

Identification of Immune-Associated Genes in Diagnosing Aortic Valve Calcification With Metabolic Syndrome by Integrated Bioinformatics Analysis and Machine Learning

OPEN ACCESS

Edited by:

Shameer Khader, AstraZeneca, United States

Reviewed by:

Maria Carmela Padula, Istituto Reumatologico della Lucania, Italy Niels Hansen, University Medical Center Goettingen, Germany

*Correspondence:

Jing Wang roubaobao09@163.com Kai Wang KWKAIWANG@163.com [†]These authors have contributed equally to this work

Specialty section:

This article was submitted to Inflammation, a section of the journal Frontiers in Immunology

Received: 06 May 2022 **Accepted:** 08 June 2022 **Published:** 04 July 2022

Citation:

Zhou Y, Shi W, Zhao D, Xiao S, Wang K and Wang J (2022) Identification of Immune-Associated Genes in Diagnosing Aortic Valve Calcification With Metabolic Syndrome by Integrated Bioinformatics Analysis and Machine Learning. Front. Immunol. 13:937886. doi: 10.3389/fimmu.2022.937886

Yufei Zhou^{1†}, Wenxiang Shi^{2†}, Di Zhao^{1†}, Shengjue Xiao³, Kai Wang^{4*} and Jing Wang^{5*}

¹ Department of Cardiology, Shanghai Institute of Cardiovascular Diseases, Zhongshan Hospital and Institutes of Biomedical Sciences, Fudan University, Shanghai, China, ² Department of Pediatric Cardiology, Xinhua Hospital, The Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, ³ Department of Cardiology, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China, ⁴ Department of Cardiology, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, ⁵ Department of Geriatric Medicine, The Affiliated Jiangning Hospital With Nanjing Medical University, Nanjing, China

Background: Immune system dysregulation plays a critical role in aortic valve calcification (AVC) and metabolic syndrome (MS) pathogenesis. The study aimed to identify pivotal diagnostic candidate genes for AVC patients with MS.

Methods: We obtained three AVC and one MS dataset from the gene expression omnibus (GEO) database. Identification of differentially expressed genes (DEGs) and module gene *via* Limma and weighted gene co-expression network analysis (WGCNA), functional enrichment analysis, protein–protein interaction (PPI) network construction, and machine learning algorithms (least absolute shrinkage and selection operator (LASSO) regression and random forest) were used to identify candidate immune-associated hub genes for diagnosing AVC with MS. To assess the diagnostic value, the nomogram and receiver operating characteristic (ROC) curve were developed. Finally, immune cell infiltration was created to investigate immune cell dysregulation in AVC.

Results: The merged AVC dataset included 587 DEGs, and 1,438 module genes were screened out in MS. MS DEGs were primarily enriched in immune regulation. The intersection of DEGs for AVC and module genes for MS was 50, which were mainly enriched in the immune system as well. Following the development of the PPI network, 26 node genes were filtered, and five candidate hub genes were chosen for nomogram building and diagnostic value evaluation after machine learning. The nomogram and all five candidate hub genes had high diagnostic values (area under the curve from 0.732 to 0.982). Various dysregulated immune cells were observed as well.

1

Conclusion: Five immune-associated candidate hub genes (*BEX2*, *SPRY2*, *CXCL16*, *ITGAL*, and *MORF4L2*) were identified, and the nomogram was constructed for AVC with MS diagnosis. Our study could provide potential peripheral blood diagnostic candidate genes for AVC in MS patients.

Keywords: aortic valve calcification, metabolic syndrome, differentially expressed genes, machine learning, immune infiltration, diagnosis

1. INTRODUCTION

Aortic valve (AV) calcification (AVC) is the most common valvular cardiac disease in the aging population of the developed world. It has a high global prevalence; approximately 12.6 million cases were reported in 2017 with an estimated 102,700 deaths (1, 2). AVC is induced by several risk factors, including genetic mutations, hyperlipidemia, hyperglycemia, and infection (3). Metabolic syndrome (MS) is a pathologic condition comprised of hypertension, hyperlipidemia, abdominal obesity, and insulin resistance (4). Therefore, patients with MS have an amplified risk of AVC but also progresses rapidly from mild to severe (5). Both AVC and MS are progressive diseases constituting a global health burden for the aging population.

Previous studies have proved that early endothelial inflammation and dysfunction take part in the initiation of AVC (6). Several risk factors, including smoking, diabetes, and hyperlipidemia, contribute to AVC, at least in part, through proinflammatory molecules and lipid deposition (7). Furthermore, inflammation reaction also exists in late-stage AVC (8). According to multiple pieces of evidence, immune cells play a crucial role in the physiological dysfunction associated with MS, as well as the pathophysiology and development of subsequent chronic diseases (9, 10). As a result, immune filtration and related pro-inflammatory molecules may be useful in the early diagnosis of AVC patients with MS.

It is well recognized that MS can accelerate the AVC proceeding. After the onset of valve disease symptoms, the prognosis is dismal. Patients usually do not seek medical evaluation for AVC until they exhibit symptoms. Therefore, it is crucial to discover sensitive and specific diagnostic tools for early-stage AVC prior to irreversible heart injury, particularly for MS patients who may be insensitive to aorta stenosis (AS) symptoms. Proteomics and sequencing tools provide an opportunity for identifying potential novel biomarkers and their roles in diverse diseases (11). Machine learning is gradually maturing in bioinformatics applications and can be used to excavate underlying mechanisms, prospective biomarkers, and therapeutic targets for a variety of diseases (12).

To the best of our knowledge, limited research has been conducted on the identification of immune-associated diagnostic candidates for AVC with MS, as well as the machine learning application for AVC diagnosis. Here, we first downloaded three AVC and one MS datasets from the gene expression omnibus (GEO) database, identified differentially expressed genes (DEGs) by Limma, and selected important module genes *via* weighted gene co-expression network analysis (WGCNA). Functional enrichment analysis, construction of protein–protein interaction network, application of machine learning (least absolute shrinkage and selection operator (LASSO) and random forest (RF)) algorithms, immune cell infiltration analysis, evaluation of nomogram, and receiver operating characteristic (ROC) curve evaluation were subsequently performed to identify pivotal immune-related diagnostic biomarkers for AVC with MS. This research could lead to the identification of immune-associated potential diagnostic markers for AVC in MS patients.

