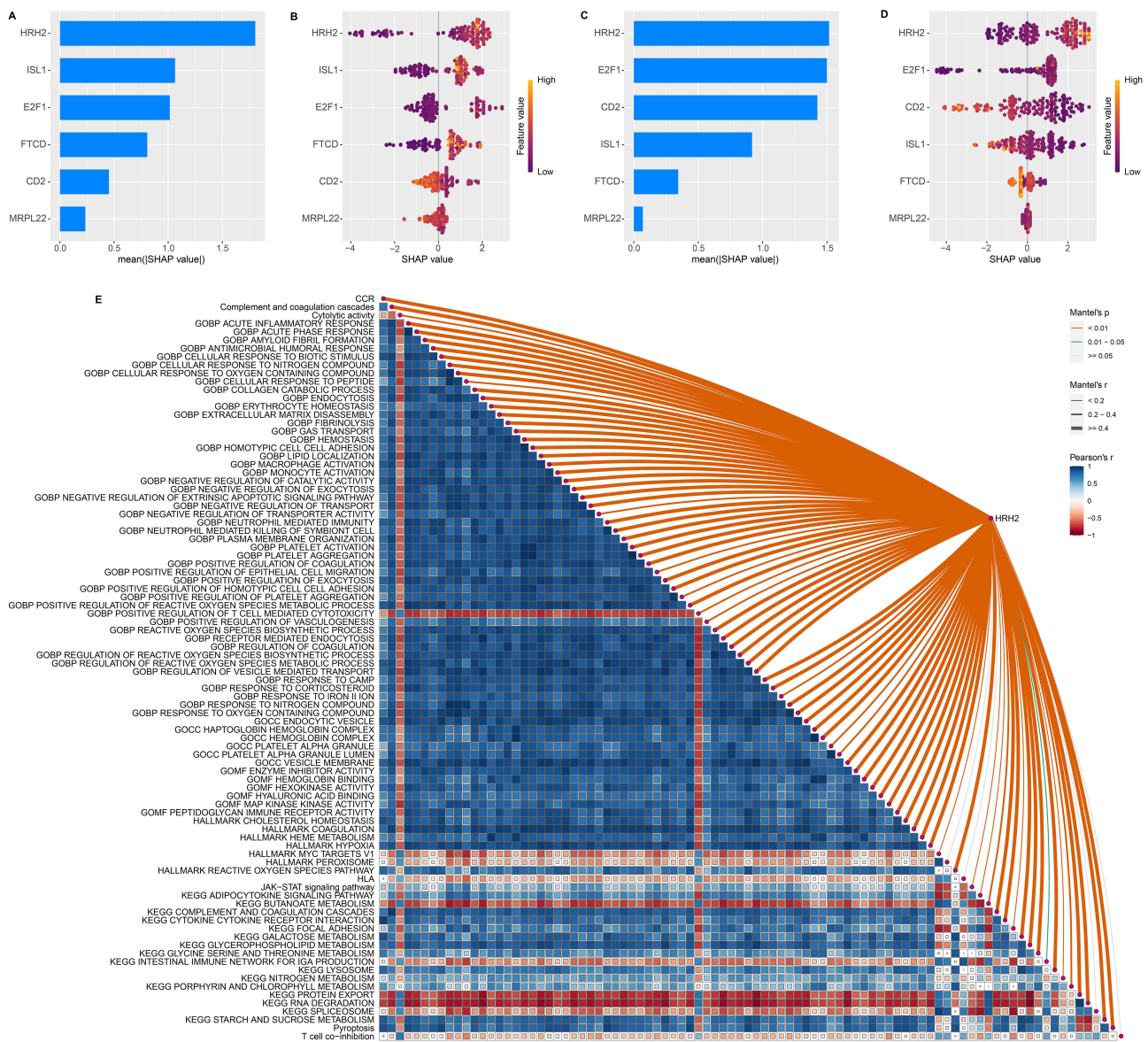


Figure 9 Identification of optimal Th2/Th17-related gene for sJRA using XGBoost algorithm and SHAP, and investigation the relationship between optimal Th2/Th17-related gene and pivotal molecular pathways using Mantel test. (A and C) Importance ranking of features in GEO-GPL96 (A) and GEO-GPL570 (C) datasets. (B and D) Visualization of SHAP variables, with the included features sorted by the average absolute value of SHAP from highest to lowest in GEO-GPL96 (B) and GEO-GPL570 (D) datasets. Yellow dots denoting a higher impact (increasing sJRA) and purple dots a lower one (reducing sJRA). (E) In GEO-GPL570 cohort, correlation between HRH2 and significant biological pathways based on Spearman correlation analysis. Correlation coefficient and *p* value were calculated by the Mantel test.

