Pharmacogenomics and Personalized Medicine Dovepress

Open Access Full Text Article ORIGINAL RESEARCH

# Immune Microenvironment Alterations and Identification of Key Diagnostic Biomarkers in Chronic Kidney Disease Using Integrated Bioinformatics and Machine Learning

Jinbao Shi, Aliang Xu, Liuying Huang, Shaojie Liu, Binxuan Wu, Zuhong Zhang

Department of Nephrology, Ningde Hospital of Traditional Chinese Medicine, Ningde, Fujian, People's Republic of China

Correspondence: Jinbao Shi, Department of Nephrology, Ningde Hospital of Traditional Chinese Medicine, No. 16 Donghu Road, Ningde, Fujian, People's Republic of China, Email shijinbao00721@aliyun.com

Pharmacogenomics and Personalized Medicine downloaded from https://www.dovepress.com/
For personal use only.

**Background:** Chronic kidney disease (CKD) involves complex immune dysregulation and altered gene expression profiles. This study investigates immune cell infiltration, differential gene expression, and pathway enrichment in CKD patients to identify key diagnostic biomarkers through machine learning methods.

**Methods:** We assessed immune cell infiltration and immune microenvironment scores using the xCell algorithm. Differentially expressed genes (DEGs) were identified using the limma package. Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA) were performed to evaluate pathway enrichment. Machine learning techniques (LASSO and Random Forest) pinpointed diagnostic genes. A nomogram model was constructed and validated for diagnostic prediction. Spearman correlation explored associations between key genes and immune cell recruitment.

**Results:** The CKD group exhibited significantly altered immune cell infiltration and increased immune microenvironment scores compared to the normal group. We identified 2335 DEGs, including 124 differentially expressed immune-related genes. GSEA highlighted significant enrichment of inflammatory and immune pathways in the high immune score (HIS) subgroup, while GSVA indicated upregulation of immune responses and metabolic processes in HIS. Machine learning identified four key diagnostic genes: RGS1, IL4I1, NR4A3, and SOCS3. Validation in an independent dataset (GSE96804) and clinical samples confirmed their diagnostic potential. The nomogram model integrating these genes demonstrated high predictive accuracy. Spearman correlation revealed positive associations between the key genes and various immune cells, indicating their roles in immune modulation and CKD pathogenesis.

**Conclusion:** This study provides a comprehensive analysis of immune alterations and gene expression profiles in CKD. The identified diagnostic genes and the constructed nomogram model offer potent tools for CKD diagnosis. The immunomodulatory roles of RGS1, IL4I1, NR4A3, and SOCS3 warrant further investigation as potential therapeutic targets in CKD.

**Keywords:** chronic kidney disease, diagnostic biomarkers, immune microenvironment, GSEA, GSVA, machine learning

## Introduction

Chronic kidney disease (CKD) is a progressive condition that poses a significant public health challenge worldwide, affecting millions of people and leading to high morbidity and mortality rates.[1] CKD is characterized by the gradual loss of kidney function over time, which can eventually lead to end-stage renal disease, necessitating dialysis or kidney transplantation.[2] The pathogenesis of CKD involves a complex interplay of genetic, environmental, and immunological factors.[3] Recent studies have highlighted the crucial role of immune dysregulation and altered gene expression profiles in the progression of CKD.[4] However, the specific mechanisms underlying these changes and their impact on disease progression remain poorly understood.

Numerous studies have explored the role of immune cells in CKD, demonstrating that immune cell infiltration and activation are key features of CKD pathology.[5,6] For instance, T cells, and macrophages have been implicated in

Received: 22 July 2024
Accepted: 4 November 2024
Published: 19 November 2024



© 2024 Shi et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (http://creativecommons.org/licenses/by-nc/3.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

promoting inflammation and fibrosis in the kidney, contributing to disease progression.[7] Despite these advances, there remains a gap in understanding the comprehensive landscape of immune cell infiltration and its relationship with gene expression changes in CKD. Furthermore, the integration of bioinformatics and machine learning methodologies remains underutilized in identifying reliable diagnostic biomarkers in CKD. Prior studies have frequently suffered from inadequate utilization of bioinformatics and machine learning methodologies for the identification of reliable diagnostic biomarkers. Additionally, recent advancements in unsupervised deep learning approaches have shown promise in uncovering complex patterns in biomedical data.[8–10] Machine learning techniques, such as the least absolute shrinkage and selection operator (LASSO) and Random Forest, have demonstrated significant potential in identifying key diagnostic genes and constructing predictive models in various diseases.[11,12] Given the complex nature of CKD and the involvement of immune dysregulation, there is a pressing need for comprehensive studies that integrate bioinformatics and machine learning approaches to unravel the intricate relationships between immune cell infiltration, gene expression alterations, and disease progression. Identifying key diagnostic biomarkers can not only enhance our understanding of CKD pathogenesis but also pave the way for the development of novel diagnostic and therapeutic strategies. This study aims to address these gaps by leveraging GEO datasets and advanced computational methods to provide a detailed analysis of the immune microenvironment in CKD.

The primary objective of this study is to investigate the alterations in the immune microenvironment and identify key diagnostic biomarkers in CKD using integrated bioinformatics and machine learning approaches. In this study, we utilized the xCell algorithm to evaluate immune cell infiltration and calculate immune microenvironment scores for CKD patients. Differential gene expression analysis was conducted using the limma package to identify DEGs between CKD patients and healthy controls. GSEA and GSVA were employed to determine the enrichment of biological pathways. Machine learning techniques, including LASSO and Random Forest, were applied to identify key diagnostic genes. A nomogram model was then constructed and validated for its diagnostic accuracy. Finally, Spearman correlation analysis was performed to investigate the relationships between the identified genes and various immune cells.

This study provides a comprehensive analysis of immune alterations and gene expression profiles in CKD, identifying critical diagnostic biomarkers and constructing a predictive model. The findings have significant implications for improving CKD diagnosis and understanding the immunomodulatory mechanisms underlying the disease. Moreover, the identified genes, RGS1, IL4I1, NR4A3, and SOCS3, present promising targets for future therapeutic interventions, potentially leading to better management and treatment of CKD.

# Materials and Methods

## Data Acquisition and Analysis

We retrieved two publicly available transcriptomic datasets from the Gene Expression Omnibus (GEO) database, namely GSE66494 and GSE96804. GSE66494 comprised 53 CKD samples and 8 normal samples obtained from kidney biopsy specimens, while GSE96804 consisted of 20 control samples and 41 samples of diabetic nephropathy obtained from glomerular specimens. The datasets were accessed via the GEO platform (https://www.ncbi.nlm.nih.gov/geo/). Before data analysis, we extracted and normalized the raw matrix files using the affy package in R. Probe IDs were converted to gene symbols based on the platform's annotation file for each dataset. If multiple probes corresponded to a single gene symbol, the average expression value of these probes was considered as the gene expression level. This approach ensures consistency and reliability in the processed data for downstream analyses.

## Immune Cell Infiltration Analysis

We employed the xCell algorithm to assess the proportions of immune cell subsets between the normal and CKD groups.[13] The results were visualized using bar plots generated with the ggplot2 package. To compare the proportions of immune cell subsets between the normal and CKD groups, we used the Student's *t*-test. The association between diagnostic genes and immune cells was assessed utilizing the ggplot2 package.

## Analysis of Differential Gene Expression

We employed the xCell algorithm to calculate immune microenvironment scores for each CKD sample. The GSE66494 dataset, consisting of 53 CKD samples, was stratified into low immune microenvironment score (LIS) and high immune microenvironment score (HIS) subgroups based on the median immune microenvironment score.[14] To identify genes associated with the immune microenvironment scores between the LIS and HIS subgroups, we used the limma package in R. The selection criteria for significant genes were an adjusted p-value (p.adj) < 0.05 and an absolute log2 fold change > 1.[15] Similarly, we identified differentially expressed genes (DEGs) between normal and CKD groups using the limma package with the same selection criteria. Volcano plots were generated using the ggplot2 package to visualize the significant DEGs and IGs.