2. MATERIALS AND METHODS

2.1 Microarray Data

Figure 1 depicts the study flowchart. Three raw datasets (GSE51472 (13), GSE12644 (14), and GSE83453 (15)) including gene expression data for AVC patients and controls and one dataset of MS (GSE98895 (16)) were downloaded from the GEO (https://www.ncbi.nlm.nih.gov/geo/) database (17).





 Table 1 presents detailed dataset information, including the microarray platform, sample groups, and numbers.

2.2 Data Processing and Differentially Expressed Gene Screening

First, background calibration, normalization, and log2 transformation were performed on the three AVC raw datasets using affy in R. When multiple probes identified the same gene, the average value was calculated to determine its expression. Following the merging of the three datasets, the Bioconductor "SVA" R package was applied to eliminate batch effects (18). Finally, $|log_2$ Fold change (FC)| > 1 (AVC filtration) or 0.585 (MS filtration) and p-value <0.05 were set as the criteria for identifying DEGs using Limma package.

2.3 Weighted Gene Co-Expression Network Analysis and Module Gene Selection

WGCNA, a system biology strategy, was adopted to explore the correlation between genes (19). First, the median absolute deviation (MAD) of each gene was determined, and 50% of genes with the smallest MAD were removed. Second, the DEG expression matrix was filtered by the goodSamplesGenes function to omit unqualified genes and samples, and a scalefree co-expression network was built. Third, adjacency was computed using the co-expression similarity-derived "soft" thresholding power (β). The adjacency was then converted into a topological overlap matrix (TOM), and the gene ratio and dissimilarity were determined. The fourth step was the detection of modules using hierarchical clustering and a dynamic tree cut function. Genes with identical expression profiles were classified into gene modules using average linkage hierarchical clustering, with a TOM-based dissimilarity metric and a minimum gene group size (n = 30) for the gene dendrogram. Fifth, the dissimilarity of module eigengenes was computed, a cut line for the module dendrogram was chosen, and several modules were combined for further investigation. The eigengene network was finally visualized. WGCNA analysis was employed to identify important modules in MS.

2.4 Functional Enrichment Analysis

The Gene Ontology (GO) system provides structured, computable information regarding the functions of genes and gene products (20). The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a widely used database for the systematic investigation of gene

functions (21). Functional enrichment analysis was conducted based on the R package clusterProfiler (22), and the results of enrichment analysis were visualized *via* the Sangerbox platform (http://vip.sangerbox.com/). p-Value <0.05 was set as the criteria. Here, GO and KEGG analyses were performed twice based on the intersection of DEGs and the most significant module genes of MS, and the intersection of DEGs for AVC and the most significant module genes of MS.

2.5 Protein–Protein Interaction Network Construction

To excavate interactions among protein-coding genes, a proteinprotein interaction (PPI) network was established using the String database (23) (version 11.5; www.string-db.org), with the minimum required interaction score set at 0.400. Cytoscape software was applied to modify images downloaded from String, and an MCODE plug-in was used to identify important interacted genes (24). All genes that could interact with each other in the PPI network were selected for subsequent analysis.

2.6 Machine Learning

Two machine learning algorithms were adopted to further filter candidate genes for AVC diagnosis. LASSO is a regression method for selecting a variable to improve the predictive accuracy and is also a regression technique for variable selection and regularization to improve the predictive accuracy and comprehensibility of a statistical model (25). RF is an appropriate approach with the benefits of no limits on variable conditions and better accuracy, sensitivity, and specificity, which can be used to predict continuous variables and provide forecasts without apparent variations (26). "glmnet" (27) and "randomForest" (28) R packages were used to perform LASSO regression and RF analysis. The intersection genes of LASSO and RF were considered as candidate hub genes in AVC diagnosis.

2.7 Nomogram Construction and Receiver Operating Characteristic Evaluation

Nomogram construction is valuable for clinical AVC diagnosis. Based on candidate genes, the "rms" R package was applied to construct the nomogram (29). "Points" indicates the score of candidate genes, and "Total Points" indicates the summation of all the scores of genes above. The ROC was subsequently established to evaluate the diagnostic value of candidate genes and nomogram regarding AVC diagnosis, and the calculation of

TABLE 1 | Basic information of GEO datasets used in the study.

GSE series	Туре	Sample size		Platform
		Control	Aortic valve calcification	
GSE51472	mRNA	5	5	GPL570
GSE12644	mRNA	10	10	GPL570
GSE83453	mRNA	8	9	GPL10558
		Control	Metabolic syndrome	
GSE98895	mRNA	20	20	GPL6947

GEO, gene expression omnibus.

area under the curve (AUC) and 95% CI was performed to quantify its value. AUC > 0.7 was considered the ideal diagnostic value. To further excavate the interrelation among the identified genes, the network was constructed using the String online tool.

2.8 Immune Infiltration Analysis

CIBERSORT, a computational approach for identifying the proportion of diverse immune cells using tissue gene expression profiles, was utilized to determine immune cell proportion in AVC and control (30). Immune cell infiltration analysis was performed using the "Cibersort" R package. The barplot was used to visualize the proportion of each type of immune cell in different samples. The comparison regarding the proportion of diverse types of immune cells between AVC and control groups was visualized *via* the vioplot. A heatmap depicting the correlation of 22 types of infiltrating immune cells was carried out using the "corrplot" R package (31).

2.9 Statistical Analysis

The establishment of the ROC curve and the calculation of AUC as well as 95% CI were constructed using SPSS Version 26.0 (IBM Corporation, Armonk, NY, USA).

Student's sample t-test was applied to compare the proportion of different immune cells between the control and AVC groups *via* GraphPad Prism Version 8.3.0 (GraphPad Software, San Diego, CA, USA). p-Value <0.05 was considered statistically significant.

