



OPEN ACCESS

EDITED BY
Jinhui Liu,
Nanjing Medical University, China

REVIEWED BY
Jinhai Deng,
King's College London,
United Kingdom
Jiankang Fang,
University of Pennsylvania,
United States
Xiuyuan Zhang,
Huazhong University of Science
and Technology, China
Chunyu Zhang,
Tongji University, China

*CORRESPONDENCE
Zhonggao Xu
zhonggao@jlu.edu.cn

†These authors have contributed
equally to this work

SPECIALTY SECTION
This article was submitted to
Nephrology,
a section of the journal
Frontiers in Medicine

RECEIVED 18 April 2022
ACCEPTED 13 September 2022
PUBLISHED 29 September 2022

CITATION
Fu S, Cheng Y, Wang X, Huang J, Su S,
Wu H, Yu J and Xu Z (2022)
Identification of diagnostic gene
biomarkers and immune infiltration
in patients with diabetic kidney
disease using machine learning
strategies and bioinformatic analysis.
Front. Med. 9:918657.
doi: 10.3389/fmed.2022.918657

COPYRIGHT
© 2022 Fu, Cheng, Wang, Huang, Su,
Wu, Yu and Xu. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Identification of diagnostic gene biomarkers and immune infiltration in patients with diabetic kidney disease using machine learning strategies and bioinformatic analysis

Shaojie Fu^{1†}, Yanli Cheng^{1†}, Xueyao Wang¹, Jingda Huang¹,
Sensen Su¹, Hao Wu¹, Jinyu Yu² and Zhonggao Xu^{1*}

¹Department of Nephrology, The First Hospital of Jilin University, Changchun, China, ²Department of Urology, The First Hospital of Jilin University, Changchun, China

Objective: Diabetic kidney disease (DKD) is the leading cause of chronic kidney disease and end-stage renal disease worldwide. Early diagnosis is critical to prevent its progression. The aim of this study was to identify potential diagnostic biomarkers for DKD, illustrate the biological processes related to the biomarkers and investigate the relationship between them and immune cell infiltration.

Materials and methods: Gene expression profiles (GSE30528, GSE96804, and GSE99339) for samples obtained from DKD and controls were downloaded from the Gene Expression Omnibus database as a training set, and the gene expression profiles (GSE47185 and GSE30122) were downloaded as a validation set. Differentially expressed genes (DEGs) were identified using the training set, and functional correlation analyses were performed. The least absolute shrinkage and selection operator (LASSO), support vector machine-recursive feature elimination (SVM-RFE), and random forests (RF) were performed to identify potential diagnostic biomarkers. To evaluate the diagnostic efficacy of these potential biomarkers, receiver operating characteristic (ROC) curves were plotted separately for the training and validation sets, and immunohistochemical (IHC) staining for biomarkers was performed in the DKD and control kidney tissues. In addition, the CIBERSORT, XCELL and TIMER algorithms were employed to assess the infiltration of immune cells in DKD, and the relationships between the biomarkers and infiltrating immune cells were also investigated.

Results: A total of 95 DEGs were identified. Using three machine learning algorithms, *DUSP1* and *PRKAR2B* were identified as potential biomarker genes for the diagnosis of DKD. The diagnostic efficacy of *DUSP1* and *PRKAR2B* was assessed using the areas under the curves in the ROC analysis of the training set (0.945 and 0.932, respectively) and validation set (0.789 and 0.709,

respectively). IHC staining suggested that the expression levels of *DUSP1* and *PRKAR2B* were significantly lower in DKD patients compared to normal. Immune cell infiltration analysis showed that B memory cells, gamma delta T cells, macrophages, and neutrophils may be involved in the development of DKD. Furthermore, both of the candidate genes are associated with these immune cell subtypes to varying extents.

Conclusion: *DUSP1* and *PRKAR2B* are potential diagnostic markers of DKD, and they are closely associated with immune cell infiltration.

KEYWORDS

diabetic kidney disease, immune infiltration, diagnostic biomarker, bioinformatic analysis, machine learning strategy

Introduction

The increase in the prevalence of diabetes mellitus (1) has had a major impact on the prevalence of diabetic kidney disease (DKD), which is the cause of approximately 50% of cases of end-stage renal disease (ESRD) in developed countries (2); the prevalence of DKD has also been increasing in China (3). Therefore, DKD is now placing a substantial burden on the healthcare system (4). Fibrosis, hypertrophy, and the accumulation of extracellular matrix around glomerular and tubular cells are typical pathological changes in DKD (5). The progression of DKD is associated with microalbuminuria in the early stages and uremia in the later stages (6). Although renal biopsy is the most accurate method of diagnosis, it is not widely used because of its invasiveness. Instead, the diagnostic criteria are typically based on the clinical manifestation of the disease and vary around the world. These criteria include albumin/creatinine ratio, estimated glomerular filtration rate, serum creatinine concentration, the results of renal ultrasonography, and the presence of retinopathy (7). However, these parameters are not highly sensitive or specific diagnostic measures and do not reflect the severity of the renal damage. It is universally accepted that proteinuria is negative in the early stage of DKD and microalbuminuria may revert back to normal urinary albumin excretion rates over declined renal function (8, 9). Therefore, more accurate biomarkers that could be used for the early diagnosis of DKD would facilitate early intervention and improve the outcomes of patients.

Machine learning is increasingly being used to help identify genes that may have diagnostic potential, and the accuracy of identification of genes that are differentially expressed on microarrays has markedly improved (10). With more and more studies revealing the role of immune system and immune cells in the occurrence and development of DKD (11), people gradually regard DKD as an immune-mediated disease (12). Currently, immune infiltrating cells mediated oxidative stress process is

considered to play an important role in the occurrence and development of DKD (13). Animal experiments also showed that the number and activity of immune cells in the kidney are related to the renal damage (14). At present, some drugs have been confirmed by experiments that influence the protein expression in renal tissues interacting with infiltrated immune cells could improve DKD fibrosis (15), which provides new therapeutic strategies for the treatment of DKD. Therefore, it is important to reveal the infiltration of immune cells in kidney tissue of DKD and explore the relationship between the identified biomarkers and immune infiltrating cells.

In the present study, we first downloaded microarray datasets obtained from patients with DKD from the Gene Expression Omnibus (GEO) database, merged them, and performed differential expression analysis. Gene set enrichment analysis (GSEA) was then performed using the obtained differentially expressed genes (DEGs), and three different machine learning algorithms were used to identify potential diagnostic markers for DKD. Subsequently, we utilized CIBERSORT, XCELL and TIMER to understand the features of the immune cell infiltration into the kidneys of patients with DKD, in comparison to normal renal tissue. Finally, we assessed the relationships between the identified candidate diagnostic markers and the immune infiltration to further investigate the molecular mechanisms of the progression of DKD.

Materials and methods

Identification of differentially expressed genes

Microarray datasets for renal glomeruli from patients with DKD and controls (GSE30528, GSE96804, and GSE99339)

were obtained from the GEO database¹ (16). The three data sets were merged and used as a training set, and the inter-batch differences were corrected for using the SVA package in R (17). The limma package in R was then used for the identification and normalization of the differentially expressed genes (DEGs) by comparing the expression levels in the glomeruli from controls and patients with DKD (18). An adjusted *P*-value < 0.05 and a |log fold change (FC)| > 1 were the criteria used to define DEGs.

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the enriched DEGs were performed through the “clusterProfiler” package in R, and an adjusted *P*-value < 0.05 was considered to represent statistical significance (19). GSEA was conducted to identify the functional terms that significantly differed between the DKD and control samples (20). We used “c2.cp.kegg.v7.4.symbols.gmt” from the Molecular Signatures Database (MSigDB) as the reference gene set and screened for significantly enriched genes using an adjusted *P*-value of <0.05.

Screening for potential diagnostic biomarkers

To identify the potential diagnostic biomarkers for DKD, three machine learning algorithms were used to predict disease status: the least absolute shrinkage and selection operator (LASSO), support vector machine-recursive feature elimination (SVM-RFE), and random forests (RF). LASSO is a regression analysis algorithm that uses regularization to improve the accuracy of prediction and was performed using the “glmnet” package in R, in which the minimum absolute shrinkage was considered optimal (21). SVM is a supervised machine learning technique that is widely used for regression and classification, and the RFE algorithm is often applied to avoid overfitting. SVM-RFE was performed through the “e1071” package in R, with fivefold cross-validation (22). RF is an ensemble learning method that is based on the construction of many classification trees and was performed using the “randomForest” package in R, with genes awarded importance scores >2.0 being screened out (23). Finally, the genes identified using all three classification models were selected as potential gene biomarkers for further analysis.