## Gene Set Enrichment Analysis (GSEA)

We conducted GSEA using the "clusterProfiler" package to explore the potential mechanisms of IGs in CKD.[16] The gene set "c2.cp.v7.2.symbols.gm" was utilized for this analysis. Gene sets with a p-value adjusted (p.adj) < 0.05 were considered to be significantly enriched. Results from the GSEA were visualized using the ggplot2 package.

## Gene Set Variation Analysis (GSVA)

In order to improve the detection of subtle changes in pathways within each sample, we employed GSVA.[17] GSVA analysis was conducted using the gene set "c2.cp.v7.2.symbols.gm" as the reference. The calculations for GSVA and subsequent analyses were performed using the tidyverse package.

## Identification of Core Differentially Immune-Related Genes (DIGs)

We employed Weighted Gene Co-expression Network Analysis (WGCNA) to identify immune microenvironment-associated modules. The soft-thresholding power was determined by calculating the maximum $R^2$ value. Subsequently, the adjacency matrix was converted into a topological overlap matrix. Utilizing hierarchical clustering analysis, a dendrogram was constructed, segmenting genes into distinct modules. The correlation coefficients and p-values between the identified modules and clinical traits were determined using the cor function. Within the most relevant modules, hub genes were determined based on module membership (MM) > 0.5 and gene significance (GS) > 0.5.[18] We identified the intersecting gene subset among IGs, DEGs, and module hub genes. These overlapping genes were identified as DIGs.

## Gene Signatures are Identified Through the Utilization of Both Random Forest (RF) and LASSO Methodologies

The glmnet package was utilized for performing LASSO regression. During this process, 10-fold cross-validation was employed to determine the optimal penalization parameters (lambda). This approach allows for the selection of a model that balances complexity and predictive performance by minimizing the mean squared error. LASSO, or Least Absolute Shrinkage and Selection Operator, enhances model interpretability by applying an L1 penalty, which shrinks the coefficients of less important variables to zero, thereby performing variable selection.[19] The randomForest package was employed to build a random forest model. Initially, this model computed the mean error rate across all DEIGs. Subsequently, a separate random forest model was developed to assess the significance of each gene, based on the decrease in prediction accuracy. Genes that exhibited an importance score exceeding the threshold of 1 were recognized as central hub genes, and these genes are slated for inclusion in the refinement of future models.[20] The identification of signature genes in the diagnosis of CKD involved determining the intersection genes between LASSO and RF techniques.

## Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR) Analysis

Renal tissue samples were gathered from 16 individuals in total, comprising 8 healthy subjects and 8 individuals diagnosed with CKD. All participants provided their written informed consent before the samples were collected,

ensuring adherence to ethical guidelines. Additionally, the study's procedures received the endorsement of the Ethics Committee at the Ningde Hospital of Traditional Chinese Medicine.

Kidney tissue specimens underwent RNA isolation through the application of TRIzol reagent (ThermoFisher Scientific, USA). For the purpose of conducting qRT-PCR assessments, the SYBR qPCR Master Mix (Bio-Rad) was used. The evaluation of the primary signature genes' expression was carried out using the Roche LC480 Real-Time PCR System (Roche). For the normalization of mRNA levels, β-actin served as the reference gene. The comparative levels of mRNA expression were quantified through the application of the $2^{-\Delta\Delta Ct}$ method.

## Development and Evaluation of a Predictive Nomogram

To develop the nomogram, we employed the rms R package, which incorporated the signature genes.[21] The diagnostic efficacy of the signature genes and the nomogram in detecting CKD was determined by constructing a Receiver Operating Characteristic (ROC) curve. Additionally, the precision of the model was gauged through Decision Curve Analysis (DCA) and the use of calibration curves.

# Results

## Enhanced Immune Microenvironment and Altered Gene Expression Profiles in Chronic Kidney Disease (CKD)

Using the xCell algorithm, we assessed the levels of immune cell infiltration between the normal group and patients with CKD. In Figure 1A, a significant increase in the infiltration levels of aDC ($p < 0.05$), basophils ($p < 0.01$), cDC ($p < 0.001$), DC ($p < 0.01$), iDC ($p < 0.001$), NKT cells ($p < 0.001$), pDC ($p < 0.01$), pro B cells ($p < 0.05$), Th2 cells ($p < 0.001$), Tregs ($p < 0.001$), fibroblasts ($p < 0.001$), and endothelial ($p < 0.01$) was observed in the CKD group compared to the normal group. Conversely, a significant reduction was noted for CD4+ Tcm ($p < 0.001$), CD8+ naive T-cells ($p < 0.05$), macrophages M2 ($p < 0.001$), and Th1 cells ($p < 0.001$). No significant differences were detected for other immune cell types. Figure 1B illustrates a comparative evaluation of the immune microenvironment scores between the normal and CKD groups using the xCell algorithm. The CKD group demonstrated a markedly higher immune microenvironment score in comparison to the normal group ($p < 0.01$), indicating a pronounced alteration in the overall immune landscape in CKD. The DEGs between the normal and CKD groups were identified using the limma package. A total of 2335 DEGs were identified, applying a threshold of an adjusted p-value less than 0.05 and an absolute log2 fold change exceeding 1, as illustrated in Figure 1C. According to the median immune microenvironment score, 53 CKD patients were stratified into low (LIS) and high (HIS) immune microenvironment score subgroups. A comparative analysis of the identified subgroups revealed 620 immune-related genes (IGs) that met the criteria of an adjusted p-value < 0.05 and an absolute log2 fold change exceeding 1, as illustrated in Figure 1D. Through venny online analysis, 124 differentially expressed immune microenvironment-related genes were recognized. The protein-protein interaction (PPI) network of these DIGs is illustrated in Figure S1. This network illustrates the interconnected nature of these DIGs, suggesting key regulatory genes that may play significant roles in the immune response and progression of CKD. The detailed evaluation of immune cell infiltration and immune microenvironment scores highlights significant alterations in the immune profiles of CKD patients.

## GSEA Between LIS and HIS Subgroups

GSEA was performed to identify significant pathways differentially enriched between the LIS and HIS in CKD patients. As shown in Figure 2, several pathways related to inflammatory responses and immune system function were significantly enriched in the HIS subgroup (adjusted p-value < 0.05). These pathways include the Inflammatory Response Pathway, the Systemic Lupus Erythematosus Pathway, the Interleukin 10 Signaling pathway, the Extrafollicular B Cell Activation by Sarscov2, Primary Immunodeficiency, the PID Cd8 Tcr Pathway, and various cytokine and receptor interaction networks such as Interleukin 4 and Interleukin 13 Signaling, and Cytokine-Cytokine Receptor Interaction. These findings suggest a pronounced alteration in immune signaling and response mechanisms in the HIS subgroup relative to the LIS subgroup. Overall, the GSEA results emphasize the heightened immune and inflammatory activity in

