



Diagnostic Value of Immune-Related Genes in Kawasaki Disease

Dong Liu^{1,2,3}, Meixuan Song⁴, Fengchuan Jing^{1,3}, Bin Liu² and Qijian Yi^{1,3*}

¹Department of Cardiovascular Medicine, Ministry of Education Key Laboratory of Child Development and Disorders, National Clinical Research Center for Child Health and Disorders, China International Science and Technology Cooperation Base of Child Development and Critical Disorders, Children's Hospital of Chongqing Medical University, Chongqing, China, ²Department of Pediatrics, Sichuan Clinical Research Center for Birth Defects, The Affiliated Hospital of Southwest Medical University, Luzhou, China, ³Chongqing Key Laboratory of Pediatrics, Chongqing, China, ⁴Department of Gastrointestinal Surgery, The Affiliated Hospital of Southwest Medical University, Luzhou, China

OPEN ACCESS

Edited by:

Lam Cheung Tsoi,
University of Michigan, United States

Reviewed by:

Lv Haitao,
Children's Hospital of Soochow
University, China
Kenichi Shimada,
Cedars Sinai Medical Center,
United States
Lijian Xie,
Shanghai Children's Hospital, China

*Correspondence:

Qijian Yi
qjyi2003@allyun.com

Specialty section:

This article was submitted to
Applied Genetic Epidemiology,
a section of the journal
Frontiers in Genetics

Received: 24 August 2021

Accepted: 11 November 2021

Published: 10 December 2021

Citation:

Liu D, Song M, Jing F, Liu B and Yi Q
(2021) Diagnostic Value of Immune-
Related Genes in Kawasaki Disease.
Front. Genet. 12:763496.
doi: 10.3389/fgene.2021.763496

Kawasaki disease (KD) is a systemic vasculitis that predominantly damages medium- and small-sized vessels, and mainly causes coronary artery lesions (CALs). The diagnostic criterion of KD mainly depends on clinical features, so children could be easily misdiagnosed and could suffer from CALs. Through analysis, a total of 14 immune-related DEGs were obtained, of which *IL1B*, *ADM*, *PDGFC*, and *TGFA* were identified as diagnostic markers of KD. Compared with the non-KD group, KD patients contained a higher proportion of naive B cells, activated memory CD4 T cells, gamma delta T cells, and neutrophils, while the proportions of memory B cells, CD8 T cells, activated memory CD4 T cells, and activated NK cells were relatively lower. In conclusion, immune-related genes can be used as diagnostic markers of KD, and the difference in immune cells between KD and non-KD might provide new insight into understanding the pathogenesis of KD.

Keywords: Kawasaki disease, immune-related genes, immune cells, diagnostic, CIBERSORTx

INTRODUCTION

Kawasaki disease (KD) is a systemic vasculitis that predominantly damages medium- and small-sized vessels, and mainly causes coronary artery lesions (CALs). KD has become the main cause of acquired heart disease in children; 80% of these cases occur in children aged 6 months to 5 years (Makino et al., 2015), and timely diagnosis and treatment with intravenous immunoglobulin (IVIG) have reduced CALs by about 84% (McCrindle et al., 2017). However, some patients, especially infants under 6 months of age and adolescents, did not meet the criterion of KD, called "incomplete KD", and it is difficult to make a diagnosis in the early stage of the disease. If there is a special index for KD, the early diagnosis will be easy. On the other hand, delayed diagnosis or missed diagnosis may also cause excessive drug treatment or invasive procedures in KD children. Therefore, an efficient diagnostic index of KD is pursued.

Diagnosis guidelines of KD provided by AHA and the Japanese Ministry of Health are mainly based on clinical symptoms and signs (Ayusawa et al., 2005; McCrindle et al., 2017). Some laboratory data, such as erythrocyte sedimentation rate (ESR), C-reactive protein (CRP), albumin and N-terminal pro-B-type natriuretic peptide (NT-proBNP) level, thrombocytosis, leukocytosis, raised transaminases, hyponatremia, and pyuria, may be helpful in the diagnosis of KD (Cho et al., 2011; Lin et al., 2015; Reddy et al., 2016; Dionne et al., 2017), but all of them are yet to find the same change in other immune and infectious disease, and cannot act as special diagnostic indexes for KD.

The pathogenesis of KD have not been clear so far. However, studies have demonstrated that both immune cells and immune-related genes are involved in the pathogenesis of KD (Sánchez-Manubens et al., 2014; Sakurai et al., 2019). Evaluating the difference of immune cell composition in KD from the perspective of the immune system may have great value in elucidating the molecular mechanism and developing new diagnostic algorithms of KD. Bioinformatics has now been widely used in revealing the molecular mechanism of diseases. CIBERSORTx is an analysis tool to infer the expression of immune cells and obtain various immune cell proportions from samples based on gene expression or sequencing data (Newman et al., 2019).

In this study, we downloaded the microarray datasets of KD from Gene Expression Omnibus (GEO) database and performed differential expression immune-related gene analysis; least absolute shrinkage and selection operator (Lasso) regression algorithm and support vector machine-recursive feature elimination (SVM-RFE) algorithm were used to further screen the diagnostic markers of KD. Based on the diagnostic markers, we have developed and verified a diagnostic model. Subsequently, CIBERSORTx was used to assess the relative content of 22 kinds of immune cell subsets in whole blood in children with KD. In addition, the relationship between diagnostic markers and immune cells was analyzed to better understand the molecular immune mechanism of KD.

MATERIALS AND METHODS

Microarray Data

The raw data of gene chip of GSE73461, GSE68004, GSE18606, GSE73463, and GSE63881 were downloaded from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) database. GSE73461 contains 459 samples (all children <17 years of age), including acute KD ($n = 78$), healthy ($n = 55$), bacterial infection ($n = 52$), viral infection ($n = 94$), juvenile idiopathic arthritis ($n = 66$), Henoch-Schönlein purpura ($n = 18$), and infections of uncertain bacterial or viral etiology ($n = 96$). GSE68004 contains 162 samples, and its raw data were divided into two parts, part 1 including acute KD ($n = 62$), healthy ($n = 14$), adenovirus ($n = 9$), and group A *streptococcus* ($n = 16$), and part 2 including acute KD ($n = 27$), healthy ($n = 23$), adenovirus ($n = 10$), and group A *streptococcus* ($n = 1$). GSE18606 contains 48 samples, including acute KD ($n = 20$) and healthy ($n = 9$). GSE73463 contains 233 samples, namely, acute KD ($n = 146$) and convalescent KD ($n = 87$). GSE63881 contains 341 samples, namely, acute KD ($n = 171$) and convalescent KD ($n = 170$). The GSE73461 was used as the training set; GSE68004, GSE18606, GSE73463, and GSE63881 were used as the independent validation set.

Data Preprocessing and Identification of Differentially Expressed Genes (DEGs)

Data analysis used R software (version 4.0.2, <https://www.r-project.org/>) and bioconductor packages (<http://www.bioconductor.org/>). Raw data of GSE73461, GSE68004,

GSE18606, GSE73463, and GSE63881 datasets were read, background correction and data normalization through the “limma” package (<http://www.bioconductor.org/packages/release/bioc/html/limma.html>). Considering the possibility of information loss during the removing batch difference processing, the datasets have not been merged. Linear models of “limma” package was used to screen DEGs of GSE73461 by comparing the expression values between KD and non-KD, and $|\log_2FC| > 1$ and adjusted $p < 0.01$ were considered statistically significant. The heatmap (<https://bioconductor.org/packages/release/bioc/html/heatmaps.html>) and ggplot2 (<https://cran.r-project.org/web/packages/ggplot2/index.html>) packages were used to draw a heatmap and volcano map to show the differential expression of DEGs.

Immune-related genes were downloaded from the Immunology Database and Analysis Portal database (ImmPort, <https://www.immport.org/shared/genelists>), in which 1,509 genes contain 17 immune categories based on molecular function (Bhattacharya S et al., 2014). The immune-related DEGs were obtained by the DEGs and immune-related genes were overlapped.

Functional Analysis of Immune-Related Differentially Expressed Genes

To assess the potential biologic functions of immune-related DEGs, Disease Ontology (DO), Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed by the “clusterProfiler” package (Yu G et al., 2012), respectively. Gene set enrichment analysis (GSEA) was performed on the all immune-related genes expression matrix by the “clusterProfiler” package. $p < 0.05$ was considered statistically significant.