Diagnostic value of the candidate biomarkers for diabetic kidney disease

To test the predictive value of the identified gene biomarkers, we plotted receiver operating characteristic (ROC) curves for each using the mRNA expression data for 64 samples from patients with DKD and 44 samples from controls that comprised the training set. The GSE47185 and GSE30122 data set was then used as the validation set to test the efficacy of prediction using the identified gene biomarkers in more depth. The area under the ROC curve (AUC) was calculated to evaluate the diagnostic efficacy of the gene biomarkers identified using the algorithms. ROC curves were plotted using the “pROC” package in R, and a two-sided *P*-value of <0.05 was taken to indicate statistical significance.

Renal immunohistochemical staining

To further understand the difference in expression of identified gene biomarkers in DKD and control kidney tissues, immunohistochemical (IHC) staining was conducted. The DKD kidney tissues were obtained from the patients with DKD diagnosed by renal biopsy pathology, while the normal kidney tissues were obtained from normal tissue next to kidney cancer. For IHC staining, all 50 DKD and 30 normal tissues from the kidney were fixed in 10% formalin for 24 hours and embedded in paraffin, then were cut into 5 μm thickness sliced sections. Paraffin-embedded tissue slides were deparaffinized in xylene and hydrated by increasing the alcohol concentrations. Tissue sections were treated with sodium citrate (pH = 6) and hydrogen peroxide for 15 min, then washed with PBS for antigen retrieval. Tissue sections were incubated in 5% BSA for 2 h (24). Subsequently, the tissue sections were treated with primary anti-DUSP1 and anti-PRKAR2B (1:200 dilution, the antibodies were purchased from Abcam) and incubated overnight at 4°C. Subsequently, the sections were incubated with horseradish peroxidase-conjugated secondary antibodies for 60 min at room temperature and stained through a horseradish peroxidase-DAB kit. All the stained sections were examined using an OLYMPUS cellSens Entry microscopy system. IHC staining was analyzed by Image-Pro Plus 6.0 software using a semiquantitative scoring method based on the staining intensity and the percentage of positively-stained cells (25).

GSEA and GSA analyses for the candidate biomarkers

To further illustrate the biological processes related to the candidate biomarkers, we divided the DKD samples into

¹ <http://www.ncbi.nlm.nih.gov/geo/>

high and low expression groups based on the median values of the candidate biomarkers' expressions respectively, and the GSVA and GSEA analyses were performed. The gene sets of "c5.go.bp.v7.4.symbols" were downloaded from the MSigDB database to enrich the biological process pathways (26). GSVA algorithm was used to analyze the data set to find biological process pathways with significant differences between samples (27).

Immune cells infiltration

The CIBERSORT, XCELL, and TIMER algorithms were used to analyze the immunological characteristics of the DKD groups and control groups. CIBERSORT is the most widely used immune infiltration algorithm. It uses the linear support vector regression to deconvolute the tissue expression matrix and can accurately quantify the abundance scores of 22 types of immune cells for each sample (28). XCELL is a new method for cell-type enrichment analysis using single-sample gene set enrichment analysis. It could calculate the enrichment scores for 64 cell types, including not only 34 types of immune cells but also 30 stromal and other cells (29). Tumor immune estimation resource (TIMER) is a tool which integrates multiple state-of-the-art algorithms for immune infiltration estimation (30). Subsequently, the "ggplot2" package in R was used for principal components analysis (PCA) clustering studies of the immune cell infiltration matrix. A heat map of the immune cell infiltration matrix by different methods was plotted using the "pheatmap" package to show the allocation of immune cell infiltrates in the respective kidney specimens. The correlations among the infiltrating immune cells identified were also plotted by the "pheatmap" package in the form of a heatmap. The "vioplot" package in R was used to draw a violin plot to aid visualization of the difference in immune cell infiltration between samples from patients with DKD and controls. Spearman's rank correlation analysis in R was used to further analyze the relationships between the key diagnostic genes and immune infiltrating cells, and the "ggstatsplot" and "ggplot2" packages were used to aid visualization of the results.

Results

Identification of DEGs in samples from patients with diabetic kidney disease and controls

The flowchart of this study was shown in **Figure 1**. Data from a total of 64 samples from patients with DKD and 44 controls obtained from three GEO data sets (GSE30528, GSE96804, and GSE99339) were retrospectively analyzed. The DEGs were identified using the limma package after

the elimination of batch effects. A total of 95 DEGs were identified; 39 genes were significantly upregulated, and 56 were significantly downregulated (**Figure 2**).

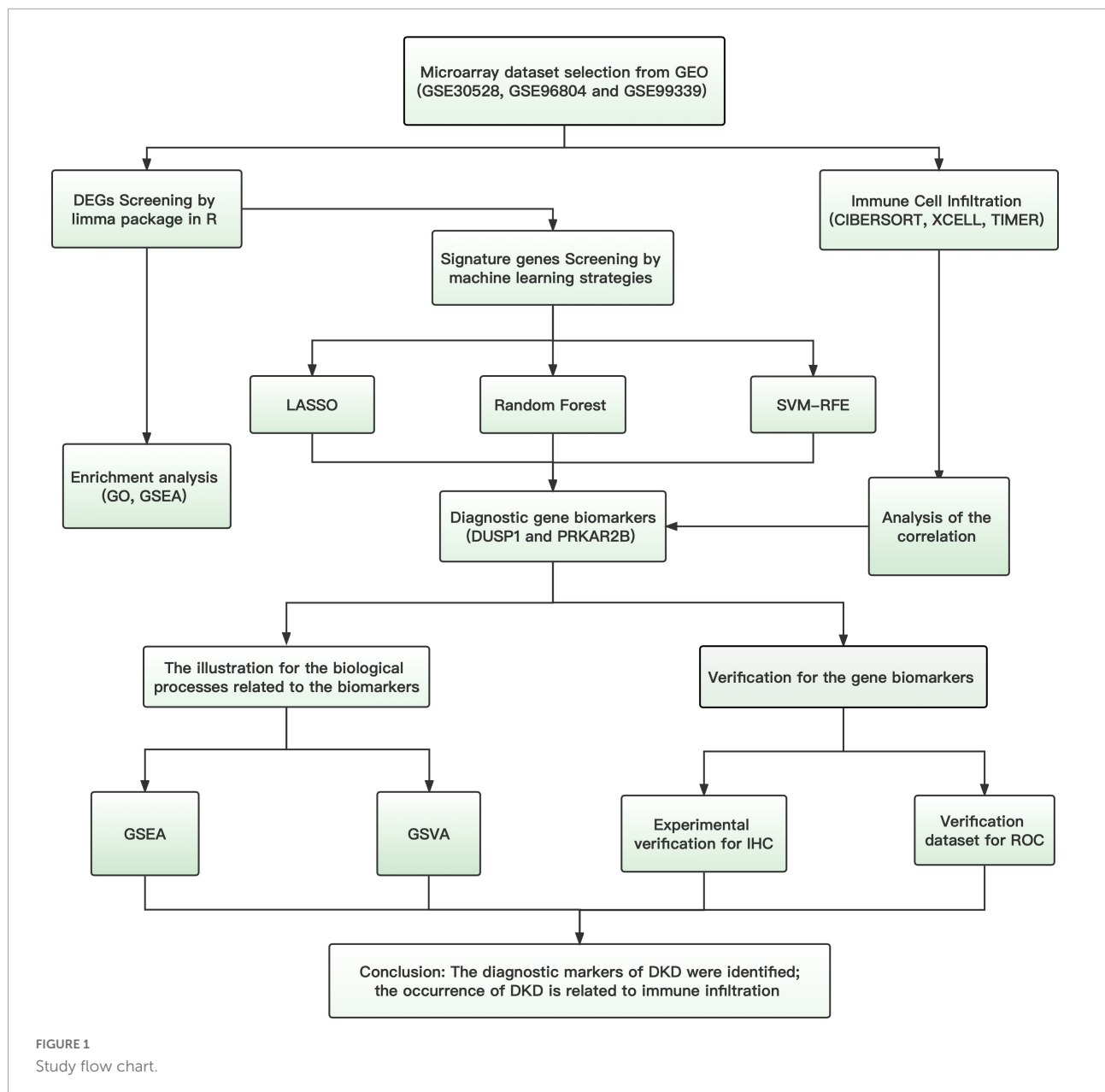
Functional correlation analysis

Gene Ontology functional enrichment, KEGG pathway enrichment, and GSEA were used to investigate the mechanisms involved in the pathogenesis of DKD. The ten most significant biological process (BP), cellular component (CC), and molecular function (MF) terms in the GO functional enrichment for the DEGs were identified (**Figure 3A**). KEGG pathway enrichment revealed that the DEGs principally represented extracellular matrix (ECM)-receptor interaction, the interleukin (IL)-17 signaling pathway, focal adhesion, the advanced glycation end-product (AGE)-receptor for AGEs signaling pathway in diabetic complications, protein digestion and absorption, and glycolysis/gluconeogenesis (**Figure 3B**). Furthermore, GSEA identified the following principal enriched pathways: asthma, cytokine-cytokine receptor interaction, ECM receptor interaction, focal adhesion, and systemic lupus erythematosus (**Figure 3C**). All these findings suggest that the immune response plays a significant role in DKD.