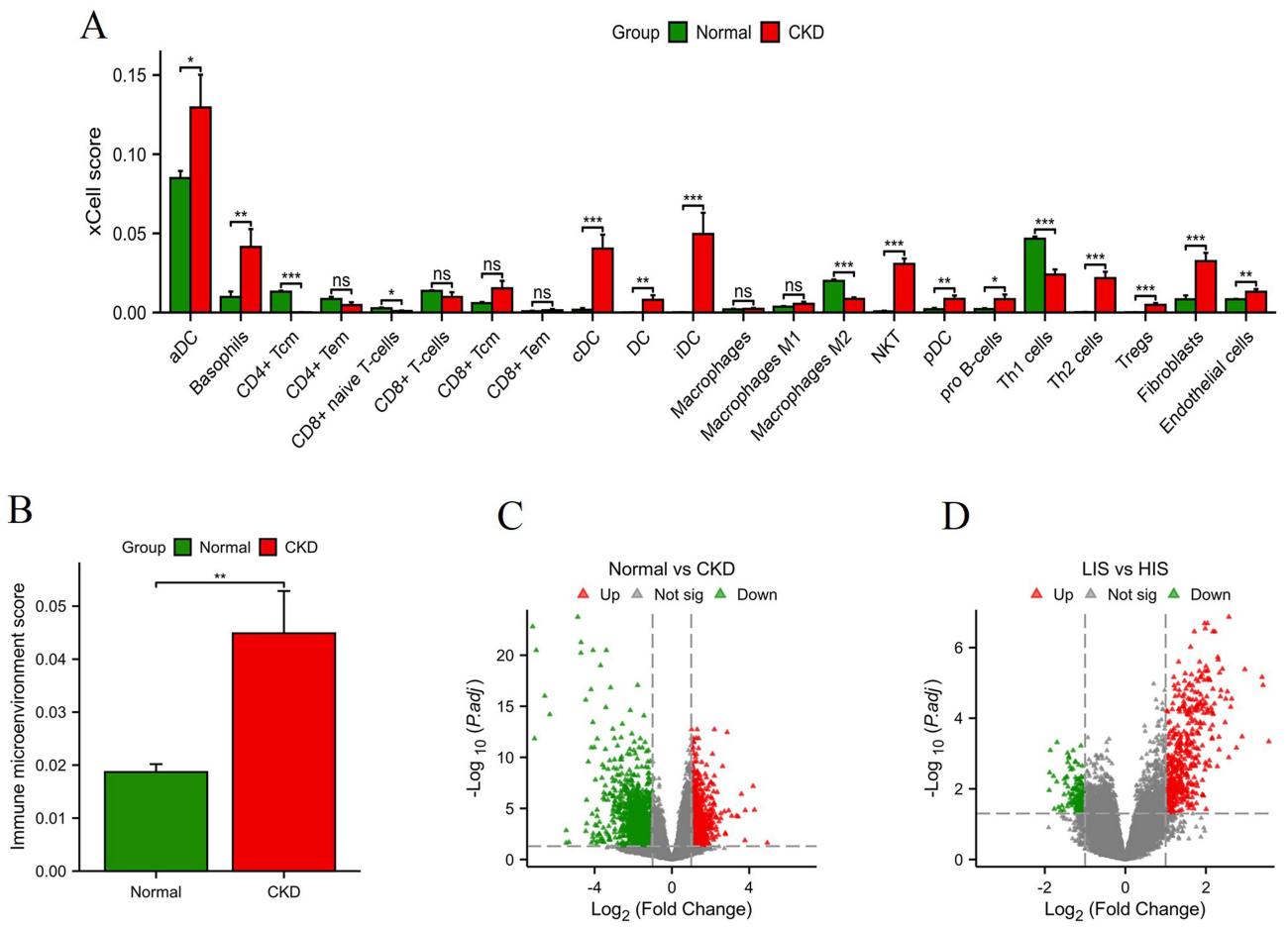


**Figure 1** Integrated immunological and transcriptomic analysis reveals distinct immune cell infiltration and differential gene expression in CKD. (**A**) Comparison of immune cell infiltration levels between normal (green) and CKD (red) groups using the xCell algorithm. Statistical significance: *p < 0.05, **p < 0.01, ***p < 0.001, ns: not significant. (**B**) Measurement of immune microenvironment scores between normal (green) and CKD (red) groups using the xCell algorithm. Statistical significance: **p < 0.01. (**C**) The volcano plot of DEGs identified using the limma package shows the comparison between normal and CKD groups. (**D**) Volcano plot comparing LIS and HIS within CKD patients.

patients with high immune microenvironment scores, reflecting the complex interplay between immune cells and signaling pathways in exacerbating CKD pathology.

## GSVA Enrichment Analysis Between LIS and HIS Subgroups

GSVA was conducted to compare the enrichment of biological processes and pathways between the LIS and HIS in CKD patients. Figure 3A presents the differential enrichment of GO terms between LIS and HIS subgroups. Several GO terms were significantly upregulated in the HIS subgroup, including processes related to immune responses such as negative regulation of T-cell activation, positive regulation of regulatory T-cell differentiation, and regulation of B-cell differentiation. Metabolic processes, including polyamine oxidase activity, tryptophan catabolic process, and amino acid catabolic process, were also upregulated in HIS. Conversely, a few GO terms were downregulated in the HIS subgroup, which included functions like smooth endoplasmic reticulum, intracellular receptor signaling pathway, etc. Figure 3B illustrates the differential enrichment of KEGG pathways between the LIS and HIS subgroups. Pathways significantly upregulated in the HIS subgroup included various metabolic pathways (eg, phenylalanine, tyrosine, and tryptophan biosynthesis, phenylalanine metabolism, and tyrosine metabolism) as well as cancer-related pathways (eg, Wnt signaling pathway, JAK-STAT signaling pathway, and pathways related to various cancers such as acute myeloid leukemia and small cell lung cancer). In contrast, several immune signaling pathways were downregulated in the HIS subgroup, including the B cell receptor signaling pathway, the T cell receptor signaling pathway, and Toll-like receptor signaling

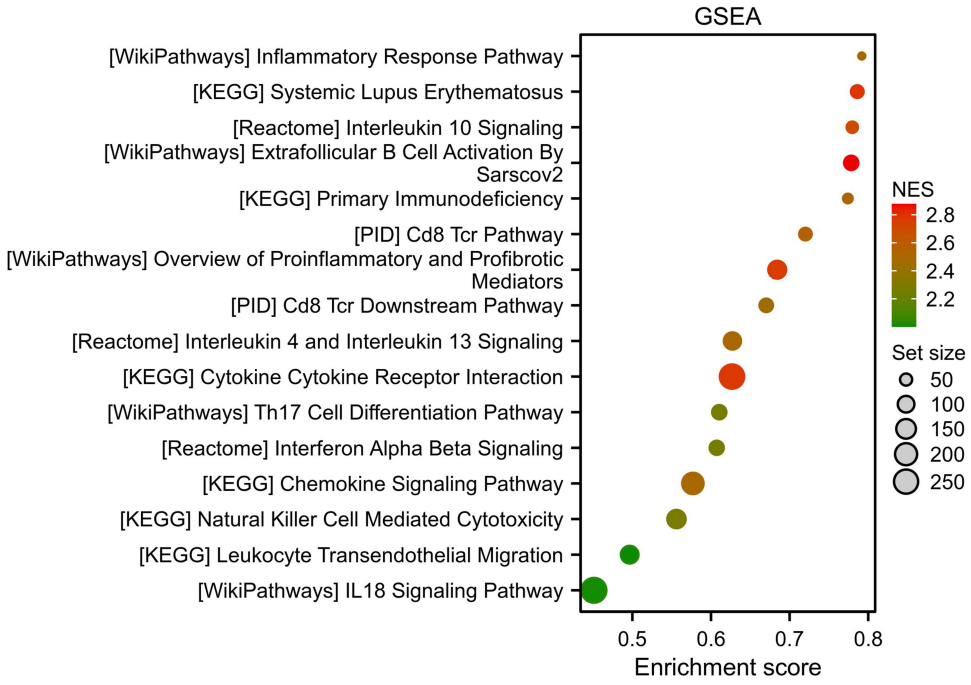


**Figure 2** GSEA identifies key immune and inflammatory pathways distinguishing LIS and HIS subgroups in CKD patients.

pathway. The GSVA results highlight robust immune and metabolic reprogramming in the HIS subgroup, with pronounced upregulation in immune and metabolic activities along with significant engagement in oncogenic pathways. These findings provide critical insights into the molecular distinctions between differing immune microenvironments within CKD patients.