Screening and Verification of Diagnostic Markers

Based on immune-related DEGs, candidate diagnostic markers for KD were selected through integrated analysis of two algorithms including Lasso logistic regression (Tibshirani R et al., 1996) with penalty conducted by 10-fold cross-validation and SVM-RFE (Huang M et al., 2014). A logistic regression model was used to construct a diagnostic signature based on the candidate immune-related genes. Expression matrices of the GSE68004, GSE18606, GSE73463, and GSE63881 were used as independent datasets to verify each candidate gene and the diagnostic model. Receiver operating characteristic (ROC) curves were used to evaluate the accuracy and efficiency of the candidate genes and the diagnostic model.

Evaluation of Immune Cell Level

CIBERSORTx (Newman AM et al., 2019) algorithm was used to quantify the proportions of immune cells in the samples of GSE73461. This algorithm contains 22 kinds of immune cells. Barplot and violin diagrams were drawn to visualize the differences of immune cell in different diagnostic samples;

TABLE 1 | Demographic features of each dataset.

Dataset	Diagnosis	Age (years)			Gender		χ^2	p^a
		Age	F	p^a	Male	Female		
GSE73461	KD	<17	—	—	43 (9.37%)	35 (7.63%)	0.016	0.900
	Bacterial infection	<17	—	—	22 (4.79%)	30 (6.54)	—	—
	Viral infection	<17	—	—	66 (14.38%)	28 (6.10%)	—	—
	Juvenile idiopathic arthritis (JIA)	<17	—	—	25 (5.45%)	41 (8.93%)	—	—
	Henoch-Schönlein purpura (HSP)	<17	—	—	9 (1.96%)	9 (1.96%)	—	—
	Uncertain	<17	—	—	62 (13.51%)	34 (7.41%)	—	—
GSE68004	Healthy	<17	—	—	29 (6.32%)	26 (5.66%)	—	—
GSE68004	Complete KD	3.68 ± 2.50	15.889	<0.001	32 (19.75%)	44 (27.16%)	0.103	0.748
	Incomplete KD	4.65 ± 3.52	—	—	9 (5.56%)	4 (2.47)	—	—
	Adenovirus (HAdV)	4.14 ± 2.48	—	—	10 (6.17%)	9 (5.56%)	—	—
	Group A streptococcal disease (GAS)	5.46 ± 4.12	—	—	10 (6.17%)	7 (4.32%)	—	—
	Healthy	7.14 ± 4.62	—	—	15 (9.26%)	22 (13.58%)	—	—
GSE18606	KD	2.98 ± 2.13	1.121	0.299	10 (34.48%)	10 (34.48%)	0.697	0.404
	Healthy	2.11 ± 1.46	—	—	3 (10.34%)	6 (20.69%)	—	—
GSE73463	Acute KD	<17	—	—	60 (25.75%)	86 (36.91%)	0.684	0.408
	Convalescent KD	<17	—	—	31 (13.30%)	56 (24.03%)	—	—
GSE63881	Acute KD	3.32 ± 2.83	0.001	0.975	69 (20.23%)	102 (29.91%)	0.002	0.964
	Convalescent KD	3.32 ± 2.84	—	—	69 (20.23%)	101 (29.62%)	—	—

^aComparison between the KD group and the non-KD group, or between the acute KD group and the convalescent KD group.

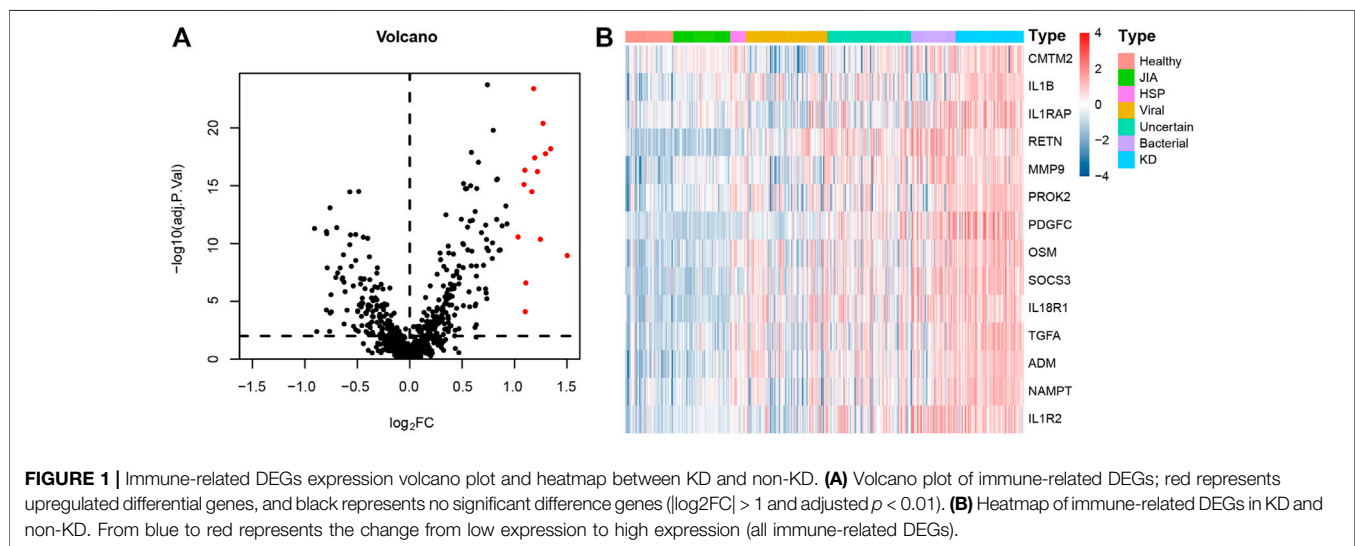


FIGURE 1 | Immune-related DEGs expression volcano plot and heatmap between KD and non-KD. **(A)** Volcano plot of immune-related DEGs; red represents upregulated differential genes, and black represents no significant difference genes ($|\log_2FC| > 1$ and adjusted $p < 0.01$). **(B)** Heatmap of immune-related DEGs in KD and non-KD. From blue to red represents the change from low expression to high expression (all immune-related DEGs).

correlation heatmap was drawn to visualize the correlation of immune cells.

Correlation Analysis Between Diagnostic Markers and Immune Cells

Spearman correlation analysis on diagnostic markers and immune cells was performed, and the analysis results were visualized.

Statistical Analysis

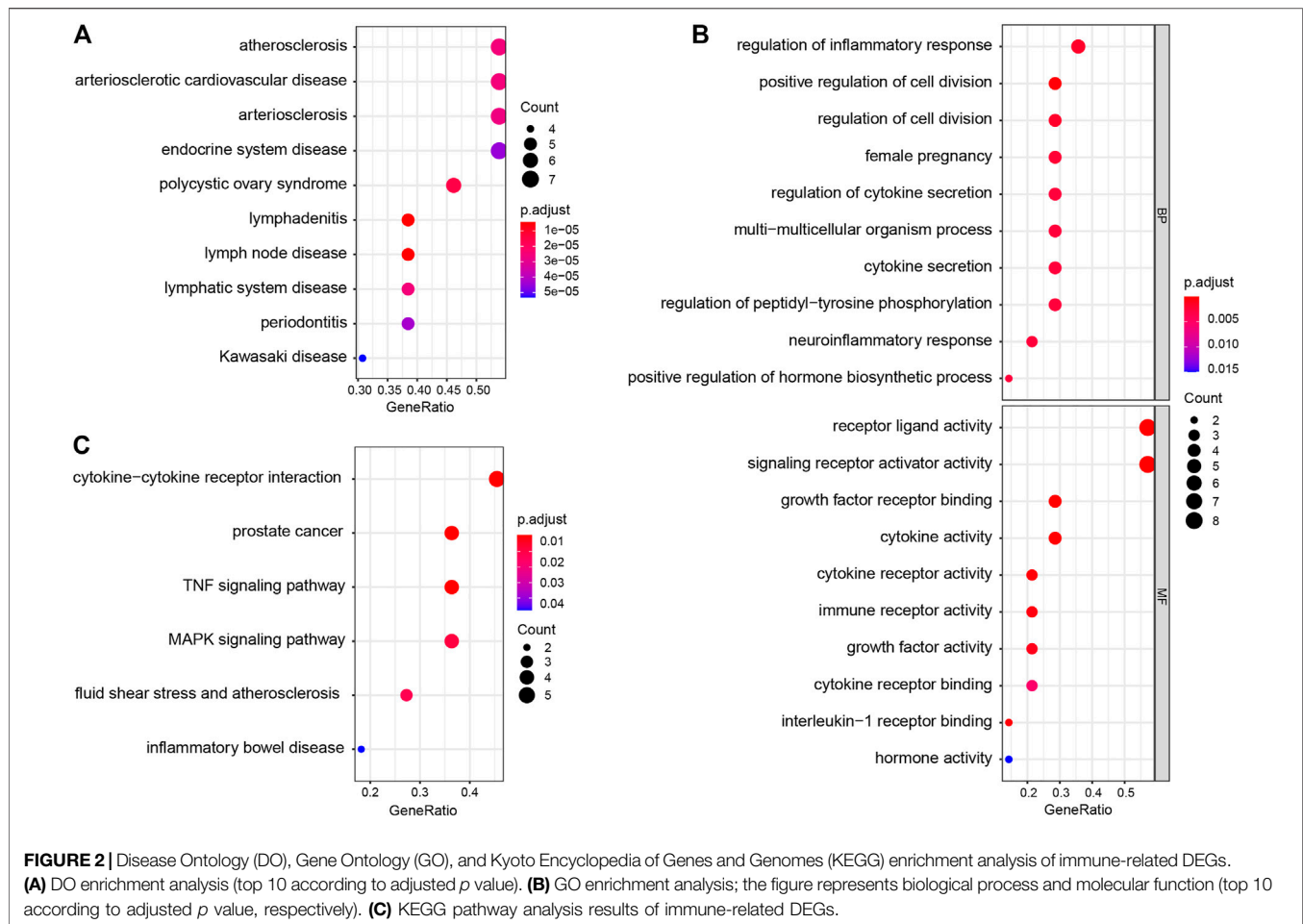
Statistical analyses were completed using the SPSS 22.0 software for Windows (SPSS, Chicago). All values are shown as mean ± standard deviation, or number and percentage (n , %). To