Identification and verification of diagnostic markers

Three different algorithms were used to identify potential diagnostic markers for DKD. Nineteen genes among the DEGs were identified as potential biomarkers using the LASSO logistic regression algorithm (**Figure 4A**), eight genes were identified as potential biomarkers using the SVM-RFE algorithm (**Figure 4B**), and six genes were identified as potential biomarkers using the RF algorithm (**Figure 4C**). Two genes, *DUSP1* and *PRKAR2B*, were identified using all three algorithms and were therefore regarded as strong candidate predictors of DKD (**Figure 4D**).

As shown in **Figures 5A,B**, the expression of these two genes was effective at discriminating DKD from control samples, with AUCs of 0.945 (95% CI 0.891–0.988) for *DUSP1* and 0.932 (95% CI 0.879–0.976) for *PRKAR2B*. To further evaluate the diagnostic efficacy of the two candidate biomarkers, the GSE47185 and GSE30122 datasets were merged as a validation set, which comprised 33 samples from patients with DKD and 67 samples from controls. As shown in **Figures 5C,D**, ROC curves for *DUSP1* and *PRKAR2B* were also plotted for the validation set and yielded AUCs of 0.789 (95% CI 0.686–0.875) and 0.709 (95% CI 0.598–0.813), respectively. These results indicate that expression of the two genes show good potential in the diagnosis of DKD.



Renal immunohistochemical staining

Immunohistochemical staining was performed in 50 DKD kidney tissues and 30 normal kidney tissues. Compared with the normal, patients with DKD presented with significantly lower expression levels of the DUSP1 and PRKAR2B (Figures 6A,B).

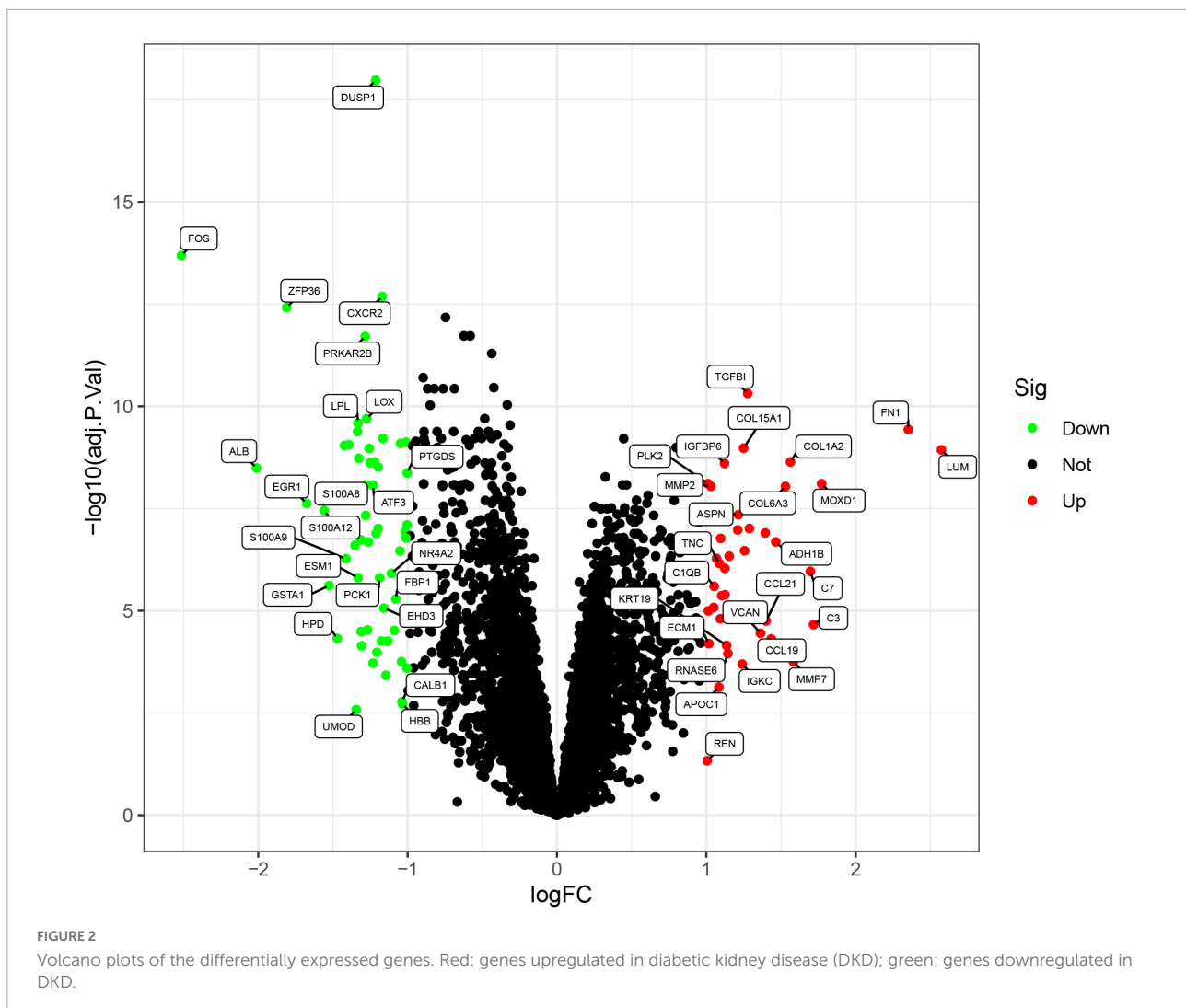
GSEA and GSEA analyses for the candidate biomarkers

The results of exploring the biological processes related to DUSP1 using GSEA and GSEA were shown in Figures 7A,B

respectively, while the results of PRKAR2B were shown in Figures 7C,D respectively. The results showed that the expression of both DUSP1 and PRKAR2B was closely related to renal function, such as the regulation of glomerular filtration, glomerular epithelial cell development, renal system process involved in regulation of blood volume, etc.

Analysis of immune cell infiltration

The heatmap of immune cells with differential infiltration between DKD patients and controls based on CIBERSORT, XCELL and TIMER algorithms is shown in Figure 8A. In



CIBERSORT algorithm, the correlation heatmap for the 22 immune cell subtypes demonstrated that two immune cell subtypes (activated mast cells and resting mast cells) were negatively correlated (**Supplementary Figure 1**), and the violin plot of the differentially infiltrated immune cells demonstrated that B memory cells, gamma delta T cells, resting natural killer (NK) cells, macrophages, and resting mast cells were present in larger numbers in the DKD samples than in the control samples, whereas follicular helper T cells, resting dendritic cells, activated mast cells, and neutrophils were present in smaller numbers (**Supplementary Figure 2**). The results of XCELL and TIMER algorithms also support the finding of CIBERSORT that macrophages infiltrate more in DKD samples and neutrophils infiltrate more in controls. Violin plots of the different infiltrated immune cells identified by XCELL and TIMER algorithms were shown in **Supplementary Figures 3, 4** respectively.

The relationships between the candidate diagnostic marker genes and the infiltrating immune cells were also analyzed.