3. RESULTS

3.1 Identification of Differentially Expressed Genes

A total of 587 DEGs were identified in the AVC combined dataset using the Limma method, of which 320 were upregulated and 267 were downregulated. The heatmap and volcano plot of AVC DEGs are shown in **Figures 2A**, **B**. Regarding the MS dataset, 62 DEGs were screened out (38 upregulated and 24 downregulated) (**Figures 3A**, **B**).

3.2 Weighted Gene Co-Expression Network Analysis and Key Module Identification

Here, WGCNA was applied to identify the most correlated module in MS. We chose $\beta = 18$ (scale-free R² = 0.9) as the "soft" threshold based on the scale independence and average connectivity (**Figures 3C, D**). **Figure 3E** depicts the clustering dendrogram of the MS and control. On the basis of this power, six gene co-expression modules (GCMs) were generated, which are presented in **Figures 3F, G** in different colors. The correlation between MS and GCMs is shown in **Figure 3H**, and the green module (1,436 genes) demonstrated the highest correlation with MS (correlation coefficient = 0.71, p = 2.5×10^{-7}) and was regarded as the pivotal module for subsequent analysis. Additionally, we calculated the correlations between module membership and gene significance in the green module for

MS. As expected, a significant positive correlation was observed between them (r = 0.6) as shown in **Figure 3I**. Therefore, green module genes were most significantly associated with MS.

3.3 Functional Enrichment Analysis of Metabolic Syndrome

GSE98895 is a new MS dataset that has not been excavated before. To assess whether this dataset could reflect MS pathogenesis to a reliable extent, we further performed functional enrichment analysis based on the intersection of genes from Limma and WGCNA module genes. A total of 36 common genes (CGs) were screened out *via* the intersection of 62 DEGs and 1436 genes in the green module (**Figure 4A**).

KEGG analysis showed that CGs were primarily enriched in the "IL-17 signaling pathway" and "rheumatoid arthritis" (**Figure 4B**). GO analysis elucidated that CGs were mainly enriched in biological process (BP) terms, including "regulation/positive regulation of immune system process" and "immune system process" (**Figure 4C**). With regard to cellular component (CC) ontology, the CGs were mainly located in the "nuclear body," "glial limiting end-foot," and "nuclear speck" (**Figure 4D**). Molecular function (MF) analysis showed that "collagen binding," "small GTPase binding," and "fibronectin binding" were the most significant items in CGs (**Figure 4E**). The detailed top ten enrichment ontologies for GO and KEGG are listed in **Supplementary Table S1**.

The enrichment analysis revealed that CGs of MS were mainly related to immune response and inflammatory response, which were highly correlated with MS pathogenesis and reliable for subsequent AVC analysis.

3.4 Enrichment Analysis of Aortic Valve Calcification With Metabolic Syndrome and Node Gene Identification *via* Protein– Protein Interaction Network Construction

To further explore whether MS-associated pivotal genes could be related to AVC pathogenesis, 50 genes were identified from the intersection of DEGs from AVC and module genes from MS visualized *via* the Venn diagram (**Figure 5A**). The KEGG enrichment analysis revealed that 50 genes were primarily enriched in the "Fc epsilon RI signaling pathway," "Leukocyte transendothelial migration," and "Chemokine signaling pathway"; all of the above ontologies were intimately related to the immune system (**Figure 5D**). Also, GO analysis showed that genes were enriched in "cell activation," "immune response," and "immune system process" (BP); "specific granule membrane," "whole membrane," and "cytosol" (CC); and "non-membrane spanning protein tyrosine kinase activity," "C-C chemokine binding," and "lipid binding" (MF) (**Figures 5E-G**, **Supplementary Table S2**).

After confirming that the screened genes were closely related to immunity, we constructed a PPI network to find node genes that could interact with each other for the subsequent machine learning filtration. **Figure 5B** shows the PPI network and that 26 genes could interact with each other; the most active module was



visualized through the MCODE plug-in, and the genes were ranked by node numbers in **Figure 5C**.

3.5 Identification of Candidate Hub Genes via Machine Learning

LASSO regression and RF machine learning algorithms were applied to screen candidate genes for nomogram construction and diagnostic value evaluation. As we can see from **Figures 6A**, **B**, the LASSO regression algorithm identified six potential candidate biomarkers, and the RF algorithm ranked the genes based on the calculation of the importance of each gene (**Figures 6C**, **D**). The intersection of the top 10 most important genes from the RF and six potential candidate genes from LASSO was visualized *via* the Venn diagram (**Figure 6E**), and five genes (*BEX2*, *CXCL16*, *ITGAL*, *MORF4L2*, and *SPRY2*) were identified for the final validation. Moreover, based on the five genes, we found that *CXCL16*, *ITGAL*, *MORF4L2*, and *SPRY2* could interact with each other through intermediate molecules, while *BEX2* showed a different pattern (**Supplementary Figure S1**).

3.6 Diagnostic Value Assessment

The nomogram was constructed based on the five candidate hub genes (**Figure 7A**), and a ROC curve was established to assess the diagnostic specificity and sensitivity of each gene and the nomogram. We calculated the AUC and 95% CI for each item. The results were as follows: *BEX2* (AUC 0.746, CI 0.601–0.892), *SPRY2* (AUC 0.788, CI 0.660–0.916), *CXCL16* (AUC 0.851, CI 0.746–0.957), *ITGAL* (AUC 0.830, CI 0.711–0.948), *MORF4L2* (AUC 0.732, CI 0.587–0.877), and nomogram (AUC 0.982, CI 0.953–1.000) (**Figures 7B–G**). All the candidate genes possess a

high diagnostic value for AVC with MS, and the constructed nomogram had the highest diagnostic value.

3.7 Immune Cell Infiltration Analysis

Since we observed that MS-associated genes could regulate AVC pathogenesis and were mainly enriched in immune regulation and could be used as the potential AVC diagnostic biomarker by nomogram construction with ROC evaluation, immune cell infiltration analysis was performed to better elucidate the immune regulation of AVC.