Discussion

At present, the absence of efficient biological indicators for preliminary diagnosis and specific therapeutics remains a significant challenge in sJRA. The current investigation employed large-scale GWAS summary data (comprising 731 IC subtypes, along with two sJRA GWAS datasets from East Asian and European populations) and seven sJRA transcriptome datasets, aiming to elucidate the crucial involvement of Th2/Th17 lymphocyte imbalance in sJRA pathogenesis. Subsequently, WGCNA, differential expression analysis, and 111 combinations of 12 ML algorithms were applied to develop and confirm an AI-based diagnostic model (Th2/Th17 classifier), using criteria such as average AUC (0.844), minimal variability, and model stability across four cohorts. Validation by qRT-PCR demonstrated that, relative to controls, HRH2, FTCD, and E2F1 were notably up-regulated in sJRA patients, while CD2 and MRPL22 expression levels were markedly decreased. Additionally, ICI and FEA analyses revealed significant associations between hub Th2/Th17-related genes and various IC types (eg, megakaryocytes, M2 macrophages, erythrocytes, fibroblasts, basophils, and neutrophils) as well as enrichment scores related to pathways such as pyroptosis, ferroptosis, complement and coagulation cascades, Fc gamma R-mediated phagocytosis, T cell-mediated cytotoxicity, and ErbB signaling, all of which may contribute to the onset of sJRA. The XGBoost algorithm, supported by SHAP analysis, identified HRH2 as a key gene, suggesting its potential as an important target for sJRA.

A growing body of evidence has indicated a relationship between IC abundance and sJRA,^{6,21,22} however, direct evidence establishing a causal link remains limited. Given the heterogeneous nature of sJRA, characterized by distinct pathophysiological profiles and individual variations, MR analysis was conducted to investigate the causal link between 731 IC subtypes and two sJRA GWAS datasets (GWAS ID: ebi-a-GCST90018653, and ebi-a-GCST90018873). The genetic analysis revealed uncovered multiple possible causative links between two immunophenotypes (lymphocyte % leukocyte and IgD+ B cell Absolute Count) and sJRA. The IVW method confirmed results consistent with those derived from MR-Egger regression, WM, the simple model, the weighted model, and both East Asian and European sJRA GWAS datasets. Lymphocytes are known to consist of B lymphocytes, T lymphocytes, and NK cells. Studies have discovered that PBMCs derived from oligoarticular and polyarticular JRAs had noticeable alterations in the frequency and phenotype of circulating B cells when contrasted with healthy controls.²³ In addition, B cellular plasmablast production was triggered by naive T helper cells from sJIA patients *in vitro*, and self-antigen array data showed that the IgG reactivity profiles of sJIA patients were different from those of healthy controls.²⁴ These findings corroborate our study, reinforcing the significant function of B cells in sJRA. A case-control study found that compared to controls, total T lymphocytes, CD8+CD28^{null} cells and CD4+CD28^{null} cells were markedly elevated in active JIA patients.²⁵ Two recent reports have shown that programmed cell death 1 (PD-1) population cloning of CD4 + T cells is amplified in the synovial fluid of JIA patients, expressing genes that drive inflammation and B-cell activation.^{26,27} In this investigation, a host-genetic-driven elevate in lymphocyte %leukocyte and IgD+ B cell absolute count was found to notably elevate sJRA risk. These findings emphasized the crucial significance of disproportionate lymphocyte (B and T cells) distributions in sJRA development. To thoroughly investigate and confirm the associations among B cells, T cells, and sJRA, we utilized eight distinct computational methods (CIBERSORT, ssGSEA, TIMER, xCell, MCPcounter, ESTIMATE, EPIC, and Quantiseq) to assess sJRA patients' ICI at the transcriptomic level. Consequently, our findings suggested that lymphocyte Th17 cells (proinflammatory) were generally more abundant in sJRA, while lymphocyte Th2 cells (anti-inflammatory) levels decreased. More importantly, the Th2/Th17 proportion of T lymphocytes was considerably changed, whereas B cells demonstrated no notable differences grounded in six distinct computational methods and seven transcriptome data. These suggested that lymphocyte Th2/Th17 ratio imbalance represents dysregulation of immune inflammation in sJRA. Tomé et al²³ collected blood samples from 105 JIA patients (children and adults) and 50 age-matched healthy people, and found higher levels of Th1 and Th 17 cells in circulation when compared with controls and had notably diminished frequencies of Th2 cells. The underlying mechanisms of Th2/Th17 imbalance in lymphocyte and sJRA pathogenesis are explained below: 1 Th cells (Th1, Th2, and Th17) are enriched in the joints and the synovial fluid of JIA patients, exhibiting numerous B cell assistance-related indicators and demonstrating the ability to stimulate B cell maturation and immunoglobulin synthesis.²⁸ 2 Circulating Th cells differentiate into specific phenotypic groups, characterized by CXCR3 and CCR6 expression (Th1-like, Th2-like and Th17-like Th cells). The

