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

Yi Ren¹, Hannah Labinsky², Andriko Palmowski³, Henrik Bäcker¹, Michael Müller¹, Arne Kienzle¹,4,*

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. Disease burden is often high 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 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

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

¹Center for Musculoskeletal Surgery, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany
²Department of Internal Medicine, Rheumatology and Immunology, Friedrich-Alexander-University (FAU) Erlangen-Nürnberg and University Hospital Erlangen, Erlangen, Germany
³Department 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
⁴Laboratory of Adaptive and Regenerative Biology, Brigham and Women’s Hospital, Harvard Medical School, Boston MA, USA
⁵Corresponding author: Arne Kienzle, Center for Musculoskeletal Surgery, Clinic for Orthopedics, Charité – Universitätsmedizin Berlin, Charitéplatz 1, Berlin, Germany. E-mail: arne.kienzle@charite.de

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

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

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

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.

©The Author(s) (2021). This work is licensed under a Creative Commons Attribution 4.0 International License
interleukin-1β, 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 SJIA [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 sedimentation rate), and overt disease symptoms. Besides, we also obtained GSE12057, 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 \( \log_{2} \text{fold change (FC)} > 1.5 \). 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 GSE12057 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 \( \geq 50\% \) 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, “ggplot” and “heatmap” 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 \(CD177\), \(OLFM4\), \(ARHGEP12\), \(MMP8\), \(PLOD2\), \(CEACAM6\), and \(CEACAM8\) was highly upregulated \((P \leq 0.001\), Figure 1C). In contrast, \(TCL1A\), \(ALOX15\), and \(HLA-DQB\) were downregulated the most \((P \leq 0.001)\).

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 \(ARG1\), \(HP\), \(MMP8\), \(MPO\), \(OLFM4\), and \(PGLYRP4\) (Table 2). No downregulated genes were found. The GSE120157 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 cut-off (Figure 4A). A total of seven genes were screened out as potential prognostic markers, including $\text{TPM2}$, $\text{PRSS33}$, $\text{LTBP1}$, $\text{GZMB}$, $\text{F5}$, $\text{BMX}$, and $\text{ARHGEF12}$. Using a binary logistic regression model, significant correlation to treatment response was only found for $\text{TPM2}$ ($P = 0.019$; 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 cells and correlation of immune cells with clinical outcome. Immune reactivity was assessed using immunofluorescence staining of CD45, CD8, CD4, CD19, CD3, and CD20. The proportion of immune cells was calculated by counting the number of positive cells in the total number of stained cells. The correlation between immune reactivity and clinical outcome was assessed using Spearman's rank correlation coefficient.

The results showed that the proportion of CD45+ cells was higher in the "poor" clinical outcome group compared to the "good" clinical outcome group. The proportion of CD8+ and CD4+ cells was higher in the "poor" clinical outcome group compared to the "good" clinical outcome group. The proportion of CD19+ cells was higher in the "good" clinical outcome group compared to the "poor" clinical outcome group. The proportion of CD3+ and CD20+ cells was similar in both groups.

The correlation analysis showed a significant positive correlation between the proportion of CD45+ cells and clinical outcome (Spearman's rank correlation coefficient: 0.62, p < 0.01). A significant negative correlation was also found between the proportion of CD8+ and CD4+ cells and clinical outcome (Spearman's rank correlation coefficient: -0.58, p < 0.01). The correlation between CD19+ and clinical outcome was not significant. The correlation between the proportions of CD3+ and CD20+ cells and clinical outcome was not significant.

In conclusion, our study demonstrated that immune reactivity plays an essential role in SJIA. The proportion of CD45+, CD8+, and CD4+ cells is higher in the "poor" clinical outcome group compared to the "good" clinical outcome group. The proportion of CD19+ cells is higher in the "good" clinical outcome group compared to the "poor" clinical outcome group. Immune reactivity is significantly correlated with clinical outcome, indicating its potential as a prognostic marker in SJIA.
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 metalloproteinase-8                                                  | 3.506  | 0.0003  | 0.0095 |
| MMP9   | Matrix metalloproteinase-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; aggregation inhibitor                           |
| 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 |
| Phenanthridinedone   | −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 |

SJIJA: Systemic juvenile idiopathic arthritis

Despite treatment, SJIJA 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].