DUSP1 expression positively correlated with neutrophil infiltration ($r = 0.69$, $P < 0.001$), activated mast cell infiltration ($r = 0.41$, $P < 0.001$), and T follicular helper cell infiltration ($r = 0.32$, $P < 0.001$) and negatively correlated with plasma cell infiltration ($r = -0.23$, $P = 0.017$), resting dendritic cell infiltration ($r = -0.25$, $P = 0.010$), memory B cell infiltration ($r = -0.34$, $P < 0.001$), resting mast cell infiltration ($r = -0.44$, $P < 0.001$), and M2 macrophage infiltration ($r = -0.52$, $P < 0.001$) (**Figure 8B**). PRKAR2B expression positively correlated with neutrophil infiltration ($r = 0.68$, $P < 0.001$), active mast cell infiltration ($r = 0.34$, $P < 0.001$), resting NK cell infiltration ($r = 0.29$, $P = 0.003$), T follicular helper cell infiltration ($r = 0.26$, $P = 0.007$), resting memory CD4 + T cell infiltration ($r = 0.22$, $P = 0.020$), and naïve B cell infiltration ($r = 0.20$, $P = 0.039$) and negatively correlated with resting dendritic cell infiltration ($r = -0.20$, $P = 0.037$), resting mast cell infiltration ($r = -0.43$, $P < 0.001$), memory B cell infiltration

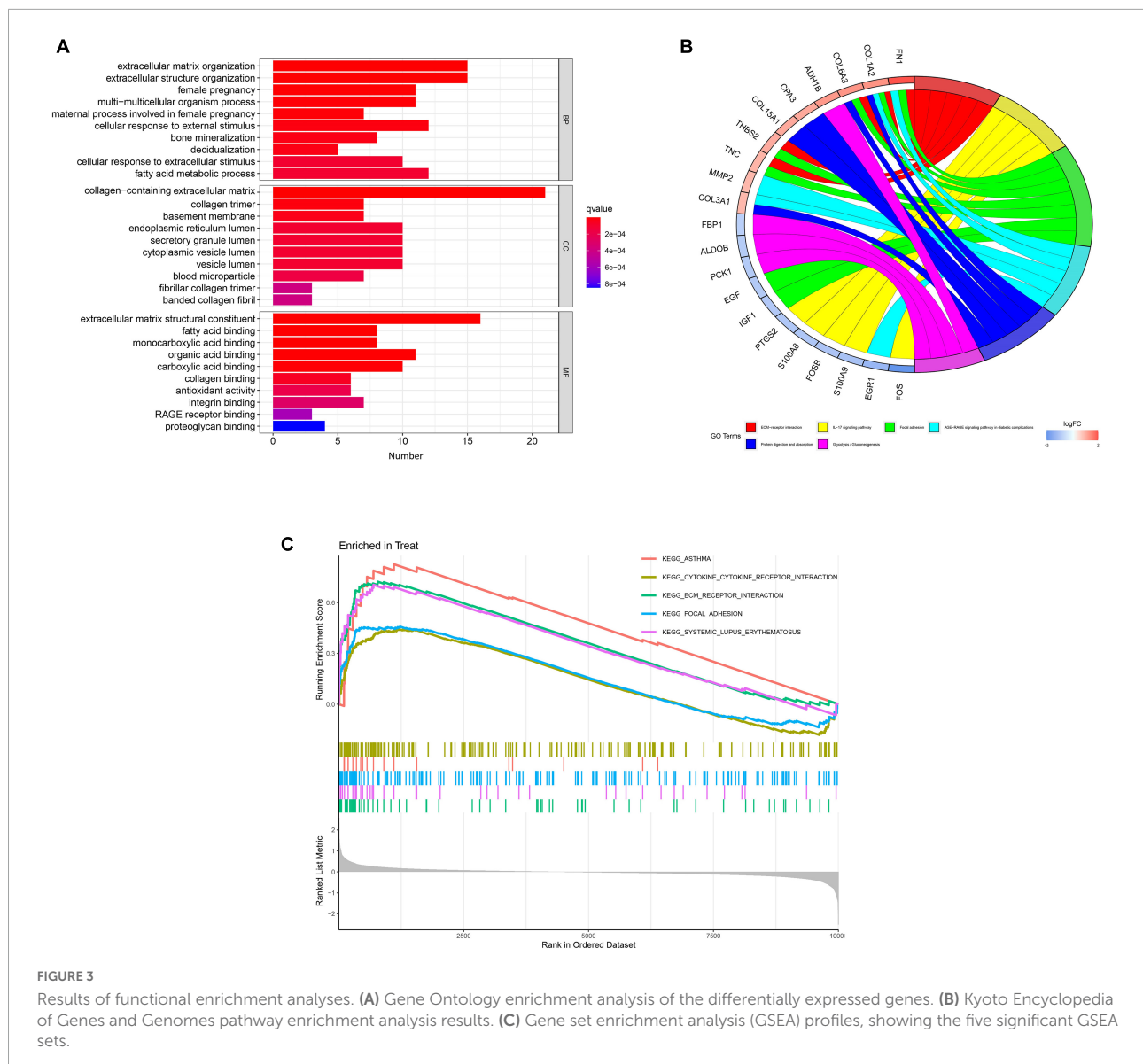


FIGURE 3

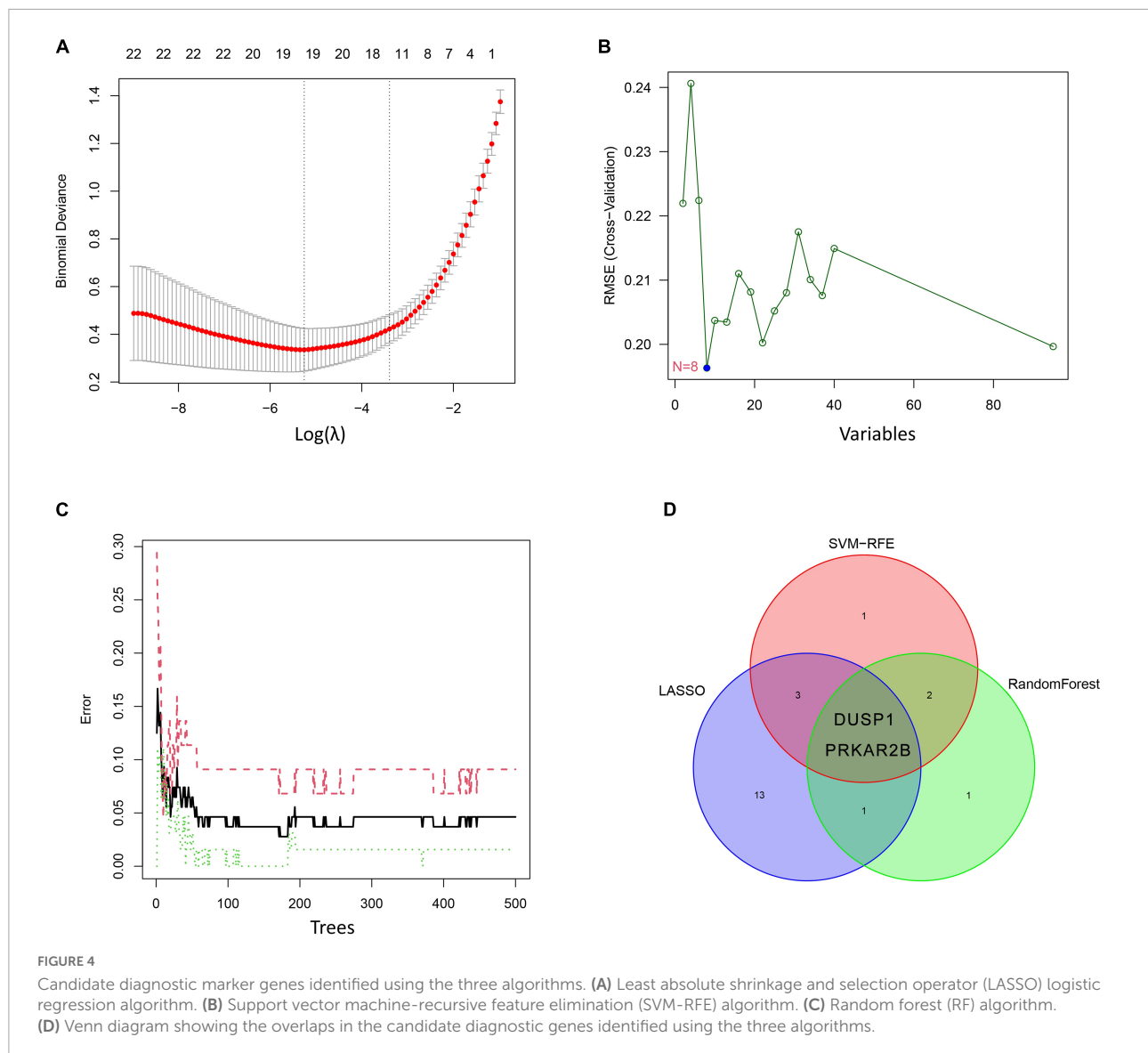
Results of functional enrichment analyses. (A) Gene Ontology enrichment analysis of the differentially expressed genes. (B) Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis results. (C) Gene set enrichment analysis (GSEA) profiles, showing the five significant GSEA sets.

