Altered molecular pathways and prognostic markers in active systemic juvenile idiopathic arthritis: Integrated bioinformatic analysis

Yi Ren[1], Hannah Labinsky[2], Andriko Palmowski[3], Henrik Bäcker[1], Michael Müller[1], Arne Kienzle[1,4*]

ABSTRACT

Systemic juvenile idiopathic arthritis (SJIA) is a severe childhood-onset inflammatory disease characterized by arthritis accompanied by systemic auto-inflammation and extra-articular symptoms. While recent advances have unraveled a range of risk factors, the pathomechanisms involved in SJIA and potential prognostic markers for treatment success remain partly unknown. In this study, we included 70 active SJIA and 55 healthy control patients from the National Center for Biotechnology Information to analyze for differentially expressed genes (DEGs) using R. Functional enrichment analysis, protein-protein interaction (PPI), and gene module construction were performed for DEGs and hub gene set. We additionally examined immune system cell composition with CIBERSORT and predicted prognostic markers and potential treatment drugs for SJIA. In total, 94 upregulated and 24 downregulated DEGs were identified. Two specific modules of interest and eight hub genes (*ARG1*, *DEFA4*, *HP*, *MMP8*, *MMP9*, *MPO*, *OLFM4*, and *PGLYRP1*) were screened out. Functional enrichment analysis suggested that complex neutrophil-related functions play a decisive role in the disease pathogenesis. CIBERSORT indicated neutrophils, Mo macrophages, CD8+ T cells, and naïve B cells to be relevant drivers of disease progression. In addition, we identified *TPM2* and *GZMB* as potential prognostic markers for treatment response to canakinumab. Moreover, sulindac sulfide, (-)-catechin, and phenanthridinone were identified as promising treatment agents. This study provides a new insight into molecular and cellular pathogenesis of active SJIA and highlights potential targets for further research.

KEYWORDS: Systemic juvenile idiopathic arthritis; hub genes; neutrophil; prognostic marker

INTRODUCTION

Systemic juvenile idiopathic arthritis (SJIA) is a childhood-onset inflammatory disease that is characterized by arthritis accompanied by systemic auto-inflammation and extra-articular symptoms. Disease burden is often high and therapeutic options limited, making diseases treatment highly challenging for patients and physicians alike. While recent advances have unraveled a range of risk factors, the pathomechanisms involved in SJIA remain complex and partly unknown [1,2].

Besides arthritis, SJIA is characterized by quotidian fever of ≥39°C that persists for longer than 2 weeks, evanescent erythematous skin rashes, and at least one of the following clinical features: Lymphadenopathy, pericarditis, pleuritis, or hepatosplenomegaly [3]. In addition, approximately 10% of all patients are prone to developing macrophage activation syndrome (MAS), a life-threatening complication [4,5].

Disease activity in affected patients can vary greatly and may range from clinically inactive to high disease activity [6,7]. Recent advances have identified prominent pro-inflammatory activation independent of disease activity [6]. In particular, patients present with leukocytosis, thrombocytosis, and highly elevated erythrocyte sedimentation rate and C-reactive protein concentration. In contrast to other subtypes of juvenile idiopathic arthritis, dysregulation of the innate immune system plays a significant role in disease progression [8]. In SJIA, toll-like receptor (TLR) signaling pathways mediate aberrant activation of phagocytes including monocytes, macrophages, and neutrophils [9]. These key players in the innate immune system are responsible for the subsequent release of pivotal pro-inflammatory cytokines [9]. In particular, expression of interleukin-1β,

[1]Center for Musculoskeletal Surgery, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität Zu Berlin, and Berlin Institute of Health, Berlin, Germany
[2]Department of Internal Medicine, Rheumatology and Immunology, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and University Hospital Erlangen, Erlangen, Germany
[3]Department of Rheumatology and Clinical Immunology, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany
[4]Laboratory of Adaptive and Regenerative Biology, Brigham and Women's Hospital, Harvard Medical School, Boston MA, USA

*Corresponding author: Arne Kienzle, Center for Musculoskeletal Surgery, Clinic for Orthopedics, Charité – Universitätsmedizin Berlin, Charitéplatz 1, Berlin, Germany. E-mail: arne.kienzle@charite.de

DOI: https://dx.doi.org/10.17305/bjbms.2021.6016

Submitted: 08 May 2021/Accepted: 07 August 2021/
Published online: 03 September 2021

Conflicts of interest: The author(s) declare no conflicts of interest.

Funding: The current study was supported by grants from China Scholarship Council (No.202006210082). Dr. Arne Kienzle is a participant in the BIH-Charité Junior Clinician Scientist Program funded by the Charité – Universitätsmedizin Berlin and the Berlin Institute of Health.

interleukin-6, and interleukin-18 was found to be significantly elevated in SJIA, leading to current targeted treatment approaches with biologicals such as anakinra or tocilizumab [10,11]. However, high rates of refractory and recurrent disease suggest other pathways to be involved in SIJA [4].

The use of traditional methods such as polymerase chain reaction, immunohistochemistry, and flow cytometry for immune system composition is not optimal for high throughput. Microarray technology, a powerful strategy to test expression of thousands of genes simultaneously, has widely gained attention for the profiling of differentially expressed genes (DEGs). Bioinformatic analysis on transcription level is capable of defining hub genes, significant signaling pathways, and immune composition patterns. Comprehensive evaluation of the involved immune cells may offer new treatment approaches for affected patients.

In this study, we analyzed immune cell composition, key genes, pathways, and protein-protein interactions in active SJIA by utilizing a bioinformatic analysis method. Subsequently, we analyzed potential drugs as new treatment approaches based on identified DEGs.

MATERIALS AND METHODS

Microarray data

All primary data analyzed in this study were accessed from public data repositories. Dataset GSE17590 and GSE80060 were downloaded from Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) of the National Center for Biotechnology Information. The datasets consist of gene expression data of whole peripheral blood samples from 44 patients diagnosed with active SJIA (placebo treated), from 43 healthy control patients, and from 77 patients treated with canakinumab [12,13]. Active SJIA was defined by a juvenile arthritis disease activity score (JADAS) above 8.5, laboratory parameters (C-reactive protein and erythrocyte sediment rate), and overt disease symptoms. Besides, we also obtained GSE112057, a RNA-sequencing high-throughput dataset, as a verification cohort with a total of 38 patients (26 active SJIA patients and 12 healthy controls) [14].

Primary data processing and identification of DEGs

Raw data from GSE17590 and GSE80060 were loaded into R software (R Development Core Team; version: 3.6.3). After combining the expression matrices of these two datasets, the interbatch difference was removed using the "sva" package [15]. Results before and after batch effect removal were plotted using a two-dimensional PCA cluster graph. "limma" package was employed to identify differences in DEGs in SJIA and in healthy controls [16]. Cutoff values for DEGs were set as adjusted $p < 0.05$ and |log2 fold change (FC)| >1.5. DEGs were visualized as heatmap and volcano plots using R packages "pheatmap" and "ggplot2" [17,18].

Functional and pathway enrichment analysis

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for DEGs were performed using the "clusterprofiler" package to determine alterations of enriched pathways [19]. Functional enrichment is identified by comparing genes with a predefined group of genes that share localization, pathways, functions, or other features. GO term enrichment analysis was conducted for three sub-ontologies: Cellular component (CC), molecular function (MF), and biological process (BP). $P \leq 0.05$ was considered statistically significant.

Protein-protein interaction (PPI) network analysis and hub gene identification

PPI network analysis was conducted using the Search Tool for the Retrieval of Interacting Genes (STRING) online database (STRING, http://string-db.org). An interaction score over 0.4 (medium confidence) was set as threshold value. The PPI network was visualized by employing the "Molecular Complex Detection (MCODE)" plugin in Cytoscape software (version 3.8.2). For detection of typical gene modules, default parameters were set to degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max. depth = 100. For modules of interest, a score >4 was set as the cutoff threshold. Subsequently, the 10 genes with the highest degree of connectivity were selected. In addition, we selected the 10 highest ranking genes with the MCC algorithm using the "cytohubba" plug-in. These two sets of genes were then intersected to generate a list of hub genes. Expression profiles of hub genes were verified in the GSE112057 cohort using the "limma" package. Modules and hub genes went through GO and KEGG pathway analysis to predict pathological impact.

Prognostic markers screened with DEGs

To predict prognosis of active SJIA with DEGs, we established a logistic regression model based on data from 77 patients in GSE80060. These patients diagnosed with active SJIA were all treated with canakinumab, a neutralizing monoclonal antibody (mAb) against IL-1beta. Clinical response to treatment was evaluated using the American College of Rheumatology (ACR) response criteria [20] at day 15 after initiation of mAb therapy. In our study, we defined an improvement ≥50% as the cutoff for a "good" versus "poor" response. We allocated all DEGs into least absolute shrinkage and selection operator (LASSO) regression using package "glmnet" [21] to perform a feature selection process to determine the most suitable prognostic markers and compared the expression level of each gene with

two different responses. Binary logistic regression model analysis was employed after selecting a set of genes. Subsequently, the receiver operating characteristics (ROC) curve was used to examine the predictive value of selected genes.

Assessment of immune cells

CIBERSORT (http://cibersort.stanford.edu) is an online tool to distinguish 22 different immune cell types and their respective composition in samples using a deconvolution algorithm. Gene expression data were uploaded to CIBERSORT with LM22 to generate a data matrix. Subsequently, "ggplot2" and "pheatmap" packages were employed to visualize the composition of immune cells. Correlation coefficients were calculated using Pearson correlation analysis and plotted as a heatmap using the "corrplot" package [22]. Immune cells with highest |log2FC| values were selected, correlation with hub genes calculated using Pearson correlation analysis, and results visualized in R software.