proportion disparity between Th2 and Th17 lymphocytes possesses the capability to stimulate B cells and enhance immunoglobulin synthesis.²⁹ However, the potential effects of Th2 and Th2/Th17 in sJRA have not been clarified. We found a new direction in which T lymphocyte Th2/Th17 imbalance may mediate sJRA, which is worthy of further exploration.

Considering the results from MR and ICI analyses at the transcriptome level, a comprehensive understanding of the potential roles of Th2/Th17-associated molecules is of great importance for enhancing early sJRA detection, improving the understanding of its pathogenesis, and developing potential therapeutic agents for sJRA patients. The diagnosis of sJRA remains a challenging issue, as specific immunological markers associated with other systemic autoimmune diseases are absent, and typical signs of active arthritis usually emerge later in the disease progression. These can easily lead to missed diagnosis or misdiagnosis early. SJA is mainly a diagnosis made by exclusion. Several immune markers (eg, MRP8/14, IL-18, IL-1, IL-6, and S100A12) are known to correlate with sJRA; however, consensus regarding their diagnostic application is lacking, and their insufficient specificity limits their utility in clinical diagnosis.³⁰ The complexity and diverse manifestations of sJRA's pathophysiological mechanisms contribute notably to this challenge. On a pathophysiological level, numerous genes exhibit differential expression patterns throughout sJRA onset and advancement, with specific genetic markers potentially serving as crucial indicators for diagnostic and prognostic purposes. Our investigation conducted a comprehensive analysis of existing sJRA bulk RNA-seq data (*Homo sapiens*) from the earliest records to current findings, emphasizing the significance of Th2/Th17-associated components in sJRA. Consequently, we utilized WGCNA and differential analysis across two large, merged cohorts to identify 64 Th2/Th17-related genes as potential biomarkers for sJRA. Additionally, previous studies have indicated that individuals often select modeling algorithms based on their preferences and expertise. We compiled a set of 12 machine learning (ML) methods that could be utilized to create diagnostic biomarkers, thus addressing the current limitations in the field. These methods were further integrated into 111 algorithm combinations, with variable selection and data dimensionality reduction performed using LASSO, RF, StepAIC, and glmBoost. The final model was identified as the combination of RF and glmBoost, which produced the highest average AUC (0.844), with reduced variability and increased model stability across the four sJRA cohorts. One of the main challenges in developing biomedical models utilizing artificial intelligence (AI) and ML is overfitting, which emerges when a model fits the training data too well but fails to generalize to external validation datasets. A final six-gene signature, the Th2/Th17 classifier, was generated using RF and glmBoost to minimize redundancy in the data. This classifier demonstrated exceptional diagnostic performance, achieving an AUC of 1.000 in the training cohort and AUCs of 0.818, 0.794, 0.761, and 0.840 in the test cohorts (GEO-GPL570, GSE13501, GSE17590, and GSE112057, respectively). Furthermore, it effectively distinguished between sJRA patients and healthy controls. According to the DCA results, treatment decisions guided by the Th2/Th17 classifier provided greater net benefits than strategies that either treated all patients or none. These findings suggest that the Th2/Th17 classifier exhibits value as a diagnostic tool for sJRA and could aid in clinical decision-making.