($r = -0.44$, $P < 0.001$), gamma delta T cell infiltration ($r = -0.48$, $P < 0.001$), and M2 macrophage infiltration ($r = -0.65$, $P < 0.001$) (Figure 8C).

Discussion

Diabetic kidney disease is the most common cause of CKD and ESRD worldwide (31) and involves the hyperplasia or hypertrophy of various cell types in the glomerulus and tubules, thickening of the glomerular and tubular basement membranes, and expansion of the tubulointerstitial and mesangial compartments (32). Because the early clinical manifestations of DKD are easy to ignore, it has usually progressed to an advanced stage by the time it is recognized,

which is associated with a poor prognosis. In addition, immune cell infiltration plays a significant role in the development of DKD (33). Therefore, the prognosis of patients with DKD could be improved by the identification and use of specific diagnostic markers and a fuller characterization of the pattern of immune cell infiltration into the glomerular region of the kidneys of patients with DKD. Recently, bioinformatics and CIBERSORT tools have been increasingly used to identify novel diagnostic markers and to analyze the patterns of immune cell infiltration in tissues (34, 35). However, few studies have focused on characterizing the relationships between candidate gene biomarkers and the immune cell infiltration in DKD (36). Therefore, we aimed to identify candidate diagnostic biomarkers for DKD and further investigate the role of immune cell infiltration in DKD.

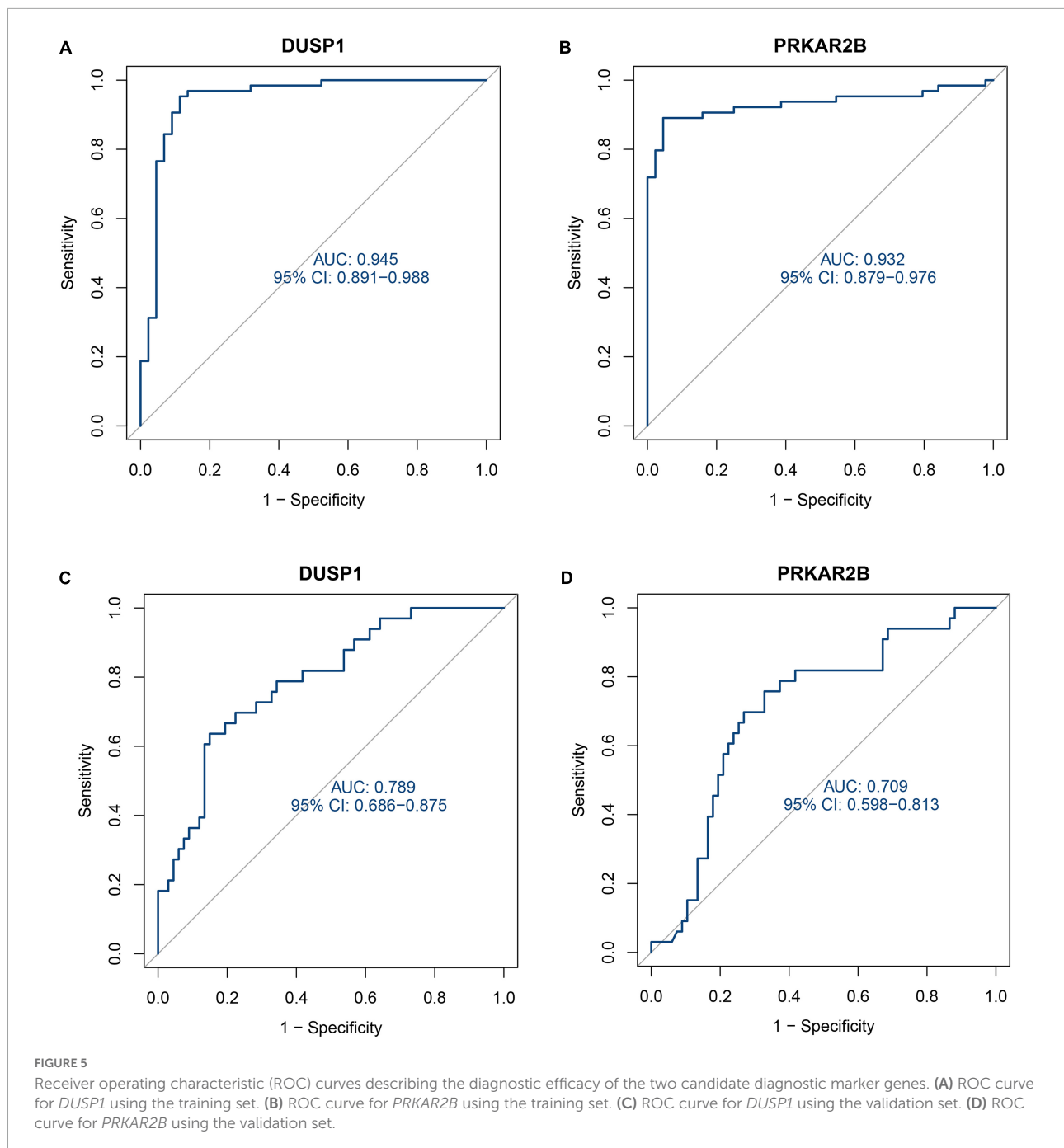


We downloaded expression profile datasets from the GEO database and identified a total of 95 DEGs. GSEA showed that the following principal pathways were enriched: asthma, cytokine-cytokine receptor interaction, ECM receptor interaction, focal adhesion, and systemic lupus erythematosus. Cui et al. (37) showed that the DEGs in glomerular and tubular samples from patients with DKD were primarily associated with pathways involved in ECM-receptor interactions and cytokine-cytokine receptor interactions, which is consistent with the present findings and suggests that these pathways may play significant roles in the progression of DKD.

Least absolute shrinkage and selection operator logistic regression is a machine learning method that determines the variable by finding λ when the classification error is the smallest (38). SVM is a machine learning method of pattern recognition and estimation of function that operates within

the framework of statistical learning theory and structural risk minimization (39). RF is an algorithm that uses an ensemble decision tree, in which random subsets are drawn from the data, with replacement (40). These three algorithms are principally used to identify the most specific DEGs in order to facilitate construction of a superior diagnostic model. In the present study, we identified *DUSP1* and *PRKAR2B* as candidate diagnostic marker genes using the above three algorithms, and to further evaluate their diagnostic efficacy, we used a validation set. The results of this analysis suggested that our integration strategy was reliable.