## Application of Machine Learning Methods for the Identification of Characteristic Genes

The heatmap displays the correlation coefficients between module eigengenes and immune scores (Figure S2). Modules are color-coded, with each row representing a different module and its corresponding correlation value. The yellow module exhibited the highest correlation with the immune microenvironment score as revealed by WGCNA analysis. Consequently, this led to the identification of 327 critical genes within this module, adhering to the criteria of MM > 0.5 and GS > 0.5. Then, we obtained 22 core DIGs through the Veen tool (Figure 4A). To pinpoint distinctive genes in patients with CKD, machine learning methods were deployed. The application of the LASSO technique revealed a set of six distinct genes, as depicted in Figure 4B. Concurrently, the employment of the RF algorithm discerned another set of six distinct genes, each exhibiting a relative significance exceeding the threshold of 1, as illustrated in Figure 4C. The intersecting set illustrated in the Venn diagram (Figure 4D) depicts the commonality among the six key genes pinpointed by the RF method and the six probable candidate genes determined through the LASSO technique. Within this shared subset, four genes (RGS1, IL4I1, NR4A3, and SOCS3) were singled out for ultimate validation.

In the GSE66494 dataset, the genes IL4I1 and SOCS3 were observed to have notably elevated expression in the group with CKD ($p < 0.001$). On the other hand, the expression of RGS1 and NR4A3 was reduced in the CKD group ($p < 0.001$) (Figure 5A). Analysis of the Receiver Operating Characteristic (ROC) curves for these genes revealed that with an Area Under the Curve (AUC) exceeding 0.9, they possess substantial potential as diagnostic biomarkers (Figure 5B). Furthermore, we used an independent GSE96804 dataset (Figure 5C and D) and clinical samples (Figure 5E and F) to validate the analysis results, which were consistent with our findings.

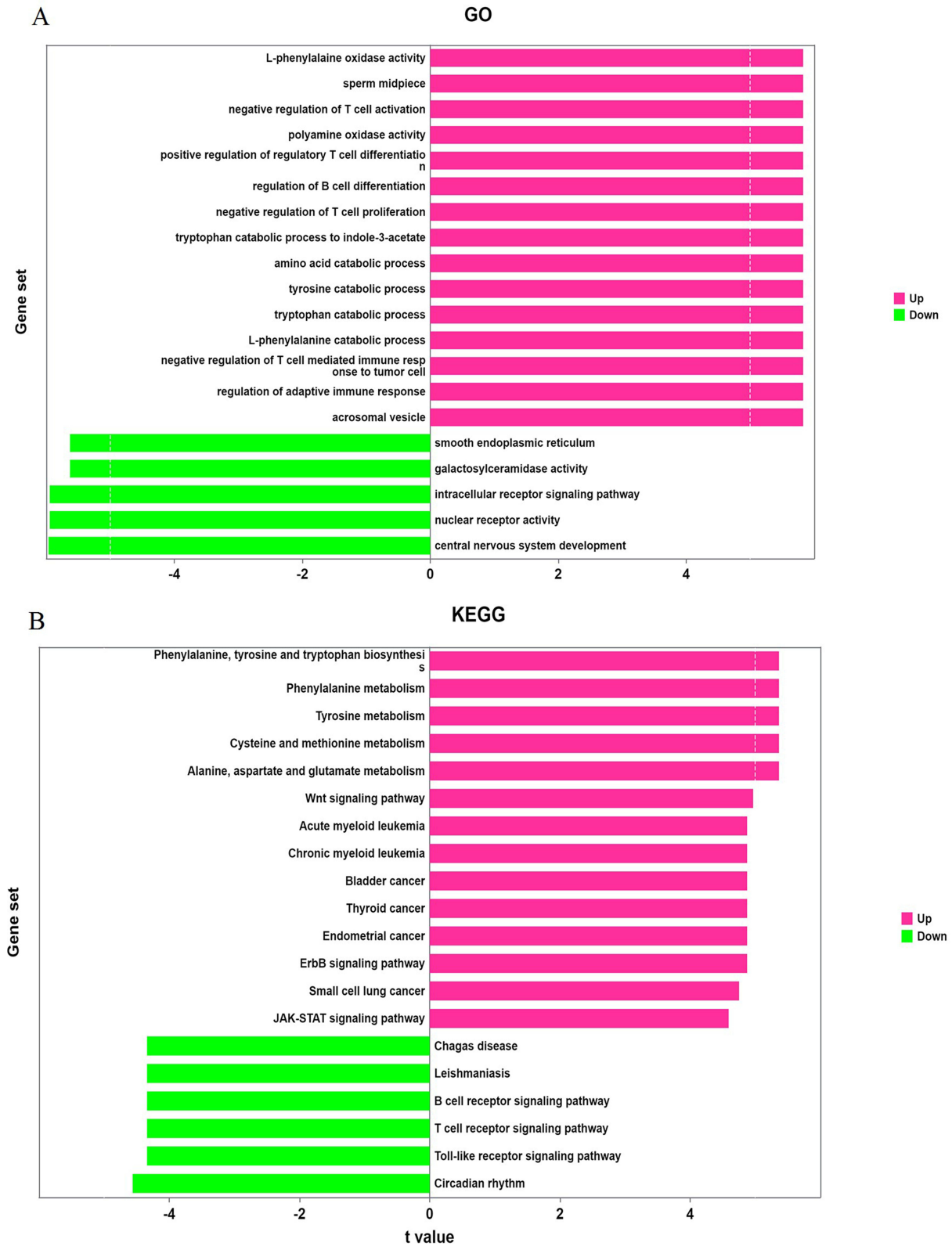


**Figure 3** GSVA reveals differential enrichment of biological processes and pathways in LIS and HIS subgroups. Enrichment analysis of GO (**A**) and KEGG (**B**) pathways using GSVA between LIS and HIS subgroups.