compare the differences between groups, Student's t -tests were used for Continuity variables. Chi-square test was used to compare frequencies between groups. A two-tailed p value < 0.05 was used as a threshold for determining statistical significance.

RESULTS

Demographic Features and Differentially Expressed Genes Screening in the Training Set

The numbers of children in each diagnostic category and demographic features are summarized in **Table 1**. A total of 100



DEGs were identified between 78 KD and 381 non-KD samples in the training set with the cutoff criteria of $|\log_2FC| > 1$ and adjusted $p < 0.01$. Fourteen immune-related DEGs were obtained by intersections with immune-related genes, as shown in the volcano plot and heatmap (**Figure 1**). All the screened genes were upregulated.

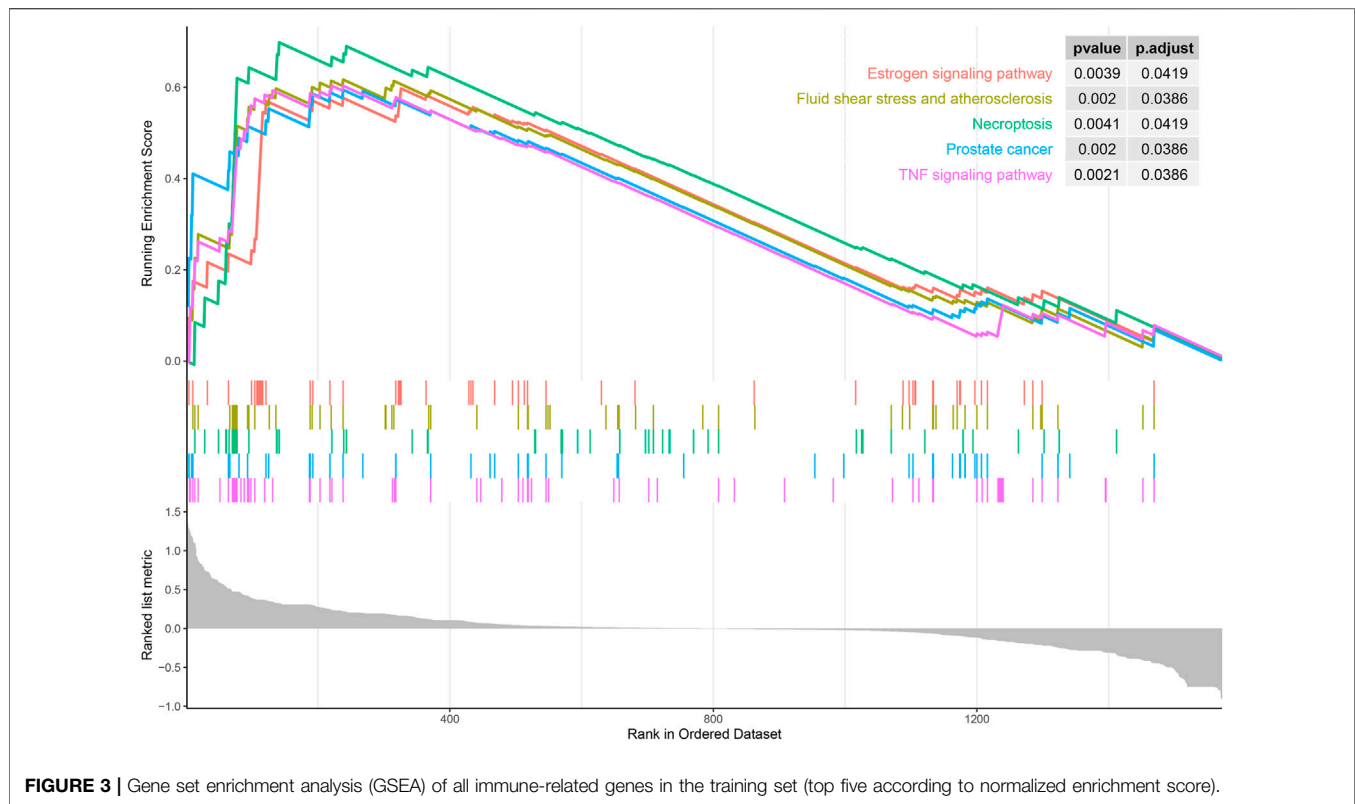
Functional Correlation Analysis of Immune-Related Differentially Expressed Genes

DO enrichment analysis showed that the immune-related DEGs were mainly involved atherosclerosis, arteriosclerotic cardiovascular disease, arteriosclerosis, endocrine system disease, polycystic ovary syndrome, lymphadenitis, lymph node disease, lymphatic system disease, periodontitis, and Kawasaki disease (**Figure 2A**). GO analysis revealed that mainly functional categories in the biological processes (BP) involved regulation of inflammatory response, positive regulation of cell division, regulation of cell division, female pregnancy, regulation of cytokine secretion, multi-multicellular organism process, cytokine secretion, regulation of peptidyl-tyrosine phosphorylation, neuroinflammatory response, and positive regulation of hormone biosynthetic process; for molecular

function (MF), the main terms involved receptor ligand activity, signaling receptor activator activity, growth factor receptor binding, cytokine activity, cytokine receptor activity, immune receptor activity, growth factor activity, cytokine receptor binding, interleukin-1 receptor binding, and hormone activity; for cellular components (CC), no term was enriched at the same cutoff criteria of BP and MF (**Figure 2B**). KEGG pathway analysis revealed that the immune-related DEGs mainly involved cytokine-cytokine receptor interaction, prostate cancer, TNF signaling pathway, MAPK signaling pathway, fluid shear stress and atherosclerosis, and inflammatory bowel disease (**Figure 2C**). GSEA analysis showed that the enriched pathways mainly involved TNF signaling pathway, prostate cancer, necroptosis, fluid shear stress and atherosclerosis, estrogen signaling pathway, IL-17 signaling pathway, NOD-like receptor signaling pathway, neuroactive ligand-receptor interaction, osteoclast differentiation, and Toll-like receptor signaling pathway (**Figure 3**).

Screening and Verification of Diagnostic Markers

Four genes were identified from immune-related DEGs by using the Lasso algorithm (**Figures 4A,B**), and twelve genes were



selected by using the SVM-RFE algorithm (Figure 4C). After overlapping the gene markers obtained by the two algorithms, four candidate diagnosis-related genes were obtained (Figure 4D). In order to further test the diagnostic efficacy of the four candidate genes, a logistic regression model was established (Table 2), and validation was performed through the GSE68004, GSE18606, GSE73463, and GSE63881 datasets. The area under the curve (AUC) of ROC for the logistic regression model was 0.882 in the training set (Figure 5A), and reached a high level in all validation sets (Figures 5B–F).