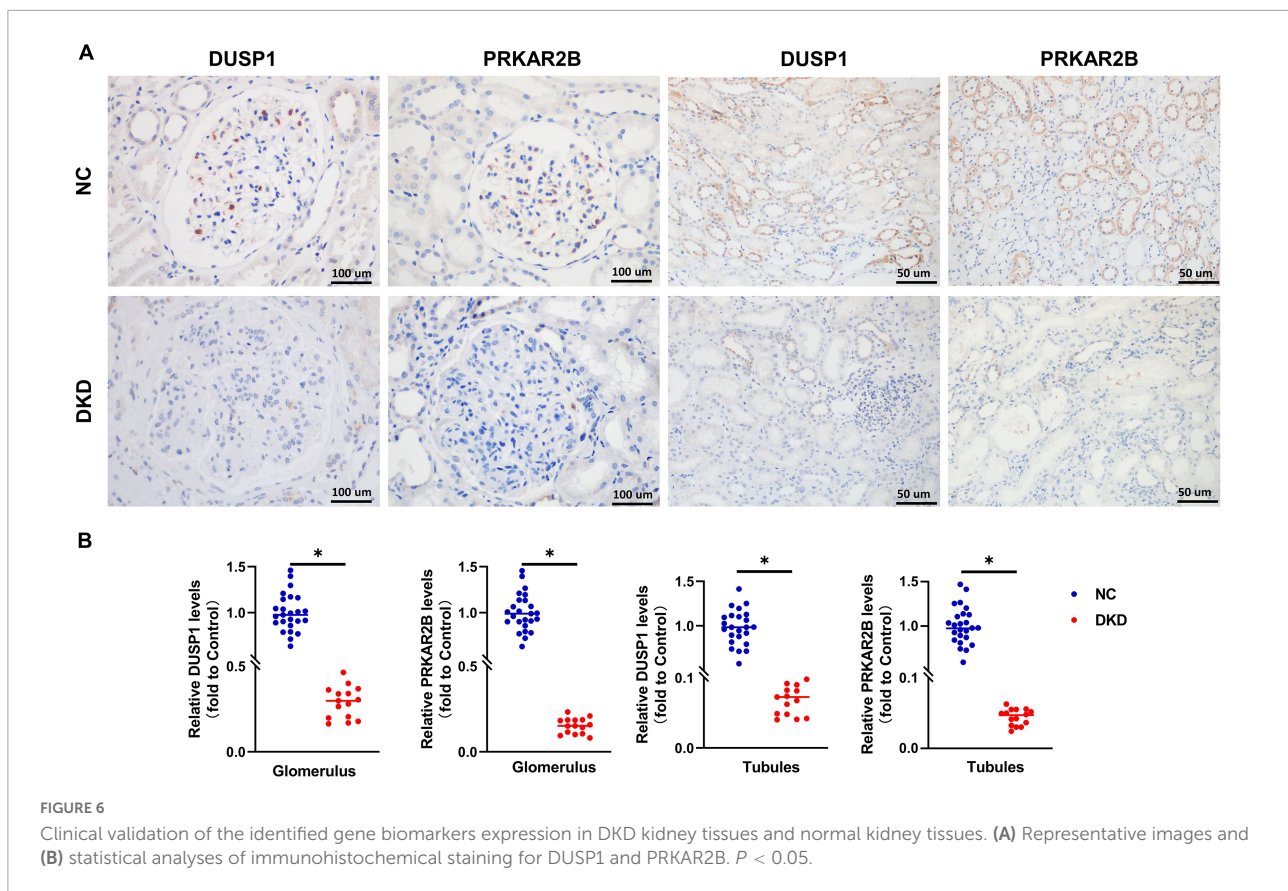
The dual-specificity phosphatase (DUSP) family is the largest group of protein phosphatases that specifically regulate mitogen-activated protein kinase (MAPK) activity in human cells (41). One study of the mechanism of fibrosis in DKD showed that *DUSP1* plays an anti-fibrotic role in the HK-2



proximal tubular cell line by inactivating both the p38MAPK and extracellular signal-related protein kinase 1/2 pathways (42). Angiotensin II-stimulated proteasome activity results in DUSP1 degradation and the subsequent activation of signal transducer and activator of transcription 1 in T cells, leading to the induction of T helper 1 differentiation (43). Moreover, the inhibition of DUSP1 expression or function in T cells may be beneficial for the treatment of T cell-mediated autoimmune diseases such as multiple sclerosis (44). Thus, DUSP1 influences

the activity of immune cells and the process of fibrosis in tissues by regulating MAPK, and it has therefore become a potential therapeutic target in some immune diseases. Therefore, DUSP1 may well influence the pathogenesis of DKD.

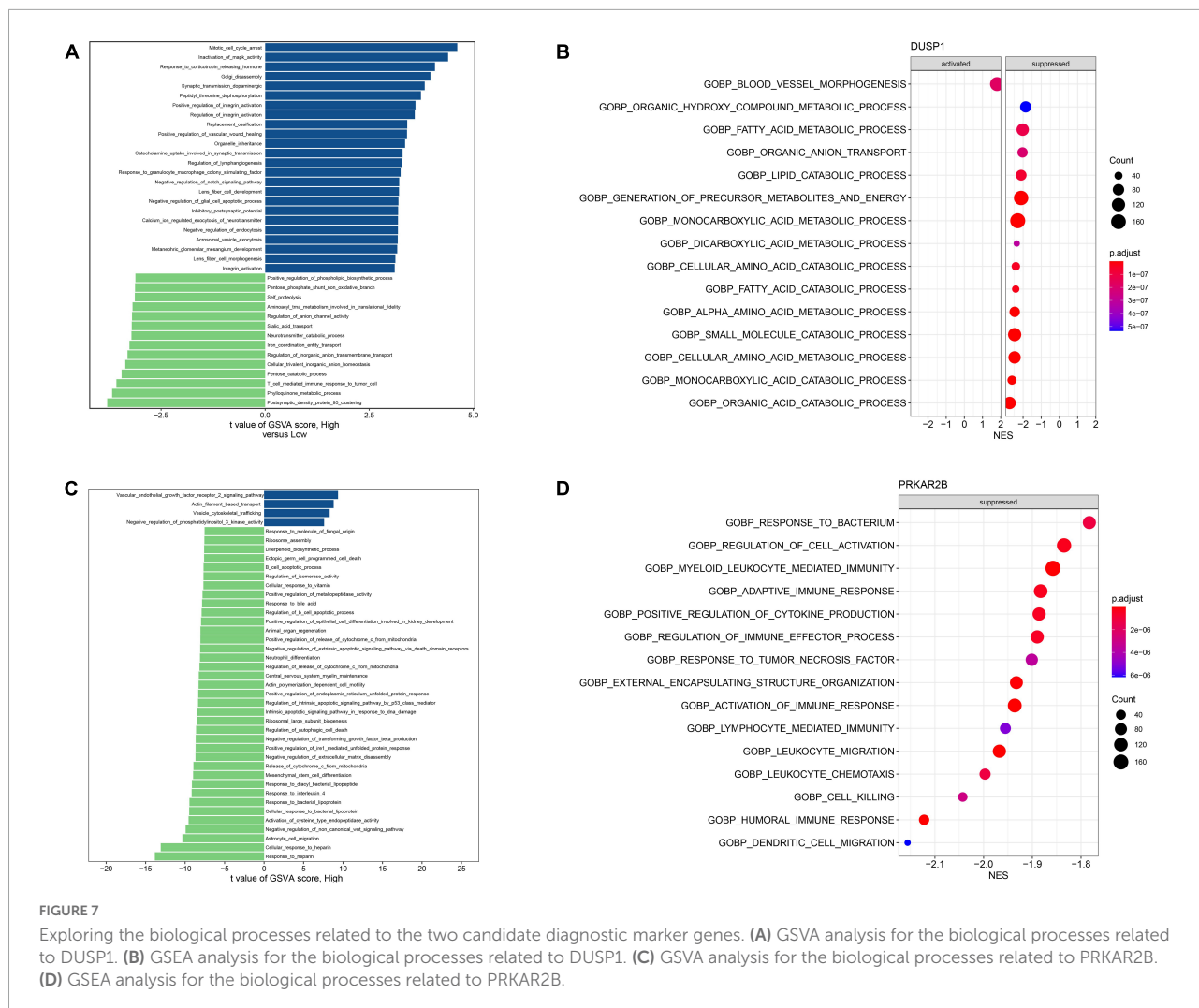
cAMP-dependent protein kinase type II-beta (PRKAR2B) is a member of the protein kinase A (PKA) family that is primarily expressed in the brain and adipose tissue (45) and phosphorylates Ser/Thr residues of target proteins (46). In a previous study, it was shown that a hypoxia-induced reduction



in *PRKAR2B* transcription is sufficient to increase PKA activity *via* hypoxia-inducible factor (HIF)-1 α , without affecting intracellular cAMP concentration (47). In another study, it was shown that *PRKAR2B* increases the expression of HIF-1 α , which provides a growth advantage in prostate cancer by increasing the Warburg effect (48). In DKD, ischemia and inflammation in the glomeruli and vascular lesions reduce the oxygen supply, resulting in an increase in HIF-1 α expression, which assists cells to cope with hypoxia. However, hyperglycemia can reduce the stability of HIF-1 α , which leads to tissue fibrosis (49). Therefore, we speculate that *PRKAR2B* may be involved in the progression of DKD under hypoxic conditions. The results of previous studies have suggested that *DUSP1* and *PRKAR2B* play significant roles in the activities of immune cells and in the progression of disease in tissues (50–52). Therefore, these two genes may represent useful diagnostic markers of DKD, but further clinical studies are needed to confirm the efficacy of their use for diagnostic purposes.

