

# Identification of Shared Signature Genes and Immune Microenvironment Subtypes for Heart Failure and Chronic Kidney Disease Based on Machine Learning

Xuefu Wang<sup>1,\*</sup>, Jin Rao<sup>2,\*</sup>, Xiangyu Chen<sup>2,\*</sup>, Zhinong Wang<sup>2</sup>, Yufeng Zhang<sup>1,2</sup>

<sup>1</sup>School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, People's Republic of China; <sup>2</sup>Department of Cardiothoracic Surgery, Shanghai Changzheng Hospital, Naval Medical University, Shanghai, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Zhinong Wang; Yufeng Zhang, Email wangzn007@smmu.edu.cn; zhyf19810824@163.com

**Background:** A complex interrelationship exists between Heart Failure (HF) and chronic kidney disease (CKD). This study aims to clarify the molecular mechanisms of the organ-to-organ interplay between heart failure and CKD, as well as to identify extremely sensitive and specific biomarkers.

**Methods:** Differentially expressed tandem genes were identified from HF and CKD microarray datasets and enrichment analyses of tandem perturbation genes were performed to determine their biological functions. Machine learning algorithms are utilized to identify diagnostic biomarkers and evaluate the model by ROC curves. RT-PCR was employed to validate the accuracy of diagnostic biomarkers. Molecular subtypes were identified based on tandem gene expression profiling, and immune cell infiltration of different subtypes was examined. Finally, the ssGSEA score was used to build the ImmuneScore model and to assess the differentiation between subtypes using ROC curves.

**Results:** Thirty-three crosstalk genes were associated with inflammatory, immune and metabolism-related signaling pathways. The machine-learning algorithm identified 5 hub genes (PHLDA1, ATP1A1, IFIT2, HLTF, and MPP3) as the optimal shared diagnostic biomarkers. The expression levels of tandem genes were negatively correlated with left ventricular ejection fraction and glomerular filtration rate. The CIBERSORT results indicated the presence of severe immune dysregulation in patients with HF and CKD, which was further validated at the single-cell level. Consensus clustering classified HF and CKD patients into immune and metabolic subtypes. Twelve immune genes associated with immune subtypes were screened based on WGCNA analysis, and an ImmuneScore model was constructed for high and low risk. The model accurately predicted different molecular subtypes of HF or CKD.

**Conclusion:** Five crosstalk genes may serve as potential biomarkers for diagnosing HF and CKD and are involved in disease progression. Metabolite disorders causing activation of a large number of immune cells explain the common pathogenesis of HF and CKD.

**Keywords:** heart failure, chronic kidney disease, immunity and inflammation, metabolic disorder, bioinformatics

## Introduction

Heart failure (HF) constitutes a major cause of morbidity and mortality in patients with chronic kidney disease (CKD), with over 50% of CKD patients succumbing to HF.<sup>1,2</sup> Furthermore, chronic kidney dysfunction stands as a potent independent risk factor for the progression of cardiovascular diseases.<sup>3,4</sup> Clinical research findings indicate that approximately 40% to 50% of HF patients also exhibit concomitant CKD.<sup>5</sup> The heart and the kidneys are two vital organs in the human body, and the pathological and physiological mechanisms between them are intricate and bidirectional. The interdependence of these two organs can lead to a vicious cycle, wherein the deterioration of either organ may

result in acute or chronic dysfunction of the other.<sup>6</sup> The pathogenesis of the cardiac-kidney organ crosstalk is intricately complex, and remains insufficiently elucidated to date. This is characterized by a high incidence and mortality rate. Despite the existence of certain diagnostic and therapeutic approaches, clinicians continue to confront substantial challenges. Hence, investigating the pathophysiological interplay between hyperlipidemia and chronic kidney disease, and identifying novel targets for shared diagnosis and intervention, remains a paramount research direction.

An increasing body of research suggests that hemodynamic factors,<sup>7,8</sup> endothelial dysfunction,<sup>9</sup> oxidative stress,<sup>10</sup> chronic inflammatory response<sup>11</sup> and immune imbalance<sup>12</sup> are associated with the onset and progression of both HF and CKD. Chronic heart failure is characterized by a reduction in cardiac output and effective circulating volume, leading to a decrease in renal perfusion pressure. This, in turn, triggers the activation of the renin-angiotensin-aldosterone system (RAAS) and pressure receptors, ultimately resulting in renal tissue ischemia, hypoxia, renal cell apoptosis, and necrosis, leading to renal dysfunction.<sup>13</sup> Immune dysregulation and inflammatory responses may be associated with its pathogenesis. Toll-like receptor (TLR) signaling plays a rapid response role in tissue damage and is linked to the induction of immune responses and upregulation of inflammatory cytokine expression in both cardiac and renal diseases.<sup>14–16</sup> Furthermore, inflammation has been proposed as a potential stressor in chronic cardiac and renal injury. Reduced renal blood flow due to heart failure stimulates the overactivation of RAAS, promoting inflammatory factors, and potentially leading to adverse outcomes.<sup>17</sup> Currently, there is a growing awareness of patients with HF combined with CKD. The interplay of immunity and inflammation in the context of HF and CKD may provide novel insights into their pathophysiological processes.

Therefore, in this study, we utilized differential expression crosstalk genes identified from HF and CKD microarray datasets. We employed machine learning algorithms to select shared diagnostic biomarkers for both diseases and assessed the accuracy of the diagnostic model through ROC curve analysis. Furthermore, we delved into the biological functions of these biomarkers, their relationship with immune cell infiltration, and analyzed their clinical correlations with heart and kidney functions. We also identified potential molecular subtypes among HF and CKD patients based on crosstalk gene expression profiles and explored differential expression patterns between these subtypes. Finally, we constructed an ImmuneScore model based on single-sample Gene Set Enrichment Analysis (ssGSEA) scores and evaluated the model's ability to distinguish different subtypes using ROC curves.

## Materials and Methods

### Download and Preprocessing of Raw Data

The original microarray data for both HF and CKD used in this study were downloaded from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>).<sup>18</sup> The GSE57338 microarray data were obtained from the GPL11532 sequencing platform, encompassing 136 healthy controls and 177 patients with HF.<sup>19</sup> The GSE66494 microarray data were sourced from the GPL6480 sequencing platform, comprising renal biopsy specimens from 53 patients with CKD and 8 healthy controls.<sup>20</sup> Additionally, we downloaded the datasets GSE26887, based on the GPL5244 sequencing platform, and GSE200818, based on the GPL19983 sequencing platform, as external validation sets to further assess the accuracy of the diagnostic model.<sup>21</sup> All datasets were acquired from GEO as original CEL files. Subsequently, the background correction and normalization processes were carried out using the “affy” R package with the Robust Multiarray Average (RMA) algorithm. Afterward, the expression values of multiple probes targeting the same gene were averaged to obtain the expression matrix necessary for subsequent analysis in this study.<sup>22</sup> To further explore the heterogeneity in the expression of key crosstalk genes at the single-cell level, we downloaded scRNA-seq data GSE222144, which includes one normal sample and one HF sample. Subsequently, we performed quality control on the scRNA-seq data using the “Seurat” package. We excluded low-quality cells (those with fewer than 300 expressed genes or with mitochondrial gene expression exceeding 25%). Next, we used the “harmony” package to integrate the data from different processing batches and remove batch effects. Principal Component Analysis (PCA) was employed for dimensionality reduction to explore heterogeneity. Marker genes specific to particular clusters were identified using the FindAllMarkers function. Cell subtypes were annotated as known cell types using marker genes from CellMarker 2.0 and literature sources.

## Identification of Co-Expressed Genes and Functional Enrichment Analysis

The “limma” R package was utilized to filter differentially expressed genes (DEGs) in the GSE57338 and GSE66494 datasets. The criteria for DEG selection in both datasets were set as  $|\log FC| \geq 0.5$  and an adjusted p-value less than 0.05 (adj. p-value < 0.05). The cross-talk genes of the two diseases were obtained by the intersection of up-regulated and down-regulated DEGs in HF and CKD cohorts by Venn diagram, and the results were displayed by cluster heat maps and volcano maps. Subsequently, crosstalk genes were extracted for enrichment analysis to determine their gene functions. GO enrichment analysis was performed using the “clusterProfiler” R package.<sup>23</sup> The significance threshold for defining differentially enriched GO terms was set at  $p < 0.05$ . GSVA was performed using the “GSVA” R package to calculate gene set variations for the hallmark gene set (h.all.v2023.1.Hs.symbols.gmt) under different treatment conditions.<sup>24</sup> We used  $p < 0.05$  as the threshold for pathway significance differences. Additionally, GSEA was conducted using the “clusterProfiler” R package to identify biological attributes and gene functions.<sup>25</sup> The statistical significance level was set at  $p < 0.05$ .

## Identification and Validation of Shared Diagnostic Genes

Based on the crosstalk genes in the HF and CKD cohorts, the Boruta algorithm was used to create all features (shadow features) for the mixed data, adding randomness to the crosstalk genes in the two expression profiles.<sup>26</sup> The algorithm underwent over 400 iterations to obtain a stable random forest model. Subsequently, the importance of feature variables was ranked to facilitate efficient feature selection. We intersected the top ten features in both cohorts, and overlapping genes were considered the optimal shared diagnostic genes. Furthermore, we explored the intrinsic correlations of shared diagnostic genes through Spearman correlation analysis. Next, we then validated the expression levels of the shared diagnostic genes in internal and external datasets, respectively, and visualized them via box plots. The accuracy of diagnostic models in predicting the prevalence of HF and CKD was assessed in two independent external datasets by the area under the ROC curve. Additionally, based on the GSE222144 dataset of HF patients at the single-cell level, we explored the expression levels of shared diagnostic genes across different cell lineages.

## Immune Infiltration Analysis in Heart Failure and Chronic Kidney Disease Cohorts

To explore the extent of immune cell infiltration in HF and CKD cohorts, the percentage of immune cells in each cell subpopulation was obtained after deconvolution of gene expression profiles of 22 immune cell subtypes based on the CIBERSORT algorithm. Visualize the proportion of 22 immune cells in each sample by stacked histograms. Differences in immune cell infiltration between the disease and healthy control groups were illustrated through box plots. In addition, the correlation between immune cells and shared diagnostic genes was further explored and the results were visualized by heatmapping with the “pheatmap” package.

## Construction and Validation of Clinical Diagnostic Models

Based on the expression patterns of shared diagnostic genes, nomogram was constructed using the “rms” package. The decision-making ability of these nomogram for disease diagnosis was assessed by calculating the area under the ROC curve. Calibration curves were employed to evaluate the model’s ability to predict the probability and accuracy of clinical outcomes related to HF and CKD. Finally, decision curve analysis (DCA) was used to assess the clinical utility of the model.

## Correlation of Left Ventricular Ejection Fraction and Glomerular Filtration Rate with Characterization Genes

The GSE42664 dataset includes clinical data on left ventricular ejection fraction for 15 HF patients, which was used for subsequent analysis of the correlation between feature genes and heart function. The NephroseqV5 webpage online tool was utilized to determine the correlation between characterized genes and kidney function. We downloaded the expression profiles of the two disease signature genes and the cardiac or renal function information, performed correlation analyses using Spearman, and plotted scatter plots using “ggplot2”.

## Identification of Correlated Subtypes Based on Crosstalk Gene Expression Patterns

Consensus clustering was performed using the “ConsensuClusterPlus” R package. Consensus clustering based on euclidean distance was performed on HF and CKD patient samples using the K-means algorithm to identify relevant molecular subtypes. The clustering process was iterated 500 times, with each iteration including 80% of the samples. The number of clusters was set to range from 2 to 10, and the optimal number of clusters was determined by the cumulative distribution function (CDF) of the consensus score. Based on the results of consensus clustering analysis, we assessed the levels of immune cell infiltration in different clustering subgroups by the ssGSEA algorithm. Finally, GSVA was performed using the “GSVA” R package to calculate the gene set enrichment scores of the signature gene set (h.all.v2023.1.Hs.symbols.gmt) in different clusters and to compare the gene set enrichment scores of the two clusters by the “Limma” R package to compare the differences in biological functions in the two clusters.

## Construction of Weighted Gene Co-Expression Networks in Immune Subtypes

Based on the results of GSVA enrichment analysis, we found that subgroup B of the HF cohort was mainly involved in immune-related pathways while subgroup A of the CKD subgroup was associated with immunity. Therefore, we used WGCNA to screen modules that were correlated with the immune phenotype.<sup>27</sup> We applied the Median Absolute Deviation (MAD) method to select the top 5000 genes with significant variation among samples for further analysis. First, we assessed the data quality for both samples and genes to exclude low-quality data. Second, the samples were systematically clustered using the “hclust” function, and the samples with obvious outliers are eliminated. According to the standards of scale-free networks, use the “pickSoftThreshold” function in the WGCNA software package to select an appropriate soft threshold  $\beta$  within the range of 1 to 20. When  $R^2=0.9$ , it indicates a relatively smooth connectivity, and at this point, it represents the optimal soft threshold. Subsequently, construct a weighted gene co-expression network to identify gene modules. The minimum number of genes in a module is set to 30, and the combined cut height is 0.25. Finally, select the shared immune-related genes by intersecting the modules with the highest relevance to the immune phenotype from both datasets with immune-related genes.