Accumulating evidence has demonstrated that ICs in sJRA patients serve a pivotal function in driving joint inflammation, bone and cartilage degradation, and overall joint damage by modulating the synovial fluid microenvironment.^{31,32} In this study, it was revealed that, in comparison to healthy controls, the scores for immune, ESTIMATE, and stromal microenvironments were markedly elevated in sJRA patients across two distinct cohorts. Furthermore, correlation analysis between hub genes associated with Th2/Th17 and infiltrating ICs suggested that neutrophils, fibroblasts, erythrocytes, megakaryocytes, and basophils might contribute to the development of sJRA. Conversely, the Th2/Th17 cell ratio, M2 macrophages, and osteoblasts appear to act as inhibitors of sJRA progression, based on seven ICI methods applied to two merged sJRA cohorts. In orthopedic surgery specimens, JRA synovial fibroblasts produced more IL-2, IL-6, and other CCL- and CXCL-chemokines than did healthy synovial fibroblasts.^{32,33} Compared with other immune-mediated disease such as rheumatoid arthritis, synovial fibroblasts and their role have received limited attention in JRA. In a retrospective cohort study, the author recruited 358 JRA patients and found that in patients with JIA, a higher erythrocyte sedimentation rate seems to be a predictor of the development of uveitis,^{34,35} which was consistent with our study. Do et al³⁶ found that in sJRA patients, up-regulation of miR-181c reduced the anti-inflammatory responses of CD163+ macrophages (analogous to M2 macrophages) to hemoglobin complexes or high mobility group box 1. According to the study, patients with JRA exhibited lower osteoblastogenesis from synovial fluid-

derived progenitors than healthy individuals. Patients with severe forms of JRA also showed decreased osteoblast differentiation, which was associated with altered chemokine/cytokine expression patterns, which was in line with our study.³⁷ However, the impact of an elevation in megakaryocytes, and basophils has not been studied in sJRA, fibroblasts and M2 macrophage infiltration in sJRA remain relatively understudied. Therefore, additional investigation of these cellular components will offer a fresh viewpoint regarding the advancement and evolution of sJRA, while contributing to the identification of prospective treatments for individuals affected by sJRA.

Afterward, ssGSEA and GSVA functional enrichment analyses were conducted to clarify the underlying biological mechanisms associated with sJRA across two transcriptomic datasets. The analyses revealed the involvement of immune-related pathways, such as T cell co-inhibition, APC co-inhibition, HLA, platelet activation, coagulation regulation, macrophage and monocyte activation, negative regulation of the extrinsic apoptotic signaling pathway, and neutrophil-mediated immunity. Furthermore, inflammatory-related pathways, including CCR, parainflammation, TNF- α /NF-kappa B signaling, PI3K AKT mTOR signaling, interferon gamma response, MAPK signaling pathway, and ErbB signaling pathway, were markedly implicated. Additionally, pathways related to cell death, such as pyroptosis, ferroptosis, and autophagy, were predominantly associated with sJRA development, while antigen processing and presentation, T cell-mediated cytotoxicity, butanoate metabolism, oxidative phosphorylation, and MYC targets v1 exhibited notable down-regulation in sJRA. These observations suggest that sJRA is marked by heightened immune and inflammatory responses, an increase in cell death, and a reduction in oxidative phosphorylation. Complement and coagulation cascades, along with the response to iron II ions and Fc gamma R-mediated phagocytosis, were also found to be markedly involved in sJRA progression. Current research into the mechanisms of sJRA predominantly focuses on immune and inflammatory signaling pathways. In contrast to random events, programmed cell death (PCD), such as ferroptosis and pyroptosis, often referred to as regulated cell death, is governed by a series of orchestrated processes. Excess iron, a contributor to the formation of reactive oxygen species, which is a hallmark of ferroptosis, further exacerbates this process. Iron metabolism serves a critical function in determining cellular sensitivity to ferroptosis by modulating the labile iron pool. Ferroptosis has been recognized as a significant factor in several inflammatory diseases, including rheumatoid arthritis; however, its specific role in sJRA remains poorly understood.³⁸ Our results suggest a strong association between sJRA and iron II ions, as well as ferroptosis. Pyroptosis, a form of programmed cell death dependent on Caspase activity and inflammasome activation, is marked by the release of multiple inflammatory cytokines. Elevated expression of inflammasome-activated cytokines IL-1 β and IL-18 was observed when serum from sJRA patients was incubated with healthy donor PBMCs,³⁹ indicating a potential role for inflammasome dysfunction in sJRA, although critical pyroptosis-associated proteins remain lacking. These findings are consistent with these results, validating the possible involvement of pyroptosis in sJRA. Notably, the precise mechanisms through which ferroptosis, pyroptosis, Fc gamma R-mediated phagocytosis, complement and coagulation cascades, and various immune/inflammatory-related pathways, such as T cell-mediated cytotoxicity, IL2 STAT5 signaling, PI3K AKT mTOR signaling, and ErbB signaling pathway, contribute to the onset of sJRA are still unclear. Further exploration in this area is essential to uncover novel therapeutic strategies for sJRA.