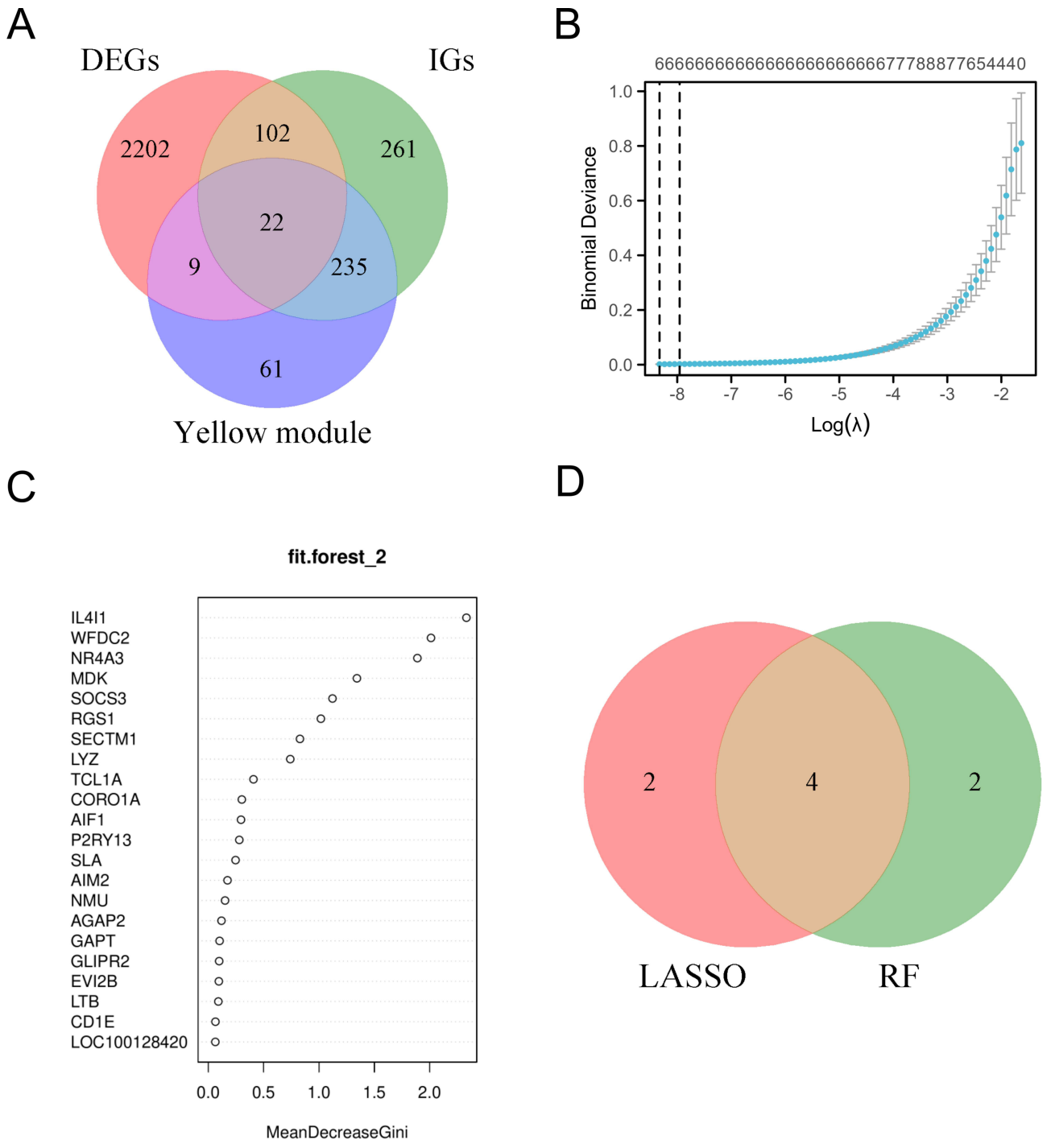


**Figure 4** Identification of signature genes using machine learning methods. (**A**) A Venn diagram depicting the overlap of genes between DEGs, IGs, and those within the yellow module. (**B**) Choosing the right lambda parameter for a LASSO regression model. (**C**) In the random forest model, IL4I1, WFDC2, NR4A3, MDK, SOCS3, and RGS1 genes exhibit a relative importance that exceeds 1. (**D**) The identification of four potential diagnostic genes was illustrated using the aforementioned two machine learning techniques, as depicted in the Venn diagram.

## Assessment of a Nomogram Model

To improve the precision of forecasting the progression of CKD, we have constructed a nomogram incorporating four key genes (Figure 6A). Evaluation of the model using the ROC curve analysis revealed a significant AUC value of 0.96, indicating high predictive performance (Figure 6B). The results from the calibration curve further substantiated the

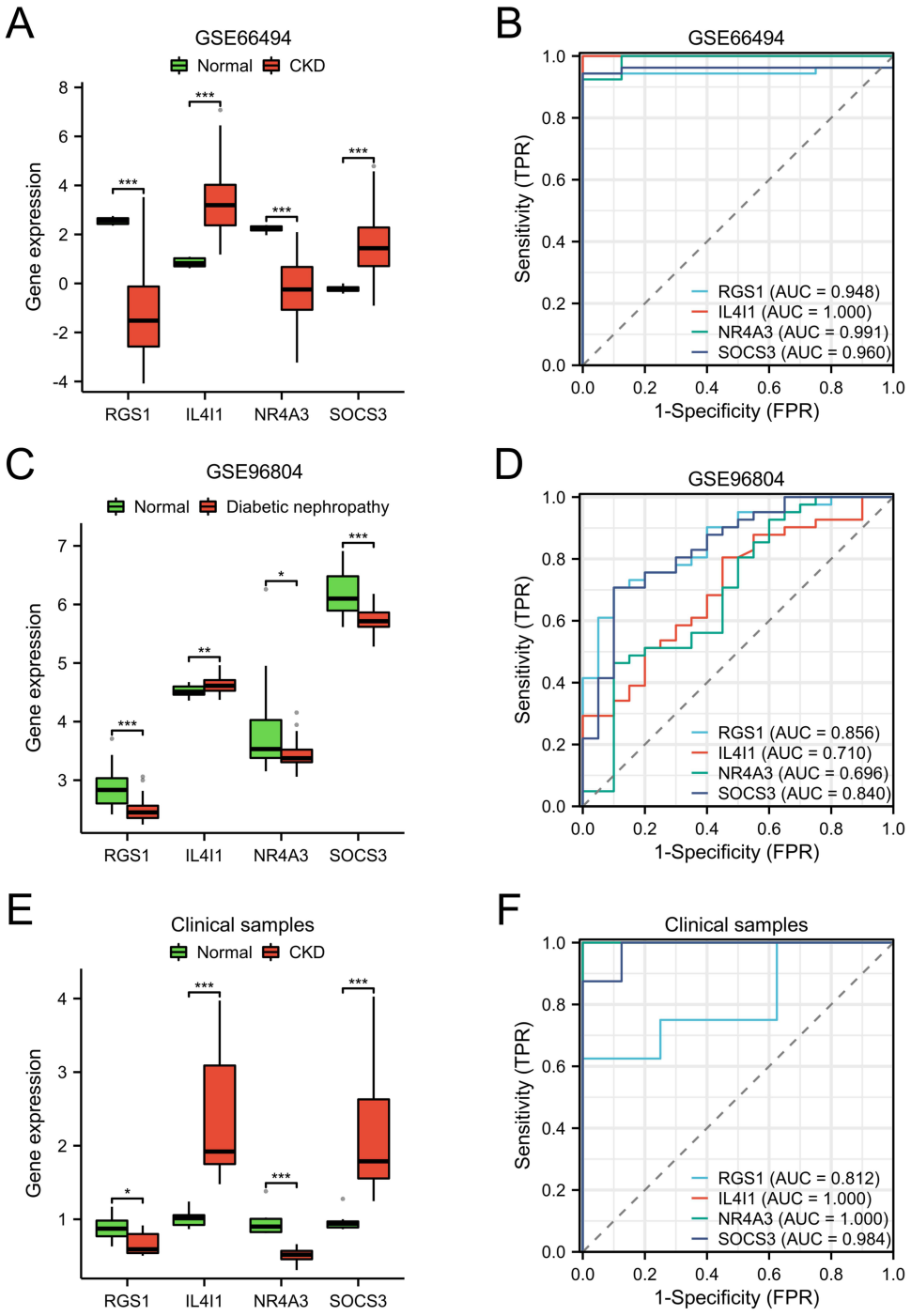


**Figure 5** Verify the clinical diagnostic potential of signature genes. Expression levels (**A**) and diagnostic effectiveness (**B**) of signature genes in the GSE66494 dataset. Expression levels (**C**) and diagnostic effectiveness (**D**) of signature genes in the GSE96804 dataset. Expression levels (**E**) and diagnostic effectiveness (**F**) of signature genes in the clinical samples.*p < 0.05, **p < 0.01 and ***p < 0.001.

nomogram model's exceptional precision in predicting the prognosis for patients with CKD (Figure 6C). In addition, the analysis of the decision curve suggested that employing the nomogram model could be advantageous for individuals suffering from CKD (Figure 6D).