Regarding the AVC and control groups, the proportion of 22 kinds of immune cells in each sample is displayed in the barplot (**Figure 8A**). The vioplot demonstrated that AVC patients had a higher level of CD8⁺ T cells, plasma cells, CD4 memory activated T cells, and M0 macrophages and a lower level of naive B cells, CD4 resting T cells, activated NK cells, and M2 macrophages (**Figure 8B**). The correlation of 22 types of immune cells revealed that CD4 memory activated T cells were positively associated with resting mast cells (r = 0.59) and that naive B cells were positively related to activated mast cells (r = 0.56), whereas resting mast cells were negatively related to resting NK cells (r = -0.50) (**Figure 8C**). In general, various kinds of immune cells were differentially infiltrated in AVC patients, which could serve as the potential regulation point for AVC treatment.

4. DISCUSSION

AVC is a major cause of cardiac dysfunction in the elderly population and leads to a great public health burden. Recent



studies have identified several new biomarkers for AVC diagnosis, such as matrix remodeling associated with protein 5, fibronectin type III domain containing 1, and miR-34a (32, 33). There have been no previous studies that have combined the two diseases. Furthermore, machine learning methods and nomogram generation have not been used in the diagnosis of AVC. Here, we used a series of integrated bioinformatics analysis and machine learning methods to construct the nomogram and to evaluate the diagnostic value for AVC in MS patients. The most noteworthy discovery is that we identified five pivotal immune-associated candidate genes (*ITGAL, CXCL16, BEX2*,

SPRY2, and *MORF4L2*) and developed a nomogram for diagnosing AVC in MS patients.

Samples regarding the MS dataset used in the study are all from peripheral blood; therefore, we only need to collect peripheral blood from MS patients and evaluate the expression of the five discovered immune-associated genes to infer the probability of AVC incidence in MS patients, which is an efficient and practical method for clinical usage. The application of peripheral blood tests in diagnosing different diseases has also been widely accepted (34, 35). Moreover, although we confirmed that gene expression level can be used



and green module genes via WGCNA. (B) KEGG pathway analysis of the intersection of genes. Different colors represent various significant pathways and related enriched genes. (C-E) GO analysis of the intersection of genes, including biological process, cellular component, and molecular function, respectively. The y-axis represents different GO terms, the x-axis represents gene ratio enriched in relative GO terms, the circle size refers to gene numbers, and the color represents pvalue. MS, metabolic syndrome; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; WGCNA, weighted gene co-expression network analysis; Limma, linear models for microarray data; DEGs, differentially expressed genes.

as an independent diagnostic marker, we intend to develop a more comprehensive diagnosis pattern by converting it into a score and taking all these five markers into consideration (36). The expression of each gene was quantified and transferred to a score, with the augmentation of the score referring to the linear predictor. We can monitor and intervene early in the MS patients when the linear predictor is high, which is more valuable for implementation in AVC with MS diagnosis.

Integrin alpha L (*ITGAL*) is the novel biomarker identified in our study in diagnosing AVC in MS patients. It belongs to the integrin family and is also known as CD11a. It serves as the receptor for the intercellular adhesion molecule (ICAM) family (*ICAM1*, *ICAM2*, *ICAM3*, and *ICAM4*) (37) as well as the secreted form of ubiquitin-like protein *ISG15* (38). It is predominantly expressed in immune cells and is involved in numerous immunological phenomena, such as leukocyteendothelial cell contact, cytotoxic T cell-mediated killing, and antibody-dependent killing by granulocytes and monocytes (39). It promotes apoptotic neutrophil phagocytosis by macrophages together with ICAM3 (40). It is involved in immunological reactions and inflammatory processes, as well as angiogenesis and cancer growth. It has been identified as a biomarker in diverse cancers, including gastric, ovarian, colorectal, and renal cancers (41-43). However, the mechanisms of ITGAL regarding immune engagement with AVC remain unclear. Some AVC risk factors enhance the production of adhesion molecules such as ICAM in valvular endothelial cells. Pulsatile shear stress caused elevated ICAM-1 levels in aortic explanted leaflets (44). Wang et al. (45) discovered that ICAM signaling is involved in decreased osteogenic bone morphogenic protein and ALP alkaline phosphatase levels in valvular interstitial cells, which are engaged in calcific valve rebuilding. As a receptor of multiple



FIGURE 5 | Enrichment analysis of common genes from AVC with MS and the identification of node genes from PPI network. (A) Venn diagram shows that 50 common genes are identified from the intersection of genes in AVC using Limma and MS using WGCNA. (B) PPI network reveals that 26 genes interact with each other, and the most significant module is visualized using MCODE plug-in. (C) The column shows the gene nodes of 26 genes in PPI network. (D) KEGG analysis of 50 common genes. (E–G) GO analysis (biological process, cellular component, and molecular function) of 50 common genes. AVC, aortic valve calcification; MS, metabolic syndrome; PPI, protein-protein interaction network; WGCNA, weighted gene co-expression network analysis; MCODE, molecular complex detection.

ICAM family members, *ITGAL* is also overexpressed in AVC. Thus, we suggest that *ITGAL* could represent a potential diagnostic target for AVC in MS patients.

C-X-C motif chemokine ligand 16 (*CXCL16*) may serve a proinflammatory function in human atherosclerosis, particularly in acute coronary syndrome (46). Its expression was shown to be considerably elevated in atherosclerotic plaque, and it participates in mechanisms that lead to increased stenosis in atherosclerotic coronary arteries (47). During inflammatory valvular heart disease, *SR-PSOX/CXCL16* is engaged in the recruitment of CD8⁺ T cells *via* activating *VLA-4* and stimulating IFN- γ production (48). Studies have demonstrated that cytokine signaling and small dense low-density lipoprotein (LDL) particles play a role in the fibrotic and calcific remodeling of AVC. Our study found that *CXCL16* was overexpressed in the AVC with MS patients; thus, we surmise that *CXCL16* may induce CD8⁺ T-cell infiltration and ox-LDL metabolism, resulting in the AVC process. It can also be used as a diagnostic biomarker.