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 SJIJA [6,11,23]. Neutrophils are major innate immune effector cells and 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 SJIJA [4]. In active SJIJA, 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 IJAs, the prevalence of suppressive CD16+ / CD62L(dim) neutrophils was lower in patients with clinically inert SJIJA suggesting neutrophil heterogeneity and unique molecular pathomechanisms to be involved [6]. Considering the significant role of the immune system in the pathogenesis of SJIJA, we explored the relative prevalence of 22 subtypes of immune cells. Confirming previous results, we found

Prediction of potentially effective drugs against active SJIJA

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), phenanthridinedone (ES = −0.847, p < 0.001), and hycanthone (ES = −0.81, p = 0.003) (Table 4).

DISCUSSION

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 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).
Yi Ren, et al: Molecular pathways and prognostic markers in active SJIA

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 SIIA models and ingenuity pathway analysis suggested

**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.
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]. ARG2, 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 ARG2 has been shown to reduce inflammation by suppressing T-cell proliferation through arginine deprivation [41,42]. Moreover, ARG2 has been shown to be upregulated in autoimmune disease such as rheumatoid arthritis [43]. In SJIA, ARG2 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, DEF4, and PGLYRP3, 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**

[4] Ombrello MJ, Remmers EF, Tachmazidou I, Grom A, Föll 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
[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):375-88. https://doi.org/10.1586/1744666X.2015.1032357

[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.41442

[4] Ter Haar NM, Tak T, Mokry M, Scholman RC, Meerding JW, 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, Paeseler 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, Yasim 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 1995;12(2):245-71.

[8] Omnadel MJ, Arthur VL, Remmers EF, Hinks A, Tachmazidou I, Grom AA, et al. Genetic architecture distinguishes systemic juvenile idiopathic arthritises from other forms of juvenile idiopathic arthritis: Clinical and therapeutic implications. Ann Rheum Dis 2017;76(5):906-13. https://doi.org/10.10039/rheumatology/khr719

[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;11:58533. https://doi.org/10.3389/fimmu.2020.58533

[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 arthritises identified by DNA microarray analysis. Ann Rheum Dis 2009;68(2):264-72. https://doi.org/10.1136/ard.2008.095987

[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 arthritises. Arthritis Rheum 2008;58(3):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 arthritises patients on canakinumab therapy. Arthritis Res Ther 2017;19(1):33. https://doi.org/10.1186/s13075-016-1212-x

[13] Quartier P, Allantaz F, Cimaz R, Pillot 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.2010.124254

[14] Mo A, Mariorta UM, Arafat D, Chan LHK, Ponder L, Jiang SR, et al. Disease-specific regulation of gene expression in a comparative analysis of juvenile idiopathic arthritises and inflammatory bowel disease. Genome Med 2018;10(1):48. https://doi.org/10.1186/s13073-018-0538-x

[15] Parker HS, Leek JT, Favorov AV, Considine M, Xia X, Chavan S, et al. Preservation of biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. Bioinformatics 2014;30(9):1475-63. 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 2015;4:279-93. https://doi.org/10.1038/psp.2015.56

[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. Mod Med Rep 2018;8(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):1209-9. 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, Baidam E, Beresford MW, et al. Mortality rates are increased in patients with systemic juvenile idiopathic arthritis. Arch Dis Child 2013;98(1):206-7. https://doi.org/10.1136/archdischild-2013-310571

[24] Kienzle A, Kurch S, Scholder I, Berges C, Ose R, Schupp J, et al. Prednisolone treatment of juvenile arthritis. J Rheumatology (Oxford) 2009;37(3):488-93. https://doi.org/10.3892/rheumatology/kep441

[25] 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):519-30. https://doi.org/10.1038/nri31024

[26] 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):227-37. https://doi.org/10.1111/imm.12245

[27] Silvestre-Roig C, Hidalgo A, Soehlein O. Neutrophil heterogeneity: Implications for homeostasis and pathogenesis. Blood 2016;127(18):2173-81. https://doi.org/10.1182/blood-2016-01-688887

[28] Taylor PR, Roy S, Meszaros EC, Sun Y, Howell SJ, Malemud CJ, et al. JAK/STAT signal pathway by bioinformatics analysis for miRNA regulation of JAK/STAT signal pathway in systemic juvenile idiopathic arthritis. J Rheumatol 2017;44(6):1178-85. https://doi.org/10.3892/rheumatology/kex441

[29] T aylor PR, Roy S, Meszaros EC, Sun Y , Howell SJ, Malemud CJ, et al. Disease-specific regulation of gene expression in a comparative analysis of juvenile idiopathic arthritises. J Rheumatol 2007;34(5):1133-8.