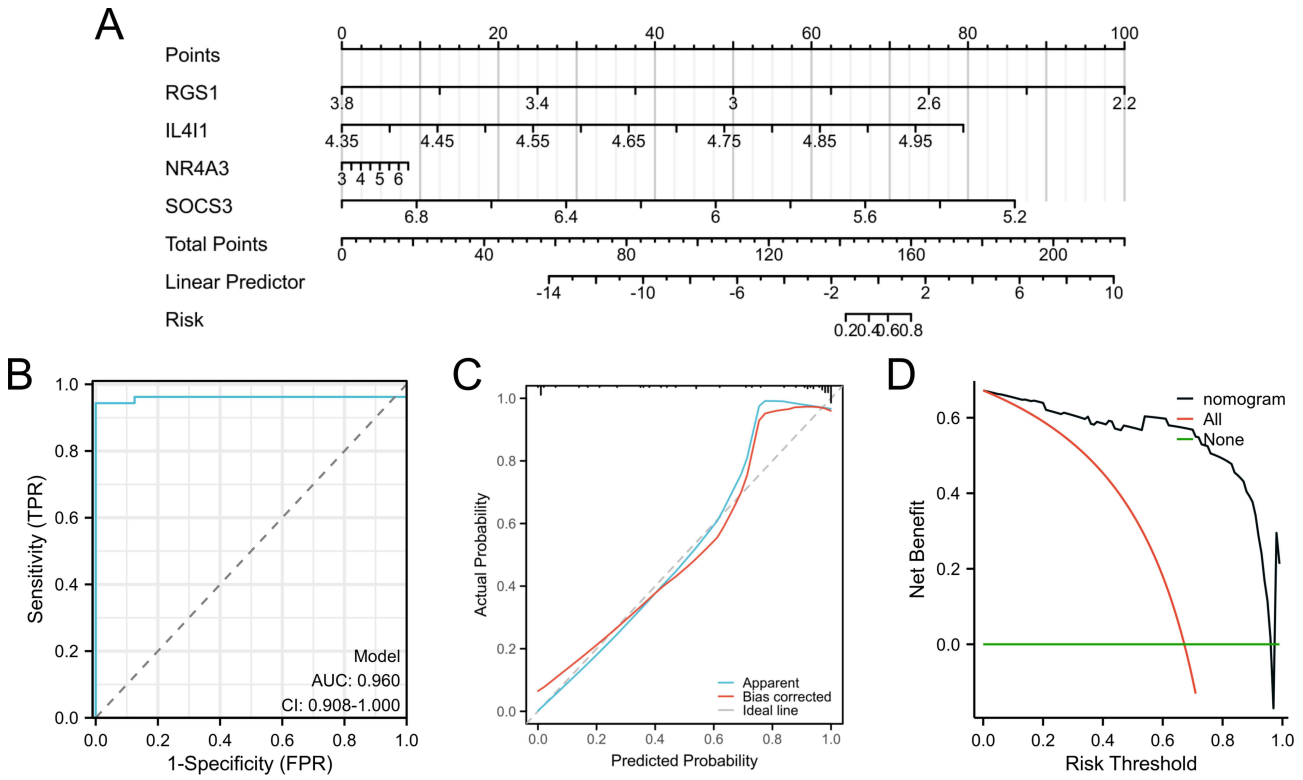


**Figure 6** Assessment of a nomogram model. (**A**) A nomogram model was utilized to evaluate the likelihood of CKD development. (**B**) The ROC curve of the nomogram model was used to diagnose CKD. The calibration curve (**C**) and the DCA (**D**) are used to assess the predictive accuracy of the nomogram model.

## The Association Between Four Diagnostic Biomarkers and the Recruitment of Immune Cells

Concurrently, the analysis using the Spearman correlation method revealed a positive association between the RGS1, IL4I1, NR4A3, and SOCS3 and various immune cells, including aDC, DC, CD8+ Tcm, CD8+ Tem, and Tgd cells (Figure 7). This correlation supported the significant role of these genes in the differentiation of these immune cell subtypes. Furthermore, the results of the GSEA indicated that the genes (RGS1, IL4I1, and SOCS3) may play a role in the development of CKD by modulating immune-related pathways, including T cell receptor signaling pathway, leukocyte transendothelial migration, and natural killer cell-mediated cytotoxicity (Figure S3). This suggests a broader immunomodulatory role for these genes beyond cell subtype differentiation, highlighting their multifaceted involvement in CKD pathogenesis.

# Discussion

CKD is characterized by complex immune dysregulation and altered gene expression profiles. The immune microenvironment plays a crucial role in the pathogenesis and progression of CKD, with immune cell infiltration and activation being key contributors to renal inflammation and fibrosis.[22,23] Understanding the alterations in the immune microenvironment and identifying related genes are essential for developing new diagnostic and therapeutic strategies for CKD. In this study, we employed a comprehensive bioinformatics approach to assess immune cell infiltration, differential gene expression, and pathway enrichment in CKD patients. Using the xCell algorithm, we observed significant changes in the immune cell composition and increased immune microenvironment scores in CKD samples compared to normal controls. These changes are indicative of heightened immune activity in CKD, which may drive chronic inflammation and subsequent renal damage. These findings align with previous studies that have reported enhanced immune cell infiltration in CKD, contributing to chronic inflammation and renal damage.[24,25] The GSEA and GSVA analyses revealed significant enrichment of inflammatory and immune pathways, particularly in the HIS subgroup. This enrichment highlights the

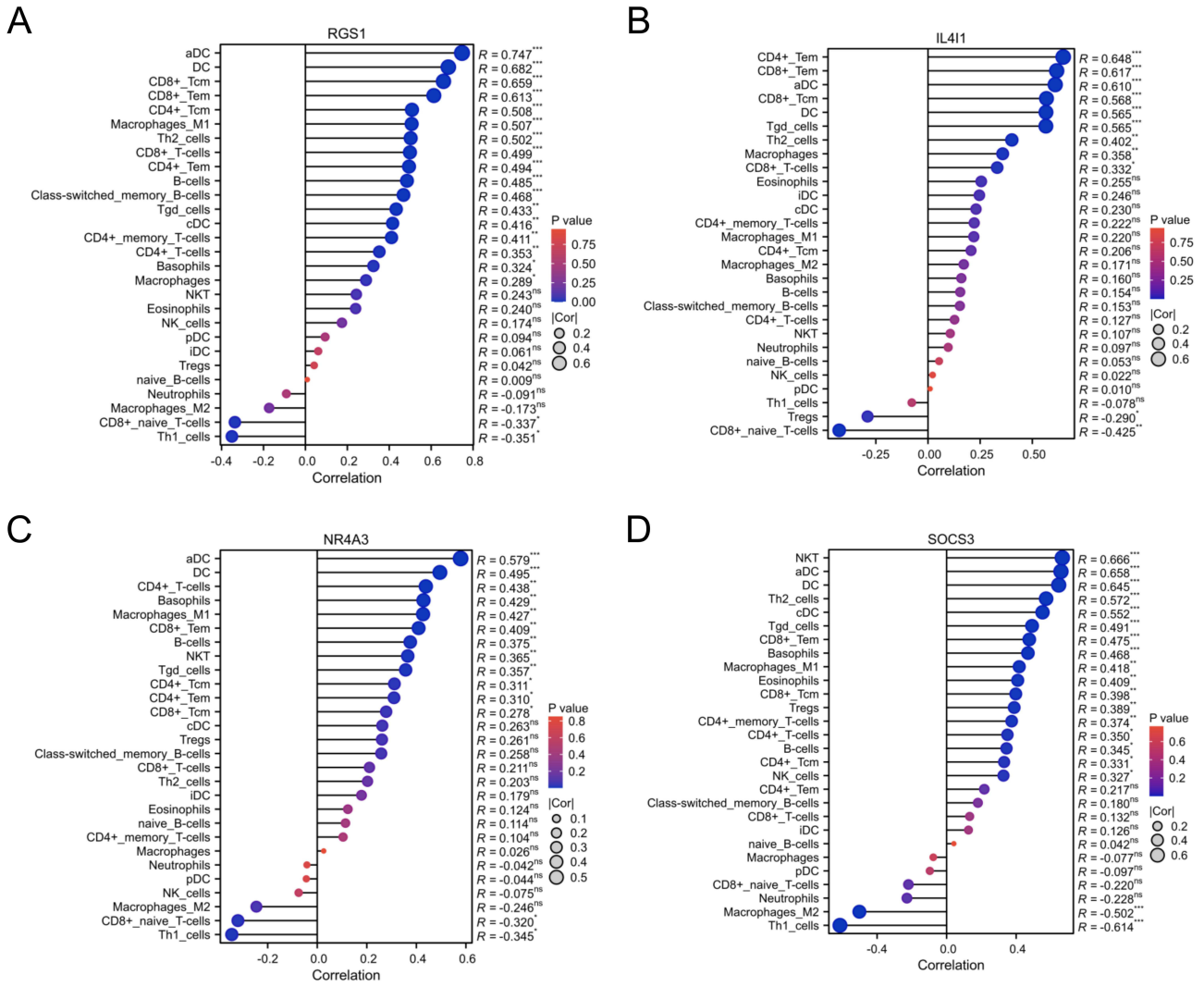


**Figure 7** The correlation between four crucial diagnostic biomarkers and the infiltration of immune cells. Correlation between infiltration of immune cells and the expression of (**A**) RGS1, (**B**) IL4I1, (**C**) NR4A3, and (**D**) SOCS3 genes. *p < 0.05, **p < 0.01, and ***p < 0.001; ns indicates no significant difference.

involvement of immune responses in CKD, suggesting that these pathways may be potential targets for therapeutic intervention. These results underscore the pivotal role of immune responses and metabolic processes in CKD pathogenesis, consistent with earlier reports highlighting the involvement of immune and metabolic dysregulation in renal diseases.[26,27] Comparing our bioinformatics approach with other techniques such as WGCNA and machine learning, our method provides a broad overview of immune activity and gene expression changes.[11] Integrating these techniques could further enhance our understanding of CKD pathogenesis.