Prediction of potential new drugs

Online database Connectivity Map (CMap; http://www.broadinstitute.org/cmap/) was employed to screen for potential drugs based on the specific gene expression signature. Over- and underexpressed DEGs were uploaded to CMap to obtain a table of predicted agents including enrichment scores. Enrichment scores (ES) ranged from −1 to 1. Drugs with negative enrichment scores have the potential capability to repress pathologically active pathways in SJIA. All drugs with an enrichment score below −0.8 were assessed. $P \leq 0.05$ was considered statistically significant.

RESULTS

Identification of DEGs

The workflows of this study are shown in Figure 1A. We recorded age-matched (median age, SJIA: 9.5, control: 9.0) and gender-matched (female, SJIA: 55%, control: 56%) patients from the two datasets. To gain an overview of altered gene expression of active SJIA, we first identified relevant DEGs. After normalization and removal of batch differences between dataset GSE17590 and GSE80060 (Supplemental Figure 1), we identified a total of 118 DEGs including 94 upregulated and 24 downregulated genes (Figure 1B, Supplemental Table 1). In particular, expression of *CD177*, *OLFM4*, *ARHGEF12*, *MMP8*, *PLOD2*, *CEACAM6*, and *CEACAM8* was highly upregulated ($P < 0.001$, Figure 1C). In contrast, *TCL1A*, *ALOX15*, and *HLA-DQB* were downregulated the most ($P < 0.001$).

Functional and pathway enrichment analysis

Based on the identified DEGs, GO and KEGG analyses were employed to assess the relevant biological functions and pathways. Using GO analysis, the top five terms with highest gene ratio in each sub-ontology were identified (Table 1 and Figure 2A). For biological processes, "neutrophil degranulation," "neutrophil activation involved in immune response," "neutrophil activation," "neutrophil-mediated immunity," and "leukocyte migration" were ranked highest. For cellular components, gene ratio was highest in "specific granule," "secretory granule lumen," "cytoplasmic vesicle lumen," "vesicle lumen," and "tertiary granule." For molecular function, "serine-type endopeptidase activity," "serine-type peptidase activity," "serine hydrolase activity," "glycosaminoglycan binding," and "protein kinase regulator activity." Five statistically significant KEGG pathways were identified, among which the Janus kinase/signal transduction and activator of transcription (JAK-STAT) signaling pathway had the highest gene ratio (Table 1 and Figure 2B).

Identification of PPI modules and hub genes

PPI analysis was used to decode the underlying interactions between proteins and the molecular pathogenesis. We obtained a PPI network with a total of 117 nodes with 199 edges from the STRING database and Cytoscape software (Figure 3A). While each node represents one gene, the edges represent an interaction between two genes. In total, five modules were identified using MCODE (Table 2). The two modules with an MCODE score ≥4 (module A and B) were analyzed further (Figure 3B-C). The genes associated with these two modules were all upregulated compared to the healthy controls. After intersecting the gene sets generated using the plug-in "cytohubba," we found a hub gene set of eight upregulated genes including *ARG1*, *DEFA4*, *HP*, *MMP8*, *MMP9*, *MPO*, *OLFM4*, and *PGLYRP1* (Table 2). No downregulated genes were found. The GSE112057 dataset was used to confirm these findings (Supplemental Figure 2). Equally, the same eight genes were significantly upregulated in SJIA patients (Table 3).

Functional enrichment analysis demonstrated these eight hub genes to be mainly involved in neutrophil function, serine peptidase pathways, and misregulation in cancer (Figure 3D). Module A was particularly related to neutrophil function with a gene ratio of 1.0 (Figure 3B). Of note, the whole DEG profile, module A, and hub gene set were consistently related to neutrophil function. In contrast, module B was associated with hemostatic and secretory functions (Figure 3C).

Screening and testing markers with prognostic value

Using ACR criteria, 44 patients with "good" and 33 patients with "poor" response to treatment were identified. We input all 118 DEGs into the LASSO regression algorithm and used the minimum value of lambda (lambda.min) as the cutoff (Figure 4A). A total of seven genes were screened out as potential prognostic markers, including *TPM2*, *PRSS33*, *LTBP1*, *GZMB*,

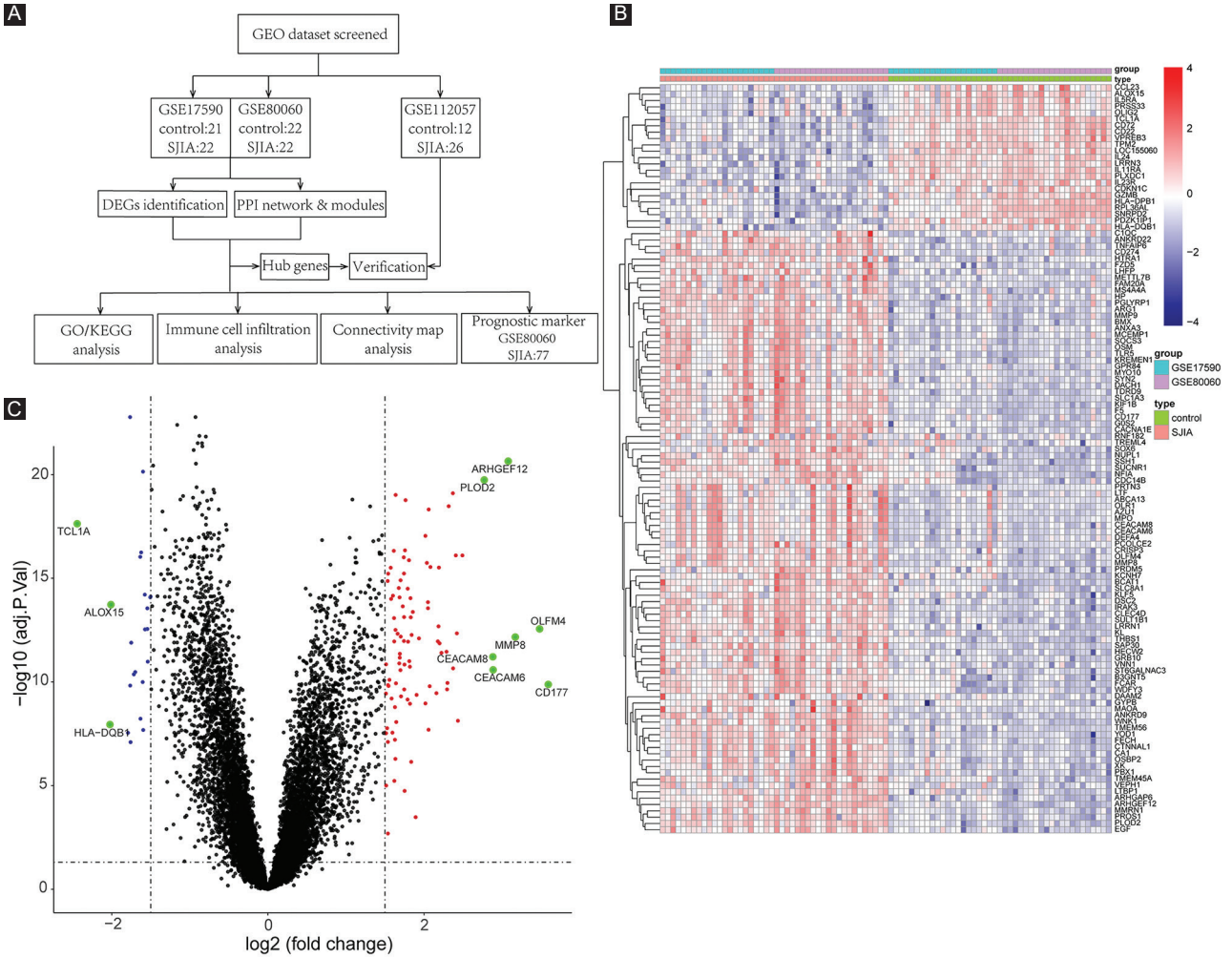


FIGURE 1. Data retrieving concept and DEG screening. (A) Flow diagram of employed bioinformatic analysis including data retrieving, processing, analyzing, and validation. (B) Heatmap of 118 DEGs for SJIA patients and healthy individuals. (C) Volcano plot of gene expression profile comparing SJIA patients and healthy individuals. Genes with highest |log2FC| values are highlighted in green. GEO: Gene expression omnibus; DEG: Differentially expressed gene; PPI: Protein-protein interaction; SJIA: Systemic juvenile idiopathic arthritis; FC: Fold change.