A key discovery in this study was the identification of six hubs Th2/Th17-related genes (HRH2, ISL1, MRPL22, FTCD, CD2, and E2F1), which could function as promising therapeutic targets in sJRA treatment. In two independent sJRA cohorts, these six hub genes demonstrated markedly differential expression. qRT-PCR analysis confirmed that, in comparison to the control group, the mRNA levels of HRH2, FTCD, and E2F1 were markedly upregulated in sJRA patients, whereas the expression of MRPL22 and CD2 was notably reduced in the sJRA cohort relative to controls. Additionally, the correlation patterns among these hub genes suggested possible intermolecular interactions (HRH2 with E2F1, HRH2 with FTCD, HRH2 with CD2, HRH2 with MRPL22), which may be implicated in the pathogenesis of sJRA. Notably, to identify the most promising diagnostic genes for sJRA, the XGBoost algorithm and SHAP were applied to visualize the importance of these feature variables. In both sJRA cohorts, HRH2 was identified as a key determinant. HRH2 may contribute to the heightened occurrence of sJRA, potentially mediating its onset and acting as a novel therapeutic target. Further investigation into the molecular role of HRH2 in sJRA revealed through Spearman correlation analysis and Mantel tests that HRH2 may influence the progression of sJRA by activating pathways related to coagulation, heme metabolism, acute phase response, erythrocyte homeostasis, neutrophil-mediated immunity, platelet

aggregation, MAP kinase activity, and response to iron II ions. Nevertheless, the possible impacts of HRH2 on sJRA progression remain unexplored and necessitate additional examination through in vitro and in vivo experiments. Additionally, the clinical significance of HRH2 as a potential target can be explored in more depth and analyzed in combination with current therapeutic strategies.

Our study presents notable variations from previously published research in several key aspects. (1) By utilizing large-scale GWAS summary data (comprising 731 IC subtypes and two sJRA GWAS datasets from East Asian and European populations), we conducted a MR analysis between immune subtypes and sJRA, establishing a robust causal relationship between lymphocyte %leukocytes, IgD+ B cell absolute count, and sJRA. (2) At the transcriptomic level, we employed eight distinct algorithms (CIBERSORT, ssGSEA, TIMER, xCell, MCPcounter, ESTIMATE, EPIC, and Quantiseq) to infer the ICI of sJRA patients from seven transcriptome datasets, revealing the critical role of the imbalance in Th2/Th17 lymphocyte proportions in sJRA development. (3) We systematically compiled sJRA transcriptome data and constructed a Th2/Th17 classifier using ML algorithms, selecting those with the highest average AUC, minimal variability, and superior model generalizability across all cohorts. This approach markedly enhanced the model's stability and diagnostic accuracy. (4) A total of 12 ML algorithms were integrated into 111 distinct combinations, with the most accurate model selected to avoid subjective biases in choosing modeling methods. (5) Our study identified key factors involved in sJRA, including megakaryocytes, basophils, M2 macrophages, pyroptosis, ferroptosis, T cell-mediated cytotoxicity, complement and coagulation cascades, Fc gamma R-mediated phagocytosis, IL2-STAT5 signaling, PI3K-AKT-mTOR signaling, and the ErbB signaling pathway, providing fresh perspectives on the molecular mechanisms that drive sJRA development. (6) The XGBoost algorithm and SHAP analysis identified HRH2 as an optimal gene, which is anticipated to serve as a new potential target for sJRA therapy. While our study was as thorough and meticulous as possible, certain limitations should be acknowledged. First, in the MR analysis, our findings did not survive the stringent Bonferroni correction for multiple comparisons. However, despite the Bonferroni-adjusted *P*-values, the MR study, supported by some biological evidence, was employed as a hypothesis-driven approach to examine epidemiologically proven relationships. It is also worth noting that our results were validated across two sJRA GWAS datasets and seven sJRA transcriptome datasets. Two, even though our model performed remarkably well in predicting, the performance of the Th2/Th17 classifier was only analyzed in seven European population cohorts, which limits the generality of the models to other races. It is not yet appropriate for widespread application until external datasets with larger sample sizes in different ethnic populations have been validated. Three, we found six hub Th2/Th17-related genes implicated in sJRA progression. However, subsequent in vivo and in vitro validation should be executed to examine the underlying molecular pathways, particularly regarding HRH2. Finally, the intricate molecular mechanisms, both in vivo and in vitro, require further investigation in future studies to identify the involvement of megakaryocytes, basophils, M2 macrophages, pyroptosis, ferroptosis, T cell-mediated cytotoxicity, complement and coagulation cascades, Fc gamma R-mediated phagocytosis, IL2-STAT5 signaling, PI3K-AKT-mTOR signaling, and the ErbB signaling pathway. These efforts may pave the way for new therapeutic strategies in the treatment of sJRA.