## Construction of the ImmuneScore Model

Based on the results from WGCNA, 12 shared immune-related genes from both datasets were included in constructing the ImmuneScore model. Using the immune gene set as the background genes, quantitatively analyze the enrichment scores of shared immune genes in each sample in the HF and CKD cohorts using the ssGSEA algorithm. Constructing ImmuneScore models based on ssGSEA scores. Patients from both cohorts are divided into low-risk and high-risk groups based on the median of ssGSEA scores, and the heterogeneity of the high-risk and low-risk groups is validated through PCA. Furthermore, we assessed the discriminatory capability of the ImmuneScore model for immune subtypes and evaluated the predictive ability of the ImmuneScore model through ROC curve.

## Collection of Human Tissues and Real-Time PCR

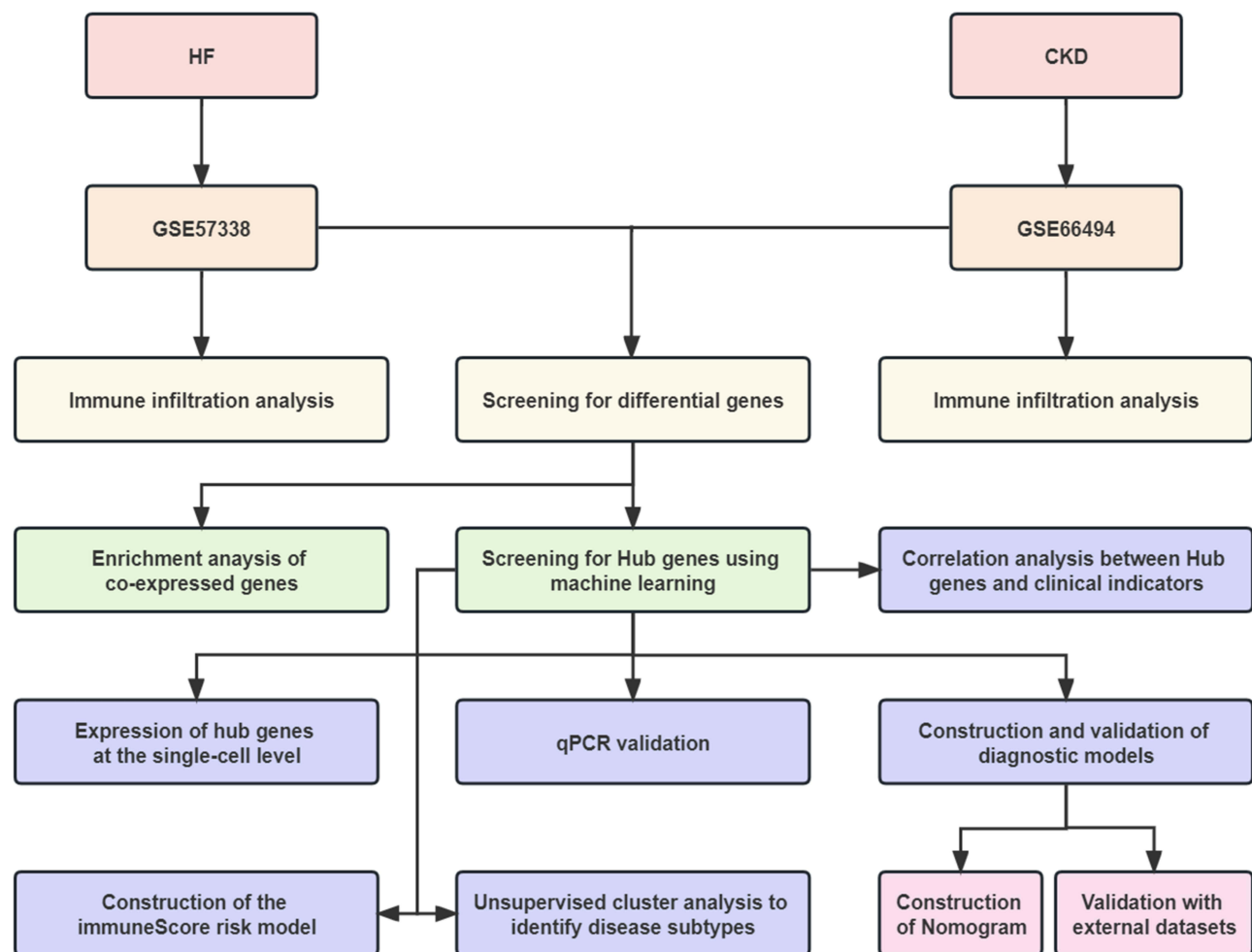
Tissues from patients with HF and CKD were obtained from the Laboratory Specimen Bank and Department of Pathology, Division of Cardiovascular Surgery and Nephrology, Changzheng Hospital, Shanghai, China. Informed consent for the use of specimens was obtained from family members or the patients themselves. The study procedure was also approved by the Ethics Committee of Shanghai Changzheng Hospital, China. Total RNA was extracted using the Trizol method, and cDNA was synthesized through reverse transcription. The SYBR Green reagent was prepared according to the instructions, and real-time PCR reactions were conducted using the ABI Q7 fluorescent quantitative PCR instrument. Student's *t*-test was used to statistically analyze and compare the groups. Data were expressed as mean values.  $p<0.05$  was considered significant.

## Results

### Identification of Crosstalk Genes in HF and CKD Cohorts

The overall layout of this study is illustrated in [Figure 1](#). In the HF cohort, a total of 428 DEGs were identified in the GSE57338 dataset, including 234 upregulated genes and 194 downregulated genes ([Figure 2A](#)). In the CKD cohort, the



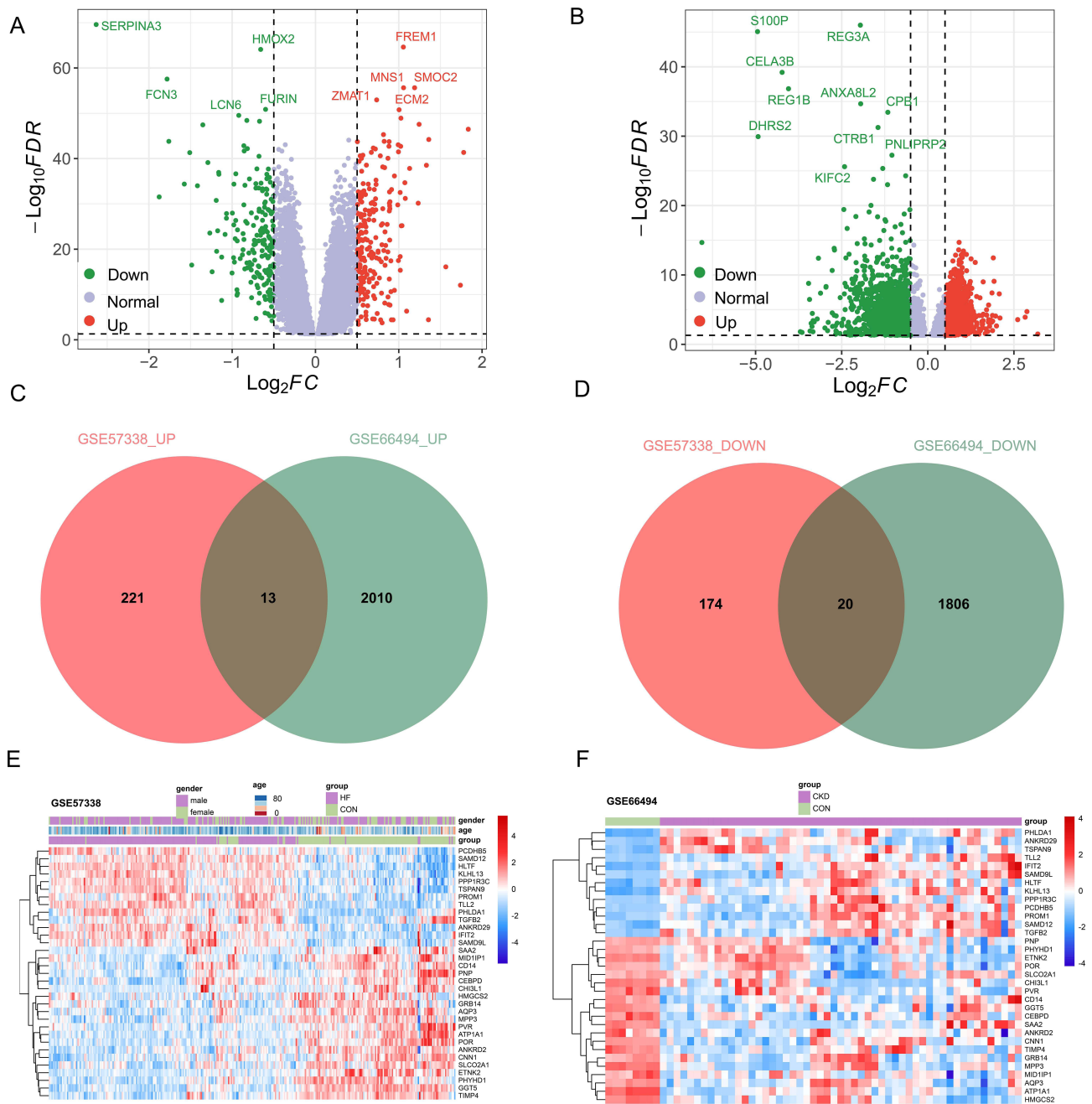


**Figure 1** Flowchart of this research.

GSE66494 dataset revealed 3849 DEGs, comprising 2023 upregulated genes and 1826 downregulated genes (Figure 2B). The crosstalk genes were obtained by intersecting genes with consistent expression trends in the HF and CKD samples, which included 13 upregulated genes and 20 downregulated genes (Figure 2C and D). The expression trends of these 33 crosstalk genes are shown in Figure 2E and F.

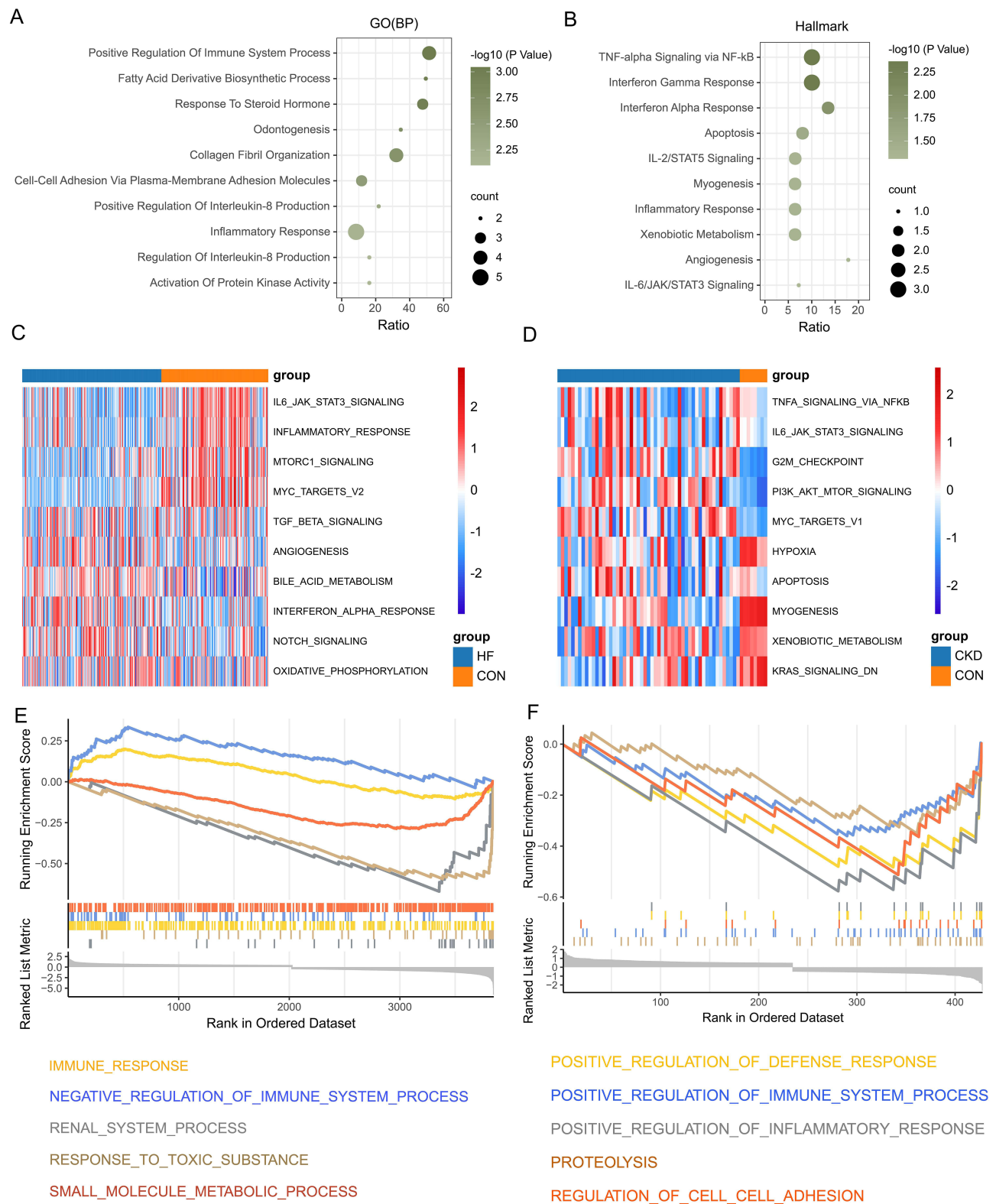
## Enrichment Analysis of Crosstalk Genes