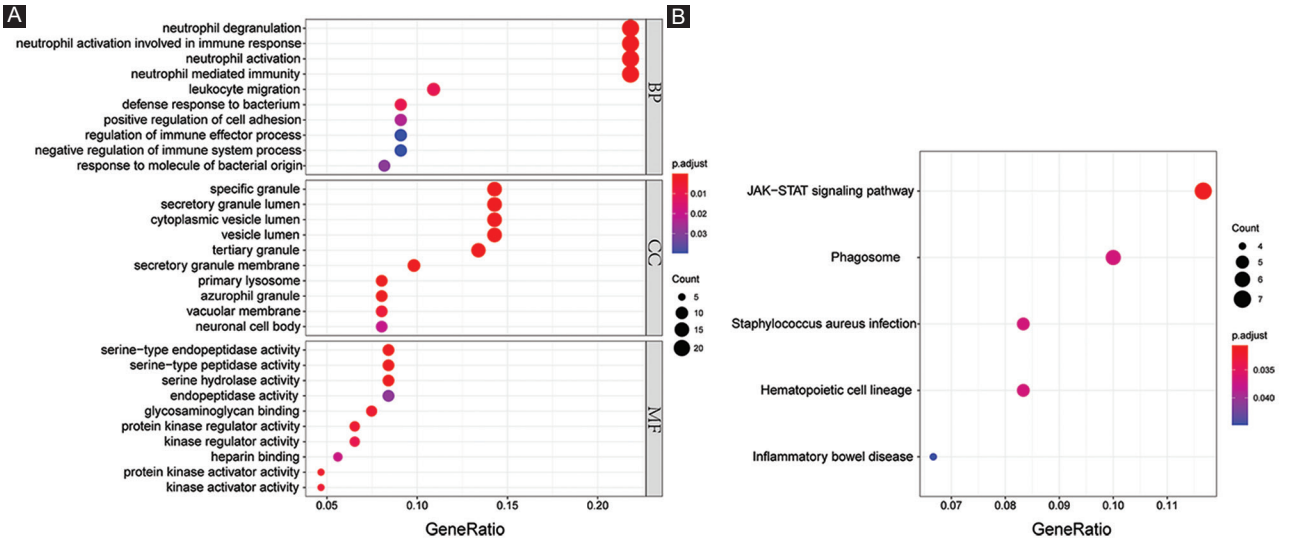


FIGURE 2. Functional and pathway enrichment analysis. (A) Top 10 results of each sub-term in GO functional enrichment analysis of the DEGs set. (B) Top 10 results of KEGG pathway enrichment analysis of DEGs. GO: Gene ontology; DEG: Differentially expressed gene; KEGG: Kyoto Encyclopedia of Genes and Genomes; BP: Biological process; CC: Cellular component; MF: Molecular function.

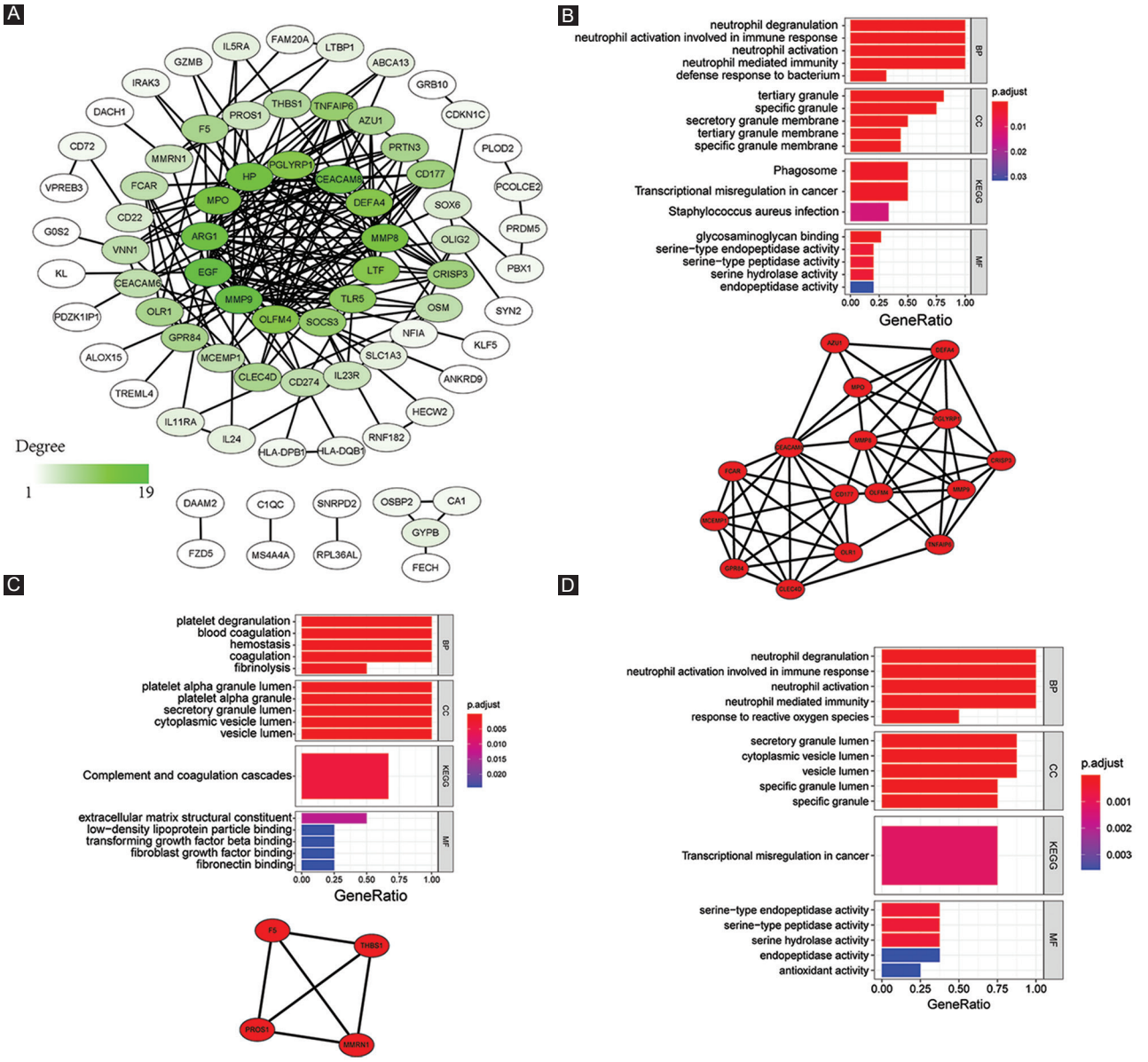


FIGURE 3. PPI network. (A) Overview of PPI network of DEGs in SJIA patients and healthy controls. Color darkness intensity represents gene connectivity. (B) PPI network and GO/KEGG analysis of gene set of module A. (C) PPI network and GO/KEGG analysis of gene set of module B. (D) GO/KEGG analysis of the hub gene set. BP: Biological process; CC: Cellular component; MF: Molecular function; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: Protein-protein interaction; DEG: Differentially expressed gene; SJIA: Systemic juvenile idiopathic arthritis; GO: Gene ontology.

F5, *BMX*, and *ARHGEF12*. Using a binary logistic regression model, significant correlation to treatment response was only found for *TPM2* (P = 0.019; coefficient = −3.56; Figure 4B). Five of the analyzed genes showed differential expression levels for "good" and "poor" clinical outcome (Figure 4C). ROC analysis for these genes suggested a good predictive value of *TPM2* for short-term prognosis with an area under the curve (AUC) of 0.803 and *GZMB* with an AUC of 0.824. A combination of *GZMB* and *TPM2* improved AUC to 0.846 (Figure 4D).

Proportion of immune cells and correlation

Immune reactivity plays an essential role in SJIA. Therefore, we investigated the constitution of the immune landscape surrounding the disease development. Relative numbers of 22 types of immune cells were calculated using the CIBERSORT algorithm (Figure 5A). Compared to the composition in healthy individuals, neutrophils, Mo macrophages, and activated dendritic cells were upregulated, while the percentage of resting mast cells, M2 macrophages, naïve B cells, and CD8+ T cells was decreased (Figure 5C). Of these cells, neutrophils were most prominently upregulated (P < 0.001) and CD8+ T cells (P < 0.001) most severely downregulated. Correlation coefficients for all 22 types of immune cells were calculated and displayed as a heat map (Figure 5B). Activated mast cells and gamma delta T cells showed the strongest positive correlation (r = 0.54, p < 0.001), while neutrophils and CD8+ T cells were

TABLE 1. Top 10 enriched terms in each sub-ontology of GO/KEGG analysis of DEGs

Ontology	Description	Gene ratio	p. adjust
BP	Neutrophil degranulation	0.2182	<0.0001
BP	Neutrophil activation involved in immune response	0.2182	<0.0001
BP	Neutrophil activation	0.2182	<0.0001
BP	Neutrophil-mediated immunity	0.2182	<0.0001
BP	Leukocyte migration	0.1091	0.0106
CC	Specific granule	0.1429	<0.0001
CC	Tertiary granule	0.1339	<0.0001
CC	Secretory granule lumen	0.1429	<0.0001
CC	Cytoplasmic vesicle lumen	0.1429	<0.0001
CC	Vesicle lumen	0.1429	<0.0001
MF	Serine-type endopeptidase activity	0.0841	0.0002
MF	Serine-type peptidase activity	0.0841	0.0002
MF	Serine hydrolase activity	0.0841	0.0002
MF	Endopeptidase activity	0.0841	0.0283
MF	Glycosaminoglycan binding	0.0748	0.0056
KEGG	JAK-STAT signaling pathway	0.1167	0.0305
KEGG	Phagosome	0.1000	0.0367
KEGG	Staphylococcus aureus infection	0.0833	0.0367
KEGG	Hematopoietic cell lineage	0.0833	0.0367
KEGG	Inflammatory bowel disease	0.0667	0.0448

BP: Biological process; CC: Cellular component; MF: Molecular function; KEGG: Kyoto Encyclopedia of Genes and Genomes

TABLE 2. Properties and node IDs of modules identified by PPI network analysis

Module	Score	Nodes	Edges	Node IDs
A	7.467	16	56	GPR84, CRISP3, CD177, CLEC4D, OLR1, MCEMP1, FCAR, CEACAM8, MMP9, PGLYRP1, OLFM4, TNFAIP6, MPO, AZU1, MMP8, DEFA4
B	4.000	4	6	MMRN1, PROS1, THBS1, F5
C	3.333	4	5	LTF, PRTN3, ARG1, HP
D	3.333	4	5	EGF, NFIA, OLIG2, SLC1A3
E	3.000	3	3	CA1, OSBP2, GYPB

PPI: Protein-protein interaction

most negatively correlated (r = −0.72, $p < 0.001$). Correlation analysis also demonstrated significantly positive correlation for the proportion of neutrophils and Mo macrophages with expression of all hub genes except *MPO*. In addition, the percentage of CD8+ T cells and naïve B cells was significantly negatively correlated with expression of all hub genes (Figure 5D).