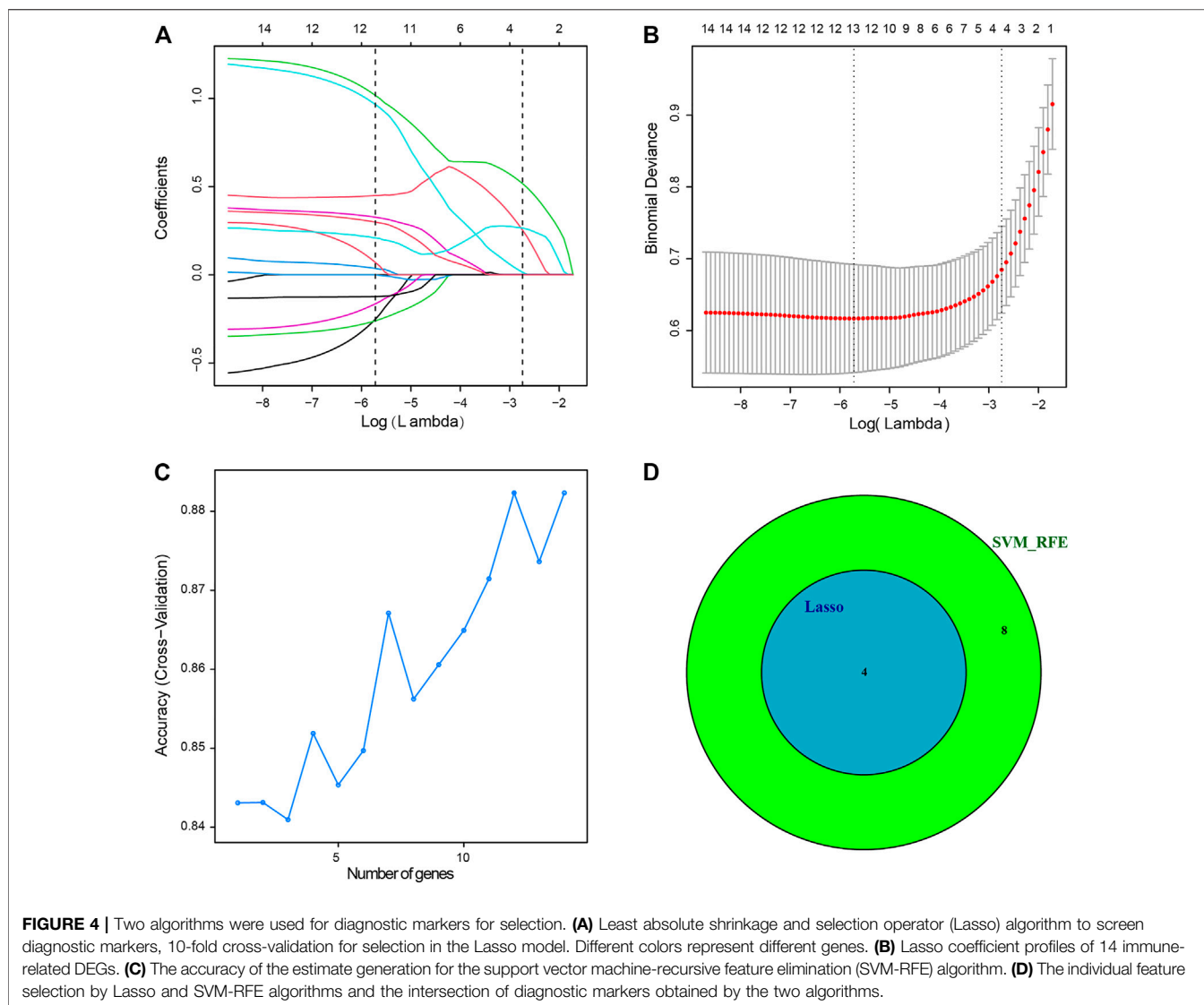
Immune Cell Analyses Results

Using CIBERSORTx algorithm, we first investigated the differences of 22 immune cell subsets between KD and non-KD whole blood samples; results from different diagnostic classifications were averaged (Figure 6A). The correlation plot of 18 immune cells (4 types of immune cells had an estimated abundance of 0) in training set KD samples showed that CD8 T cells had a significant low positive correlation with resting NK cells ($r = 0.426$, $p = 0.001$), a significant low negative correlation with macrophages M0 ($r = -0.378$, $p < 0.001$), and a significant moderate negative correlation with monocytes, activated dendritic cells, and neutrophils ($r = -0.561$, -0.591 , and -0.624 ; all $p < 0.001$). Regulatory T cells (Tregs) had a significant low negative correlation with monocytes ($r = -0.318$, $p = 0.003$). Resting NK cells had a low negative correlation with monocytes and activated dendritic cells ($r = -0.462$ and -0.432 , $p = 0.008$ and <0.001). Activated dendritic cells had a low positive correlation with monocytes and resting

mast cells ($r = -0.309$ and 0.386 , $p = 0.042$ and <0.001) (Figure 6B). Compared with non-KD, whole blood of KD patients contained a higher proportion of naive B cells, activated memory CD4 T cells, gamma delta T cells, monocytes, M0 macrophages, activated dendritic cells, activated mast cells, and neutrophils, while the proportions of memory B cells, CD8 T cells, activated memory CD4 T cells, activated NK cells, M1 and M2 macrophages, and resting mast cells were relatively lower (Figure 6C).

Correlation Analysis Between Diagnostic Markers and Immune Cells in Kawasaki Disease

Correlation analysis showed that *IL1B* was positively correlated with neutrophils, activated mast cells, and M0 macrophages, and negatively correlated with CD8 T cells, resting mast cells, activated NK cells, and naive CD4 T cells (Figure 7A). *ADM* was positively correlated with neutrophils, gamma delta T cells, M0 macrophages, monocytes, and activated mast cells, and negatively correlated with CD8 T cells, resting NK cells, and naive CD4 T cells (Figure 7B). *PDGFC* was positively correlated with monocytes, gamma delta T cells, activated dendritic cells, M0 macrophages, and neutrophils, and negatively correlated with CD8 T cells, resting NK cells, and regulatory T cells (Tregs) (Figure 7C). *TGFA* was positively correlated with neutrophils, gamma delta T cells, M0 macrophages, and monocytes, and negatively correlated with CD8 T cells, regulatory T cells (Tregs), naive CD4 T cells, and resting NK cells (Figure 7D).