Brain expressed X-linked 2 (*BEX2*) has been shown to control mitochondrial apoptosis and the G1 cell cycle in breast cancer (49) and to increase the proliferation of human glioblastoma cells (50). Sprouty RTK Signaling Antagonist 2 (*SPRY2*) is involved in cell proliferation and differentiation and can modulate receptor



FIGURE 6 | Machine learning in screening candidate diagnostic biomarkers for AVC with MS. (**A**, **B**) Biomarkers screening in the Lasso model. The number of genes (n = 6) corresponding to the lowest point of the curve is the most suitable for AVC with MS diagnosis. (**C**, **D**) The random forest algorithm shows the error in AVC; control group and genes are ranked based on the importance score. (**E**) Venn diagram shows that five candidate diagnostic genes are identified *via* the above two algorithms. AVC, aortic valve calcification; MS, metabolic syndrome.

tyrosine kinase signaling. By inhibiting FGFR2-induced ERK phosphorylation, Xu et al. (51) found that *SPRY2* was associated with a favorable prognosis for intrahepatic cholangiocarcinoma. Mortality Factor 4 Like 2 (*MORF4L2*) is a component of the NuA4 histone acetyltransferase complex, which is involved in transcriptional activation of select genes primarily by acetylating nucleosomal histone H4 and H2A (52).

Subsequently, we explored the interrelation among the five identified immune-associated genes. As depicted in **Supplementary Figure S1**, Intercellular Adhesion Molecule 1 (*ICAM1*) acts as a bridge by binding *CXCL16* and *ITGAL* directly. Additionally, previous research has clarified the role of the *CXCL16/ICAM1/ITGAL* pathway in the regulation of inflammation and immunological diseases. First, *ICAM1* is the ligand for *ITGAL*, and stabilizing the *ITGAL/ICAM1* complex is

crucial for the regulation of the adaptive immunological process (53). Second, it has been confirmed that the regulation of *CXCL16* is closely related to *ICAM1* in numerous diseases. Abu et al. (54) reported that treatment of human retinal microvascular endothelial cells (HRMECs) with *CXCL16* led to increased production of *ICAM-1* and increased leukocyte adherence to HRMECs, which led to an inflammatory response. Zhao et al. (55) discovered that *CXCL16* could impact the establishment of atherosclerotic lesions by targeting *ICAM1*. Taking into account the network and prior findings, *CXCL16* may regulate many immunological and inflammatory processes *via* indirect interactions with *ITGAL*. The relationship between *MORF4L2* and *SPRY2* is not as close as the relationship between *CXCL16* and *ITGAL*. Based on the shortest link path, the bridge is composed of Lysine Acetyltransferase 5 (*KAT5*),



FIGURE 7 | Nomogram construction and the diagnostic value evaluation. (A) The visible nomogram for diagnosing AVC with MS. (B–G) The ROC curve of each candidate gene (*BEX2*, *SPRY2*, *CXCL16*, *ITGAL*, and *MORF4L2*) and nomogram show the significant AVC with MS diagnostic value. AVC, aortic valve calcification; MS, metabolic syndrome; BEX2, brain expressed X-linked 2; SPRY2, sprouty RTK signaling antagonist 2; CXCL16, C-X-C motif chemokine ligand 16; ITGAL, integrin subunit alpha L; MORF4L2, mortality factor 4 like 2; AUC, area under the curve.

Tumor Protein P53 (*TP53*), and Erb-B2 Receptor Tyrosine Kinase 3 (*ERBB3*). *KAT5* is a histone acetylase that is crucial for regulating apoptosis, autophagy, RNA transcription, and circadian rhythms (56) and for activating *TP53* in relation to tumor DNA repair (57). *ERBB3* is a receptor tyrosine kinase of the epidermal growth factor receptor (EGFR) family. Zhang et al. (58) discovered that activating *ERBB3* could mitigate myocardial ischemia/reperfusion damage. Due to the lengthy pathway, the association between *MORF4L2* and *SPRY2* is inconclusive; further research is required to determine whether the two genes regulate the immunological process in a comparable pattern.

Previous studies have demonstrated that immunological modulation and inflammatory modulation appear in all stages of AVC. Abdelbaky et al. (59) discovered that early AV inflammation may predispose patients to AV sclerosis based on 111 participants. Coté et al. (60) performed a histological analysis in 285 patients with AVC undergoing AV replacement and analyzed the presence of chronic inflammatory infiltrates. They found that dense inflammatory infiltrates within AVC are associated with the severity of aortic stenosis. Additionally, Mazzone et al. (61) found that neo-angiogenesis, T-lymphocyte infiltration, and heat shock protein-60 are biological hallmarks of an immune-mediated inflammatory process in end-stage AVC. According to our results, AVC patients had a higher level of CD8⁺ T cells, plasma cells, CD4 memory activated T cells, and M0 macrophages and a lower level of naive B cells, CD4 resting T cells, activated NK cells, and M2 macrophages, which are



consistent with the previous studies. Above all, understanding inflammatory signaling mechanisms can pave the way for the development of diagnosis and targeted therapeutics of AVC.

5. LIMITATION

Our study had several limitations. First, although we pooled three AVC datasets, the samples still remained few, and the diagnostic value of the nomogram was rather high due to the limited sample size. Also, we aimed to choose another dataset for validating the diagnostic value. However, only one dataset with only six samples was available. Thus we were unable to validate the findings. The results should be subsequently confirmed in a more large-scale study with a large sample size. Second, although the five candidate hub genes were mainly enriched in regulating immune pathways, the interaction between candidate hub genes and dysregulated immune cells was still worth investigating.

6. CONCLUSION

Our study systematically discovered five immune-associated candidate hub genes (*ITGAL*, *CXCL16*, *MORF4L2*, *SPRY2*, and *BEX2*) and provided the nomogram for diagnosing AVC with MS by various bioinformatics analysis and machine learning algorithms. We also point out the dysregulated immune cell proportion in AVC with MS. Our study could provide potential peripheral blood diagnostic candidate genes for AVC in MS patients.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

JW and KW contributed to the hypothesis development and manuscript preparation. YZ, WS, and DZ contributed to the study design and data analysis. SX contributed to the image processing. All the authors drafted the manuscript and approved its submission.