Prediction of potentially effective drugs against active SJIA

A list of potential small molecular drugs targeting the up- and downregulated DEGs was generated. In total, eight drugs were identified: Sulindac sulfide (ES = −0.975, $p < 0.001$), phenyl biguanide (ES = −0.963, $p < 0.001$), (-)-catechin (ES = −0.957, $p < 0.001$), splitomicin (ES = −0.903, $p < 0.001$), methocarbamol (ES = −0.901, $p = 0.002$), Gly-His-Lys (ES = −0.861, $p = 0.005$), phenanthridinone (ES = −0.847, $p < 0.001$), and hycanthone (ES = −0.81, $p = 0.003$) (Table 4).

TABLE 3. Hub genes identified by PPI network analysis

Gene	Description	Log FC	p value	FDR
ARG1	Arginase, liver	2.896	0.0001	0.0046
DEFA4	Defensin, alpha 4, corticostatin	3.175	0.0001	0.0065
HP	Haptoglobin	2.443	0.0009	0.0150
MMP8	Matrix metallopeptidase 8	3.506	0.0003	0.0095
MMP9	Matrix metallopeptidase 9	2.017	0.0003	0.0098
MPO	Myeloperoxidase, nuclear gene encoding mitochondrial protein	2.888	0.0002	0.0073
OLFM4	Olfactomedin 4	4.472	0.0001	0.0053
PGLYRP1	Peptidoglycan recognition protein 1	1.558	0.0023	0.0250

PPI: Protein-protein interaction; FC: Fold change; FDR: False discovery rate

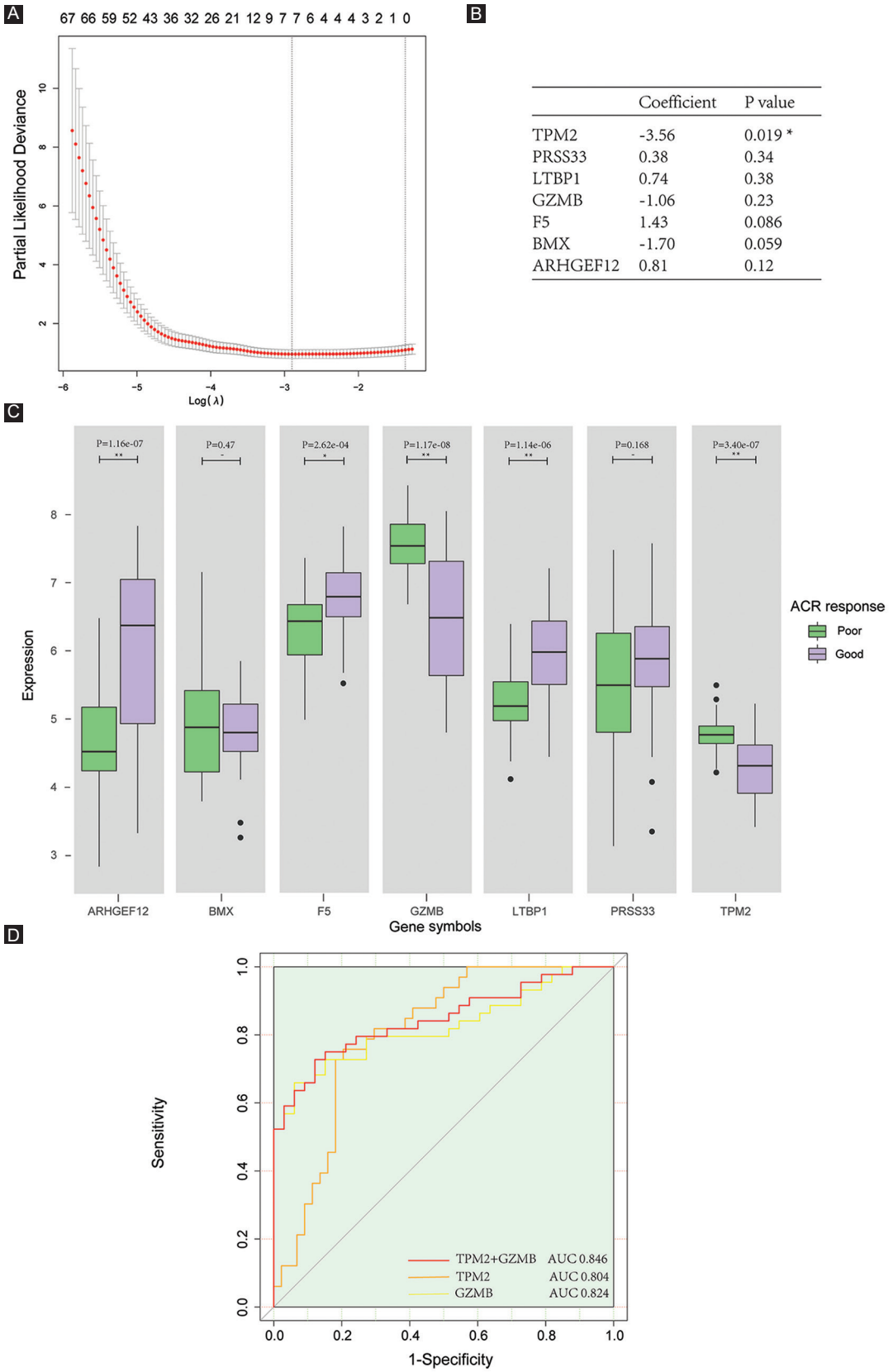
TABLE 4. Characteristics of the most significant predicted drugs for potential treatment of SJIA

Drug name	Enrichment score	p value	Description
sulindac sulfide	−0.9750	<0.0001	Nonsteroidal anti-inflammatory drug; anti-cancer activity; apoptosis inducer; anti-neoplastic agent
phenyl biguanide	−0.9630	<0.0001	5-HT3 agonist; triggers dopamine release in the nucleus accumbens
(-)-catechin	−0.9570	<0.0001	Natural phenol; antioxidant;
splitomicin	−0.9030	<0.0001	Sir2 inhibitor; platelet aggregation inhibitor
methocarbamol	−0.9010	0.0018	Centrally acting muscle relaxant; inhibitor of acetylcholinesterase at synapses in the autonomic nervous system, neuromuscular junction, and central nervous system
Gly-His-Lys	−0.8610	0.0055	Aids wound healing; modulator of lung tissue destruction in chronic obstructive pulmonary disease
Phenanthridinone	−0.8470	<0.0001	Poly (ADP-ribose) polymerase (PARP) inhibitor; immunosuppressive activity
Hycanthone	−0.8100	0.0025	Schistosomicide; anti-schistosomal activity; potential anti-neoplastic agent

SJIA: Systemic juvenile idiopathic arthritis

DISCUSSION

Despite treatment, SJIA is a refractory and recurrent disease that is associated with systemic inflammation and high mortality rates [8,15,23]. The identification of molecular changes and alterations in immune profile composition is paramount for a better understanding of the pathomechanisms involved and may provide new therapeutic targets [10,24].



	Coefficient	P value
TPM2	-3.56	0.019 *
PRSS33	0.38	0.34
LTBP1	0.74	0.38
GZMB	-1.06	0.23
F5	1.43	0.086
BMX	-1.70	0.059
ARHGEF12	0.81	0.12

FIGURE 4. Prognostic markers screening and testing. (A) Partial likelihood deviance plot for LASSO regression analysis. (B) Binary logistic regression model with most potential genes. (C) Different expression patterns compared between two response types. (D) ROC curve for testing prognostic value. ACR: American College of Rheumatology.

GO enrichment demonstrated DEGs enrichment in neutrophil-related processes, including neutrophil degranulation, neutrophil-mediated immunity, and neutrophil activation. In accordance with our results, previous research found neutrophilia to be closely linked to the activity of SJIA [6,11,25]. Neutrophils are major innate immune effector cells and