Conclusion

This investigation constitutes the initial effort to integrate MR and transcriptomic analyses to explore the causal relationships between Th2/Th17 lymphocyte populations and sJRA. A sum of 111 combinations of 12 ML algorithms was employed to develop and validate a consensus diagnostic signature (Th2/Th17 classifier) for sJRA based on six Th2/Th17-associated genes. Additionally, the identification of key cell types, including megakaryocytes, M2 macrophages, erythrocytes, fibroblasts, and basophils, as well as the enrichment scores for pyroptosis, ferroptosis, Fc gamma R-mediated phagocytosis, complement and coagulation cascades, and T cell-mediated cytotoxicity, may provide critical therapeutic targets for sJRA, offering valuable insights into the disease's underlying pathogenesis.

Data Sharing Statement

The raw data utilized in this research can be found within the manuscript/Supplementary files. Additional questions may be addressed to the lead correspondent (Yongsheng Xu).

Ethics Statement

The Inner Mongolia Autonomous Region People's Hospital's Research Ethics Committee examined and approved the experiments involving human blood samples, ensuring that they adhered to the Declaration of Helsinki. Before specimen acquisition, formal written authorization was secured from family members of qualified participants.

Acknowledgments

We thank Clinical Medical Research Center of Inner Mongolian People's Hospital for the support of this study. We thank Bullet Edits Limited for the linguistic editing and proofreading of the manuscript.

Funding

This work was supported by Autonomous Region Health Science and Technology Plan (No.202202034), Science and Technology Program of the Joint Fund for Scientific Research in Public Hospitals (No.2023GLLH0087) and Key Laboratory of Affiliated Hospital of Inner Mongolia Medical University Open Fund (No.2022NYFYSYS010).

Disclosure

The researchers affirm that no conflicts of interest exist in this investigation.