One of the key contributions of our study is the integration of machine learning techniques to identify four diagnostic genes: RGS1, IL4I1, NR4A3, and SOCS3. These genes were validated in an independent dataset (GSE96804) and clinical samples, confirming their diagnostic potential. Among these genes, a regulator of G-Protein Signalling-1 (RGS1) enhances the activity of Gαi GTPase and functions to decrease the cellular response to prolonged chemokine stimulation.[28,29] RGS1 exhibits a significant level of expression in various types of immune cells, such as monocytes, dendritic cells, natural killer cells, B cells, T cells, and B cells. This suggests that RGS1 plays a crucial role in the regulation of immune cell activity.[30] Inhibition of RGS1 expression effectively suppressed the inflammatory response observed in both renal interstitial fibrosis cell cultures and mouse models. Consequently, targeting RGS1 could represent a promising therapeutic approach for the treatment of renal interstitial fibrosis.[31] The gene known as Interleukin-4-induced gene 1 (IL4I1) was initially discovered in B splenocytes of mice that had been activated by interleukin 4.[32]

Expression of IL4I1 has been described mainly in cells of the human immune system, with cells of myeloid origin showing the highest production, particularly after stimulation with inflammatory and T-helper type 1 stimulus.[33,34] Recent research has indicated that IL4I1 may serve as a promising biomarker and a target for therapy in oncological conditions.[32] The elevated expression of IL4I1 in immune cells in CKD patients further supports its role as a marker of immune activation and potential therapeutic target. NR4A3 belongs to the nuclear receptor subfamily 4, playing a crucial role in the regulation of cell function and inflammatory responses.[35] A current research finding indicates that the NR4A3 gene is implicated in promoting inflammation during the progression of osteoarthritis.[36] The involvement of NR4A3 in inflammation corroborates its potential role in contributing to the inflammatory environment in CKD. SOCS3 suppresses inflammatory responses and triggers the initial stage of immune defense across various mouse models.[37] SOCS3 plays a role in inhibiting the M1 proinflammatory state, thus reducing the activity of inflammatory reactions within macrophages.[38] Elevated levels of SOCS in peripheral blood mononuclear cells have been associated with a decline in kidney function.[39] This suggests that SOCS3 might be a critical regulator in mitigating the inflammatory responses that exacerbate CKD.

In our study, we developed a predictive nomogram for CKD diagnosis utilizing these four key genes. The model demonstrated considerable prognostic efficacy, indicating its potential usefulness in clinical settings. To address real-world applicability, this model can aid in early diagnosis and risk stratification in clinical practice, allowing for prompt and personalized treatment strategies. Regular monitoring of these gene expressions can also help in tracking disease progression, thereby enhancing patient management. It is important to compare our results with other existing techniques. For instance, traditional biomarkers for CKD, such as serum creatinine and estimated glomerular filtration rate (eGFR), provide a general assessment of kidney function but lack specificity regarding underlying immune mechanisms.[40] In contrast, our gene-based approach not only offers diagnostic accuracy but also sheds light on the immune dysregulation in CKD, providing a more comprehensive understanding of disease pathogenesis. Spearman correlation analysis revealed positive associations between the key diagnostic genes and various immune cells, indicating their roles in immune modulation and CKD pathogenesis. This finding is in line with previous research that has highlighted the complex interactions between immune cells and gene expression in CKD.[41,42] These correlations suggest that the identified genes not only serve as biomarkers but also play active roles in shaping the immune landscape in CKD, potentially offering new avenues for therapeutic intervention. The immunomodulatory roles of RGS1, IL4I1, NR4A3, and SOCS3 suggest they may influence the immune landscape in CKD, contributing to disease progression.

## Conclusions

In conclusion, our study provides a comprehensive analysis of immune alterations and gene expression profiles in CKD. The identified diagnostic genes and the constructed nomogram model offer potent tools for CKD diagnosis, with implications for both clinical practice and biomedical engineering. The insights gained from the roles of RGS1, IL4I1, NR4A3, and SOCS3 in immune modulation open avenues for the development of novel therapeutic strategies. The physical interactions and molecular mechanisms underlying the immunomodulatory roles of RGS1, IL4I1, NR4A3, and SOCS3 warrant further investigation as potential therapeutic targets in CKD. However, our study has limitations, such as dependence on bioinformatics predictions and the necessity for experimental validation in larger, more diverse groups. These issues may affect how we interpret the results and should be addressed in future research to enhance the clinical relevance of our findings.

## Data Sharing Statement

All dataset in the present study are available in GSE66494 (https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE66494) and GSE96804 (https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE96804).

## Ethics Approval

This study was approved by the Ethics Committee of the Ningde Hospital of Traditional Chinese Medicine.

## Funding

There is no funding to report.

## Disclosure

All authors declared that they have no competing interests.