[30] Li HW, Zeng HS. Regulation of JAK/STAT signal pathway by bioinformatics analysis for miRNA regulation of JAK/STAT signal pathway in systemic juvenile idiopathic arthritis. J Rheumatol 2017;44(6):1178-85. https://doi.org/10.3892/rheumatology/kex441
Related articles published in BJBJMS

1. PCAT1 is a poor prognostic factor in endometrial carcinoma and associated with cancer cell proliferation, migration and invasion Editors’ Choice. Xiaohuan Zhao et al., BJBJMS, 2019

2. Integrated profiling identifies ITGB3BP as prognostic biomarker for hepatocellular carcinoma Qiuli Liang et al., BJBJMS, 2020
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 | 2.0E-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    | Rhoguanine 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-acetylgalcosaminyl transferase 5              | 1.8400 | 5.45E-13   |
| PLOD2       | Procollagen-lysine, 2-oxoglutarate 5-dioxynegase 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   |
| DEFNA       | Defensin, alpha 4, corticotatin                                             | 2.4300 | 7.50E-09   |
| ARG1        | Arginase, liver                                                            | 2.4200 | 4.57E-13   |
| CACNA1E     | Calcium channel, voltage dependent, alpha 1 subunit                        | 2.3700 | 2.23E-11   |
| MMP9        | Matrix metallopeptidase 9                                                  | 2.3700 | 7.98E-20   |
| CA1         | Carbonic anhydrase 1                                                        | 2.3000 | 2.30E-10   |
| CD274       | CD274 antigen                                                               | 2.0200 | 1.09E-09   |
| MYO10       | Myosin X                                                                   | 2.3000 | 1.12E-10   |
| AZU1        | Azurin 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 Dac2b                                     | 2.0500 | 2.80E-14   |
| FECH        | Ferrichelatase, nuclear gene encoding mitochondrial protein, transcript 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   |

(Contd...)
TABLE S1. (Continued)

| Gene symbol | Definition | Log FC | p. adjust |
|-------------|------------|--------|-----------|
| TMEM45A     | Transmembrane protein 45A          |  1.6300 | 1.65E-09 |
| DAAV2       | 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       | S3-associated polypeptide, 30kDa   |  1.6100 | 4.77E-17 |
| FZD5        | Frizzled homolog 5                 |  1.6000 | 6.97E-15 |
| PRN3        | 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 |
| NUP1L1      | Nucleoporin like 1, transcript variant 3 |  1.5400 | 6.04E-16 |
| SLC8A1      | Solute carrier family 8, member 1 |  1.5400 | 1.43E-10 |
| SULT1B1     | Sulfoxidase family, cytosolic, 1B, member 1 |  1.5400 | 1.36E-15 |
| OSM         | Oncostatin M                       |  1.5100 | 1.43E-11 |
| PRX1        | Pre-B-cell leukemia transcription factor 1 |  1.5100 | 3.65E-10 |
| PRSS3       | Protease, serine, 33               |  1.5000 | 2.86E-08 |
| PRDM5       | PR domain containing 5             |  1.5000 | 3.01E-08 |
| RNF182      | Ring finger protein 182            |  1.5000 | 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-DRB1    | 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 16.3kDa, 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 Ckbetaab-1 |  1.6300 | 5.97E-09 |
| CD72        | CD72 antigen                       |  1.6400 | 9.33E-17 |
| PDZK1P1     | 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       | Lysine rich repeat neuronal 5      |  1.7500 | 1.29E-12 |
| TREML4      | Triggering receptor expressed on myeloid cells-like 4 |  1.8900 | 3.35E-04 |
| VEPH1       | 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...)