The results of immune cell infiltration found that there was greater infiltration with B memory cells, gamma delta T cells, resting NK cells, macrophages, and resting mast cells, and less infiltration with T follicular helper cells, resting dendritic cells, activated mast cells, and neutrophils, which may be involved in the development and progression of DKD. Current studies generally demonstrated that the mechanism of

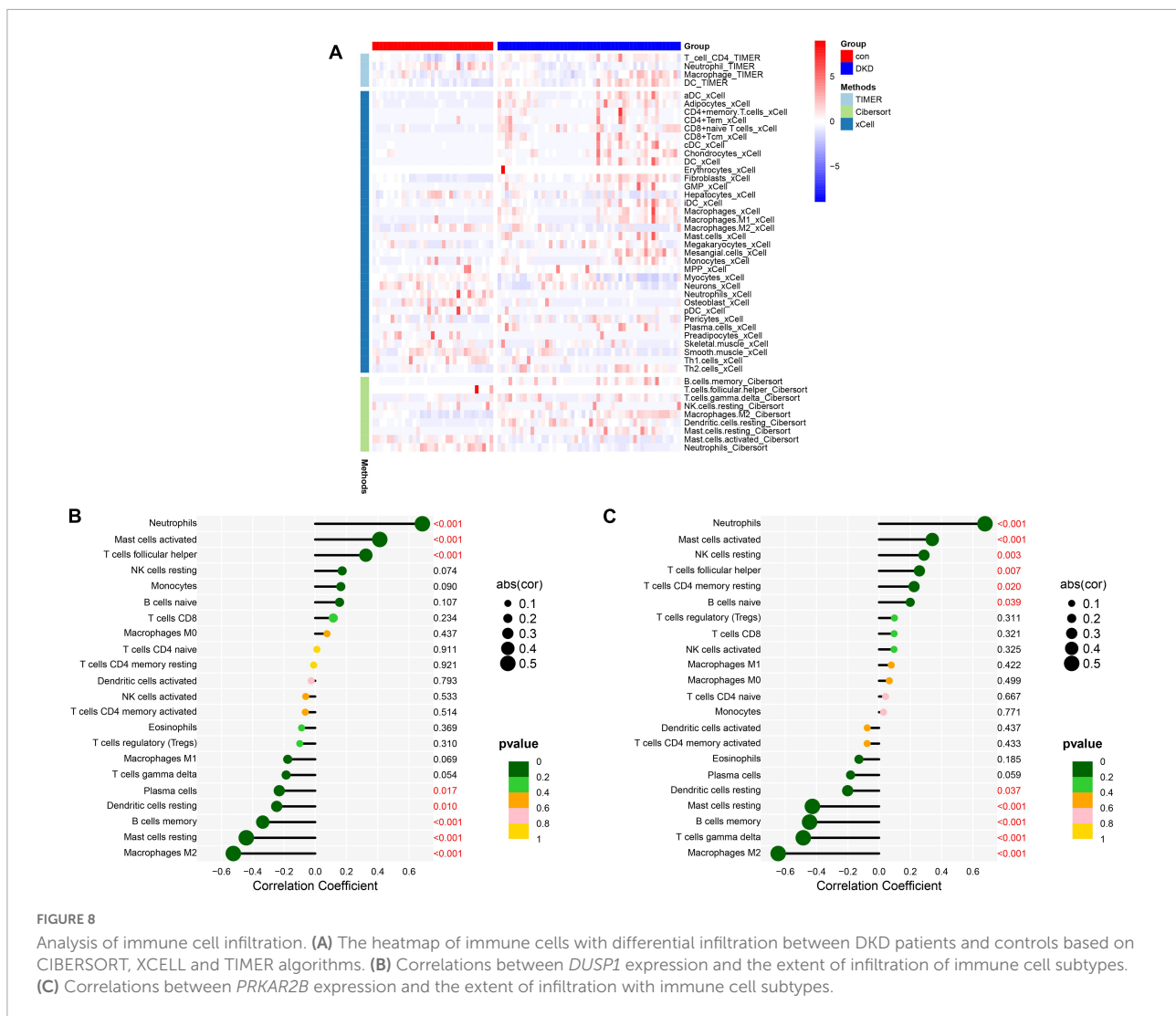
neutrophils in regulation the pathogenesis of DKD is mediated by inflammatory response. A study including 22 renal biopsies of DKD proved that the infiltration of neutrophils increased in interstitial, peritubular, and capillary regions (53). Neutrophil extracellular traps (NETs) are the bactericidal mechanism after neutrophil death. A previous study showed that neutrophils isolated from diabetic humans and mice were primed to produce NETs (54). The increase of Biomarkers of NETs correlates with DKD severity, which NETs promote NLRP3 inflammasome activation and glomerular endothelial dysfunction under high glucose stress *in vitro* and *in vivo* (14). Additionally, neutrophils may also migrate into the kidneys of patients with DKD because their spontaneous adhesion increases (55), which is followed by abnormal activation and the secretion of proinflammatory cytokines, degranulation, and the release of reactive oxygen species (56). To sum up, neutrophils may involve in the occurrence of DKD. In our study, we found that the proportion of neutrophils was relatively lower in DKD samples, which is consistent with the results of Wang et al. on immune infiltration in DKD (57). We suggested that it may be a potential limitation of the current immune infiltration algorithms, because the higher proportion of macrophages in DKD patients makes the proportion of other immune cells, including neutrophils, appear lower (57). The mechanism of other immune cells in regulation the pathogenesis of DKD have also been showed in previous



studies. One previous animal study showed that CD4 + CD8 + T cells and dendritic cells are present in the glomeruli of diabetic mice, along with immunoglobulin (Ig)G and IgG + B cells, which may suggest that the IgG present in the glomeruli may be produced by the infiltrating B cells *in situ* (58). Some features of DKD, such as the high expression of intracellular adhesion molecule 1 and monocyte chemoattractant protein 1 by renal tubular cells, may increase the recruitment of macrophages (59, 60), which then respond to the locally high concentrations of glucose, AGEs, and oxidized low-density lipoprotein by secreting proinflammatory cytokines (61, 62). Eller et al. (63) demonstrated that the depletion of regulatory T cells using an anticluster of differentiation 25 monoclonal antibody accelerates the progression of renal injury, involving increases in glomerular hyperfiltration and albuminuria. One study of streptozotocin-induced diabetes in mice showed that T helper 17 cells may protect diabetic kidneys by reducing and modifying the inflammatory response (64). Mast cells have also been shown to be present in large numbers in the renal

interstitium of patients with DKD, and their number has been shown to correlate with serum creatinine concentration (65). The degranulation of mast cells has been demonstrated in renal biopsies from patients with type 2 diabetes and various stages of nephropathy (66).

An analysis of the correlations between the expression of the two candidate diagnostic marker genes and the extent of infiltration with subtypes of immune cells revealed close negative correlations of *DUSP1* and *PRKAR2B* expression with the extent of M2 macrophages infiltration and positive correlations with neutrophil infiltration. Studies on renal biopsies showed that M1 and M2 macrophages infiltrating in the renal tissue of DKD were increased in a process-dependent manner, and the activation status of M1 and M2 macrophages was positively correlated with the progression of DKD (67). Another previous *in vitro* cell experiment proved that M2 macrophage could improve high glucose-induced podocytes injury via secreting exosomal miR-25-3p to activate autophagy of the cells through suppressing *DUSP1* expression



(68). The above indicates that with the progression of DKD, M2 macrophages act as a protective factor to inhibit the expression of *DUSP1*, thereby ameliorating the renal damage of DKD. Unfortunately, few study revealed the interaction between *PRKAR2B* and M2 macrophages in DKD. Previous studies have shown that neutrophils and mast cells play significant roles in the progression of DKD (55, 56, 66). Macrophages and certain types of T cells may be increasingly recruited during the progression of DKD (58–60). Therefore, we infer that *DUSP1* and *PRKAR2B* may be involved in regulating the infiltration of immune cells in the kidney, and thus influence the development and progression of DKD. However, this hypothesis requires further testing in preclinical and clinical contexts.

We have used novel bioinformatic methods to filter potential diagnostic marker genes and validated them through IHC, and also investigated the relationship between the identified markers and immune cell infiltration. Nevertheless, there were a number of limitations to the present study. First,

clinical information regarding the samples and the contributing patients was not available for the datasets obtained from the GEO database; therefore, the influences of other disease factors on the expression of genes could not be evaluated. Second, our study performed IHC in DKD patients and controls but did not detect the expression of these two biomarkers in the blood samples, the early detection potential of *DUSP1* and *PRKAR2B* of DKD needs to be further explored. Third, the public database we used did not provide follow up information for us to explore relationship of these biomarkers with disease prognosis. Therefore, further prospective clinical studies on the prognostic relevance of *DUSP1* and *PRKAR2B* to DKD need to be conducted in the future.