References

1. Martini A, Lovell DJ, Albani S, et al. Juvenile idiopathic arthritis. *Nat Rev Dis Primers*. 2022;8(1):5. doi:10.1038/s41572-021-00332-8
2. Lee JY, Schneider R. Systemic Juvenile Idiopathic Arthritis. *Pediatr Clin North Am*. 2018;65(4):691–709. doi:10.1016/j.pcl.2018.04.005
3. Al-Mayouf SM, Al Mutairi M, Bouayed K, et al. Epidemiology and demographics of juvenile idiopathic arthritis in Africa and Middle East. *Pediatr Rheumatol Online J*. 2021;19(1):166. doi:10.1186/s12969-021-00650-x
4. Liu W, Wen D, Liu Z, Wang K, Wang J. Erythropoiesis signature and ubiquitin-mediated proteolysis are enriched in systematic juvenile idiopathic arthritis. *Int J Immunogenet*. 2022;49(3):193–201. doi:10.1111/iji.12573
5. Martini A, Ravelli A, Avcin T, et al. Toward new classification criteria for juvenile idiopathic arthritis: first steps, pediatric rheumatology international trials organization international consensus. *J Rheumatol*. 2019;46(2):190–197. doi:10.3899/jrheum.180168
6. Mazzoni A, Annunziato F, Maggi L. T lymphocytes-related cell network in the pathogenesis of juvenile idiopathic arthritis: a key point for personalized treatment. *Curr Opin Rheumatol*. 2024;36(1):40–45. doi:10.1097/BOR.0000000000000991
7. Omoyinmi E, Hamaoui R, Pesenacker A, et al. Th1 and Th17 cell subpopulations are enriched in the peripheral blood of patients with systemic juvenile idiopathic arthritis. *Rheumatology*. 2012;51(10):1881–1886. doi:10.1093/rheumatology/kes162
8. Henderson LA, Hoyt KJ, Lee PY, et al. Th17 reprogramming of T cells in systemic juvenile idiopathic arthritis. *JCI Insight*. 2020;5(6):e132508. doi:10.1172/jci.insight.132508
9. Kim JY, Song M, Kim MS, et al. An atlas of associations between 14 micronutrients and 22 cancer outcomes: Mendelian randomization analyses. *BMC Med*. 2023;21(1):316. doi:10.1186/s12916-023-03018-y
10. Chen Z, Zeng L, Liu G, et al. Construction of autophagy-related gene classifier for early diagnosis, prognosis and predicting immune micro-environment features in sepsis by machine learning algorithms. *J Inflamm Res*. 2022;15:6165–6186. doi:10.2147/JIR.S386714
11. Chen Z, Chen R, Ou Y, et al. Construction of an HLA classifier for early diagnosis, prognosis, and recognition of immunosuppression in sepsis by multiple transcriptome datasets. *Front Physiol*. 2022;13:870657. doi:10.3389/fphys.2022.870657
12. Orrù V, Steri M, Sidore C, et al. Complex genetic signatures in immune cells underlie autoimmunity and inform therapy [published correction appears in nat genet. *Nat Genet*. 2020;52(10):1036–1045. doi:10.1038/s41588-020-0684-4
13. 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
14. Cui J, Chen Y, Ou Y, et al. Cancer germline antigen gene MAGEB2 promotes cell invasion and correlates with immune microenvironment and immunotherapeutic efficiency in laryngeal cancer. *Clin Immunol*. 2022;240:109045. doi:10.1016/j.clim.2022.109045
15. Zeng D, Ye Z, Shen R, et al. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front Immunol*. 2021;12:687975. doi:10.3389/fimmu.2021.687975
16. Hu J, Wu Y, Zhang D, et al. Regulatory T cells-related gene in primary sclerosing cholangitis: evidence from Mendelian randomization and transcriptome data. *Genes Immun*. 2024;2024:1–22. doi:10.1038/s41435-024-00304-4
17. Zhu Y, Chen B, Zu Y. Identifying OGN as a biomarker covering multiple pathogenic pathways for diagnosing heart failure: from machine learning to mechanism interpretation. *Biomolecules*. 2024;14(2):179. doi:10.3390/biom14020179
18. Hou Y, Chen Z, Wang L, et al. Characterization of immune-related genes and immune infiltration features in epilepsy by multi-transcriptome data. *J Inflamm Res*. 2022;15:2855–2876. doi:10.2147/JIR.S360743
19. Chen Z, Dong X, Liu G, et al. Comprehensive characterization of costimulatory molecule gene for diagnosis, prognosis and recognition of immune microenvironment features in sepsis. *Clin Immunol*. 2022;245:109179. doi:10.1016/j.clim.2022.109179
20. Zhong Y, He J, Luo J, et al. A machine learning algorithm-based model for predicting the risk of non-suicidal self-injury among adolescents in western China: a multicentre cross-sectional study. *J Affect Disord*. 2024;345:369–377. doi:10.1016/j.jad.2023.10.110
21. Zhang W, Cai Z, Liang D, et al. Immune cell-related genes in juvenile idiopathic arthritis identified using transcriptomic and single-cell sequencing data. *Int J Mol Sci*. 2023;24(13):10619. doi:10.3390/ijms241310619