**TABLE 2 |** Diagnostic markers for KD.

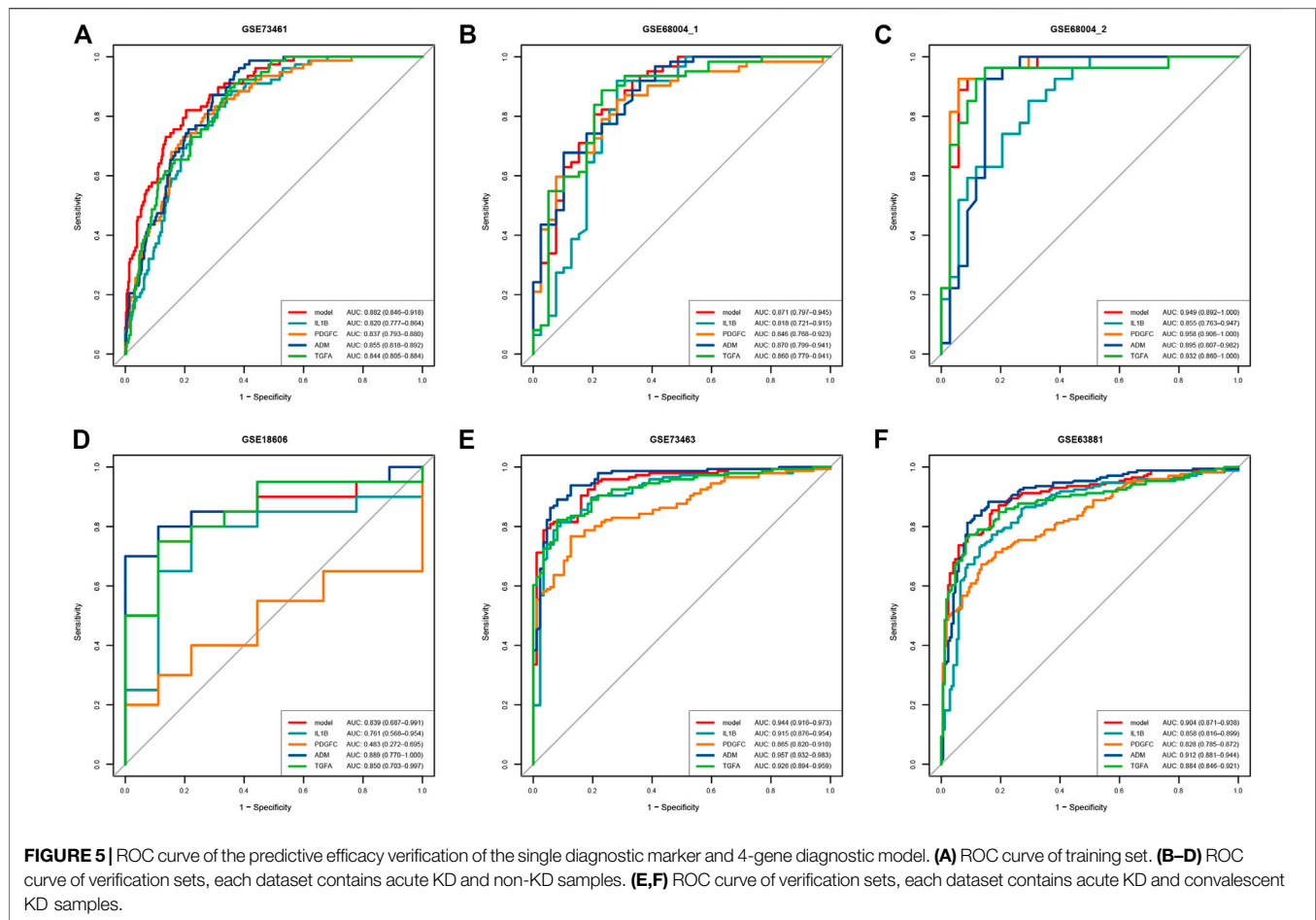
Intercept and genes	β	Odds ratio (95% CI)	p
Intercept	-23.5334	—	—
<i>IL1B</i>	0.7655	2.150 (1.192–4.019)	0.013
<i>ADM</i>	0.6158	1.851 (0.785–4.555)	0.169
<i>PDGFC</i>	0.7697	2.159 (1.503–3.161)	<0.001
<i>TGFA</i>	0.2468	1.280 (0.736–2.238)	0.383

DISCUSSION

The diagnostic criterion of KD mainly depends on clinical symptoms and signs, but some patients, especially infants under 6 months old and adolescents, did not meet the criterion in the early stage of disease, and thus it is difficult to make a timely diagnosis. Delayed diagnosis and treatment may cause serious adverse outcomes such as CAAs. Abnormal gene

expression level has been studied in diagnosis of KD (Kuo et al., 2016; Jaggi et al., 2018; Wright et al., 2018). However, these studies proposed that biomarkers were too many or not further validated by the external cohort. In this study, we developed a simple and concise diagnosis model, composed of four genes, by analyzing expression levels of immune-related genes. Verification of this model was performed by three other external datasets including KD and non-KD, and shows a great capacity for diagnosis of KD. At the same time, this model was validated with two datasets including acute KD and convalescent KD.

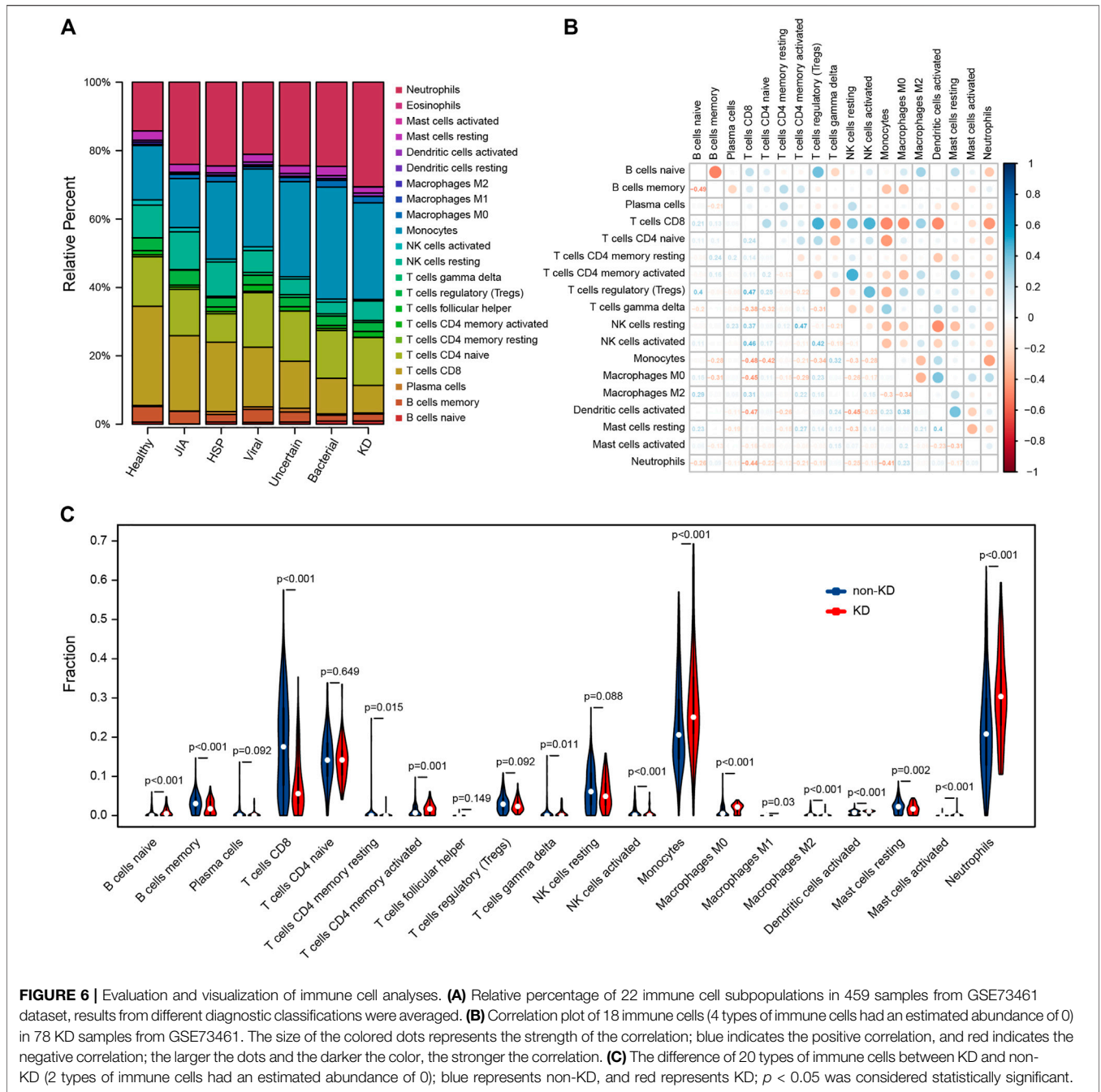
The training set was used to identify 14 immune-related DEGs. DO enrichment analysis showed that the DEGs were mainly related to KD, atherosclerosis, arteriosclerosis, and arteriosclerotic cardiovascular disease. This enrichment result is consistent with previous studies that have shown an increased risk of early atherosclerosis in children with KD (Cho et al., 2014; Zhang et al., 2016). In addition, both KEGG and GSEA were enriched in fluid shear stress and atherosclerosis,



and tumor necrosis factor (TNF) signaling pathway. Masaki et al. (Xie et al., 2018). revealed that gene polymorphisms of TNF may affect KD susceptibility, and TNF- α is associated with the development of CALs in KD patients (Hu et al., 2017; Guo et al., 2020). The results indicated that the analysis results of this study are accurate. The etiology of KD is considered to have an infectious trigger in a genetically susceptible individual, followed by the activation of the immune system (Yim et al., 2013). The main trigger pathogens suspected were the viral agents (Usta Guc et al., 2008; Alexoudi et al., 2011; Chang et al., 2014; Shulman et al., 2015; Kim et al., 2016); some studies have linked triggers to bacteria agents (Galeotti et al., 2010; Alexoudi et al., 2011). However, in our study, the levels of immune-related differentially expressed genes (DEGs) in bacterial infection samples were more similar to those of KD than viral infection samples. It is speculated that the bacteria triggering may be the main factor of the activation of the KD immune system.