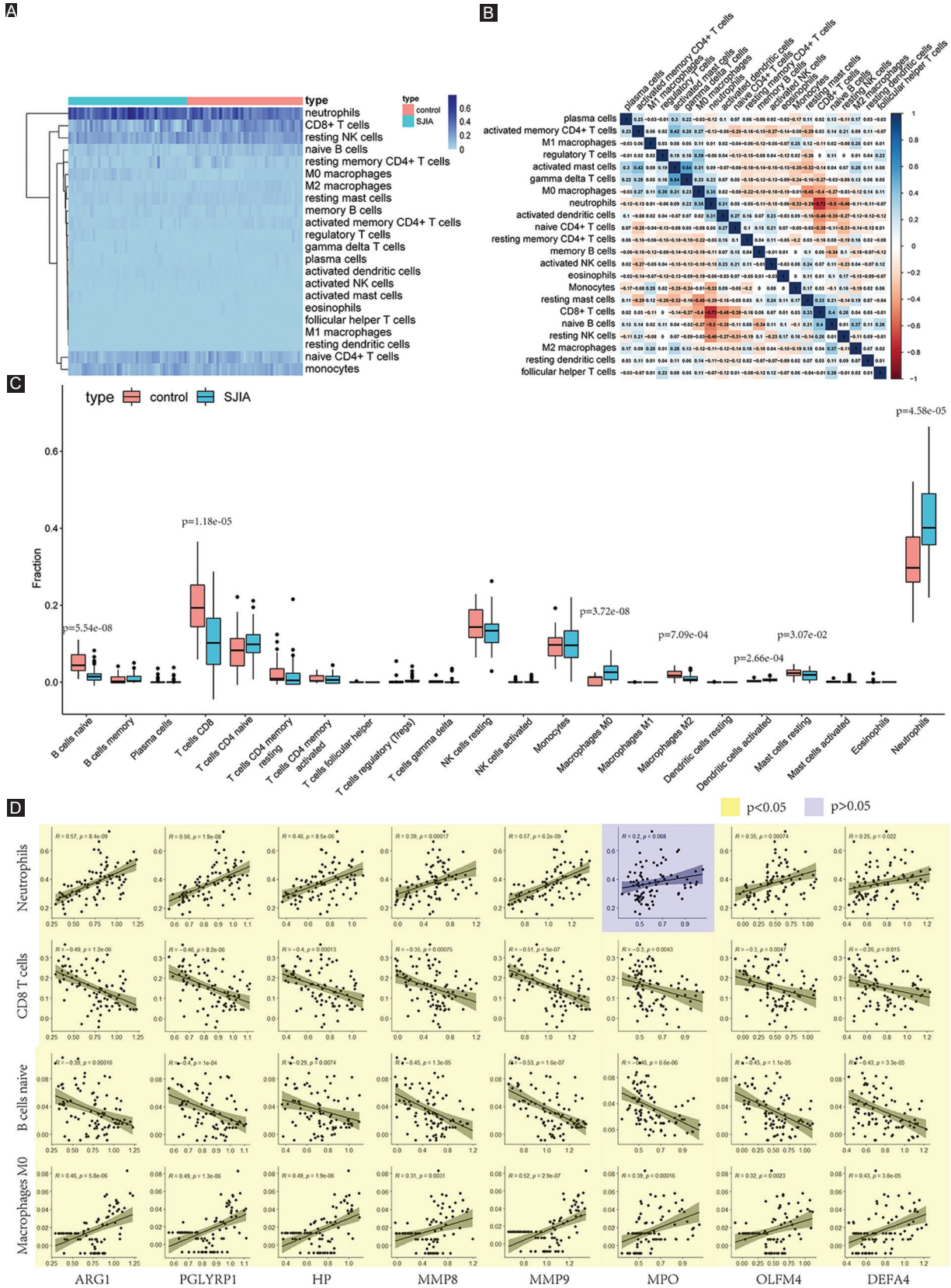


FIGURE 5. Comparison of 22 different immune system cells in peripheral blood specimens. (A) Heatmap of immune system cells in SJIA patients and healthy controls. (B) Correlation heatmap plot of 22 immune system cells. Color darkness intensity represents correlation between two cells. (C) Boxplot distribution of immune system cells. Significant differences were marked with p values. (D) Linear regression plot between eight hub genes and the four immune system cells with largest shift in prevalence in SJIA. SJIA: Systemic juvenile idiopathic arthritis.

play a key role in the response against microorganisms [26]. Immunological functions including degranulation and excretion of reactive oxygen species are upregulated in these effector cells in inflammation [26]. Reverting these sepsis-like features of neutrophils by current treatment strategies showed satisfactory outcomes in patients with SJIA [4]. In active SJIA, neutrophil function is significantly upregulated and outlasts clinical symptoms suggesting neutrophils to play a decisive role in long-term disease progression [6]. In addition, in contrast to other JIAs, the prevalence of suppressive CD16+/CD62L(dim) neutrophils was lower in patients with clinically inert SJIA suggesting neutrophil heterogeneity and unique molecular pathomechanisms to be involved [6]. Considering the significant role of the immune system in the pathogenesis of SJIA, we explored the relative prevalence of 22 subtypes of immune cells. Confirming previous results, we found elevated prevalence of neutrophils [6,25]. In contrast to a previous study, we found lower levels of CD8+ T cells [27]. We hypothesize that this difference is due to a dynamic change in cell prevalence depending on disease progression. Silvestre-Roig et al. suggested CD16+/CD62L(dim) neutrophil-driven suppression of T-cell proliferation through integrins and ROS production as a potential mechanism [28]. Further evidence is lent to this theory by our results showing negative correlation of neutrophil and CD8+ T-cell prevalence suggesting disease stage-dependent regulatory interplays between innate and adaptive immune cells.

Previous research demonstrated neutrophil-related protease activity to be regulated by the JAK-STAT pathway [29]. Using KEGG analysis, the JAK-STAT pathway was identified as the most enriched pathway. Similarly, previous research using SJIA models and ingenuity pathway analysis suggested inhibition of the JAK-STAT pathway to be a viable therapy and an ongoing randomized controlled trial (NCT03000439) is anticipated to provide therapeutical evidence for patients with SJIA [25,30,31].

In total, we identified eight upregulated hub genes in patients with active SJIA: *HP*, *MPO*, *MMP8*, *MMP9*, *ARG1*, *OLFM4*, *DEFA4*, and *PGLYRP1*. We postulate these genes to be significantly involved in the inflammatory environment and subsequent disease progression in patients with active SJIA. HP encodes haptoglobin, which binds free hemoglobin in the plasma by forming a haptoglobin-hemoglobin complex that can be taken up by CD163+ macrophages [32]. In accordance with our results, *HP* was found to be overexpressed in peripheral blood samples in SJIA patients with subclinical MAS [33]. Together with the MMP family, MPO mediates antimicrobial activity and pro-inflammatory response by induction of oxidative tissue damage and neutrophil respiratory burst [34,35]. Despite its localized effectiveness, we found expression of *MPO*, *MMP8*, and *MMP9* to be elevated in systemically affected SJIA patients. In accordance with these results, the previous studies have suggested elevated expression of *MPO*, *MMP8*, and *MMP9* in several inflammatory morbidities such as SJIA [25,31,36]. Further, *MMP9* has been suggested as a prognostic plasma biomarker in patients with SJIA [37]. *ARG1*, arginase-encoding gene, was reported to be upregulated in the peripheral blood and cancer tissue of various cancer patients [38-40]. In tumor-associated myeloid cells, upregulation of *ARG1* has been shown to reduce inflammation by suppressing T-cell proliferation through arginine deprivation [41,42]. Moreover, *ARG1* has been shown to be upregulated in autoimmune disease such as rheumatoid arthritis [43]. In SJIA, *ARG1* was reported to be linked to anti-inflammatory M2 macrophage polarization [44] – however, there is lack of more detailed knowledge on its role in disease progression. We also found increased expression of *OLFM4*, *DEFA4*, and *PGLYRP1*, which encode olfactomedin-4, defensin alpha 4, and peptidoglycan recognition protein 1, respectively. Albeit these genes have been linked to autoimmune diseases [45-47], elevated expression in SJIA has not been reported and their function is yet to be determined. Upregulation of both pro- and anti-inflammatory genes indicates neutrophils to play a complex, dual role in SJIA.

GO and KEGG enrichment analysis was repeated for modules and the hub gene set to identify their own specific effects and to exclude potentially confounding genes. Conversely, KEGG analysis on module B related to hemostatic and secretory functions. Correlation analysis demonstrated all eight upregulated genes described above to positively correlate with neutrophil and Mo macrophage and to negatively correlate with CD8+ T cells and naïve B cells. In consideration of these results and our hub genes analysis, alteration of neutrophil prevalence and function appears to be a decisive pathological deviation in SJIA. Further, analysis of neutrophil heterogeneity and function in SJIA is necessary to improve our understanding of SJIA pathogenesis and to potentially develop novel neutrophil-specific treatment approaches.

To bring our results into a translational context, we investigated markers to predict the prognosis and treatment response. Canakinumab is often prescribed at an early stage of SJIA yet is not universally effective in all patients [48]. In our research, we utilized LASSO regression and a binary logistic regression model to identify predictive markers for treatment response to canakinumab. We found *TPM2* to have a significant predictive value. Tropomyosin 2, encoded by *TPM2*, is an intracellular protein isoform in the tropomyosin family that binds to and stabilizes actin filaments in muscle fibers. However, its role in SJIA is unknown. In addition, we found *GZMB* encoding granzyme B, a serine proteinase secreted by cytotoxic T cells and natural killer cells, to be a promising predictive marker. *GZMB* has been linked to progression of JIA previously, even though it was not screened out by logistic regression model [49]. Both markers were significantly decreased in patients with a good treatment response and can potentially be utilized clinically.

Using CMap, we evaluated potential drugs with activity against the differentially expressed genes found in this study. Of the identified drugs, sulindac sulfide, (-)-catechin, and phenanthridinone were the most promising based on their known anti-inflammatory properties. Phenyl biguanide, splitomicin, methocarbamol, Gly-His-Lys, and hycanthone have not been reported to be effective against inflammation. However, further research on these drugs might provide novel therapeutic approaches. Sulindac sulfide is a non-selective anti-inflammatory drug that inhibits the chlorinating activity of MPO [50]. In addition, it displays mild COX-2 inhibiting effects and possesses potent ability to decrease ROS levels [51,52]. Catechin has been shown to inhibit production of pro-inflammatory factors IL-1, TNF-alpha, and prostaglandin E2 in adjuvant arthritis [53]. However, there is a paucity of knowledge on the specific effect of the (-)-catechin enantiomer. Phenanthridinone is an inhibitor of the nuclear enzyme poly(ADP-ribose) polymerase-1 (PARP-1) [54]. In a rodent acute lung inflammation model, phenanthridinone downregulated MPO activity and subsequent inflammatory cytokine excretion [55]. In addition, phenanthridinone was shown to suppress neutrophil infiltration in local inflammation [54].