ACKNOWLEDGMENTS

We thank Dr. Liu (nucleobase translocation of bioinformatics) for sharing his bioinformatics experience generously. We also thank the Sangerbox for providing us with the platform as part of the data analysis.

REFERENCES

- Yadgir S, Johnson CO, Aboyans V, Adebayo OM, Adedoyin RA, Afarideh M, et al. Global, Regional, and National Burden of Calcific Aortic Valve and Degenerative Mitral Valve Diseases, 1990-2017. *Circulation* (2020) 141 (21):1670–80. doi: 10.1161/CIRCULATIONAHA.119.043391
- Yutzey KE, Demer LL, Body SC, Huggins GS, Towler DA, Giachelli CM, et al. Calcific Aortic Valve Disease A Consensus Summary From the Alliance of Investigators on Calcific Aortic Valve Disease. *Arterioscler Thromb Vasc Biol* (2014) 34(11):2387–93. doi: 10.1161/ATVBAHA.114.302523
- Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic Heart Disease. Lancet (2012) 379(9819):953–64. doi: 10.1016/S0140-6736(11)61171-9
- Saklayen MG. The Global Epidemic of the Metabolic Syndrome. Curr Hypertension Rep (2018) 20(2):12. doi: 10.1007/s11906-018-0812-z
- Banovic M, Athithan L, McCann GP. Aortic Stenosis and Diabetes Mellitus: An Ominous Combination. *Diabetes Vasc Dis Res* (2019) 16(4):310–23. doi: 10.1177/1479164118820657
- Rajamannan NM. Oxidative-Mechanical Stress Signals Stem Cell Niche Mediated Lrp5 Osteogenesis in eNOS-/- Null Mice. J Cell Biochem (2012) 113(5):1623–34. doi: 10.1002/jcb.24031
- Zmyslowski A, Szterk A. Current Knowledge on the Mechanism of Atherosclerosis and Pro-Atherosclerotic Properties of Oxysterols. *Lipids Health Dis* (2017) 16(1):188. doi: 10.1186/s12944-017-0579-2
- Natorska J, Marek G, Sadowski J, Undas A. Presence of B Cells Within Aortic Valves in Patients With Aortic Stenosis: Relation to Severity of the Disease. J Cardiol (2016) 67(1-2):80–5. doi: 10.1016/j.jjcc.2015.05.002
- 9. Kanneganti T-D, Dixit VD. Immunological Complications of Obesity. Nat Immunol (2012) 13(8):707–12. doi: 10.1038/ni.2343
- Andersen CJ, Murphy KE, Fernandez ML. Impact of Obesity and Metabolic Syndrome on Immunity. *Adv Nutr* (2016) 7(1):66–75. doi: 10.3945/ an.115.010207
- Zhao E, Xie H, Zhang Y. Identification of Differentially Expressed Genes Associated With Idiopathic Pulmonary Arterial Hypertension by Integrated Bioinformatics Approaches. J Comput Biol (2021) 28(1):79–88. doi: 10.1089/ cmb.2019.0433
- Kumar N, Narayan Das N, Gupta D, Gupta K, Bindra J. Efficient Automated Disease Diagnosis Using Machine Learning Models. J Healthc Eng (2021) 2021 9983652. doi: 10.1155/2021/9983652
- Rysä J. Gene Expression Profiling of Human Calcific Aortic Valve Disease. Genom Data (2016) 7:107–8. doi: 10.1016/j.gdata.2015.12.015
- Bossé Y, Miqdad A, Fournier D, Pépin A, Pibarot P, Mathieu P. Refining Molecular Pathways Leading to Calcific Aortic Valve Stenosis by Studying Gene Expression Profile of Normal and Calcified Stenotic Human Aortic Valves. *Circ Cardiovasc Genet* (2009) 2(5):489–98. doi: 10.1161/ CIRCGENETICS.108.820795
- Guauque-Olarte S, Droit A, Tremblay-Marchand J, Gaudreault N, Kalavrouziotis D, Dagenais F, et al. RNA Expression Profile of Calcified Bicuspid, Tricuspid, and Normal Human Aortic Valves by RNA Sequencing. *Physiol Genomics* (2016) 48(10):749–61. doi: 10.1152/physiolgenomics. 00041.2016

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.937886/ full#supplementary-material

Supplementary Table 1 | Functional enrichment analysis of intersection genes via Limma and WGCNA in metabolic syndrome.

Supplementary Table 2 | Functional enrichment analysis of differentially expressed genes in aortic valve calcification with metabolic syndrome.

Supplementary Figure 1 | The interaction among five identified immune associated genes in diagnosing aortic valve calcification with metabolic syndrome.