## References

1. Stevens PE, Levin A. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann Intern Med*. 2013;158(11):825–830.
2. Hill NR, Fatoba ST, Oke JL, et al. Global prevalence of chronic kidney disease—a systematic review and meta-analysis. *PLoS One*. 2016;11(7):e0158765. doi:10.1371/journal.pone.0158765
3. Law JP, Pickup L, Pavlovic D, Townend JN, Ferro CJ. Hypertension and cardiomyopathy associated with chronic kidney disease: epidemiology, pathogenesis and treatment considerations. *J Human Hypertens*. 2023;37(1):1–19. doi:10.1038/s41371-022-00751-4
4. Tecklenborg J, Clayton D, Siebert S, Coley SM. The role of the immune system in kidney disease. *Clin Exp Immunol*. 2018;192(2):142–150. doi:10.1111/cei.13119
5. Rodríguez-Iturbe B, Pons H, Herrera-Acosta J, Johnson RJ. Role of immunocompetent cells in nonimmune renal diseases. *Kidney Int*. 2001;59(5):1626–1640. doi:10.1046/j.1523-1755.2001.0590051626.x
6. Mertowska P, Mertowski S, Wojnicka J, et al. A link between chronic kidney disease and gut microbiota in immunological and nutritional aspects. *Nutrients*. 2021;13:10. doi:10.3390/nu13103637
7. Sato Y, Yanagita M. Immune cells and inflammation in AKI to CKD progression. *Am J Physiol Renal Physiol*. 2018;315(6):F1501–f1512. doi:10.1152/ajprenal.00195.2018
8. Asif D, Bibi M, Arif MS, Mukheimer A. Enhancing heart disease prediction through ensemble learning techniques with hyperparameter optimization. *Algorithms*. 2023;16(6):308. doi:10.3390/a16060308
9. Arif MS, Mukheimer A, Asif D. Enhancing the early detection of chronic kidney disease: a robust machine learning model. *Big Data Cogn Comput*. 2023;7(3):144. doi:10.3390/bdcc7030144
10. Awan MZ, Arif MS, Ul Abideen MZ, Abodayeh K. Comparative analysis of machine learning models for breast cancer prediction and diagnosis: a dual-dataset approach. *Indones J Electr Eng Comput Sci*. 2024;34(3):13.
11. Hu Z, Liu Y, Zhu Y, Cui H, Pan J. Identification of key biomarkers and immune infiltration in renal interstitial fibrosis. *Ann Transl Med*. 2022;10(4):190. doi:10.21037/atm-22-366
12. Fan J, Shi S, Qiu Y, Liu M, Shu Q. Analysis of signature genes and association with immune cells infiltration in pediatric septic shock. *Front Immunol*. 2022;13:1056750. doi:10.3389/fimmu.2022.1056750
13. Aran D. Cell-type enrichment analysis of bulk transcriptomes using xCell. *Methods Mol Biol*. 2020;2120:263–276.
14. Wufuer D, Li Y, Aierken H, Zheng J. Bioinformatics-led discovery of ferroptosis-associated diagnostic biomarkers and molecule subtypes for tuberculosis patients. *Eur J Med Res*. 2023;28(1):445. doi:10.1186/s40001-023-01371-5
15. Han W, Li C, Wang Y, Huo B, Li W, Shi W. Heme metabolism-related gene TENT5C is a prognostic marker and investigating its immunological role in colon cancer. *Pharmacogenomics Pers Med*. 2023;16:1127–1143. doi:10.2147/PGPM.S433790
16. Canzler S, Hackermüller J. multiGSEA: a GSEA-based pathway enrichment analysis for multi-omics data. *BMC Bioinf*. 2020;21(1):561. doi:10.1186/s12859-020-03910-x
17. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;14:7. doi:10.1186/1471-2105-14-7
18. Li S, Ma L, Cui R. Identification of novel diagnostic biomarkers and classification patterns for osteoarthritis by analyzing a specific set of genes related to inflammation. *Inflammation*. 2023;46(6):2193–2208. doi:10.1007/s10753-023-01871-w
19. Kang J, Choi YJ, Kim IK, et al. LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer. *Cancer Res Treat*. 2021;53(3):773–783. doi:10.4143/crt.2020.974
20. Wang F, Wang Y, Ji X, Wang Z. Effective macrosomia prediction using random forest algorithm. *Int J Environ Res Public Health*. 2022;19:6.
21. Zhou Y, Shi W, Zhao D, Xiao S, Wang K, Wang J. Identification of immune-associated genes in diagnosing aortic valve calcification with metabolic syndrome by integrated bioinformatics analysis and machine learning. *Front Immunol*. 2022;13:937886. doi:10.3389/fimmu.2022.937886
22. Xu L. The role of myeloid cells in acute kidney injury and kidney repair. *Kidney360*. 2021;2(11):1852–1864. doi:10.34067/KID.0000672021
23. Zeng J, Zhang Y, Huang C. Macrophages polarization in renal inflammation and fibrosis animal models (Review). *Mol Med Rep*. 2024;29:2. doi:10.3892/mmr.2023.13125
24. Speer T, Dimmeler S, Schunk SJ, Fliser D, Ridker PM. Targeting innate immunity-driven inflammation in CKD and cardiovascular disease. *Nat Rev Nephrol*. 2022;18(12):762–778. doi:10.1038/s41581-022-00621-9
25. Diaz-Ricart M, Torramade-Moix S, Pascual G, et al. Endothelial damage, inflammation and immunity in chronic kidney disease. *Toxins*. 2020;12:6. doi:10.3390/toxins12060361
26. Li Y, Sha Z, Peng H. Metabolic reprogramming in kidney diseases: evidence and therapeutic opportunities. *Int J Nephrol*. 2021;2021:5497346. doi:10.1155/2021/5497346
27. Grayson PC, Eddy S, Taroni JN, et al. Metabolic pathways and immunometabolism in rare kidney diseases. *Ann Rheumatic Dis*. 2018;77(8):1226–1233. doi:10.1136/annrheumdis-2017-212935
28. Han SB, Moratz C, Huang NN, et al. Rgs1 and Gnai2 regulate the entrance of B lymphocytes into lymph nodes and B cell motility within lymph node follicles. *Immunity*. 2005;22(3):343–354. doi:10.1016/j.immuni.2005.01.017
29. Moratz C, Hayman JR, Gu H, Kehrl JH. Abnormal B-cell responses to chemokines, disturbed plasma cell localization, and distorted immune tissue architecture in Rgs1-/- mice. *Mol Cell Biol*. 2004;24(13):5767–5775. doi:10.1128/MCB.24.13.5767-5775.2004

30. Bai Y, Hu M, Chen Z, Wei J, Du H. Single-cell transcriptome analysis reveals RGS1 as a new marker and promoting factor for T-cell exhaustion in multiple cancers. *Front Immunol*. 2021;12:767070. doi:10.3389/fimmu.2021.767070
31. Lu T, Chen S, Xu J. RGS1 mediates renal interstitial fibrosis through activation of the inflammatory response. *Arch Biochem Biophys*. 2023;750:109744. doi:10.1016/j.abb.2023.109744
32. Rao D, Yu C, Wang T, et al. Pan-cancer analysis combined with experimental validation revealed IL4I1 as an immunological and prognostic biomarker. *Int Immunopharmacol*. 2022;111:109091. doi:10.1016/j.intimp.2022.109091
33. Marquet J, Lasoudris F, Cousin C, et al. Dichotomy between factors inducing the immunosuppressive enzyme IL-4-induced gene 1 (IL4I1) in B lymphocytes and mononuclear phagocytes. *Eur J Immunol*. 2010;40(9):2557–2568. doi:10.1002/eji.201040428
34. Molinier-Frenkel V, Prévost-Blondel A, Castellano F. The IL4I1 enzyme: a new player in the immunosuppressive tumor microenvironment. *Cells*. 2019;8:7. doi:10.3390/cells8070757
35. Jiang Y, Feng YP, Tang LX, Yan YL, Bai JW. The protective role of NR4A3 in acute myocardial infarction by suppressing inflammatory responses via JAK2-STAT3/NF-κB pathway. *Biochem Biophys Res Commun*. 2019;517(4):697–702. doi:10.1016/j.bbrc.2019.07.116
36. Ma C, Wu L, Song L, et al. The pro-inflammatory effect of NR4A3 in osteoarthritis. *J Cell Mol Med*. 2020;24(1):930–940. doi:10.1111/jcmm.14804
37. Carow B, Rottenberg ME. SOCS3, a major regulator of infection and inflammation. *Front Immunol*. 2014;5:58. doi:10.3389/fimmu.2014.00058
38. Qin H, Holdbrooks AT, Liu Y, Reynolds SL, Yanagisawa LL, Benveniste EN. SOCS3 deficiency promotes M1 macrophage polarization and inflammation. *J Immunol*. 2012;189(7):3439–3448. doi:10.4049/jimmunol.1201168
39. Rastmanesh MM, Bluyssen HA, Joles JA, Boer P, Willekes N, Braam B. Increased expression of SOCS3 in monocytes and SOCS1 in lymphocytes correlates with progressive loss of renal function and cardiovascular risk factors in chronic kidney disease. *Eur J Pharmacol*. 2008;593(1–3):99–-104. doi:10.1016/j.ejphar.2008.07.013
40. Vassalotti JA, Centor R, Turner BJ, Greer RC, Choi M, Sequist TD. Practical approach to detection and management of chronic kidney disease for the primary care clinician. *Am J Med*. 2016;129(2):153–162.e7. doi:10.1016/j.amjmed.2015.08.025
41. Holdsworth SR, Kitching AR. Immune-mediated kidney disease in 2017: progress in mechanisms and therapy for immunological kidney disease. *Nat Rev Nephrol*. 2018;14(2):76–78. doi:10.1038/nrneph.2017.171
42. Seethapathy H, Mistry K, Sise ME. Immunological mechanisms underlying clinical phenotypes and noninvasive diagnosis of immune checkpoint inhibitor-induced kidney disease. *Immunol Rev*. 2023;318(1):61–69. doi:10.1111/imr.13243

**Pharmacogenomics and Personalized Medicine** Dovepress

## Publish your work in this journal

Pharmacogenomics and Personalized Medicine is an international, peer-reviewed, open access journal characterizing the influence of genotype on pharmacology leading to the development of personalized treatment programs and individualized drug selection for improved safety, efficacy and sustainability. This journal is indexed on the American Chemical Society's Chemical Abstracts Service (CAS). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

**Submit your manuscript here:** https://www.dovepress.com/pharmacogenomics-and-personalized-medicine-journal