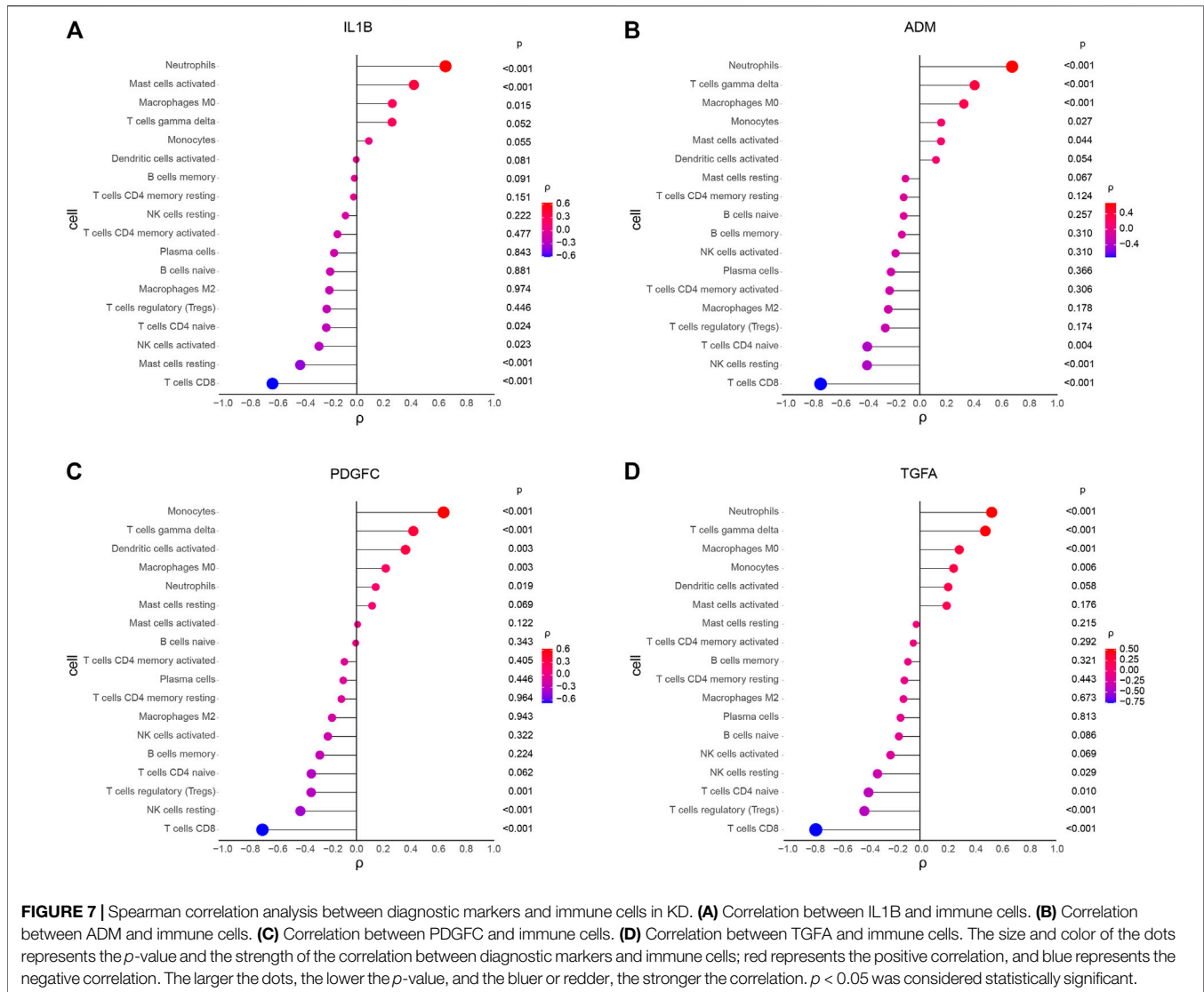
Based on immune-related DEGs and machine learning algorithms, Lasso and SVM-RFE were performed to screen KD diagnostic markers. By integrating the features of the two algorithms, *IL1B*, *ADM*, *PDGFC*, and *TGFA* were selected as diagnostic biomarkers. Although the single biomarker for diagnosis and prognosis has been widely reported, value and robustness are also a major concern. So, we further developed a

logistics prediction model with these 4 markers, and further external verification shows that the model was reliable. In our validation sets, two sets only contain acute KD and convalescent KD, but both diagnostic markers and model showed well differentiation. Interleukin-1 β (IL-1 β) is a potential inflammation-causing cytokine released predominantly by immune-derived cells. Studies indicate that IL-1 β has a significant role in several aspects of vascular inflammation (Chamberlain et al., 2006); there is a high level of IL-1 β in the endothelium of atherosclerotic coronary arteries (Chibana et al., 2017) and might be one of the targets of atherosclerotic protective therapy (Ridker et al., 2017). Similarly, there are reports that the gene polymorphisms of IL-1 β may be associated with initial IVIG treatment failure (Weng et al., 2010) and significantly impact the risk of CAL formation in children with KD (Fu et al., 2019). IL-1 β may also increase the production of neutrophil extracellular traps by activating neutrophils, thereby inducing the progression of coronary artery ectasia (Guo et al., 2020). Adrenomedullin (*ADM*) is a vasoactive peptide involved in vasodilation and regulation of endothelial function. *ADM* inhibits the apoptosis of cultured endothelial cells and the high permeability of human vascular smooth muscle cells induced by agonists (Hippenstiel et al., 2002). Plasma *ADM* levels can reflect the



degree of endothelial damage in atherosclerosis patients (Bunton et al., 2004). It is a good indicator of prognosis in patients with coronary artery disease (Wong et al., 2012). In KD, the expression level of *ADM* gene in plasma has been reported to be elevated (Nishida et al., 2001; Nomura et al., 2005); plasma *ADM* levels were higher in KD patients developed CAAs than in those did not (Nishida K et al., 2001). Platelet-derived growth factors (PDGFs) play an important role in vascular pathologies such as atherosclerosis, restenosis, and aortic aneurysm. PDGF-C can promote the migration and proliferation of macrophages

(Boor et al., 2010), endothelial cells (Li et al., 2005), and vascular smooth muscle cells (Crawford et al., 2009). Karvinen et al. (2009) have revealed that PDGF-C is expressed in different stages of atherosclerosis. However, as a new member of the PDGF family, PDGF-C has been less studied in cardiovascular diseases since its discovery. However, *PDGFC* failed in GSE18606 when independently distinguishing KD from non-KD, which may be related to the small sample size of this dataset. Transforming growth factor α (TGF- α) as a member of the epidermal growth factor (EGF) family has been associated with cerebrovascular diseases (Dai et al., 2020) and



pulmonary vascular disease (Le Cras et al., 2003; Bouzina et al., 2018). Similarly, TGF- α has been less studied in cardiovascular disease.

CIBERSORTx evaluation showed that the abundance of naive B cells, activated memory CD4 T cells, gamma delta T cells, monocytes, M0 macrophages, activated dendritic cells, activated mast cells, and neutrophils was increased, and the abundance of memory B cells, CD8 T cells, activated memory CD4 T cells, activated NK cells, M1 and M2 macrophages, and resting mast cells was decreased. Studies have shown that, CD8 T cells, neutrophils, IgA plasma cells, and macrophages infiltrate and destroy the internal elastic lamina of medium-sized vessels (Takahashi et al., 2005; Alexoudi et al., 2011; Orenstein et al., 2012). However, compared to CD8 T cells, CD4 T cells predominate in peripheral blood (Alexoudi et al., 2011), which may be due to the CD8 T-cell transmission to tissues during the development of KD. In the acute phase of KD, the increase in the number of circulating neutrophils and

activation eventually cause endothelial dysfunction (Sakurai et al., 2019). In addition, our results revealed that CD8 T cells are closely related to resting NK cells, macrophages M0, monocytes, activated dendritic cells, and neutrophils; regulatory T cells (Tregs) are closely related to monocytes; resting NK cells are closely related to monocytes and activated dendritic cells; and activated dendritic cells are closely related to monocytes and resting mast cells. Considering the complexity of the immune system, it is difficult to study the immune cells of KD. Our analysis provides reference for the study of immune response in the pathogenesis of KD.

We used novel scientific algorithms to identify the diagnostic markers for KD and verified them with multiple external datasets. CIBERSORTx was used to analyze the proportion of various immune cells in KD whole blood for the first time. Although the diagnostic model in this study is composed of only four genes with significant diagnostic impact, which makes it handy to use in clinical practice, it

still has some limitations. Firstly, our research is the second mining and analysis of previously published datasets, and the sample size of two validation sets is small. Secondly, CIBERSORTx analysis is based on limited genetic data to infer cell-type abundance and cell-type-specific gene expression, which may deviate from the actual situation. In addition, the reliability of the results of our study needs further experiments to verify.