Enrichment analysis was performed using GO and HALLMARK gene sets. GO analysis revealed that crosstalk genes are primarily involved in biological processes related to the positive regulation of the immune system, apoptosis, fatty acid derivative biosynthetic process, response to steroid hormone, cell-cell adhesion via plasma-membrane adhesion molecules, positive regulation of interleukin-8 production, and inflammatory response, among other biological processes (Figure 3A and Supplementary Table 1). Enrichment analysis based on the HALLMARK gene set indicates significant enrichment in pathways including TNF-alpha Signaling via NF-kB, Interferon Gamma Response, Xenobiotic Metabolism, Apoptosis, IL-2/STAT5 Signaling, and Inflammatory Response (Figure 3B and Supplementary Table 2). In addition, GSVA results showed that IL6-JAK-STAT3-SIGNALING, INFLAMMATORY-RESPONSE, MTORC1-SIGNALING, and MYC-TARGETS-V2 were significantly enriched in the healthy control group, whereas metabolism-related pathways, such as oxidative phosphorylation, were mainly enriched in the HF group (Figure 3C). In the CKD cohort, GSVA results showed significant enrichment of pathways such as TNFA-SIGNALING-VIA-NFKB, IL6-JAK-STAT3-SIGNALING, G2M-CHECKPOINT, and PI3K-AKT-MTOR-SIGNALING in the CKD group. However, the



**Figure 2** Identification of Crosstalk Genes in Differential Expression. (A and B) Volcano plots illustrating differentially expressed genes in the GSE57338 and GSE66434 datasets. (C) Upregulated crosstalk genes in the GSE57338 and GSE66434 datasets. (D) Downregulated crosstalk genes in the GSE57338 and GSE66434 datasets. (E and F) Heatmaps depicting the expression patterns of crosstalk genes in the heart failure and chronic kidney disease datasets.

activity of metabolism-related processes like XENOBIOTIC-METABOLISM was lower in the CKD group compared to the healthy control group (Figure 3D). GSVA results suggest potential differences in immune response patterns and metabolic processes between the two diseases. Additionally, GSEA enrichment analysis was used to explore essential biological processes associated with crosstalk genes. In the HF cohort, positive regulation of defense responses, positive regulation of immune system processes, positive regulation of inflammatory responses, and regulation of cell adhesion were negatively associated with crosstalk genes (Figure 3E), and in the CKD cohort, negative regulation of immune responses and immune system processes were negatively associated with crosstalk genes, whereas renal system processes, response to toxic substances, and small molecule metabolism processes were positively associated with



**Figure 3** Enrichment Analysis Results of Crosstalk Genes. **(A and B)** Results of GO enrichment analysis and enrichment analysis based on Hallmark gene sets. **(C and D)** GSEA enrichment analysis results in heart failure and chronic kidney disease cohorts. **(E and F)** Enrichment analysis based on the GSE57338 and GSE66434 datasets using GSEA.

crosstalk genes (Figure 3F). Based on above enrichment analysis results, it was observed that crosstalk genes are primarily involved in immune responses and metabolic processes in HF and CKD.

## Identify the Optimal Shared Diagnostic Genes

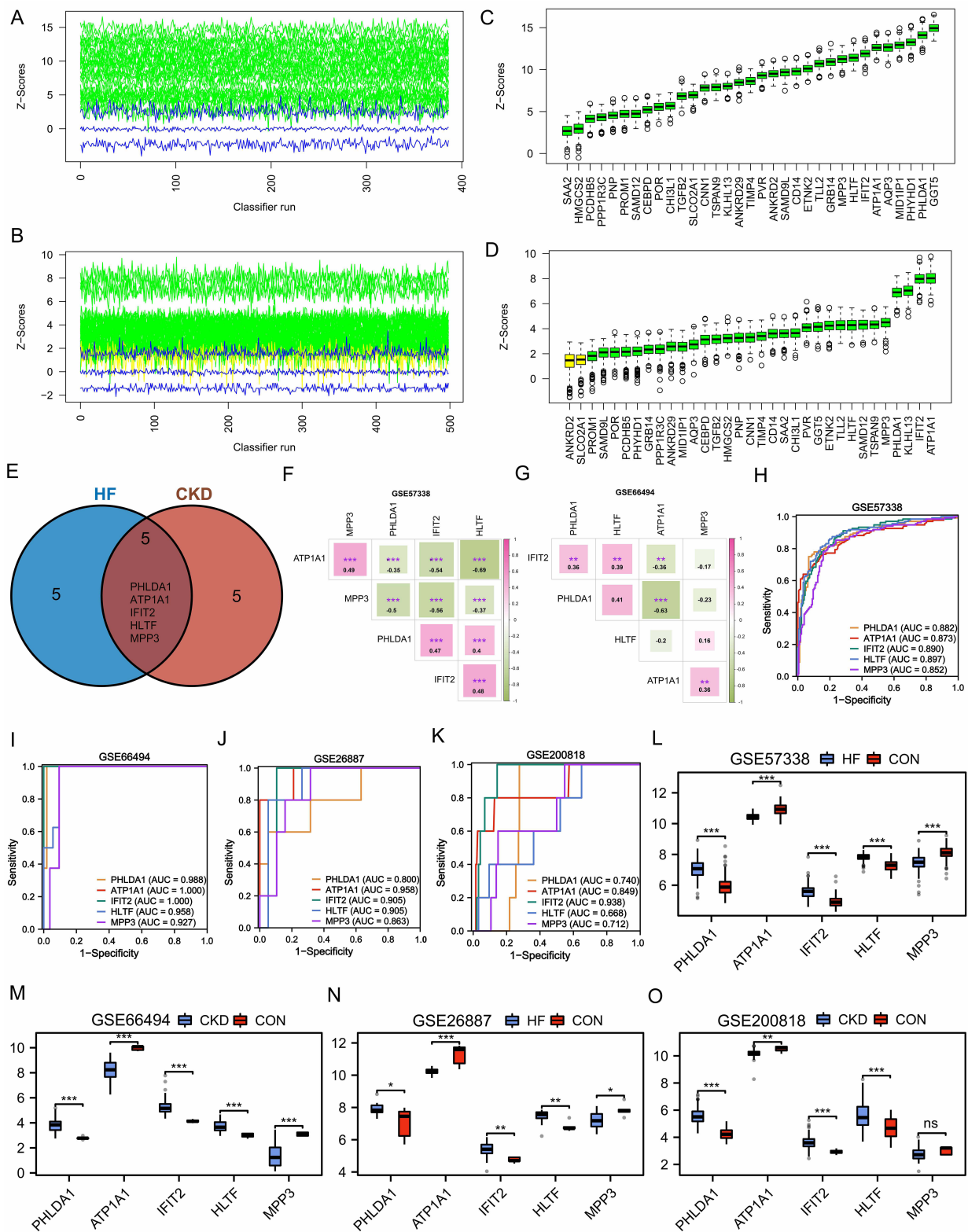
In the HF cohort, 33 crosstalk genes obtained rankings of gene importance after 400 iterations (Figure 4A). The Z-scores of each gene were compared with the maximum Z-score among shadow attributes (MZSA), and it was evident that the Z-scores of all genes were significantly higher than the MZSA of shadow attributes (Figure 4B). In the CKD cohort, 33 crosstalk genes were ranked by 500 iterations to obtain the importance of the genes (Figure 4C). The Z-Score of each gene was compared with the MZSA, which was significantly higher than the MZSA of shadow attributes, except for ANKRD2 and SLC22A1, which were considered to be to-be-confirmed variables (Figure 4D). Five overlapping crosstalk genes (PHLDA1, ATP1A1, IFIT2, HLTF, and MPP3) were screened by intersecting the top ten genes in the Z-Score scores of the two cohorts (Figure 4E). The intrinsic correlations among these five genes are shown in Figure 4F and G. In the GSE57338 dataset, ATP1A1 is significantly positively correlated with MPP3, while it is significantly negatively correlated with HLTF. In the GSE66494 dataset, PHLDA1 is positively correlated with HLTF and negatively correlated with ATP1A1. We utilized ROC curves to assess the diagnostic capability of the overlapping perturbed genes. In both the HF dataset (GSE57338) and the CKD dataset (GSE66494), the five genes demonstrated strong diagnostic performance, with all genes having areas under the curve (AUC) exceeding 0.8 (Figure 4H and I). Further confirmation of the diagnostic accuracy of the five potential genes was verified by pertinent external datasets, GSE26887 and GSE200818 (Figure 4J and K). Figures 4L-O illustrate the differential gene expression patterns of the five candidate biomarkers in the HF (GSE57338 and GSE26887) and CKD (GSE66494 and GSE200818) datasets. Compared to the control group, there were significant upregulations of PHLDA1, IFIT2 and HLTF, while ATP1A1 and MPP3 were downregulated. This analysis supports the diagnostic potential and clinical relevance of these five biomarkers in disease diagnosis and treatment.

## Expression of Shared Genes in Single-Cell Datasets

After processing the single-cell data for HF and healthy individuals, a total of 18,104 cells were obtained (Supplementary Figure 1A-C). After eliminating the batch effect of the samples, a resolution of 0.3 was selected as the most suitable for segmenting the single-cell subpopulations, resulting in the acquisition of a total of 14 single-cell subpopulations (Figure 5A and Supplementary Figure 1D and E). Depending on the marker genes of the different cells, the cell populations were classified into nine distinct cell lineages, including fibroblasts, endothelial cells, B cells, neutrophils, NK cells, macrophages, smooth muscle cells, neuronal cells, and cardiomyocytes (Figure 5B). The expression of marker genes in each cell subgroup is depicted in Figure 5C. Additionally, we investigated the expression levels of the five common diagnostic genes across various cellular subpopulations. Fibroblasts, smooth muscle cells, NK cells and macrophages exhibited predominant expression of PHLDA1 (Figure 5D). ATP1A1 was detected in all types of cell subpopulations (Figure 5E), whereas IFIT2 expression was primarily observed in endothelial cells and macrophages (Figure 5F). HLTF expression was not markedly detected in the majority of cells (Figure 5G). MPP3 expression, on the other hand, was prominently observed in endothelial cells and smooth muscle cells (Figure 5H).

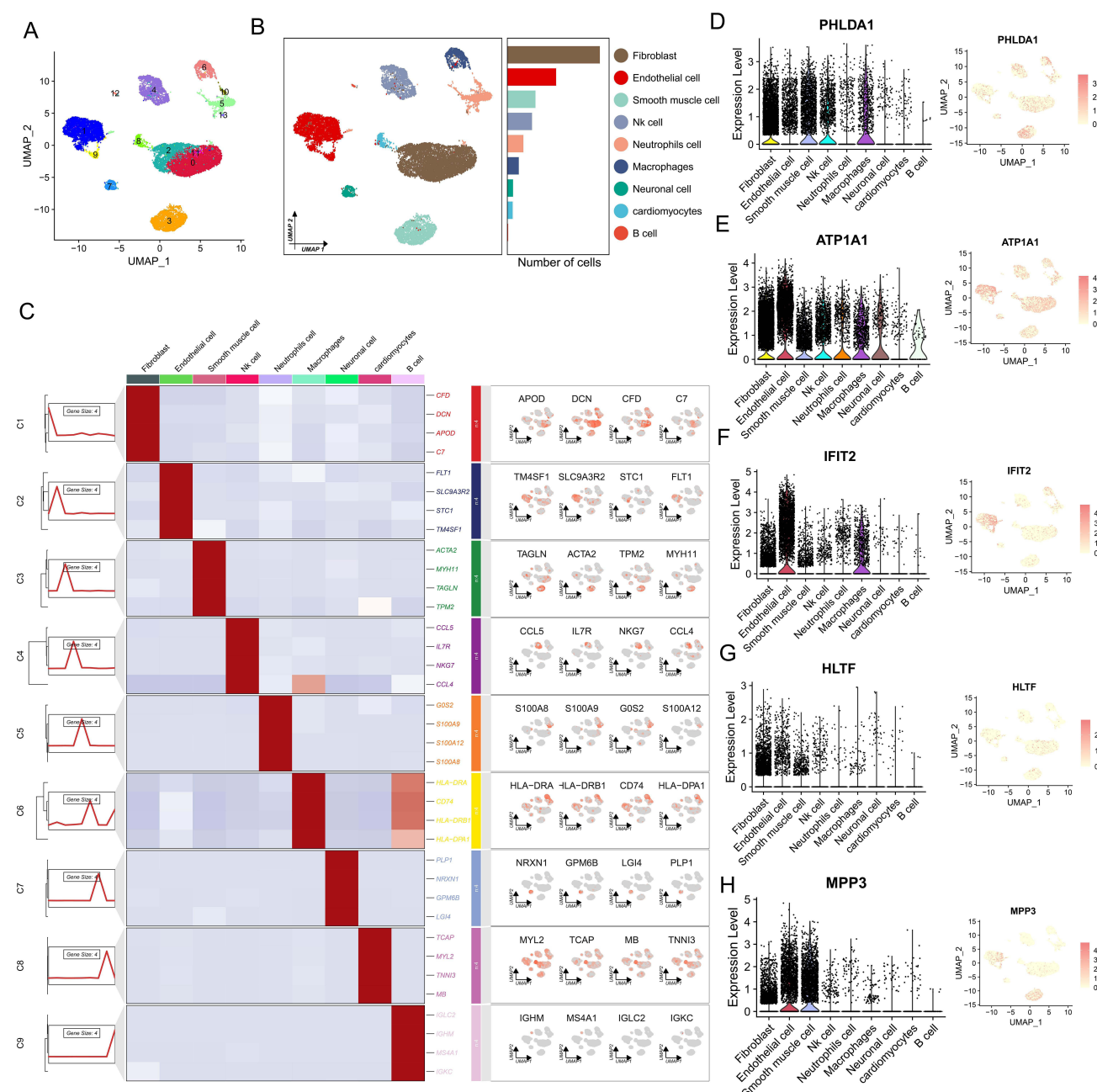
## Comparison of the Immune Microenvironment in HF and CKD Cohorts

To investigate the immune profile in samples with HF and CKD, the CIBERSORT-based algorithm was employed to forecast differences in immune infiltration between HF and healthy controls, as well as between CKD and healthy controls. The findings demonstrate higher percentages of Plasma cells, T cells CD8, and resting Mast cells in HF patients than in normal controls. In contrast, the percentages of T cells CD4 memory resting, T cells regulatory (Tregs), Monocytes, and Macrophages M2 were lower than those in normal controls (see Figure 6A). The percentage of activated NK cells and Macrophages M2 is lower in CKD patients than in healthy patients, while that of Monocytes is higher (Figure 6B). Thus, there are similarities and differences in the activity of immune cells among patients with HF and CKD. Furthermore, the use of box plots illustrated dissimilarities in immune cell infiltration amongst 22 patients with HF. Specifically, individuals with HF displayed an immune-activated state in plasma cells, T cells CD8 and Mast cells resting.