22. Parackova Z, Zentsova I, Bloomfield M, et al. Expanded population of low-density neutrophils in juvenile idiopathic arthritis. *Front Immunol.* 2023;14:1229520. doi:10.3389/fimmu.2023.1229520
23. Tomé C, Oliveira-Ramos F, Campanilho-Marques R, et al. Children with extended oligoarticular and polyarticular juvenile idiopathic arthritis have alterations in B and T follicular cell subsets in peripheral blood and a cytokine profile sustaining B cell activation. *RMD Open.* 2023;9(3):e002901. doi:10.1136/rmdopen-2022-002901
24. Kuehn J, Schleifenbaum S, Hending M, et al. Aberrant naive CD4-positive T cell differentiation in systemic juvenile idiopathic arthritis committed to B cell help. *Arthritis Rheumatol.* 2023;75(5):826–841. doi:10.1002/art.42409
25. Zahran AM, Abdallah AM, Saad K, et al. Peripheral blood B and T cell profiles in children with active juvenile idiopathic arthritis. *Arch Immunol Ther Exp.* 2019;67(6):427–432. doi:10.1007/s00005-019-00560-7
26. Maschmeyer P, Heinz GA, Skopnik CM, et al. Antigen-driven PD-1+ TOX+ BHLHE40+ and PD-1+ TOX+ EOMES+ T lymphocytes regulate juvenile idiopathic arthritis in situ. *Eur J Immunol.* 2021;51(4):915–929. doi:10.1002/eji.202048797
27. Vanni A, Mazzoni A, Semeraro R, et al. Clonally expanded PD-1-expressing T cells are enriched in synovial fluid of juvenile idiopathic arthritis patients. *Eur J Immunol.* 2023;53(7):e2250162. doi:10.1002/eji.202250162
28. Julé AM, Lam KP, Taylor M, et al. Disordered T cell-B cell interactions in autoantibody-positive inflammatory arthritis. *Front Immunol.* 2023;13:1068399. doi:10.3389/fimmu.2022.1068399
29. Morita R, Schmitt N, Bentebibel SE, et al. Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion [published correction appears in *Immunity.* 2011;34(1):108–121. doi:10.1016/j.immuni.2010.12.012
30. Park C, Miranda-Garcia M, Berendes R, et al. MRP8/14 serum levels as diagnostic markers for systemic juvenile idiopathic arthritis in children with prolonged fever. *Rheumatology.* 2022;61(7):3082–3092. doi:10.1093/rheumatology/keab729
31. Simonds MM, Schlefman AR, McCahan SM, Sullivan KE, Rose CD, Brescia AMC. The culture microenvironment of juvenile idiopathic arthritis synovial fibroblasts is favorable for endochondral bone formation through BMP4 and repressed by chondrocytes. *Pediatr Rheumatol Online J.* 2021;19(1):72. doi:10.1186/s12969-021-00556-8
32. Liu M, Gong Y, Lin M, Ma Q. Comprehensive analysis of juvenile idiopathic arthritis patients' immune characteristics based on bulk and single-cell sequencing data. *Front Mol Biosci.* 2024;11:1359235. doi:10.3389/fmolb.2024.1359235
33. Lomholt S, Pedersen MJ, Glerup M, Kragstrup TW. Synovial fibroblasts in juvenile idiopathic arthritis: a scoping review. *Semin Arthritis Rheum.* 2023;58:152159. doi:10.1016/j.semarthrit.2022.152159
34. Brescia AC, Simonds MM, Sullivan KE, Rose CD. Secretion of pro-inflammatory cytokines and chemokines and loss of regulatory signals by fibroblast-like synoviocytes in juvenile idiopathic arthritis. *Proteomics Clin Appl.* 2017;11(5–6):1600088. doi:10.1002/prca.201600088
35. Haasnoot AJ, van Tent-Hoeve M, Wulffraat NM, et al. Erythrocyte sedimentation rate as baseline predictor for the development of uveitis in children with juvenile idiopathic arthritis. *Am J Ophthalmol.* 2015;159(2):372–7.e1. doi:10.1016/j.ajo.2014.11.007
36. Do T, Tan R, Bennett M, et al. MicroRNA networks associated with active systemic juvenile idiopathic arthritis regulate CD163 expression and anti-inflammatory functions in macrophages through two distinct mechanisms. *J. Leukoc Biol.* 2018;103(1):71–85. doi:10.1002/JLB.2A0317-107R
37. Lazić E, Jelušić M, Grčević D, Marušić A, Kovačić N. Osteoblastogenesis from synovial fluid-derived cells is related to the type and severity of juvenile idiopathic arthritis. *Arthritis Res Ther.* 2012;14(3):R139. doi:10.1186/ar3872.
38. Sun J, Huang L, Wang J, Hu Y, Wang W, Zhu H. Programmed cell death in autoimmune diseases: ferroptosis. *Ann Biol Clin.* 2024;82(1):33–42. doi:10.1684/abc.2024.1866
39. Pascual V, Allantaz F, Arce E, Punaro M, Banchereau J. Role of interleukin-1 (IL-1) in the pathogenesis of systemic onset juvenile idiopathic arthritis and clinical response to IL-1 blockade. *J Exp Med.* 2005;201(9):1479–1486. doi:10.1084/jem.20050473

International Journal of General Medicine

Dovepress

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>