CONCLUSION

In the present study, 4 immune-related DEGs were identified, and a diagnostic model for KD was established and verified. The biological functions and pathways of the immune-related genes provide a detailed molecular mechanism for understanding the pathogenesis of KD. By CIBERSORTx algorithms, we found a difference in immune cells between KD and non-KD. These immune cells may play a key role in the pathogenesis and development of KD. Further analysis of the relation between immune-related DEGs and immune cells will be helpful to further explore the pathogenesis of KD and may determine the immunotherapy targets of KD (Bajolle et al., 2014).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. These data can be found here: <https://www.ncbi.nlm.nih.gov/gds/> (Gene Expression Omnibus) GSE73461, GSE68004, GSE18606, GSE73463, and GSE63881.

REFERENCES

- Alexoudi, I., Kanakis, M., Kapsimali, V., and Vaiopoulos, G. (2011). Kawasaki Disease: Current Aspects on Aetiopathogenesis and Therapeutic Management. *Autoimmun. Rev.* 10, 544–547. doi:10.1016/j.autrev.2011.04.005
- Ayusawa, M., Sonobe, T., Uemura, S., Ogawa, S., Nakamura, Y., Kiyosawa, N., et al. (2005). Revision of Diagnostic Guidelines for Kawasaki Disease (The 5th Revised Edition). *Pediatr. Int.* 47, 232–234. doi:10.1111/j.1442-200x.2005.02033.x
- Bajolle, F., Meritet, J.-F., Rozenberg, F., Chalumeau, M., Bonnet, D., Gendrel, D., et al. (2014). Markers of a Recent Bocavirus Infection in Children with Kawasaki Disease: "A Year Prospective Study". *Pathologie Biologie* 62, 365–368. doi:10.1016/j.patbio.2014.06.002
- Bhattacharya, S., Andorf, S., Gomes, L., Dunn, P., Schaefer, H., Pontius, J., et al. (2014). ImmPort: Disseminating Data to the Public for the Future of Immunology. *Immunol. Res.* 58, 234–239. doi:10.1007/s12026-014-8516-1
- Boor, P., van Roeyen, C. R. C., Kunter, U., Villa, L., Bücher, E., Hohenstein, B., et al. (2010). PDGF-C Mediates Glomerular Capillary Repair. *Am. J. Pathol.* 177, 58–69. doi:10.2353/ajpath.2010.091008
- Bouzina, H., and Rådegran, G. (2018). Low Plasma Stem Cell Factor Combined with High Transforming Growth Factor- α Identifies High-Risk Patients in Pulmonary Arterial Hypertension. *ERJ. Open Res.* 4, 00035–02018. doi:10.1183/23120541.00035-2018
- Bunton, D. C., Petrie, M. C., Hillier, C., Johnston, F., and McMurray, J. J. V. (2004). The Clinical Relevance of Adrenomedullin: a Promising Profile? *Pharmacol. Ther.* 103, 179–201. doi:10.1016/j.pharmthera.2004.07.002
- Chamberlain, J., Evans, D., King, A., Dewberry, R., Dower, S., Crossman, D., et al. (2006). Interleukin-1 β and Signaling of Interleukin-1 in Vascular Wall and Circulating Cells Modulates the Extent of Neointima Formation in Mice. *Am. J. Pathol.* 168, 1396–1403. doi:10.2353/ajpath.2006.051054

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

DL, FJ, BL, and QY conceived and designed the study. DL performed the bioinformatics analyses. DL and FJ prepared figures and/or tables. DL and MXS authored drafts of the paper. MS organized the clinical and gene expression data. BL reviewed drafts of the paper. QY critically revised the drafts of the paper, and approved the final draft.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant Number 81270412).

ACKNOWLEDGMENTS

The authors thank the National Natural Science Foundation of China (No.81270412) for its support and GEO database for its generous sharing of data.

- Chang, L.-Y., Lu, C.-Y., Shao, P.-L., Lee, P.-I., Lin, M.-T., Fan, T.-Y., et al. (2014). Viral Infections Associated with Kawasaki Disease. *J. Formos. Med. Assoc.* 113, 148–154. doi:10.1016/j.jfma.2013.12.008
- Chibana, H., Kajimoto, H., Ueno, T., Yokoyama, S., Sasaki, K.-i., Ohtsuka, M., et al. (2017). Interleukin-1 β Is Associated with Coronary Endothelial Dysfunction in Patients with mTOR-Inhibitor-Eluting Stent Implantation. *Heart Vessels* 32, 823–832. doi:10.1007/s00380-017-0947-x
- Cho, H. J., Yang, S. I., Kim, K. H., Kim, J. N., and Kil, H. R. (2014). Cardiovascular Risk Factors of Early Atherosclerosis in School-Aged Children after Kawasaki Disease. *Korean J. Pediatr.* 57, 217–221. doi:10.3345/kjp.2014.57.5.217
- Cho, S. Y., Kim, Y., Cha, S. H., Suh, J. T., Han, M. Y., and Lee, H. J. (2011). Adjuvant Laboratory Marker of Kawasaki Disease; NT-Pro-BNP or Hs-CRP? *Ann. Clin. Lab. Sci.* 41, 360–363.
- Crawford, Y., Kasman, I., Yu, L., Zhong, C., Wu, X., Modrusan, Z., et al. (2009). PDGF-C Mediates the Angiogenic and Tumorigenic Properties of Fibroblasts Associated with Tumors Refractory to Anti-VEGF Treatment. *Cancer Cell* 15, 21–34. doi:10.1016/j.ccr.2008.12.004
- Dai, X., Chen, J., Xu, F., Zhao, J., Cai, W., Sun, Z., et al. (2020). TGF α Preserves Oligodendrocyte Lineage Cells and Improves white Matter Integrity after Cerebral Ischemia. *J. Cereb. Blood Flow Metab.* 40, 639–655. doi:10.1177/0271678x19830791
- Dionne, A., Meloche-Dumas, L., Desjardins, L., Turgeon, J., Saint-Cyr, C., Autmizguine, J., et al. (2017). N-terminal Pro-B-type Natriuretic Peptide Diagnostic Algorithm versus American Heart Association Algorithm for Kawasaki Disease. *Pediatr. Int.* 59, 265–270. doi:10.1111/ped.13154
- Fu, L. Y., Qiu, X., Deng, Q. L., Huang, P., Pi, L., Xu, Y., et al. (2019). The IL-1B Gene Polymorphisms Rs16944 and Rs1143627 Contribute to an Increased Risk of Coronary Artery Lesions in Southern Chinese Children with Kawasaki Disease. *J. Immunol. Res.* 2019, 4730507 doi:10.1155/2019/4730507