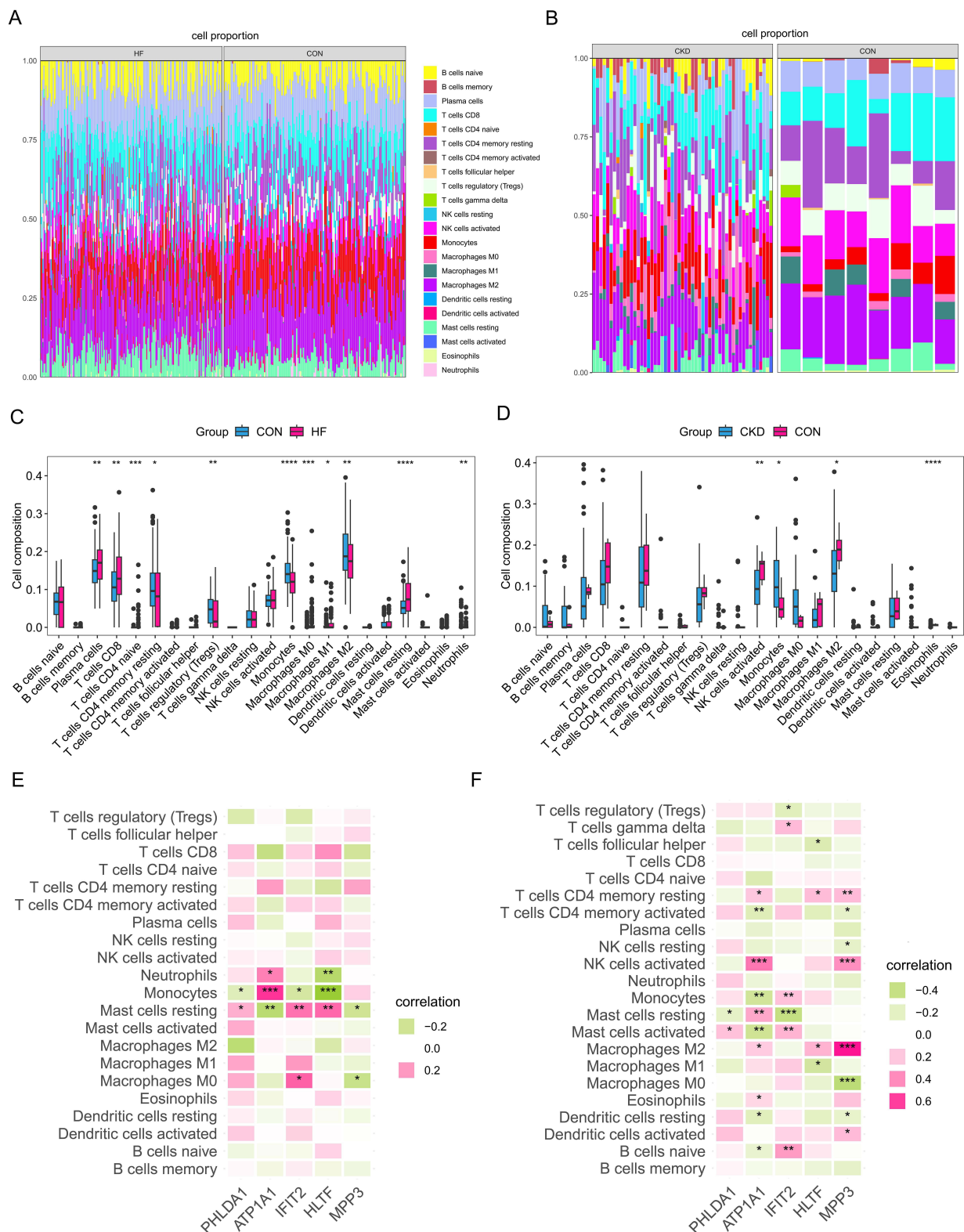
**Figure 4** Identification and Validation of Shared Diagnostic Genes. (A and B) Changes in Z-scores during the Boruta algorithm run. Green lines correspond to confirmed features, yellow represents pending features, and blue indicates the importance of shadow features in minimum, average, and maximum. (C and D) Variable importance ranking results from the Boruta algorithm for feature selection. (E) Venn diagram displaying the best variables with diagnostic value. (F and G) Correlations of five shared crosstalk genes in the GSE57338 and GSE66434 datasets. (H-K) ROC curves assessing the diagnostic value of shared genes in the GSE57338, GSE66494, GSE200818, and GSE26887 datasets. (L-O) Expression levels of five shared diagnostic genes in the GSE57338, GSE66494, GSE200818, and GSE26887 datasets. ns,  $p>0.05$ ; \* $p<0.05$ ; \*\* $p<0.01$ ; \*\*\* $p<0.001$ .





**Figure 5** Identification of Cell Subtypes in Single-Cell Dataset. **(A)** Shared Nearest Neighbor (SNN) modularized optimized clustering algorithm divides cells into 14 subgroups, with the UMAP plot illustrating the clustering of different cell types. **(B)** The 14 subgroups are annotated as 9 cell types. **(C)** Expression levels of marker genes in different subgroups. **(D-H)** Expression levels of shared diagnostic genes in different cell subtypes.

Conversely, T cells CD4 memory resting, T cells regulatory (Tregs), Monocytes and Macrophages M2 exhibited an immunosuppressed state (Figure 6C). NK cells activated and Macrophages M2 exhibited an immunosuppressed state while Monocytes were in an activated state in CKD patients (Figure 6D). These analyses indicate that immune cells and responses have a significant impact on the development of HF or CKD, with particular importance placed on monocytes and macrophages. The study investigated the correlation between shared diagnostic genes and the infiltration of 22 immune cells. The results indicated a significant association between the expression of five biomarkers with Mast cells resting and monocytes, primarily in HF patients (Figure 6E). Additionally, in the CKD group, the expression of five biomarkers showed significant association with the activity of most immune cells (Figure 6F). In conclusion, our study



**Figure 6** Immune Cell Infiltration Landscape in Heart Failure and Chronic Kidney Disease. (**A** and **B**) Immune cell abundance in each sample from the GSE57338 and GSE66494 datasets. (**C** and **D**) Differences in immune cell infiltration between disease and healthy groups in the GSE57338 and GSE66494 datasets based on CIBERSORT. (**E** and **F**) Correlation of hub Crosstalk Genes with 22 types of immune cells. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ .

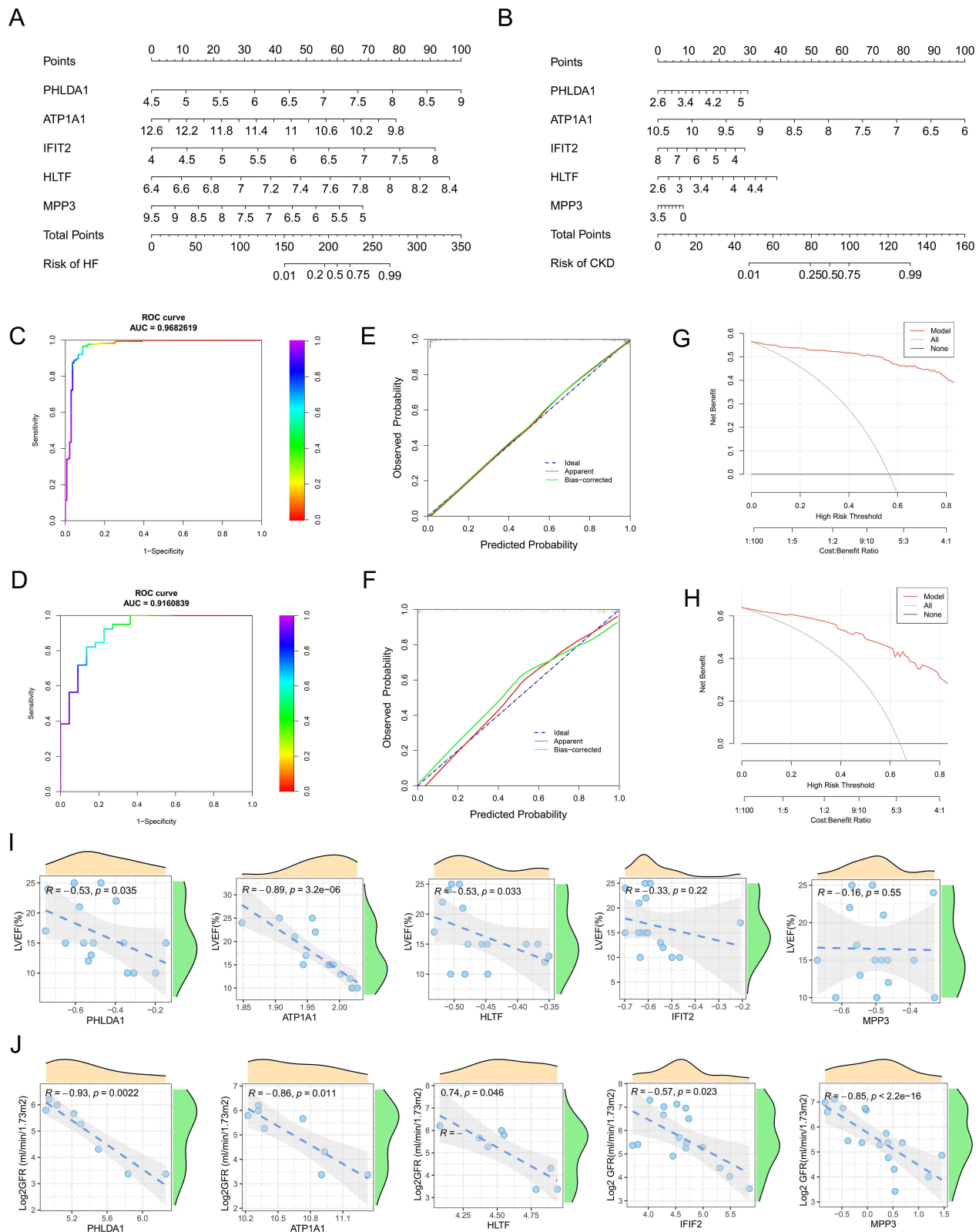
indicates a strong association between monocytes, NK cells, and macrophages with HF and CKD. The expression of PHLDA1, ATP1A1, IFIT2, HLTF, and MPP3 may either inhibit or promote immune cell infiltration.