CONCLUSION

This study provides a new insight into molecular and cellular pathogenesis of active SJIA and highlights potential targets for further research. We used bioinformatic analyses to demonstrate the essential role of neutrophils and to identify hub genes involved in the pathogenesis of active SJIA. In addition, we identified *TPM2* and *GZMB* as potential prognostic markers.

The observations in this study also have implications for future investigations. For a deeper understanding of disease progression and to identify new molecular targets, the involved and highly complex cytokine-mediated cell-cell interactions will have to be unraveled in more detail. In our study, we identified several new molecular agents as potential therapeutic candidates for patients with active SJIA. Whether these drugs can successfully translate into clinical treatment regimens remain to be investigated.

ACKNOWLEDGMENTS

We are thankful for the grants from China Scholarship Council (No.202006210082) and BIH-Charité Junior Clinician Scientist Program.

REFERENCES

[1] Ombrello MJ, Remmers EF, Tachmazidou I, Grom A, Foell D, Haas JP, et al. HLA-DRB1*11 and variants of the MHC class II locus are strong risk factors for systemic juvenile idiopathic arthritis. Proc Natl Acad Sci U S A 2015;112(52):15970-5. https://doi.org/10.1186/1546-0096-13-s1-075

[2] Woerner A, von Scheven-Gete A, Cimaz R, Hofer M. Complications of systemic juvenile idiopathic arthritis: Risk factors and management recommendations. Expert Rev Clin Immunol 2015;11(5):575-88. https://doi.org/10.1586/1744666x.2015.1032257

[3] Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International league of associations for rheumatology classification of juvenile idiopathic arthritis: Second revision, Edmonton, 2001. J Rheumatol 2004;31(2):390-2. https://doi.org/10.1002/art.40442

[4] Ter Haar NM, Tak T, Mokry M, Scholman RC, Meerding JM, de Jager W, et al. Reversal of sepsis-like features of neutrophils by interleukin-1 blockade in patients with systemic-onset juvenile idiopathic arthritis. Arthritis Rheumatol 2018;70(6):943-56.

[5] Behrens EM, Beukelman T, Paessler M, Cron RQ. Occult macrophage activation syndrome in patients with systemic juvenile idiopathic arthritis. J Rheumatol 2007;34(5):1133-8.

[6] Brown RA, Henderlight M, Do T, Yasin S, Grom AA, DeLay M, et al. Neutrophils from children with systemic juvenile idiopathic arthritis exhibit persistent proinflammatory activation despite long-standing clinically inactive disease. Front Immunol. 2018;9:2995. https://doi.org/10.3389/fimmu.2018.02995

[7] Schneider R, Laxer RM. 6 Systemic onset juvenile rheumatoid arthritis. Baill Clin Rheumatol 1998;12(2):245-71.

[8] Ombrello MJ, Arthur VL, Remmers EF, Hinks A, Tachmazidou I, Grom AA, et al. Genetic architecture distinguishes systemic juvenile idiopathic arthritis from other forms of juvenile idiopathic arthritis: Clinical and therapeutic implications. Ann Rheum Dis 2017;76(5):906-13. https://doi.org/10.1093/rheumatology/keh710

[9] Jung JY, Kim JW, Suh CH, Kim HA. Roles of interactions between toll-like receptors and their endogenous ligands in the pathogenesis of systemic juvenile idiopathic arthritis and adult-onset still's disease. Front Immunol 2020;11:583513. https://doi.org/10.3389/fimmu.2020.583513

[10] Ishikawa S, Mima T, Aoki C, Yoshio-Hoshino N, Adachi Y, Imagawa T, et al. Abnormal expression of the genes involved in cytokine networks and mitochondrial function in systemic juvenile idiopathic arthritis identified by DNA microarray analysis. Ann Rheum Dis 2009;68(2):264-72. https://doi.org/10.1136/ard.2007.079533

[11] Gattorno M, Piccini A, Lasiglie D, Tassi S, Brisca G, Carta S, et al. The pattern of response to anti-interleukin-1 treatment distinguishes two subsets of patients with systemic-onset juvenile idiopathic arthritis. Arthritis Rheum 2008;58(5):1505-15. https://doi.org/10.1002/art.23437

[12] Brachat AH, Grom AA, Wulffraat N, Brunner HI, Quartier P, Brik R, et al. Early changes in gene expression and inflammatory proteins in systemic juvenile idiopathic arthritis patients on canakinumab therapy. Arthritis Res Ther 2017;19(1):13. https://doi.org/10.1186/s13075-016-1212-x

[13] Quartier P, Allantaz F, Cimaz R, Pillet P, Messiaen C, Bardin C, et al. A multicentre, randomised, double-blind, placebo-controlled trial with the interleukin-1 receptor antagonist anakinra in patients with systemic-onset juvenile idiopathic arthritis (ANAJIS trial). Ann Rheum Dis 2011;70(5):747-54. https://doi.org/10.1136/ard.2010.134254

[14] Mo A, Marigorta UM, Arafat D, Chan LHK, Ponder L, Jang SR, et al. Disease-specific regulation of gene expression in a comparative analysis of juvenile idiopathic arthritis and inflammatory bowel disease. Genome Med 2018;10(1):48. https://doi.org/10.1186/s13073-018-0558-x

[15] Parker HS, Leek JT, Favorov AV, Considine M, Xia X, Chavan S, et al. Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. Bioinformatics 2014;30(19):2757-63. https://doi.org/10.1093/bioinformatics/btu375

[16] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47.