- Galeotti, C., Bayry, J., Kone-Paut, I., and Kaveri, S. V. (2010). Kawasaki Disease: Aetiopathogenesis and Therapeutic Utility of Intravenous Immunoglobulin. *Autoimmun. Rev.* 9, 441–448. doi:10.1016/j.autrev.2009.12.004
- Guo, C., Tan, C., Xia, X., Yuan, Y., Zhao, M., Yuan, Z., et al. (2020). Tumour Necrosis Factor- α and Myoglobin Associated with the Recovery Time of Coronary Artery Lesions in Kawasaki Disease Patients. *J. Paediatr. Child. Health* 56, 1382–1387. doi:10.1111/jpc.14942
- Guo, Y., Liu, R., Chen, L., Wu, W., and Zhang, S. (2020). Neutrophil Activation and Neutrophil Derived Neutrophil Extracellular Trap Formation in Patients with Coronary Artery Ectasia. *BMC Cardiovasc. Disord.* 20, 101. doi:10.1186/s12872-020-01398-0
- Hippenstiel, S., Witzentrath, M., Schmeck, B., Hocke, A., Krisp, M., Krull, M., et al. (2020). Adrenomedullin Reduces Endothelial Hyperpermeability. *Circ. Res.* 91, 618–625. doi:10.1161/01.res.0000036603.61868.f9
- Hu, P., Jiang, G. M., Wu, Y., Huang, B. Y., Liu, S. Y., Zhang, D. D., et al. (2017). TNF- α Is Superior to Conventional Inflammatory Mediators in Forecasting IVIG Nonresponse and Coronary Arteritis in Chinese Children with Kawasaki Disease. *Clinica Chim. Acta* 471, 76–80. doi:10.1016/j.cca.2017.05.019
- Huang, M. L., Hung, Y. H., Lee, W. M., Li, R. K., and Jiang (2014). BR, SVM-RFE Based Feature Selection and Taguchi Parameters Optimization for Multiclass SVM Classifier. *ScientificWorldJournal*, 2014. 795624. doi:10.1155/2014/795624
- Jaggi, P., Mejias, A., Xu, Z., Yin, H., Moore-Clingenpeel, M., Smith, B., et al. (2018). Whole Blood Transcriptional Profiles as a Prognostic Tool in Complete and Incomplete Kawasaki Disease. *PLoS One* 13. . e0197858. doi:10.1371/journal.pone.0197858
- Karvinen, H., Rutanen, J., Leppänen, O., Lach, R., Levonen, A.-L., Eriksson, U., et al. (2009). PDGF-C and -D and Their Receptors PDGFR- α and PDGFR- β in Atherosclerotic Human Arteries. *Eur. J. Clin. Invest.* 39, 320–327. doi:10.1111/j.1365-2362.2009.02095.x
- Kim, K. Y., and Kim, D. S. (2016). Recent Advances in Kawasaki Disease. *Yonsei. Med. J.* 57, 15–21. doi:10.3349/ymj.2016.57.1.15
- Kuo, H.-C., Hsieh, K.-S., Ming-Huey Guo, M., Weng, K.-P., Ger, L.-P., Chan, W.-C., et al. (2016). Next-generation Sequencing Identifies Micro-RNA-based Biomarker Panel for Kawasaki Disease. *J. Allergy Clin. Immunol.* 138, 1227–1230. doi:10.1016/j.jaci.2016.04.050
- Le Cras, T. D., Hardie, W. D., Fagan, K., Whitsett, J. A., and Korfhagen, T. R. (2003). Disrupted Pulmonary Vascular Development and Pulmonary Hypertension in Transgenic Mice Overexpressing Transforming Growth Factor- α . *Am. J. Physiology-Lung Cell Mol. Physiol.* 285, L1046–L1054. doi:10.1152/ajplung.00045.2003
- Li, X., Tjwa, M., Moons, L., Fons, P., Noel, A., Ny, A., et al. (2005). Revascularization of Ischemic Tissues by PDGF-CC via Effects on Endothelial Cells and Their Progenitors. *J. Clin. Invest.* 115, 118–127. doi:10.1172/jci19189
- Lin, K.-H., Chang, S.-S., Yu, C.-W., Lin, S.-C., Liu, S.-C., Chao, H.-y., et al. (2015). Usefulness of Natriuretic Peptide for the Diagnosis of Kawasaki Disease: a Systematic Review and Meta-Analysis. *BMJ Open* 5. e006703. doi:10.1136/bmjopen-2014-006703
- Makino, N., Nakamura, Y., Yashiro, M., Ae, R., Tsuboi, S., Aoyama, Y., et al. (2015). Descriptive Epidemiology of Kawasaki Disease in Japan, 2011–2012: From the Results of the 22nd Nationwide Survey. *J. Epidemiol.* 25, 239–245. doi:10.2188/jea.20140089
- McCordle, B. W., Rowley, A. H., Newburger, J. W., Burns, J. C., Bolger, A. F., Gewitz, M., et al. (2017). Diagnosis, Treatment, and Long-Term Management of Kawasaki Disease: A Scientific Statement for Health Professionals from the American Heart Association. *Circulation* 135, e927–e999. doi:10.1161/CIR.0000000000000484
- Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining Cell Type Abundance and Expression from Bulk Tissues with Digital Cytometry. *Nat. Biotechnol.* 37, 773–782. doi:10.1038/s41587-019-0114-2
- Nishida, K., Watanabe, K., Echigo, S., Mayumi, M., and Nishikimi, T. (2001). Increased Plasma Adrenomedullin Levels in Kawasaki Disease with Coronary Artery Involvement. *Am. J. Med.* 111, 165–166. doi:10.1016/s0002-9343(01)00781-1
- Nomura, I., Abe, J., Noma, S., Saito, H., Gao, B., Wheeler, G., et al. (2005). Adrenomedullin Is Highly Expressed in Blood Monocytes Associated with Acute Kawasaki Disease: a Microarray Gene Expression Study. *Pediatr. Res.* 57, 49–55. doi:10.1203/01.pdr.0000147745.52711.dd
- Orenstein, J. M., Shulman, S. T., Fox, L. M., Baker, S. C., Takahashi, M., Bhatti, T. R., et al. (2012). Three Linked Vasculopathic Processes Characterize Kawasaki Disease: a Light and Transmission Electron Microscopic Study. *PLoS One* 7. e38998. doi:10.1371/journal.pone.0038998
- Reddy, M., Singh, S., Rawat, A., Sharma, A., Suri, D., and Rohit, M. K. (2016). Pro-brain Natriuretic Peptide (ProBNP) Levels in North Indian Children with Kawasaki Disease. *Rheumatol. Int.* 36, 551–559. doi:10.1007/s00296-016-3430-6
- Ridker, P. M., Everett, B. M., Thuren, T., MacFadyen, J. G., Chang, W. H., Ballantyne, C., et al. (2017). Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N. Engl. J. Med.* 377, 1119–1131. doi:10.1056/nejmoa1707914
- Sakurai, Y. (2019). Autoimmune Aspects of Kawasaki Disease. *J. Investig. Allergol. Clin. Immunol.* 29, 251–261. doi:10.18176/jiaci.0300
- Sánchez-Manubens, J., Bou, R., and Anton, J. (2014). Diagnosis and Classification of Kawasaki Disease. *J. Autoimmun.* 48–49, 113–117. doi:10.1016/j.jaut.2014.01.010
- Shulman, S. T., Rowley, A. H., and Kawasaki, d. (2015). Kawasaki Disease: Insights into Pathogenesis and Approaches to Treatment. *Nat. Rev. Rheumatol.* 11, 475–482. doi:10.1038/nrrheum.2015.54
- Takahashi, K., Oharaseki, T., Naoe, S., Wakayama, M., and Yokouchi, Y. (2005). Neutrophilic Involvement in the Damage to Coronary Arteries in Acute Stage of Kawasaki Disease. *Pediatr. Int.* 47, 305–310. doi:10.1111/j.1442-200x.2005.02049.x
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B (Methodological)* 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x
- Usta Guç, B., Cengiz, N., Yildirim, S. V., and Uslu, Y. (2008). Cytomegalovirus Infection in a Patient with Atypical Kawasaki Disease. *Rheumatol. Int.* 28, 387–389. doi:10.1007/s00296-007-0440-4
- Weng, K.-P., Hsieh, K.-S., Ho, T.-Y., Huang, S.-H., Lai, C.-R., Chiu, Y.-T., et al. (2010). IL-1B Polymorphism in Association with Initial Intravenous Immunoglobulin Treatment Failure in Taiwanese Children with Kawasaki Disease. *Circ. J.* 74, 544–551. doi:10.1253/circj.09-0664
- Wong, H. K., Cheung, T. T., and Cheung, B. M. Y. (2012). Adrenomedullin and Cardiovascular Diseases. *JRSM Cardiovasc. Dis* 1, 1–7. doi:10.1258/cvd.2012.012003
- Wright, V. J., Herberg, J. A., Kaforou, M., Shimizu, C., Eleftherohorinou, H., Shailes, H., et al. (2018). Diagnosis of Kawasaki Disease Using a Minimal Whole-Blood Gene Expression Signature. *JAMA Pediatr.* 172. e182293. doi:10.1001/jamapediatrics.2018.2293
- Xie, X., Shi, X., and Liu, M. (2018). The Roles of Genetic Factors in Kawasaki Disease: A Systematic Review and Meta-Analysis of Genetic Association Studies. *Pediatr. Cardiol.* 39, 207–225. doi:10.1007/s00246-017-1760-0
- Yim, D., Curtis, N., Cheung, M., and Burgner, D. (2013). Update on Kawasaki Disease: Epidemiology, Aetiology and Pathogenesis. *J. Paediatr. Child. Health* 49, 704–708. doi:10.1111/jpc.12172
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16, 284–287. doi:10.1089/omi.2011.0118
- Zhang, H., Xu, M.-G., Xie, L.-J., Huang, M., Shen, J., and Xiao, T.-T. (2016). Meta-analysis of Risk Factors Associated with Atherosclerosis in Patients with Kawasaki Disease. *World J. Pediatr.* 12, 308–313. doi:10.1007/s12519-016-0023-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Liu, Song, Jing, Liu and Yi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.