## Construction and Validation of Clinical Diagnostic Models and Correlation of Cardiac and Renal Function with Characterized Genes

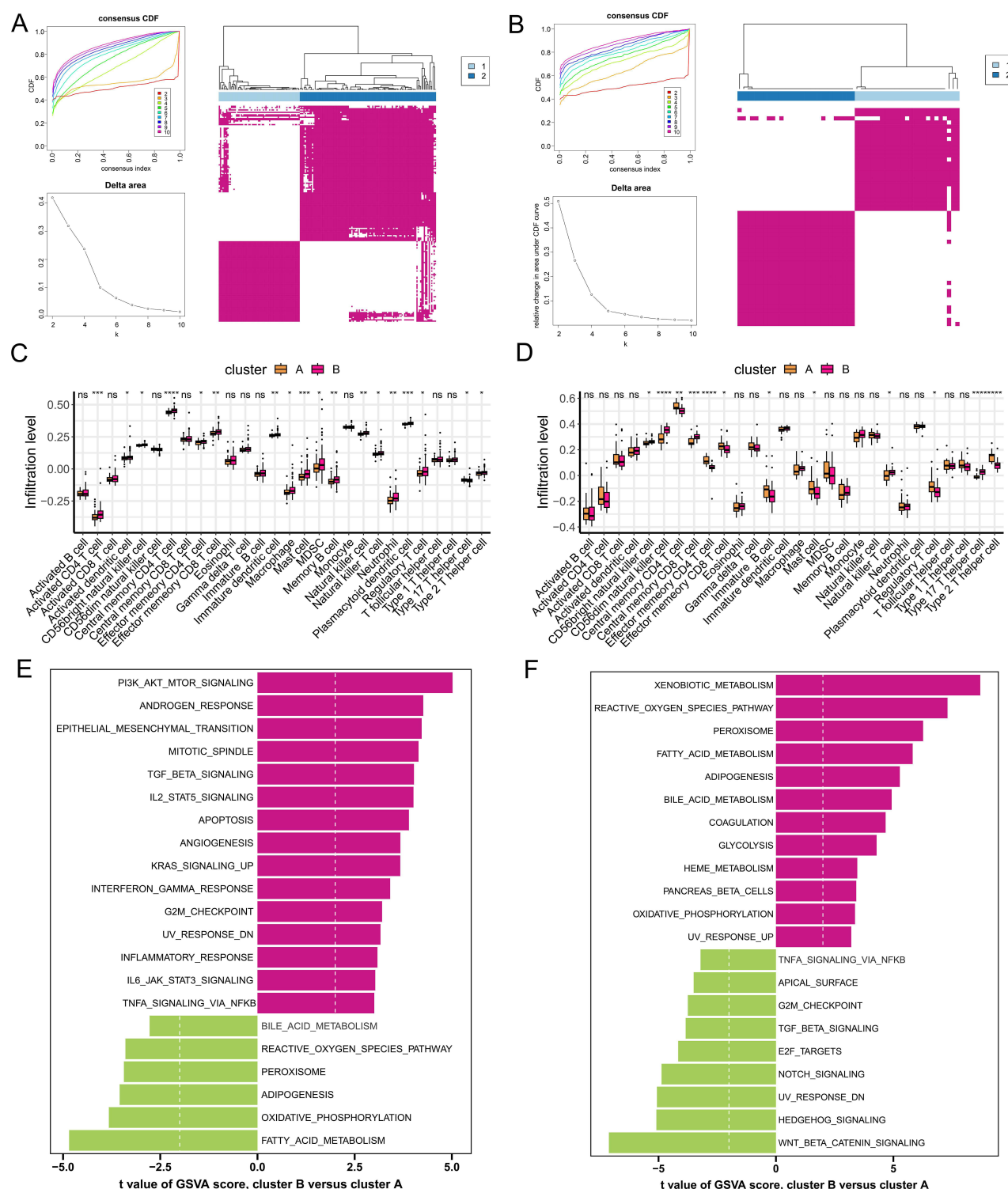
To explore the value of shared diagnostic genes in clinical applications, we established diagnostic models utilizing the “rms” R package, which rely on the expression levels of five biomarkers drawn from the GSE57338 and GSE66464 datasets (Figure 7A and B). The model’s diagnostic effectiveness was evaluated using the ROC curve, calibration curve and DCA curve. The ROC curves revealed that both models had AUC values exceeding 0.95, highlighting the high diagnostic efficiency of the model (Figure 7C and D). The calibration curves demonstrated immaterial disparities between predicted and actual risks, showcasing the model’s usefulness in predicting the occurrence of diseases (Figure 7E and F). DCA analysis demonstrates that the advantageous curve of the predictive model exceeds that of the reference strategy, yielding clinical usefulness at the treatment decision threshold. Moreover, patients stand to gain from the model’s implementation (Figure 7G and H). The diagnostic model developed using shared genes demonstrates potent diagnosis and aids precise clinical decisions with regards to disease prognostication. Additionally, the correlation analysis between the quintet of genes and the left ventricular ejection fraction and glomerular filtration rate was further scrutinized for diagnostic purposes. Spearman correlation analysis demonstrated a negative correlation between the expression levels of PHLDA1, ATP1A1, and HLTF with left ventricular ejection fraction. However, there was no statistically significant correlation found between the expression levels of IFIT2 and MPP3 with left ventricular ejection fraction (Figure 7I). Moreover, all five genes - PHLDA1, ATP1A1, IFIT2, HLTF, and MPP3, were found to have a negative correlation with glomerular filtration rate (Figure 7J). Of these genes, ATP1A1 demonstrated the strongest correlation with left ventricular ejection fraction ( $R = -0.89$ ,  $P = 3.2e-5$ ), whilst PHLDA1 exhibited the strongest correlation with glomerular filtration rate ( $R = -0.93$ ,  $P = 0.0022$ ). These findings indicate that the expression levels of PHLDA1, ATP1A1, and HLTF could accurately determine the extent of heart failure. Additionally, the expression of all five genes may be indicative of the severity of renal injury and could potentially contribute to the progression of the condition. The clinical importance and accuracy of common biomarkers in predicting and diagnosing various diseases have been further validated.

## Identification of Molecular Subtypes in HF and CKD Cohorts

According to the cumulative distribution function curve and the Delta curve, the ideal number of clusters for the HF and CKD samples is  $K=2$ . The patients with HF were divided into two subgroups based on the expression pattern of crosstalk genes, with cluster A containing 65 patients and cluster B containing 112 patients (Figure 8A). Similarly, CKD patients were categorized into two subgroups according to the expression pattern of crosstalk genes. Cluster A comprised 25 patients, while Cluster B included 28 patients (Figure 8B). To explore the immunological characteristics of immune cells in HF and CKD at the molecular level, we quantitatively analyzed the relative abundance of 28 immune cell types in each sample using the ssGSEA algorithm. For HF patients, cluster B exhibited a higher degree of immune cell infiltration than cluster A. Furthermore, immune cells such as Central memory CD4 T cells, Macrophages, and Mast cells were in an activated state (Figure 8C). In contrast, for CKD patients, the level of immune infiltration exhibited lower levels in cluster B than in cluster A (Figure 8D). Therefore, we categorized cluster B associated with HF as the high immunity group and cluster A associated with CKD as the low immunity group. Gene enrichment analysis in the two clusters was performed using GSVA and differences in pathways or functions were analyzed using the “Limma” package. Cluster A in the HF cohort is primarily engaged in biological processes like fatty acid metabolism and oxidative phosphorylation. Cluster B was primarily involved in immune, inflammatory and cell development pathways such as PI3K AKT MTOR signaling, Apoptosis, Inflammatory response, IL6 JAK STAT3 signaling and TNFA signaling via NFkB (Figure 8E). Both subgroups of CKD patients displayed the same expression patterns as those of HF patients. It was observed that Cluster A was significantly enriched in pathways related to immunity and cell development, whilst Cluster B was mainly involved in amino acid metabolism, fatty acid metabolism, and metabolic processes of small molecules (Figure 8F).



**Figure 7** Construction of Nomogram and Clinical Functional Analysis. (**A** and **B**) Nomogram predicting the probability of heart failure and chronic kidney disease occurrence. (**C** and **D**) ROC curves assessing the clinical discriminatory ability of the model. (**E** and **F**) Calibration curves assessing the model's calibration ability. (**G** and **H**) Decision curve analysis (DCA) curves assessing the clinical decision benefits of the model. (**I**) Scatter plot showing the correlation of shared Crosstalk Genes with left ventricular ejection fraction. (**J**) Scatter plot showing the correlation of shared Crosstalk Genes with glomerular filtration rate.



**Figure 8** Consensus Clustering Analysis Identifying Molecular Subtypes Associated with Heart Failure and Chronic Kidney Disease. (**A** and **B**) Cumulative distribution function curves of consensus scores, changes in area under the Delta curve, and consensus clustering heatmap at K=2-10. (**C** and **D**) Analysis of immune cell distribution in Crosstalk Gene subtypes Cluster 1 and Cluster 2 based on the ssGSEA algorithm. (**E** and **F**) GSVA reveals significant differences in biological pathways between subtypes of heart failure and chronic kidney disease. ns,  $p>0.05$ ; \* $p<0.05$ ; \*\* $p<0.01$ ; \*\*\* $p<0.001$ ; \*\*\*\* $p<0.0001$ .

Taken together, the consistent clustering clusters further confirmed the link between crosstalk genes and immune cell infiltration and metabolic landscape. Additionally, we discovered consistent expression patterns of subclustered genes in HF and CKD. Cluster A for HF was for metabolic subtypes, whereas cluster B was for immune subtypes. In contrast,



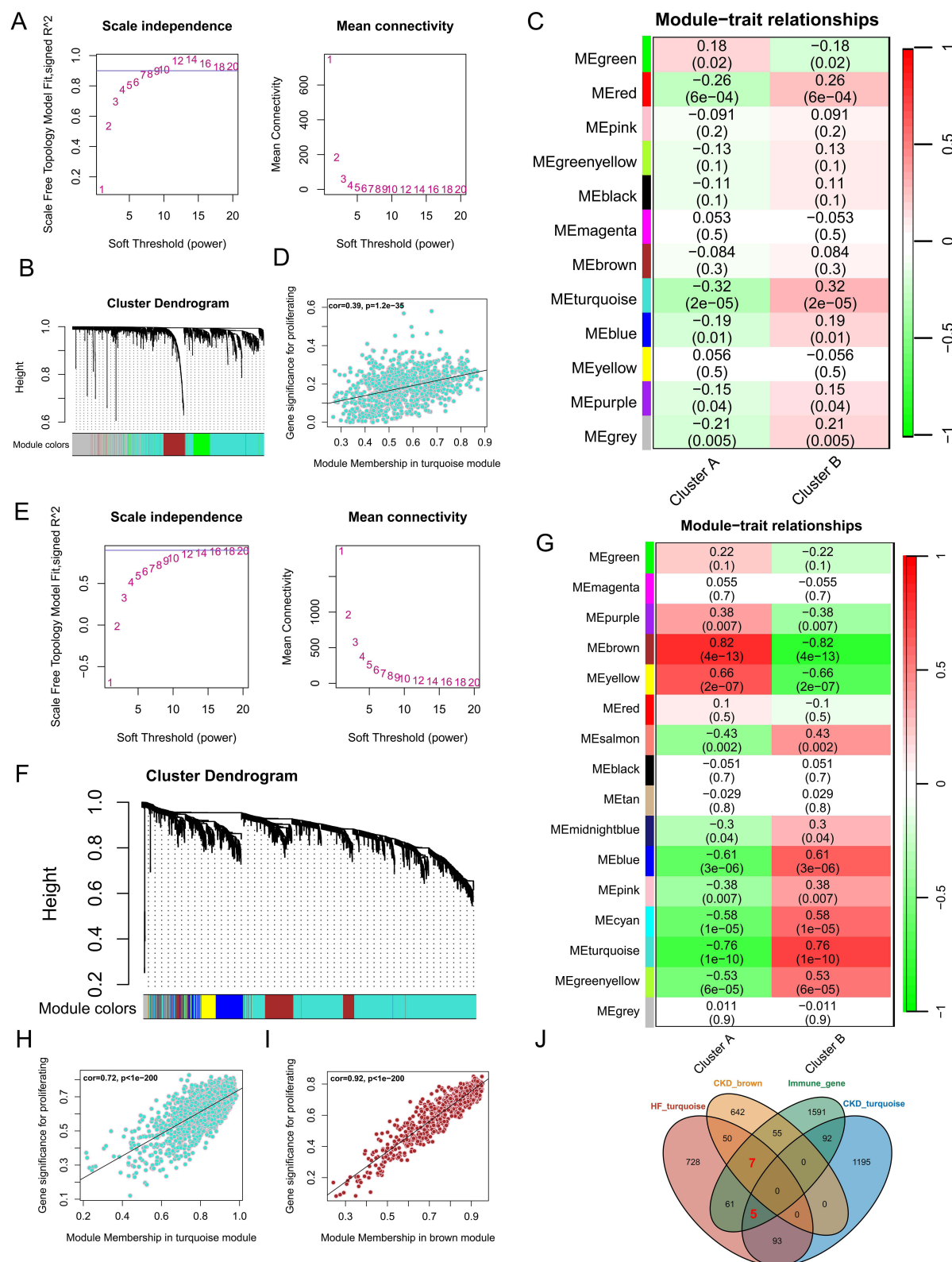
cluster A for CKD was the immune subtype, while cluster B was the metabolic subtype. The interplay between HF and CKD in the pathogenesis at a molecular level was further authenticated. The diagnosis and treatment of the ailment could be made more precise by identifying various subtypes.