https://doi.org/10.1093/nar/gkv007
[17] Ito K, Murphy D. Application of ggplot2 to Pharmacometric Graphics. CPT Pharmacometrics Syst Pharmacol 2013;2:e79. https://doi.org/10.1038/psp.2013.56
[18] Wu Y, Zhang L, Zhang Y, Zhen Y, Liu S. Bioinformatics analysis to screen for critical genes between survived and nonsurvived patients with sepsis. Mol Med Rep 2018;18(4):3737-43. https://doi.org/10.3892/mmr.2018.9408
[19] Yu G, Wang LG, Han Y, He QY. clusterProfiler: An R package for comparing biological themes among gene clusters. OMICS 2012;16(5):284-7. https://doi.org/10.1089/omi.2011.0118
[20] Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. Arthritis Rheum 1997;40(7):1202-9. https://doi.org/10.1002/art.1780400703
[21] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010;33(1):1-22. https://doi.org/10.18637/jss.v033.i01
[22] Friendly M. Corrgrams. Am Stat 2002;56(4):316-24.
[23] Davies R, Southwood T, Kearsley-Fleet L, Lunt M, Baildam E, Beresford MW, et al. Mortality rates are increased in patients with systemic juvenile idiopathic arthritis. Arch Dis Child 2017;102(2):206-7. https://doi.org/10.1136/archdischild-2016-311571
[24] Kienzle A, Kurch S, Schloder J, Berges C, Ose R, Schupp J, et al. Dendritic mesoporous silica nanoparticles for pH-stimuli-responsive drug delivery of TNF-alpha. Adv Healthc Mater 2017;6(13):12. https://doi.org/10.1002/adhm.201700012
[25] Ramanathan K, Glaser A, Lythgoe H, Ong J, Beresford MW, Midgley A, et al. Neutrophil activation signature in juvenile idiopathic arthritis indicates the presence of low-density granulocytes. Rheumatology (Oxford) 2018;57(3):488-98. https://doi.org/10.1093/rheumatology/kex441
[26] Mantovani A, Cassatella MA, Costantini C, Jaillon S. Neutrophils in the activation and regulation of innate and adaptive immunity. Nat Rev Immunol 2011;11(8):519-31. https://doi.org/10.1038/nri3024
[27] Wu Q, Pesenacker AM, Stansfield A, King D, Barge D, Foster HE, et al. Immunological characteristics and T-cell receptor clonal diversity in children with systemic juvenile idiopathic arthritis undergoing T-cell-depleted autologous stem cell transplantation. Immunology 2014;142(2):227-36. https://doi.org/10.1111/imm.12245
[28] Silvestre-Roig C, Hidalgo A, Soehnlein O. Neutrophil heterogeneity: Implications for homeostasis and pathogenesis. Blood 2016;127(18):2173-81. https://doi.org/10.1182/blood-2016-01-688887
[29] Taylor PR, Roy S, Meszaros EC, Sun Y, Howell SJ, Malemud CJ, et al. JAK/STAT regulation of Aspergillus fumigatus corneal infections and IL-6/23-stimulated neutrophil, IL-17, elastase, and MMP9 activity. J Leukoc Biol 2016;100(1):213-22. https://doi.org/10.1189/jlb.4a1015-483r
[30] Li HW, Zeng HS. Regulation of JAK/STAT signal pathway by miR-21 in the pathogenesis of juvenile idiopathic arthritis. World J Pediatr 2020;16(5):502-13. https://doi.org/10.1007/s12519-019-00268-w
[31] Jiang H, Zhu M, Wang H, Liu H. Suppression of lncRNA MALAT1 reduces pro-inflammatory cytokines production by regulating miR-150-5p/ZBTB4 axis through JAK/STAT signal pathway in systemic juvenile idiopathic arthritis. Cytokine 2021;138:155397. https://doi.org/10.1016/j.cyto.2020.155397
[32] Madsen M, Graversen JH, Moestrup SK. Haptoglobin and CD163: Captor and receptor gating hemoglobin to macrophage lysosomes. Redox Rep 2001;6(6):386-8. https://doi.org/10.1179/135100001101536490
[33] Fall N, Barnes M, Thornton S, Luyrink L, Olson J, Ilowite NT, et al. Gene expression profiling of peripheral blood from patients with untreated new-onset systemic juvenile idiopathic arthritis reveals molecular heterogeneity that may predict macrophage activation syndrome. Arthritis Rheum 2007;56(11):3793-804. https://doi.org/10.1002/art.22981
[34] Kobus A, Baginska J, Lapinska-Antonczuk J, Lawicki S, Kierklo A. Levels of selected matrix metalloproteinases, their inhibitors in saliva, and oral status in juvenile idiopathic arthritis patients vs. healthy controls. Biomed Res Int 2019;2019:7420345. https://doi.org/10.1155/2019/7420345
[35] Aratani Y. Myeloperoxidase: Its role for host defense, inflammation, and neutrophil function. Arch Biochem Biophys 2018;640:47-52. https://doi.org/10.1016/j.abb.2018.01.004
[36] Ilisson J, Zagura M, Zilmer K, Salum E, Heilman K, Piir A, et al. Increased carotid artery intima-media thickness and myeloperoxidase level in children with newly diagnosed juvenile idiopathic arthritis. Arthritis Res Ther 2015;17:180. https://doi.org/10.1186/s13075-015-0699-x
[37] Ling XB, Lau K, Deshpande C, Park JL, Milojevic D, Macaubas C, et al. Urine peptidomic and targeted plasma protein analyses in the diagnosis and monitoring of systemic juvenile idiopathic arthritis. Clin Proteomics 2010;6(4):175-93. https://doi.org/10.1007/s12014-010-9058-8
[38] Wu CW, Chi CW, Tsay SH, Lui WY, P'Eng FK, Chang KL, et al. The effects of arginase on neoplasm. I. The role of arginase in the immunosuppressive effects of extract from gastric cancer. Zhonghua Min Guo Wei Sheng Wu Ji Mian Yi Xue Za Zhi 1987;20(4):279-89.
[39] Wu CW, Chi CW, Lin EC, Lui WY, P'Eng F K, Wang SR. Serum arginase level in patients with gastric cancer. J Clin Gastroenterol 1994;18(1):84-5.
[40] Porembska Z, Luboiński G, Chrzanowska A, Mielczarek M, Magnuska J, Barańczyk-Kuźma A. Arginase in patients with breast cancer. Clin Chim Acta 2003;328(1-2):105-11. https://doi.org/10.1016/s0009-8981(02)00391-1
[41] Vasquez-Dunddel D, Pan F, Zeng Q, Gorbounov M, Albesiano E, Fu J, et al. STAT3 regulates arginase-I in myeloid-derived suppressor cells from cancer patients. J Clin Invest 2013;123:1580-9. https://doi.org/10.1172/jci60083
[42] Miret JJ, Kirschmeier P, Koyama S, Zhu M, Li YY, Naito Y, et al. Suppression of myeloid cell arginase activity leads to therapeutic response in a NSCLC mouse model by activating anti-tumor immunity. J Immunother Cancer 2019;7(1):32. https://doi.org/10.1186/s40425-019-0504-5
[43] Panfili E, Gerli R, Grohmann U, Pallotta MT. Amino acid metabolism in rheumatoid arthritis: Friend or foe? Biomolecules 2020;10(9):1280. https://doi.org/10.3390/biom10091280
[44] Li D, Duan M, Feng Y, Geng L, Li X, Zhang W. MiR-146a modulates macrophage polarization in systemic juvenile idiopathic arthritis by targeting INHBA. Mol Immunol 2016;77:205-12. https://doi.org/10.1016/j.molimm.2016.08.007
[45] O'Hanlon TP, Rider LG, Gan L, Fannin R, Paules RS, Umbach DM, et al. Gene expression profiles from discordant monozygotic twins suggest that molecular pathways are shared among multiple systemic autoimmune diseases. Arthritis Res Ther 2011;13(2):R69. https://doi.org/10.1186/ar3330
[46] Amirbeagi F, Thulin P, Pullerits R, Pedersen B, Andersson BA, Dahlgren C, et al. Olfactomedin-4 autoantibodies give unusual c-ANCA staining patterns with reactivity to a subpopulation of neutrophils. J Leukoc Biol 2015;97(1):181-9. https://doi.org/10.1189/jlb.5a0614-311r
[47] Luo Q, Li X, Zhang L, Yao F, Deng Z, Qing C, et al. Serum PGLYRP1 is a highly discriminatory biomarker for the diagnosis of rheumatoid arthritis. Mol Med Rep 2019;19(1):589-94.
[48] Toplak N, Blazina Š, Avčin T. The role of IL-1 inhibition in systemic juvenile idiopathic arthritis: Current status and future perspectives. Drug Des Dev Ther 2018;12:1633-43. https://doi.org/10.2147/dddt.s114532
[49] Jahrsdörfer B, Vollmer A, Blackwell SE, Maier J, Sontheimer K, Beyer T, et al. Granzyme B produced by human plasmacytoid dendritic cells suppresses T-cell expansion. Blood 2010;115(6):1156-65. https://doi.org/10.1182/blood-2009-07-235382

[50] Neve J, Parij N, Moguilevsky N. Inhibition of the myeloperoxidase chlorinating activity by non-steroidal anti-inflammatory drugs investigated with a human recombinant enzyme. Eur J Pharmacol 2001;417(1-2):37-43. https://doi.org/10.1016/s0014-2999(01)00895-0
[51] Fernandes E, Toste SA, Lima JL, Reis S. The metabolism of sulindac enhances its scavenging activity against reactive oxygen and nitrogen species. Free Radic Biol Med 2003;35(9):1008-17. https://doi.org/10.1016/s0891-5849(03)00437-4
[52] Ko CJ, Lan SW, Lu YC, Cheng TS, Lai PF, Tsai CH, et al. Inhibition of cyclooxygenase-2-mediated matriptase activation contributes to the suppression of prostate cancer cell motility and metastasis. Oncogene 2017;36(32):4597-609. https://doi.org/10.1038/onc.2017.82
[53] Tang LQ, Wei W, Wang XY. Effects and mechanisms of catechin for adjuvant arthritis in rats. Adv Ther 2007;24(3):679-90.
[54] Mabley JG, Jagtap P, Perretti M, Getting SJ, Salzman AL, Virag L, et al. Anti-inflammatory effects of a novel, potent inhibitor of poly (ADP-ribose) polymerase. Inflamm Res 2001;50(11):561-9. https://doi.org/10.1007/pl00000234
[55] Liaudet L, Pacher P, Mabley JG, Virag L, Soriano FG, Hasko G, et al. Activation of poly(ADP-Ribose) polymerase-1 is a central mechanism of lipopolysaccharide-induced acute lung inflammation. Am J Respir Crit Care Med 2002;165(3):372-7. https://doi.org/10.1164/ajrccm.165.3.2106050

Related articles published in BJBMS

1. PCAT1 is a poor prognostic factor in endometrial carcinoma and associated with cancer cell proliferation, migration and invasion
Xiaohuan Zhao et al., BJBMS, 2019
2. Integrated profiling identifies ITGB3BP as prognostic biomarker for hepatocellular carcinoma
Qiuli Liang et al., BJBMS, 2020

SUPPLEMENTAL DATA

TABLE S1. Full list of the 118 DEGs, including 94 upregulated and 24 downregulated genes