- D'Amore S, Härdfeldt J, Cariello M, Graziano G, Copetti M, Di Tullio G, et al. Identification of miR-9-5p as Direct Regulator of ABCA1 and HDL-Driven Reverse Cholesterol Transport in Circulating CD14+ Cells of Patients With Metabolic Syndrome. *Cardiovasc Res* (2018) 114(8):1154–64. doi: 10.1093/ cvr/cvy077
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: Archive for Functional Genomics Data Sets–Update. *Nucleic Acids Res* (2013) 41(Database issue):D991–995. doi: 10.1093/nar/gks1193
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* (2012) 28(6):882–3. doi: 10.1093/bioinformatics/ bts034
- Langfelder P, Horvath S. WGCNA: An R Package for Weighted Correlation Network Analysis. BMC Bioinf (2008) 9:559. doi: 10.1186/1471-2105-9-559
- The Gene Ontology Resource: 20 Years and Still GOing Strong. Nucleic Acids Res (2019) 47(D1):D330–d338. doi: 10.1093/nar/gky1055
- Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res (2000) 28(1):27–30. doi: 10.1093/nar/28.1.27
- 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. doi: 10.1089/omi.2011.0118
- Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING Database in 2021: Customizable Protein-Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res* (2021) 49(D1):D605–d612. doi: 10.1093/nar/ gkaa1074
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* (2003) 13(11):2498–504. doi: 10.1101/ gr.1239303
- Yang C, Delcher C, Shenkman E, Ranka S. Machine Learning Approaches for Predicting High Cost High Need Patient Expenditures in Health Care. *BioMed Eng Online* (2018) 17(Suppl 1):131. doi: 10.1186/s12938-018-0568-3
- Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A Random Forest Classifier for the Prediction of Energy Expenditure and Type of Physical Activity From Wrist and Hip Accelerometers. *Physiol Meas* (2014) 35 (11):2191–203. doi: 10.1088/0967-3334/35/11/2191
- Zhang M, Zhu K, Pu H, Wang Z, Zhao H, Zhang J, et al. An Immune-Related Signature Predicts Survival in Patients With Lung Adenocarcinoma. Front Oncol (2019) 9:1314. doi: 10.3389/fonc.2019.01314
- Alderden J, Pepper GA, Wilson A, Whitney JD, Richardson S, Butcher R, et al. Predicting Pressure Injury in Critical Care Patients: A Machine-Learning Model. Am J Crit Care (2018) 27(6):461–8. doi: 10.4037/ajcc2018525
- Pan X, Jin X, Wang J, Hu Q, Dai B. Placenta Inflammation Is Closely Associated With Gestational Diabetes Mellitus. *Am J Transl Res* (2021) 13 (5):4068–79.
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust Enumeration of Cell Subsets From Tissue Expression Profiles. *Nat Methods* (2015) 12(5):453–7. doi: 10.1038/nmeth.3337

- Hu K. Become Competent Within One Day in Generating Boxplots and Violin Plots for a Novice Without Prior R Experience. *Methods Protoc* (2020) 3(4):31. doi: 10.3390/mps3040064
- Bouchareb R, Guauque-Olarte S, Snider J, Zaminski D, Anyanwu A, Stelzer P, et al. Proteomic Architecture of Valvular Extracellular Matrix FNDC1 and MXRA5 Are New Biomarkers of Aortic Stenosis. *Jacc-Basic to Trans Sci* (2021) 6(1):25–39. doi: 10.1016/j.jacbts.2020.11.008
- 33. Toshima T, Watanabe T, Narumi T, Otaki Y, Shishido T, Aono T, et al. Therapeutic Inhibition of microRNA-34a Ameliorates Aortic Valve Calcification via Modulation of Notch 1-Runx2 Signalling. Cardiovasc Res (2020) 116(5):983–94. doi: 10.1093/cvr/cvz210
- Jogia T, Kopp MA, Schwab JM, Ruitenberg MJ. Peripheral White Blood Cell Responses as Emerging Biomarkers for Patient Stratification and Prognosis in Acute Spinal Cord Injury. *Curr Opin Neurol* (2021) 34(6):796–803. doi: 10.1097/WCO.00000000000995
- 35. Pecht T, Gutman-Tirosh A, Bashan N, Rudich A. Peripheral Blood Leucocyte Subclasses as Potential Biomarkers of Adipose Tissue Inflammation and Obesity Subphenotypes in Humans. *Obes Rev* (2014) 15(4):322–37. doi: 10.1111/obr.12133
- 36. ÓH B, Gransar H, Callister T, Shaw LJ, Schulman-Marcus J, Stuijfzand WJ, et al. Development and Validation of a Simple-To-Use Nomogram for Predicting 5-, 10-, and 15-Year Survival in Asymptomatic Adults Undergoing Coronary Artery Calcium Scoring. JACC Cardiovasc Imaging (2018) 11(3):450-8. doi: 10.1016/j.jcmg.2017.03.018
- Ostermann G, Weber KSC, Zernecke A, Schroder A, Weber C. JAM-1 is a Ligand of the Beta(2) Integrin LFA-1 Involved in Transendothelial Migration of Leukocytes. *Nat Immunol* (2002) 3(2):151–8. doi: 10.1038/ni755
- Swaim CD, Scott AF, Canadeo LA, Huibregtse JM. Extracellular ISG15 Signals Cytokine Secretion Through the LFA-1 Integrin Receptor. *Mol Cell* (2017) 68 (3):581-+. doi: 10.1016/j.molcel.2017.10.003
- Barber DF, Faure M, Long EO. LFA-1 Contributes an Early Signal for NK Cell Cytotoxicity. J Immunol (2004) 173(6):3653-9. doi: 10.4049/ jimmunol.173.6.3653
- Kristof E, Zahuczky G, Katona K, Doro Z, Nagy E, Fesues L. Novel Role of ICAM3 and LFA-1 in the Clearance of Apoptotic Neutrophils by Human Macrophages. *Apoptosis* (2013) 18(10):1235–51. doi: 10.1007/s10495-013-0873-z
- Zhang JC, Wang H, Yuan C, Wu J, Xu JN, Chen SY, et al. ITGAL as a Prognostic Biomarker Correlated With Immune Infiltrates in Gastric Cancer. *Front Cell Dev Biol* (2022) 10. doi: 10.3389/fcell.2022.808212
- Vendrell E, Ribas M, Valls J, Sole X, Grau M, Moreno V, et al. Genomic and Transcriptomic Prognostic Factors in R0 Dukes B and C Colorectal Cancer Patients. *Int J Oncol* (2007) 30(5):1099–107. doi: 10.3892/ijo.30.5.1099
- Boguslawska J, Kedzierska H, Poplawski P, Rybicka B, Tanski Z, Piekielko-Witkowska A. Expression of Genes Involved in Cellular Adhesion and Extracellular Matrix Remodeling Correlates With Poor Survival of Patients With Renal Cancer. J Urol (2016) 195(6):1892–902. doi: 10.1016/ j.juro.2015.11.050
- 44. Sucosky P, Balachandran K, Elhammali A, Jo H, Yoganathan AP. Altered Shear Stress Stimulates Upregulation of Endothelial VCAM-1 and ICAM-1 in a BMP-4-and TGF-Beta 1-Dependent Pathway. *Arterioscler Thromb Vasc Biol* (2009) 29(2):254–60. doi: 10.1161/ATVBAHA.108.176347
- 45. Wang D, Zeng Q, Song R, Ao L, Fullerton DA, Meng X. Ligation of ICAM-1 on Human Aortic Valve Interstitial Cells Induces the Osteogenic Response: A Critical Role of the Notch1-NF-Kappa B Pathway in BMP-2 Expression. *Biochim Et Biophys Acta-Mol Cell Res* (2014) 1843(11):2744–53. doi: 10.1016/ j.bbamcr.2014.07.017
- 46. Lehrke M, Millington SC, Lefterova M, Cumaranatunge RG, Szapary P, Wilensky R, et al. CXCL16 is a Marker of Inflammation, Atherosclerosis, and Acute Coronary Syndromes in Humans. J Am Coll Cardiol (2007) 49 (4):442–9. doi: 10.1016/j.jacc.2006.09.034
- Lundberg GA, Kellin A, Samnegard A, Lundman P, Tornvall P, Dimmeler S, et al. Severity of Coronary Artery Stenosis Is Associated With a Polymorphism in the CXCL16/SR-PSOX Gene. J Internal Med (2005) 257(5):415–22. doi: 10.1111/j.1365-2796.2005.01469.x
- Yamauchi R, Tanaka M, Kume N, Minami M, Kawamoto T, Togi K, et al. Upregulation of SR-PSOX/CXCL16 and Recruitment of CD8+ T Cells in