## WGCNA Screening of Immune-Related Shared Genes in HF and CKD Subtypes

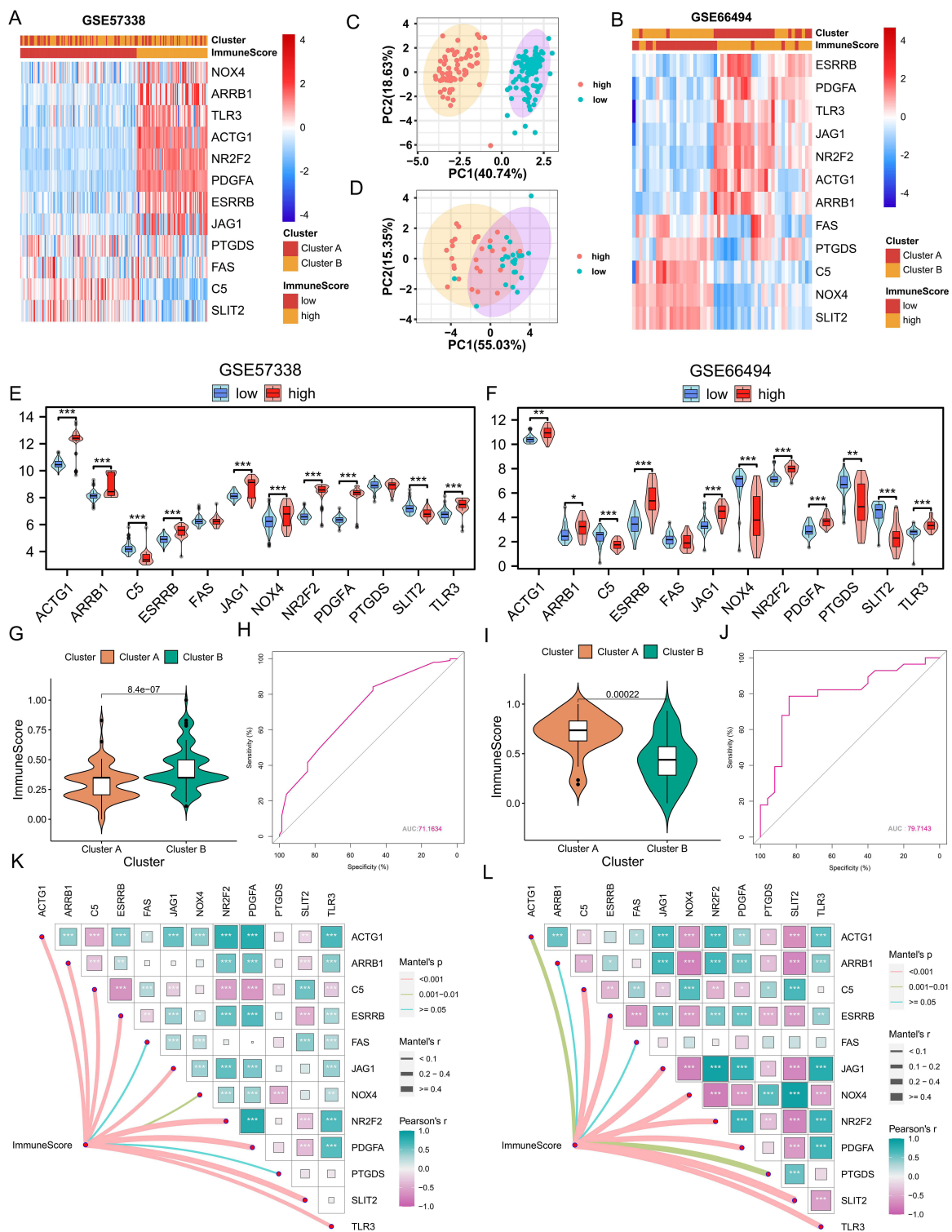
The outcomes from the unsupervised cluster analysis indicate the presence of an identical immune subtype between patients with HF and CKD. Further, we conducted a WGCNA analysis to investigate immune-related genes within the subtypes. The samples in the HF and CKD datasets were initially clustered and analysed to find outlier samples. We constructed scale-free networks after eliminating three outlier samples each from the GSE57338 dataset (GSM79815, GSM79996, and GSM80018) and the GSE66494 dataset (GSM23315, GSM23352, and GSM23356). In the dataset GSE57338 for HF, which exhibits smoother connectivity at  $R^2 = 0.9$ , we identified  $\beta = 9$  as the optimum soft threshold (Figure 9A). The clusters beneath the clustering tree for all samples are presented in Figure 9B. Twelve clusters associating with both subtypes were detected (Figure 9C). MEturquoise, one particular cluster, exhibited the highest association with immune subtype B ( $R=0.32$ ,  $p=2e-05$ ). Additionally, we identified a significant relation between module members of MEturquoise and gene significance (Figure 9D). In the CKD dataset GSE66494, which has a smoother connectivity when  $R^2 = 0.9$ ,  $\beta = 12$  was chosen as the optimal soft threshold (Figure 9E). Figure 9F shows the lower module of the clustering tree for all samples. A total of 16 modules related to the two subtypes were identified (Figure 9G). The module MEturquoise had the robust negative correlation with immune subtype A ( $R = -0.76$ ,  $P = 1e-10$ ), while the module MEbrown had the highest positive correlation with immune subtype A ( $R = 0.82$ ,  $P = 4e-13$ ). In addition, a robust correlation was observed between the membership of the module and the importance of genes for MEturquoise and MEbrown (Figure 9H and I). To filter out immune genes associated with immune subtypes, genes from the three modules were intersected with the set of immune genes to obtain 12 immune genes shared by HF and CKD (Figure 9J). These 12 immune genes were ultimately utilised to establish the ImmuneScore model.

## Construction of ImmuneScore Model Based on Expression Patterns of Immune-Related Shared Genes

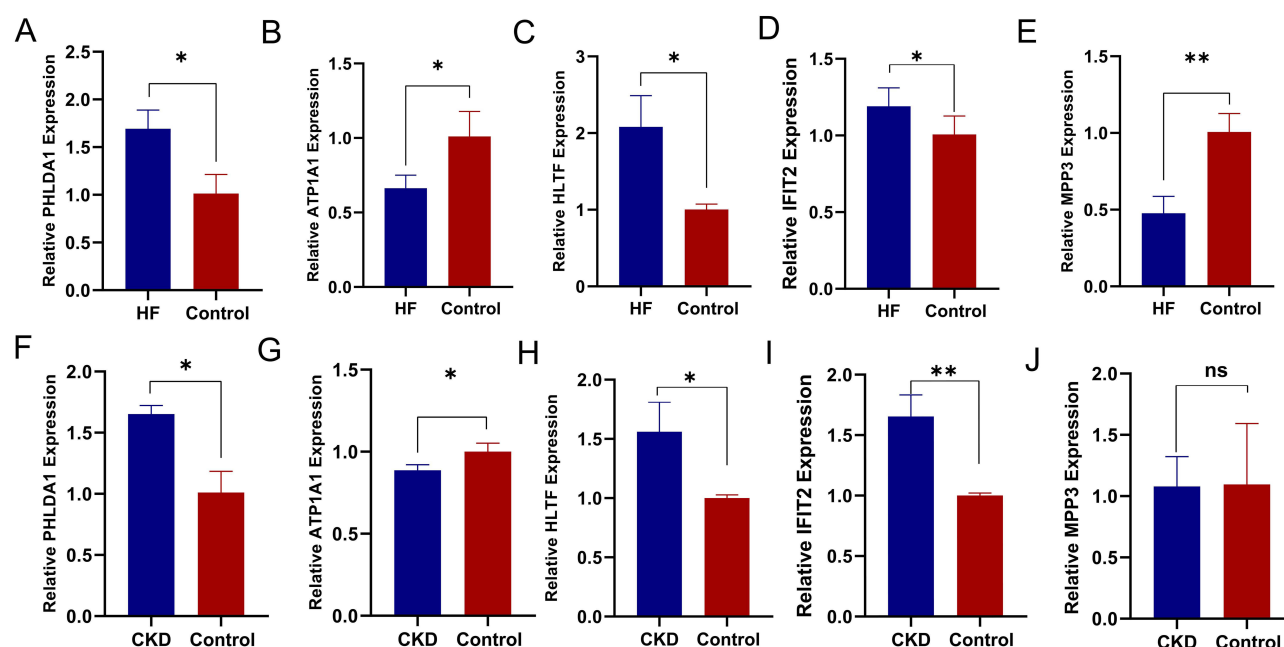
Based on the median ssGSEA score we categorized patients into high and low risk groups (Supplementary Table 3 and 4). The expression levels of the 12 overlapping immune genes in the high-risk and low-risk groups are shown in Figure 10A and B. PCA clearly shows the sample distribution of the high-risk and low-risk groups, and ImmuneScore can clearly categorize the patients into high- and low-risk groups (Figure 10C and D). Significant differences in the expression levels of most genes, with the exception of FAS and PTGDS, were observed between high- and low-risk groups of HF patients (Figure 10E). Also, the expression levels of all genes, except FAS, exhibited significant variation between the high- and low-risk groups of CKD patients (Figure 10F). The study investigated the capability of the ImmuneScore model to differentiate between immune and non-immune subtypes. Box plots revealed significant differentiation in patients with HF or CKD ( $p = 8.4e-7$ ,  $p = 0.00022$ ), with higher ImmuneScore in the immune subgroup compared to the control group (Figure 10G and I). Additionally, the ImmuneScore model's discriminative capability was evaluated using the AUC value. For the GSE57338 dataset, the AUC value was 0.71634 (Figure 10H), and for the GSE66494 dataset, it was 0.797143 (Figure 10J). The ImmuneScore model has strong diagnostic power for patients with different molecular subtypes. In addition, the network heatmap shows the correlation of ImmuneScore with 12 immune genes. In patients with HF, the findings revealed a strong association between ACTG1, PDGFA and SLIT2 with a higher likelihood of being classified in the high-risk group as per the ImmuneScore model. Furthermore, these genes may have a significant role in the development of HF (Figure 10K). Similarly, in CKD patients, C5, ESRRB, JAG1, NOX4, NR2F2, PDGFA, TLR3, PTGS2, and SLIT2 were highly associated with high-risk groups in the ImmuneScore model (Figure 10L). Based on the results, ACTG1, PDGFA, and SLIT2 emerge as significant genes with potential roles in the high-risk ImmuneScore group.



**Figure 9** Selection of Immune-Related Crosstalk Genes using the WGCNA. **(A)** The optimal soft threshold for unscaled network fit in heart failure samples is 5, which provides the smoothest connectivity. **(B)** Clustering of genes with similar functions in the co-expression network. Each line represents one gene. **(C)** Correlation heatmap of module-clustering in subtypes of heart failure patients. **(D)** Scatter plot showing the correlation between genes in the MEturquoise module and phenotypes in heart failure subtypes. **(E)** The optimal soft threshold for unscaled network fit in chronic kidney disease samples is 12, which provides the smoothest connectivity. **(F)** Clustering of genes with similar functions in the co-expression network. Each line represents one gene. **(G)** Correlation heatmap of module-clustering in subtypes of chronic kidney disease patients. **(H)** Scatter plot showing the correlation between genes in the MEturquoise module and phenotypes in chronic kidney disease subtypes. **(I)** Scatter plot showing the correlation between genes in the MEbrown module and phenotypes in chronic kidney disease subtypes. **(J)** Venn diagram showing 12 co-expressed immune-related genes in the HF and CKD subgroups.



**Figure 10** Construction of the ImmuneScore Model. (A) Expression profiles of immune-related Crosstalk Genes in high-risk and low-risk groups of ImmuneScore based on heart failure samples from GSE57338. (B) Expression profiles of immune-related Crosstalk Genes in high-risk and low-risk groups of ImmuneScore based on chronic kidney disease samples from GSE66494. (C) PCA plot of high-risk and low-risk patients in heart failure subtypes. (D) PCA plot of high-risk and low-risk patients in chronic kidney disease subtypes. (E) Boxplots showing differential expression levels of Crosstalk Genes in high-risk and low-risk groups of heart failure patients. (F) Boxplots showing differential expression levels of Crosstalk Genes in high-risk and low-risk groups of chronic kidney disease patients. (G and H) Boxplots depicting the distribution of ImmuneScore in heart failure subtypes and ROC curves assessing model accuracy. (I and J) Boxplots depicting the distribution of ImmuneScore in chronic kidney disease subtypes and ROC curves assessing model accuracy. (K) Heatmap showing the network correlation of 12 immune-related genes with ImmuneScore in heart failure cohorts. (L) Heatmap showing the network correlation of 12 immune-related genes with ImmuneScore in chronic kidney disease cohorts. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .



**Figure 11** The expression levels of key crosstalk genes in human tissues. (A-E) Expression levels of PHLDA1, ATP1A1, HLTF, IFIT2, and MPP3 in heart failure and normal samples (n=3). (F-J) Expression levels of PHLDA1, ATP1A1, HLTF, IFIT2, and MPP3 in chronic kidney disease and normal samples (n=3). \*, P < 0.05, \*\*, P < 0.01.

## Validation of the Expression Levels of Key Crosstalk Genes in Human Tissues

RT-PCR was performed on biopsy samples from normal and HF tissues as well as from healthy and CKD. The expression levels of the genes were consistent with the analysis of the bioinformatics data, except for MPP3, which did not show any significant difference between diseased and healthy patients. Expression of PHLDA1, HLTF, and IFIT2 was up-regulated in HF and CKD, whereas ATP1A1 expression was down-regulated (Figure 11A-J).