Gene symbol	Definition	Log FC	p. adjust
ABCA13	ATP-binding cassette, sub-family A, member 13	2.0600	2.03E-08
CD177	CD177 antigen	3.5900	1.33E-10
ANKRD22	Ankyrin repeat domain 22	2.2200	3.97E-12
ANKRD9	Ankyrin repeat domain 9	2.4100	7.95E-17
OLFM4	Olfactomedin 4	3.4800	2.78E-13
MMP8	Matrix metallopeptidase 8	3.1700	6.89E-13
ARHGEF12	Rho guanine nucleotide exchange factor 12	3.0800	2.27E-21
CEACAM6	Carcinoembryonic antigen-related Cell adhesion molecule 6	2.8900	2.63E-11
CEACAM8	Carcinoembryonic antigen-related Cell adhesion molecule 8	2.8800	6.19E-12
B3GNT5	UDP-GlcNAc: betaGal beta-1,3-N-acetylglucosaminyl transferase 5	1.8400	5.45E-13
PLOD2	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2, transcript variant 1	2.7700	1.81E-20
SLC1A3	Solute carrier family 1, member 3	2.4900	7.95E-17
C1QC	Complement component 1, q subcomponent, C chain	1.5300	1.71E-08
DEFA4	Defensin, alpha 4, corticostatin	2.4300	7.50E-09
ARG1	Arginase, liver	2.4200	4.57E-13
CACNA1E	Calcium channel, voltage dependent, alpha 1E subunit	2.3700	2.23E-11
MMP9	Matrix metallopeptidase 9	2.3700	7.98E-20
CA1	Carbonic anhydrase I	2.3000	2.30E-10
CD274	CD274 antigen	2.0200	1.09E-09
MYO10	Myosin X	2.3000	1.12E-10
AZU1	Azurocidin 1	2.2900	3.51E-12
LRRN1	Leucine rich repeat neuronal 1	2.1900	3.06E-16
KL	Klotho, transcript variant 1	2.1800	1.11E-12
PCOLCE2	Procollagen C-endopeptidase enhancer 2	2.1600	3.51E-10
CLEC4D	C-type lectin domain family 4, member D	2.3200	3.37E-19
MPO	Myeloperoxidase, nuclear gene encoding mitochondrial protein	2.0700	1.62E-10
YOD1	YOD1 OTU deubiquitinating enzyme 1 homolog	2.0600	4.82E-19
DSC2	Desmocollin 2, transcript variant Dsc2b	2.0500	2.80E-14
FECH	Ferrochelatase, nuclear gene encoding mitochondrial protein, transcript variant 2	2.0500	1.43E-14
FCAR	Fc fragment of IgA, transcript variant 9	2.0400	1.83E-16
THBS1	Thrombospondin 1	1.9400	7.96E-13
CRISP3	Cysteine-rich secretory protein 3	1.9100	4.25E-10
PROS1	Protein S	1.8600	4.23E-14
FAM20A	Family with sequence similarity 20, member A	1.8000	1.15E-12
OLR1	Oxidized low density lipoprotein receptor 1	1.8400	7.21E-07
CTNNAL1	Catenin, alpha-like 1	1.8300	9.17E-15
EGF	Epidermal growth factor	1.8300	7.65E-14
MS4A4A	Membrane-spanning 4-domains, subfamily A, member 4, transcript variant 1	1.8200	1.86E-11
GPR84	G protein-coupled receptor 84	2.1900	2.10E-13
SYN2	Synapsin II, transcript variant IIb	1.8200	1.13E-09
F5	Coagulation factor V	1.8100	1.36E-16
MAOA	Monoamine oxidase A, nuclear gene encoding mitochondrial protein	1.8100	1.00E-11
HECW2	HECT, C2, and WW domain containing E3 ubiquitin protein ligase 2	1.7000	1.26E-15
MMRN1	Multimerin 1	1.7900	3.20E-15
OSBP2	Oxysterol-binding protein 2, transcript variant 1	1.7900	7.10E-10
IRAK3	Interleukin-1 receptor-associated kinase 3	1.7600	1.72E-19
SOCS3	Suppressor of cytokine signaling 3	1.7600	4.51E-17
ANXA3	AnnexinA3	1.7500	1.46E-11
HP	Haptoglobin	1.7500	2.31E-10
LTF	Lactotransferrin	1.7500	1.81E-05
BCAT1	Branched chain aminotransferase 1, cytosolic	1.7400	5.94E-16
BMX	BMX non-receptor tyrosine kinase	1.7400	9.83E-17
KCNH7	Potassium voltage-gated channel, subfamily H, member 7, transcript variant 1	1.5400	7.88E-08
TLR5	Toll-like receptor 5	1.7300	2.28E-14
CDC14B	CDC14 cell division cycle 14 homolog B, transcript variant 3	1.7100	5.01E-10
G0S2	G0/G1switch 2	1.7000	6.54E-13
KREMEN1	Kringle-containing transmembrane protein 1, transcript variant 1	1.7400	1.07E-12
HTRA1	HtrA serine peptidase 1	1.7000	2.72E-11
LTBP1	Latent transforming growth factor beta binding protein 1, transcript variant 1	1.6900	6.78E-12
WDFY3	WD repeat and FYVE domain containing 3, transcript variant 1	1.6900	8.39E-15
XK	X-linked Kx blood group	1.6900	4.40E-12
SSH1	Slingshot homolog 1	1.6800	2.97E-15
VNN1	Vanin 1	1.6800	1.28E-11
GRB10	Growth factor receptor-bound protein 10, transcript variant 3	1.6600	4.77E-13
MCEMP1	Mast cell-expressed membrane protein 1	1.5600	7.99E-11
METTL7B	Methyltransferase like 7B	1.6400	8.62E-09
KIF1B	Kinesin family member 1B, transcript variant 1	1.6400	9.63E-20
PGLYRP1	Peptidoglycan recognition protein 1	1.6400	3.13E-13
ST6GALNAC3	ST6-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3	1.6400	1.17E-10
GYPB	Glycophorin B	1.6300	5.18E-07

(Contd...)

TABLE S1. *(Continued)*

Gene symbol	Definition	Log FC	p. adjust
TMEM45A	Transmembrane protein 45A	1.6300	1.65E-09
DAAM2	Dishevelled-associated activator of morphogenesis 2	1.6200	5.95E-06
NFIA	Nuclear factor I/A	1.7500	7.46E-14
DACH1	Dachshund homolog 1, transcript variant 1	1.6200	1.12E-12
TNFAIP6	Tumor necrosis factor, alpha-induced protein 6	1.6200	8.36E-14
SAP30	Sin3-associated polypeptide, 30kDa	1.6100	4.77E-17
FZD5	Frizzled homolog 5	1.6000	6.97E-15
PRTN3	Proteinase 3	1.6000	2.82E-08
KLF5	Kruppel-like factor 5	1.5900	6.36E-10
ARHGAP6	Rho GTPase- activating protein 6, transcript variant 2	1.5800	9.96E-15
WNK1	WNK lysine deficient protein kinase 1	1.5700	3.12E-16
LHFP	Lipoma HMGIC fusion partner	1.5500	1.56E-10
NUPL1	Nucleoporin like 1, transcript variant 3	1.5400	6.04E-16
SLC8A1	Solute carrier family 8, member 1	1.5400	1.43E-10
SULT1B1	Sulfotransferase family, cytosolic, 1B, member 1	1.5400	1.36E-15
OSM	Oncostatin M	1.5100	1.43E-11
PBX1	Pre-B-cell leukemia transcription factor 1	1.5100	3.65E-10
PRSS33	Protease, serine, 33	-1.7900	2.86E-08
PRDM5	PR domain containing 5	1.5100	3.01E-08
RNF182	Ring finger protein 182	1.5400	2.06E-03
RPL36AL	Ribosomal protein L36a-like	-1.5000	7.69E-15
LOC155060	Hypothetical protein LOC155060	-1.5100	8.35E-15
CD22	CD22 antigen	-1.5400	2.82E-14
HLA-DPB1	Major histocompatibility complex, class II, DP beta 1	-1.5400	1.06E-11
VPREB3	Pre-B lymphocyte gene 3	-1.5400	2.81E-13
SNRPD2	Small nuclear ribonucleoprotein D2 polypeptide 16.5kDa, transcript variant 2	-1.5700	2.93E-13
SOX6	SRY-box 6, transcript variant 2	1.5200	1.00E-05
PLXDC1	Plexin domain containing 1	-1.5800	6.20E-15
CDKN1C	Cyclin-dependent kinase inhibitor 1C	-1.6000	2.09E-08
SUCNR1	Succinate receptor 1	2.0800	2.88E-16
IL11RA	Interleukin 11 receptor, alpha, transcript variant 2	-1.6000	7.24E-21
IL23R	Interleukin 23 receptor	-1.6000	1.02E-10
IL24	Interleukin 24, transcript variant 2	-1.6200	5.71E-17
TDRD9	Tudor domain containing 9	2.0200	9.06E-18
CCL23	Chemokine ligand 23, transcript variant CKbeta8-1	-1.6300	5.97E-09
CD72	CD72 antigen	-1.6400	9.33E-17
PDZK1IP1	PDZK1 interacting protein 1	-1.7000	3.37E-11
TMEM56	Transmembrane protein 56	2.2000	1.38E-12
GZMB	Granzyme B	-1.7200	4.33E-11
LRRN3	Leucine rich repeat neuronal 3	-1.7500	1.29E-12
TREML4	Triggering receptor expressed on myeloid cells-like 4	1.8900	3.35E-04
VEPH 1	Ventricular zone expressed PH domain homolog 1	1.6300	6.18E-08
IL5RA	Interleukin 5 receptor, alpha, transcript variant 6	-1.7600	1.50E-10

(Contd...)

TABLE S1. *(Continued)*

Gene symbol	Definition	Log FC	p. adjust
OLIG2	Oligodendrocyte lineage transcription factor 2	-1.7600	8.08E-08
TPM2	Tropomyosin 2, transcript variant 2	-1.7700	1.70E-23
ALOX15	PREDICTED: Arachidonate 15-lipoxygenase	-2.0100	1.87E-14
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1	-2.0200	1.14E-08
TCL1A	T-cell leukemia/lymphoma 1A	-2.4400	2.36E-18

FC: Fold change; DEG: Differentially expressed gene; SJIA: Systemic juvenile idiopathic arthritis.

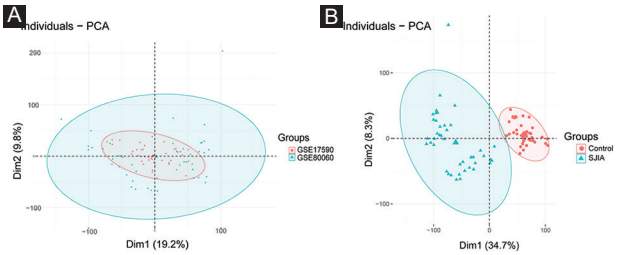


FIGURE S1. (A) Two-dimensional PCA plots show altered distribution of samples after batch removal of dataset GSE80060 and GSE17590. (B) Group difference between SJIA patients and healthy controls. PCA: Principal component analysis.

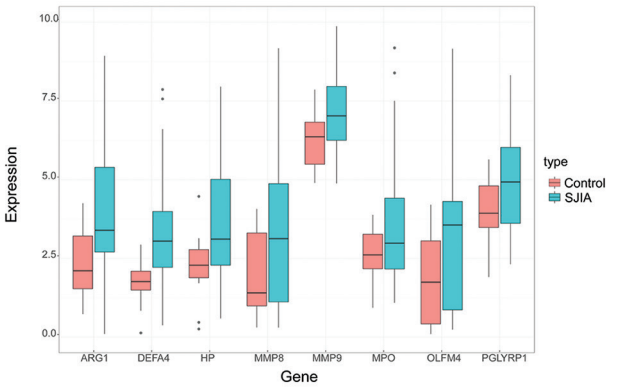


FIGURE S2. Validation of expression level of hub genes in GSE112057 dataset.