Cardiac Valves During Inflammatory Valvular Heart Disease. *Arterioscler Thromb Vasc Biol* (2004) 24(2):282-7. doi: 10.1161/01.ATV. 0000114565.42679.c6

- Naderi A, Liu J, Bennett IC. BEX2 Regulates Mitochondrial Apoptosis and G1 Cell Cycle in Breast Cancer. Int J Cancer (2010) 126(7):1596–610. doi: 10.1002/ijc.24866
- Meng Q, Zhi T, Chao Y, Nie E, Xu X, Shi Q, et al. Bex2 Controls Proliferation of Human Glioblastoma Cells Through NF-Kappa B Signaling Pathway. J Mol Neurosci (2014) 53(2):262–70. doi: 10.1007/s12031-013-0215-1
- 51. Xu YF, Liu HD, Liu ZL, Pan C, Yang XQ, Ning SL, et al. Sprouty2 Suppresses Progression and Correlates to Favourable Prognosis of Intrahepatic Cholangiocarcinoma via Antagonizing FGFR2 Signalling. J Cell Mol Med (2018) 22(11):5596–606. doi: 10.1111/jcmm.13833
- Jacquet K, Fradet-Turcotte A, Avvakumov N, Lambert JP, Roques C, Pandita RK, et al. The TIP60 Complex Regulates Bivalent Chromatin Recognition by 53BP1 Through Direct H4K20me Binding and H2AK15 Acetylation. *Mol Cell* (2016) 62(3):409–21. doi: 10.1016/j.molcel.2016.03.031
- Ma VP, Hu Y, Kellner AV, Brockman JM, Velusamy A, Blanchard AT, et al. The Magnitude of LFA-1/ICAM-1 Forces Fine-Tune TCR-Triggered T Cell Activation. *Sci Adv* (2022) 8(8):eabg4485. doi: 10.1126/sciadv.abg4485
- 54. Abu El-Asrar AM, Nawaz MI, Ahmad A, De Zutter A, Siddiquei MM, Blanter M, et al. Evaluation of Proteoforms of the Transmembrane Chemokines CXCL16 and CX3CL1, Their Receptors, and Their Processing Metalloproteinases ADAM10 and ADAM17 in Proliferative Diabetic Retinopathy. *Front Immunol* (2020) 11:601639. doi: 10.3389/fimmu.2020.601639
- Zhao J, Yang M, Wu J. CXCL16 may be a Predisposing Factor to Atherosclerosis: An Animal Study. *Mol Med Rep* (2021) 24(4):716. doi: 10.3892/mmr.2021.12355
- Li Z, Rasmussen LJ. TIP60 in Aging and Neurodegeneration. Ageing Res Rev (2020) 64:101195. doi: 10.1016/j.arr.2020.101195
- Ghobashi AH, Kamel MA. Tip60: Updates. J Appl Genet (2018) 59(2):161–8. doi: 10.1007/s13353-018-0432-y
- Zhang M, Cheng K, Chen H, Tu J, Shen Y, Pang L, et al. LncRNA AK020546 Protects Against Cardiac Ischemia-Reperfusion Injury by Sponging miR-350-3p. Aging (Albany NY) (2021) 13(10):14219–33. doi: 10.18632/aging.203038
- Abdelbaky A, Corsini E, Figueroa AL, Subramanian S, Fontanez S, Emami H, et al. Early Aortic Valve Inflammation Precedes Calcification: A Longitudinal FDG-PET/CT Study. *Atherosclerosis* (2015) 238(2):165–72. doi: 10.1016/ j.atherosclerosis.2014.11.026
- Cote N, Mahmut A, Bosse Y, Couture C, Page S, Trahan S, et al. Inflammation Is Associated With the Remodeling of Calcific Aortic Valve Disease. *Inflammation* (2013) 36(3):573–81. doi: 10.1007/s10753-012-9579-6
- Mazzone A, Epistolato MC, De Caterina R, Storti S, Vittorini S, Sbrana S, et al. Neoangiogenesis, T-Lymphocyte Infiltration, and Heat Shock Protein-60 Are Biological Hallmarks of an Immunomediated Inflammatory Process in End-Stage Calcified Aortic Valve Stenosis. *J Am Coll Cardiol* (2004) 43(9):1670–6. doi: 10.1016/j.jacc.2003.12.041

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 © 2022 Zhou, Shi, Zhao, Xiao, Wang and Wang. 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.