## Discussion

The incidence of HF combined with CKD is increasing with age. However, its pathogenesis, treatment and prognosis have not been fully and accurately understood due to its late start and large research system involving multiple complex disciplines such as heart and kidney. A growing body of evidence suggests that in addition to hemodynamic alterations as the primary drivers of both diseases, non-hemodynamic alterations such as neurohormonal activation, immune dysregulation and inflammatory responses are present throughout cardiovascular disease and CKD.<sup>28</sup> Inflammatory cytokines such as interleukin-1 $\beta$  (IL-1 $\beta$ ), tumor necrosis factor-alpha (TNF-alpha), interleukin-6 (IL-6) and C-reactive protein (CRP) are the major inflammatory mediators involved in disease development, and the various inflammatory mediators also play important synergistic roles with each other. These inflammatory mediators also play important synergistic roles among themselves.<sup>29</sup> Among them, TNF- $\alpha$  and IL-6 can cause sodium retention and renal fibrosis, resulting in left ventricular dilatation and dysfunction, leading to ventricular remodeling. Activation of the protein kinase C pathway by the RASS system generates a large amount of reactive oxygen species (ROS), resulting in the onset of oxidative stress, and an excess of reactive oxygen species can induce necrosis of the renal unit, myocardial fibrosis.<sup>30</sup> Therefore, it appears that there is a complex and intricate relationship between these two diseases. In addition, the algorithms of machine learning have been respected by many researchers due to their accurate determination of feature selection, model construction, and other applications in cardiovascular and renal diseases.<sup>31,32</sup> Therefore, this paper identifies differentially expressed tandem genes and explores the biological functions of tandem genes based on HF and CKD microarray datasets. Common diagnostic biomarkers for both diseases were screened using machine learning algorithms and their association with cardiac and renal function was analyzed. The abundance of immune cell infiltration in HF and CKD was assessed and the relationship between the biomarkers and immune cell infiltration was investigated. Consensus cluster

analysis classified HF and CKD patients into immune and metabolic subtypes. Finally, the ImmuneScore model was constructed by ssGSEA scoring, and the model can effectively distinguish different subtypes.

Genes with consistent expression trends in HF and CKD samples were intersected to obtain crosstalk genes, including 13 up-regulated and 20 down-regulated genes. To elucidate the biological functions of the crosstalk genes, we performed GO enrichment analysis. The crosstalk genes were mainly involved in positive regulation of immune system processes, apoptosis, response to steroid hormone, cell-cell adhesion via plasma membrane adhesion molecules, positive regulation of interleukin-8 production and inflammatory response, among other biological processes. We then used machine learning algorithms to solidify five common diagnostic genes, including PHLDA1, ATP1A1, IFIT2, HLTF and MPP3, and the ROC curves further confirmed their value in diagnosis. Pleckstrin homology structural domain A member 1 (PHLDA1) encodes 401 amino acids and is increasingly recognized as a key determinant of immune regulation and apoptosis.<sup>33</sup> It has been reported that PHLDA1 can induce apoptosis in T cells, myocardial cells, endothelial cells, and kidney cells.<sup>34,35</sup> So far, numerous studies have confirmed the association of PHLDA1 with various types of heart diseases, such as cardiomyopathy<sup>36,37</sup> and myocardial injury.<sup>38</sup> PHLDA1 induces renal apoptosis and promotes renal interstitial fibrosis via the TGF- $\beta$ -IRE1-XBP1 pathway in CKD patients.<sup>39</sup> In addition, by interacting with TRAF6, PHLDA1 can activate the NF- $\kappa$ B signalling pathway and promote inflammatory responses.<sup>40</sup> ATP1A1 is an important cell membrane Na<sup>+</sup>/K<sup>+</sup>-ATPase enzyme that plays a crucial role in the regulation of aldosterone production and Na/K transport.<sup>41</sup> Chronic heart failure manifests as a decrease in cardiac output and effective circulating fluid volume, leading to activation of the RAAS and pressor receptors, ultimately resulting in renal tissue ischemia, hypoxia, apoptosis and necrosis of renal cells, leading to renal insufficiency.<sup>13</sup> Furthermore, ATP1A1 can also enhance macrophage infiltration and adhesion of renal epithelial cells, ultimately promoting renal inflammation.<sup>42</sup> IFIT2 is a member of the interferon (IFN) family, which is involved in a number of signaling cascades that regulate the intracellular immune response to attenuate myocardial injury and remodeling.<sup>43</sup> Additionally, IFIT2 can also induce cell apoptosis, which may be associated with the inhibition of protein synthesis and activation of mitochondrial responses.<sup>44</sup> Overexpression of IFIT2 is associated with a worse clinical outcome in patients with arrhythmias and HF. Therefore, combining IFIT2 levels with biomarkers such as C-reactive protein may provide greater clinical utility in diagnosing the disease. The helicase-like transcription factor (HLTF), a member of the Switch/sucrose non-fermenting (SWI/SNF) family, has been found to be highly methylated in many cancers.<sup>45</sup> Previous studies have suggested that high expression of HLTF can protect cells from the effects of apoptosis, and this protective effect may be achieved by impairing the activity of lysosomal autophagy inhibitors.<sup>46</sup> HLTF can act as a tumor promoter and is associated with poor prognosis.<sup>47</sup> Most studies of HLTF have focused on the pathogenesis of tumors, with few reports in HF and CKD, which may become effective biomarkers. In response to inflammatory stimuli, MPP3 rapidly enhances the secretion of many pro-inflammatory/myeloid differentiation cytokines by granulocyte/macrophage progenitor (GMP) and myeloid cells via autocrine/paracrine signaling.<sup>48</sup> We therefore concluded that MPP3 is associated with upregulation of inflammatory cytokine expression in cardiac and renal diseases, with important synergistic roles between different inflammatory mediators. Taken together, we identified five common diagnostic genes that could serve as effective biomarkers. Immune response, apoptosis and inflammatory response may be synergistic in both diseases.

In addition, the CIBERSORT algorithm revealed differences in immune infiltration between HF and healthy controls as well as between CKD and healthy controls. Plasma cells, CD8 T cells and mast cells showed immune activation, whereas CD4 memory T cells, regulatory T cells (Tregs), monocytes and M2 macrophages showed immune suppression in HF patients. Mast cells may exacerbate HF progression by releasing protein kinase A (PKA) and myocardial fibrosis.<sup>49</sup> It has been suggested that inhibition of CD4<sup>+</sup> T-cell activity may mimic the transition from hypertrophy to heart failure during pressure overload, thereby protecting cardiac function.<sup>50</sup> Macrophage-mediated innate immune response is essential in chronic kidney injury.<sup>51</sup> In our study we found that NK cells activated and Macrophages M2 exhibited an immunosuppressed state while Monocytes were in an activated state in CKD patients. Previous studies have indicated that the activity of M2 macrophages predicts the degree of renal fibrosis, but not M1 macrophages, which is consistent with our study.<sup>52,53</sup> These analyses suggest that these immune cells and immune responses play a critical role in the progression of HF or CKD. In addition, the expression of five biomarkers was significantly correlated with the activity of most immune cells. In conclusion, our study suggests that monocytes, NK cells and macrophages are strongly associated



with HF and CKD, and the expression of PHLDA1, ATP1A1, IFIT2, HLTF and MPP3 may activate the process of immune cell infiltration through metabolic disorders.

We further classified HF and CKD patients into two subtypes based on the expression of crosstalk genes: cluster A of HF was the metabolic subtype, whereas cluster B showed higher immune cell infiltration and stronger immune activation, which we defined as the immune subtype. In contrast, CKD is clustered A for the immune subtype and B for the metabolic subtype. The different subtypes have different patterns of immune infiltration between them. We then constructed a risk score model based on the median ImmuneScore score using crosstalk genes in immune subtypes. Patients in the high ImmuneScore group showed significant immune infiltration and immune response. The ImmuneScore model can significantly discriminate between immune subtypes and non-immune subtypes in patients with HF or in patients with CKD. Correlation analysis showed that SLIT2, NR2F2 and PDGFA were significantly and positively correlated with ImmuneScore scores. SLIT2 ameliorates inflammatory response and renal fibrosis after endothelial injury.<sup>54</sup> In addition, SLIT2 is activated during cardiac fibrosis.<sup>55</sup> NR2F2 increases glycolysis in fibroblasts to accelerate renal fibrosis.<sup>56</sup> NR2F2 expression has also been linked to congenital heart disease and atherosclerosis. Chronic inflammation in heart failure patients may lead to endothelial dysfunction after loss of NR2F2.<sup>57</sup> Platelet-derived growth factor (PDGF) plays an important role in promoting renal fibrosis. Studies suggest that upregulation of PDGF is associated with chronic inflammation in CKD.<sup>58</sup> Overexpression of PDGFA causes severe myocardial fibrosis and increases ventricular volume many times over, leading to fatal HF.<sup>59</sup> Taken together, we concluded that the high-risk group predicted a higher immune, inflammatory response and fibrotic process. The expression of SLIT2, NR2F2 and PDGFA may accelerate the fibrotic process in the myocardium and kidneys through a chronic inflammatory response, ultimately leading to deterioration in cardiac and renal function.

However, our study has limitations. First, it is based on information digging from sequenced data of failing heart or kidney tissues, from which we constructed a diagnostic model for HF and CKD. Database digging for clinically relevant molecules has some reference value, but whether the diagnostic value of crosstalk genes is applicable to more readily accessible peripheral blood needs further verification. On the other hand, the results are limited to the transcriptional level and have not been further validated in *in vivo* or *in vitro* experiments.

## Conclusion

The five hub genes (PHLDA1, ATP1A1, IFIT2, HLTF, and MPP3) serve as optimal shared diagnostic biomarkers for HF and CKD. ROC curves confirmed the effectiveness of the diagnostic model. The expression levels of crosstalk genes are correlated with left ventricular ejection fraction and glomerular filtration rate. Immune pathways, inflammatory responses, metabolism, and cell apoptosis induced by various mechanisms may have strong associations with both HF and CKD. Consensus clustering analysis divides HF and CKD patients into immune and metabolic subtypes. The high and low-risk ImmuneScore models can predict different molecular subtypes of HF or CKD. However, these findings still require further validation.

## Abbreviations

HF, Heart failure; CKD, Chronic kidney disease; GO, Gene Ontology; GSEA, Gene Set Variation Analysis; GSEA, Gene Set Enrichment Analysis; ROC, Receiver Operating Characteristic; ssGSEA, Single-sample Gene Set Enrichment Analysis; RAAS, renin-angiotensin-aldosterone system; TLR, Toll-like receptor; GEO, Gene Expression Omnibus; RMA, Robust Multiarray Average; PCA, Principal Component Analysis; GEGs, Differentially expressed genes; CDF, Cumulative distribution function; MAD, Median Absolute Deviation; IL-1  $\beta$ , Interleukin-1 $\beta$ ; TNF- $\alpha$ , Tumor necrosis factor- $\alpha$ ; IL-6, Interleukin-6; CRP, C-reactive protein; ROS, Reactive oxygen species; PHLDA1, Pleckstrin homology structural domain A member 1; HLTF, Helicase-like transcription factor; SWI/SNF, Switch/sucrose non-fermenting; GMP, Granulocyte/macrophage progenitor; PKA, Protein kinase A.

## Data Sharing Statement

The data used in this study were obtained from the GEO database, which is openly accessible to the general public, and the datasets can be freely selected and downloaded for use.

## Ethics Approval and Informed Consent

Research involving human participants was reviewed and approved by the Changzheng Hospital Ethics Committee, and written informed consent was signed by the patients themselves or their families. The use of recipient tissue samples was in accordance with the guidelines of the Declaration of Helsinki.

## Acknowledgments

We thank the scholars who provided the original data and the GEO database platform.

## Author Contributions

All authors contributed significantly to the work of the report in terms of article conception, research design, execution, data acquisition, analysis and interpretation. All authors have substantially revised or critically reviewed the article and given final approval of the version to be published. All authors have agreed on the journal to which the article will be submitted and have agreed to take responsibility for all aspects of the work.

## Funding

This work was supported by the Shanghai Shuguang Project [No. 21SG37] and the National Natural Science Foundation of China [No. 81770244].

## Disclosure

The authors declare that they have no competing interests.

## References

1. Isakova T, Nickolas TL, Denburg M, et al. KDOQI US Commentary on the 2017 KDIGO Clinical Practice Guideline Update for the Diagnosis, Evaluation, Prevention, and Treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). *Am J Kidney Dis.* 2017;70(6):737–751. doi:10.1053/j.ajkd.2017.07.019
2. Porter AC, Lash JP, Xie D, et al. Predictors and Outcomes of Health-Related Quality of Life in Adults with CKD. *Clin J Am Soc Nephrol.* 2016;11(7):1154–1162. doi:10.2215/CJN.09990915
3. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, et al. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. *Lancet.* 2013;382(9889):339–352. doi:10.1016/S0140-6736(13)60595-4
4. Sud M, Tangri N, Pintilie M, Levey AS, Naimark D. Risk of end-stage renal disease and death after cardiovascular events in chronic kidney disease. *Circulation.* 2014;130(6):458–465. doi:10.1161/CIRCULATIONAHA.113.007106
5. Scheffold JC, Filippatos G, Hasenfuss G, Anker SD, von Haehling S. Heart failure and kidney dysfunction: epidemiology, mechanisms and management. *Nat Rev Nephrol.* 2016;12(10):610–623. doi:10.1038/nrneph.2016.113
6. Rahman M, Xie D, Feldman HI, et al. Association between chronic kidney disease progression and cardiovascular disease: results from the CRIC Study. *Am J Nephrol.* 2014;40(5):399–407. doi:10.1159/000368915
7. Bang C, Antoniadou C, Antonopoulos AS, et al. Intercellular communication lessons in heart failure. *Eur J Heart Fail.* 2015;17(11):1091–1103. doi:10.1002/ehf.399
8. Obi Y, Kim T, Kovesdy CP, Amin AN, Kalantar-Zadeh K. Current and Potential Therapeutic Strategies for Hemodynamic Cardiorenal Syndrome. *Cardiorenal Med.* 2016;6(2):83–98. doi:10.1159/000441283
9. Rajapakse NW, Nanayakkara S, Kaye DM. Pathogenesis and treatment of the cardiorenal syndrome: implications of L-arginine-nitric oxide pathway impairment. *Pharmacol Ther.* 2015;154:1–12. doi:10.1016/j.pharmthera.2015.05.011
10. Giam B, Kaye DM, Rajapakse NW. Role of Renal Oxidative Stress in the Pathogenesis of the Cardiorenal Syndrome. *Heart Lung Circ.* 2016;25(8):874–880. doi:10.1016/j.hlc.2016.02.022
11. Li X, Hassoun HT, Santora R, Rabb H. Organ crosstalk: the role of the kidney. *Curr Opin Crit Care.* 2009;15(6):481–487. doi:10.1097/MCC.0b013e328332f69e
12. Rosner MH, Ronco C, Okusa MD. The role of inflammation in the cardio-renal syndrome: a focus on cytokines and inflammatory mediators. *Semin Nephrol.* 2012;32(1):70–78. doi:10.1016/j.semnephrol.2011.11.010
13. Vinod P, Krishnappa V, Chauvin AM, Khare A, Raina R. Cardiorenal Syndrome: role of Arginine Vasopressin and Vaptans in Heart Failure. *Cardiol Res.* 2017;8(3):87–95. doi:10.14740/cr553w
14. White LE, Hassoun HT. Inflammatory Mechanisms of Organ Crosstalk during Ischemic Acute Kidney Injury. *Int J Nephrol.* 2012;2012:505197. doi:10.4061/2012/505197
15. Murakami M, Hirano T. The pathological and physiological roles of IL-6 amplifier activation. *Int J Biol Sci.* 2012;8(9):1267–1280. doi:10.7150/ijbs.4828
16. Tornatore L, Thotakura AK, Bennett J, Moretti M, Franzoso G. The nuclear factor kappa B signaling pathway: integrating metabolism with inflammation. *Trends Cell Biol.* 2012;22(11):557–566. doi:10.1016/j.tcb.2012.08.001
17. Kingma JG, Simard D, Rouleau JR, Drolet B, Simard C. The Pathophysiology of Cardiorenal Syndrome: a Review of the Potential Contributions of Inflammation. *J Cardiovasc Dev Dis.* 2017;4(4). doi:10.3390/jcdd4040021

18. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* 2013;41(Database issue):D991–5. doi:10.1093/nar/gks1193
19. Liu Y, Morley M, Brandimarto J, et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics.* 2015;105(2):83–89. doi:10.1016/j.ygeno.2014.12.002
20. Nakagawa S, Nishihara K, Miyata H, et al. Molecular Markers of Tubulointerstitial Fibrosis and Tubular Cell Damage in Patients with Chronic Kidney Disease. *PLoS One.* 2015;10(8):e0136994. doi:10.1371/journal.pone.0136994
21. Greco S, Fasanaro P, Castelvécchio S, et al. MicroRNA dysregulation in diabetic ischemic heart failure patients. *Diabetes.* 2012;61(6):1633–1641. doi:10.2337/db11-0952
22. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 2014;15(2):R29. doi:10.1186/gb-2014-15-2-r29
23. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* 2012;16(5):284–287. doi:10.1089/omi.2011.0118
24. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinf.* 2013;14(1):7. doi:10.1186/1471-2105-14-7
25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–15550. doi:10.1073/pnas.0506580102
26. Pudjihartono N, Fadason T, Kempa-Liehr AW, O'Sullivan JM. A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction. *Front Bioinform.* 2022;2:927312. doi:10.3389/fbinf.2022.927312
27. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 2008;9(1):559. doi:10.1186/1471-2105-9-559
28. Kitai T, Kirsop J, Tang WH. Exploring the Microbiome in Heart Failure. *Curr Heart Fail Rep.* 2016;13(2):103–109. doi:10.1007/s11897-016-0285-9
29. Colombo PC, Ganda A, Lin J, et al. Inflammatory activation: cardiac, renal, and cardio-renal interactions in patients with the cardiorenal syndrome. *Heart Fail Rev.* 2012;17(2):177–190. doi:10.1007/s10741-011-9261-3
30. Pastori S, Virzi GM, Brocca A, et al. Cardiorenal syndrome type 1: a defective regulation of monocyte apoptosis induced by proinflammatory and proapoptotic factors. *Cardiorenal Med.* 2015;5(2):105–115. doi:10.1159/000371898
31. Hayiroğlu M, Altay S. The Role of Artificial Intelligence in Coronary Artery Disease and Atrial Fibrillation. *Balkan Med J.* 2023;40(3):151–152. doi:10.4274/balkanmedj.galenos.2023.06042023
32. Yilmaz A, Hayiroğlu M, Salturk S, et al. Machine Learning Approach on High Risk Treadmill Exercise Test to Predict Obstructive Coronary Artery Disease by using P, QRS, and T waves' Features. *Curr Probl Cardiol.* 2023;48(2):101482. doi:10.1016/j.cpcardiol.2022.101482
33. Frank D, Mendelsohn CL, Ciccone E, Svensson K, Ohlsson R, Tycko B. A novel pleckstrin homology-related gene family defined by Ipl/Tssc3, TDAG51, and Tih1: tissue-specific expression, chromosomal location, and parental imprinting. *Mamm Genome.* 1999;10(12):1150–1159. doi:10.1007/s003359901182
34. Park CG, Lee SY, Kandala G, Lee SY, Choi Y. A novel gene product that couples TCR signaling to Fas(CD95) expression in activation-induced cell death. *Immunity.* 1996;4(6):583–591. doi:10.1016/S1074-7613(00)80484-7
35. Hossain GS, van Thienen JV, Werstuck GH, et al. TDAG51 is induced by homocysteine, promotes detachment-mediated programmed cell death, and contributes to the development of atherosclerosis in hyperhomocysteinemia. *J Biol Chem.* 2003;278(32):30317–30327. doi:10.1074/jbc.M212897200
36. Wang J, Wang F, Zhu J, Song M, An J, Li W. Transcriptome Profiling Reveals PHLDA1 as a Novel Molecular Marker for Ischemic Cardiomyopathy. *J Mol Neurosci.* 2018;65(1):102–109. doi:10.1007/s12031-018-1066-6
37. Liu L, Huang J, Liu Y, et al. Multiomics Analysis of Transcriptome, Epigenome, and Genome Uncovers Putative Mechanisms for Dilated Cardiomyopathy. *Biomed Res Int.* 2021;2021:6653802. doi:10.1155/2021/6653802
38. Guo Y, Jia P, Chen Y, et al. PHLDA1 is a new therapeutic target of oxidative stress and ischemia reperfusion-induced myocardial injury. *Life Sci.* 2020;245:117347. doi:10.1016/j.lfs.2020.117347
39. Carlisle RE, Mohammed-Ali Z, Lu C, et al. TDAG51 induces renal interstitial fibrosis through modulation of TGF- $\beta$  receptor 1 in chronic kidney disease. *Cell Death Dis.* 2021;12(10):921. doi:10.1038/s41419-021-04197-3
40. Han C, Yan P, He T, et al. PHLDA1 promotes microglia-mediated neuroinflammation via regulating K63-linked ubiquitination of TRAF6. *Brain Behav Immun.* 2020;88:640–653. doi:10.1016/j.bbi.2020.04.064
41. Morth JP, Pedersen BP, Buch-Pedersen MJ, et al. A structural overview of the plasma membrane Na<sup>+</sup>, K<sup>+</sup>-ATPase and H<sup>+</sup>-ATPase ion pumps. *Nat Rev Mol Cell Biol.* 2011;12(1):60–70. doi:10.1038/nrm3031
42. Khalaf FK, Tassavvor I, Mohamed A, et al. Epithelial and Endothelial Adhesion of Immune Cells Is Enhanced by Cardiotonic Steroid Signaling Through Na<sup>(+)</sup>/K<sup>(+)</sup>-ATPase- $\alpha$ -1. *J Am Heart Assoc.* 2020;9(3):e013933. doi:10.1161/JAHA.119.013933
43. Farbehi N, Patrick R, Dorison A, et al. Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *Elife.* 2019;8.
44. Stawowczyk M, Van Scoy S, Kumar KP, Reich NC. The interferon stimulated gene 54 promotes apoptosis. *J Biol Chem.* 2011;286(9):7257–7266. doi:10.1074/jbc.M110.207068
45. Liu L, Liu H, Zhou Y, et al. HLTF suppresses the migration and invasion of colorectal cancer cells via TGF- $\beta$ /SMAD signaling in vitro. *Int J Oncol.* 2018;53(6):2780–2788. doi:10.3892/ijo.2018.4591
46. Piao S, Ojha R, Rebecca VW, et al. ALDH1A1 and HLTF modulate the activity of lysosomal autophagy inhibitors in cancer cells. *Autophagy.* 2017;13(12):2056–2071. doi:10.1080/15548627.2017.1377377
47. Debauxe G, Capouillez A, Belayew A, Saussez S. The helicase-like transcription factor and its implication in cancer progression. *Cell Mol Life Sci.* 2008;65(4):591–604. doi:10.1007/s00018-007-7392-4
48. Kang YA, Paik H, Zhang SY, et al. Secretory MPP3 reinforce myeloid differentiation trajectory and amplify myeloid cell production. *J Exp Med.* 2023;220(8). doi:10.1084/jem.20230088
49. Ngkelo A, Richart A, Kirk JA, et al. Mast cells regulate myofilament calcium sensitization and heart function after myocardial infarction. *J Exp Med.* 2016;213(7):1353–1374. doi:10.1084/jem.20160081

50. Laroumanie F, Douin-Echinard V, Pozzo J, et al. CD4+ T cells promote the transition from hypertrophy to heart failure during chronic pressure overload. *Circulation*. 2014;129(21):2111–2124. doi:10.1161/CIRCULATIONAHA.113.007101
51. Lee H, Fessler MB, Qu P, Heymann J, Kopp JB. Macrophage polarization in innate immune responses contributing to pathogenesis of chronic kidney disease. *BMC Nephrol*. 2020;21(1):270. doi:10.1186/s12882-020-01921-7
52. Han Y, Ma FY, Tesch GH, Manthey CL, Nikolic-Paterson DJ. Role of macrophages in the fibrotic phase of rat crescentic glomerulonephritis. *Am J Physiol Renal Physiol*. 2013;304(8):F1043–53. doi:10.1152/ajprenal.00389.2012
53. Anders HJ, Ryu M. Renal microenvironments and macrophage phenotypes determine progression or resolution of renal inflammation and fibrosis. *Kidney Int*. 2011;80(9):915–925. doi:10.1038/ki.2011.217
54. Zhou X, Yao Q, Sun X, et al. Slit2 ameliorates renal inflammation and fibrosis after hypoxia-and lipopolysaccharide-induced epithelial cells injury in vitro. *Exp Cell Res*. 2017;352(1):123–129. doi:10.1016/j.yexcr.2017.02.001
55. Liu Y, Yin Z, Xu X, et al. Crosstalk between the activated Slit2-Robo1 pathway and TGF- $\beta$ 1 signalling promotes cardiac fibrosis. *ESC Heart Fail*. 2021;8(1):447–460. doi:10.1002/ehf2.13095
56. Li L, Galichon P, Xiao X, et al. Orphan nuclear receptor COUP-TFII enhances myofibroblast glycolysis leading to kidney fibrosis. *EMBO Rep*. 2021;22(6):e51169. doi:10.15252/embr.202051169
57. Dougherty EJ, Chen LY, Awad KS, et al. Inflammation and DKK1-induced AKT activation contribute to endothelial dysfunction following NR2F2 loss. *Am J Physiol Lung Cell Mol Physiol*. 2023;324(6):L783–L798. doi:10.1152/ajplung.00171.2022
58. Floege J, Eitner F, Alpers CE. A new look at platelet-derived growth factor in renal disease. *J Am Soc Nephrol*. 2008;19(1):12–23. doi:10.1681/ASN.2007050532
59. Gallini R, Lindblom P, Bondjers C, Betsholtz C, Andrae J. PDGF-A and PDGF-B induces cardiac fibrosis in transgenic mice. *Exp Cell Res*. 2016;349(2):282–290. doi:10.1016/j.yexcr.2016.10.022

## Journal of Inflammation Research

Dovepress

### Publish your work in this journal

The Journal of Inflammation Research is an international, peer-reviewed open-access journal that welcomes laboratory and clinical findings on the molecular basis, cell biology and pharmacology of inflammation including original research, reviews, symposium reports, hypothesis formation and commentaries on: acute/chronic inflammation; mediators of inflammation; cellular processes; molecular mechanisms; pharmacology and novel anti-inflammatory drugs; clinical conditions involving inflammation. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>