In conclusion, we have shown that *DUSP1* and *PRKAR2B* may represent potential diagnostic markers of DKD. We have also shown that immune cells such as B memory cells, macrophages, resting mast cells, follicular helper T cells, activated mast cells, and neutrophils may be involved in

the development and progression of DKD. Furthermore, the expression of DUSP1 and PRKAR2B was found to significantly correlate with the numbers of immune cells such as neutrophils and M2 macrophages in the kidneys of such patients. DUSP1 and PRKAR2B may be involved in regulating the infiltration of immune cells in the kidney, and thus influence the development and progression of DKD.

Data availability statement

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

Ethics statement

The studies involving human participants were reviewed and approved by Ethical Institutions of the First Hospital of Jilin University (Changchun, China; approval no. 2022–417). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

SF and YC conceived, designed the research, conducted the bioinformatics analysis and wrote original draft of the manuscript. XW, JH, and JY curated and analyzed the data. SS and HW helped in data collection and manuscript editing. ZX supervised the study and reviewed the manuscript. All authors contributed to the study and approved the final version.

Funding

This work was supported by the National Natural Science Foundation of China (81974094 to ZX and 81700635 to YC),

References

1. Yang W, Lu J, Weng J, Jia W, Ji L, Xiao J, et al. Prevalence of diabetes among men and women in China. *N Engl J Med.* (2010) 362:1090–101. doi: 10.1056/NEJMoa0908292
2. Tuttle KR, Bakris GL, Bilous RW, Chiang JL, de Boer IH, Goldstein-Fuchs J, et al. Diabetic kidney disease: a report from an ADA consensus conference. *Diabetes Care.* (2014) 37:2864–83. doi: 10.2337/dc14-1296
3. Jia W, Gao X, Pang C, Hou X, Bao Y, Liu W, et al. Prevalence and risk factors of albuminuria and chronic kidney disease in Chinese population with type 2 diabetes and impaired glucose regulation: shanghai diabetic complications study (SHDCS). *Nephrol Dial Transplant.* (2009) 24:3724–31. doi: 10.1093/ndt/gf p349
4. Matoba K, Takeda Y, Nagai Y, Kawanami D, Utsunomiya K, Nishimura R. Unraveling the role of inflammation in the pathogenesis of diabetic kidney disease. *Int J Mol Sci.* (2019) 20:3393. doi: 10.3390/ijms20143393
5. Kato M, Natarajan R. Epigenetics and epigenomics in diabetic kidney disease and metabolic memory. *Nat Rev Nephrol.* (2019) 15:327–45. doi: 10.1038/s41581-019-0135-6
6. Chen Y, Lee K, Ni Z, He JC. Diabetic kidney disease: challenges, advances, and opportunities. *Kidney Dis.* (2020) 6:215–25. doi: 10.1159/000506634
7. Berhane AM, Weil EJ, Knowler WC, Nelson RG, Hanson RL. Albuminuria and estimated glomerular filtration rate as predictors of diabetic end-stage renal disease and death. *Clin J Am Soc Nephrol.* (2011) 6:2444–51. doi: 10.2215/cjn.00580111

the Jilin Province Science and Technology Development Project (20200201362JC to ZX), the Postdoctoral Science Foundation of China (2022M711298 to YC), the “PRO-Run” Fund of the Nephrology Group of CEBM (KYJ202206-0003-7 to YC), the Jilin Province Department of Finance (JLSCZD2019-076 to SS), the Natural Science Foundation of Jilin Province (20210101348JC to HW), and Special fund for medical and health talents of Jilin Province (JLSWSRCZX2021 to HW).

Acknowledgments

We thank the contributors to the GEO (<http://www.ncbi.nlm.nih.gov/geo/>) database for sharing data.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2022.918657/full#supplementary-material>

8. Perkins BA, Ficociello LH, Silva KH, Finkelstein DM, Warram JH, Krolewski AS. Regression of microalbuminuria in type 1 diabetes. *N Engl J Med.* (2003) 348:2285–93. doi: 10.1056/NEJMoa021835
9. Krolewski AS, Niewczas MA, Skupien J, Gohda T, Smiles A, Eckfeldt JH, et al. Early progressive renal decline precedes the onset of microalbuminuria and its progression to macroalbuminuria. *Diabetes Care.* (2014) 37:226–34. doi: 10.2337/dcl3-0985
10. Deo RC. Machine learning in medicine. *Circulation.* (2015) 132:1920–30. doi: 10.1161/circulationaha.115.001593
11. Navarro-González JF, Mora-Fernández C. Muros de Fuentes M, García-Pérez J. Inflammatory molecules and pathways in the pathogenesis of diabetic nephropathy. *Nat Rev Nephrol.* (2011) 7:327–40. doi: 10.1038/nrneph.2011.51
12. Hickey FB, Martin F. Role of the immune system in diabetic kidney disease. *Curr Diab Rep.* (2018) 18:20. doi: 10.1007/s11892-018-0984-6
13. Jha JC, Banal C, Chow BS, Cooper ME, Jandeleit-Dahm K. Diabetes and kidney disease: role of oxidative stress. *Antioxid Redox Signal.* (2016) 25:657–84. doi: 10.1089/ars.2016.6664
14. Gupta A, Singh K, Fatima S, Ambreen S, Zimmermann S, Younis R, et al. Neutrophil extracellular traps promote NLRP3 inflammasome activation and glomerular endothelial dysfunction in diabetic kidney disease. *Nutrients.* (2022) 14:2965. doi: 10.3390/nu14142965
15. Jia Y, Xu H, Yu Q, Tan L, Xiong Z. Identification and verification of vascular cell adhesion protein 1 as an immune-related hub gene associated with the tubulointerstitial injury in diabetic kidney disease. *Bioengineered.* (2021) 12:6655–73. doi: 10.1080/21655979.2021.1976540
16. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* (2013) 41:D991–5. doi: 10.1093/nar/gks1193
17. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics.* (2012) 28:882–3. doi: 10.1093/bioinformatics/bts034
18. Liu J, Zhou S, Li S, Jiang Y, Wan Y, Ma X, et al. Eleven genes associated with progression and prognosis of endometrial cancer (EC) identified by comprehensive bioinformatics analysis. *Cancer Cell Int.* (2019) 19:136. doi: 10.1186/s12935-019-0859-1
19. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omic.* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
20. Powers RK, Goodspeed A, Pielke-Lombardo H, Tan AC, Costello JC. GSEA-incontext: identifying novel and common patterns in expression experiments. *Bioinformatics.* (2018) 34:i555–64. doi: 10.1093/bioinformatics/bty271
21. Cheng C, Hua ZC. Lasso peptides: heterologous production and potential medical application. *Front Bioeng Biotechnol.* (2020) 8:571165. doi: 10.3389/fbioe.2020.571165
22. Lin X, Li C, Zhang Y, Su B, Fan M, Wei H. Selecting feature subsets based on SVM-RFE and the overlapping ratio with applications in bioinformatics. *Molecules.* (2017) 23:52. doi: 10.3390/molecules23010052
23. Toth R, Schiffmann H, Hube-Magg C, Büschel F, Höflmayer D, Weidemann S, et al. Random forest-based modelling to detect biomarkers for prostate cancer progression. *Clin Epigenetics.* (2019) 11:148. doi: 10.1186/s13148-019-0736-8
24. Cheng Y, Zhang X, Ma F, Sun W, Wang W, Yu J, et al. The role of Akt2 in the protective effect of fenofibrate against diabetic nephropathy. *Int J Biol Sci.* (2020) 16:553–67. doi: 10.7150/ijbs.40643
25. Yang B, Su Z, Chen G, Zeng Z, Tan J, Wu G, et al. Identification of prognostic biomarkers associated with metastasis and immune infiltration in osteosarcoma. *Oncol Lett.* (2021) 21:180. doi: 10.3892/ol.2021.12441
26. Li S, Zhang Q, Chen Z, Huang Z, Zhang L, Chen F. Novel insight into ferroptosis-related genes, molecular subtypes, and immune characteristics in intracranial aneurysms. *Inflamm Res.* (2022): [Epub ahead of print]. doi: 10.1007/s00011-022-01633-8
27. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* (2013) 14:7. doi: 10.1186/1471-2105-14-7
28. Kawada JI, Takeuchi S, Imai H, Okumura T, Horiba K, Suzuki T, et al. Immune cell infiltration landscapes in pediatric acute myocarditis analyzed by CIBERSORT. *J Cardiol.* (2021) 77:174–8. doi: 10.1016/j.jcc.2020.08.004
29. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* (2017) 18:220. doi: 10.1186/s13059-017-1349-1
30. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* (2020) 48:W509–14. doi: 10.1093/nar/gkaa407
31. Doshi SM, Friedman AN. Diagnosis and management of type 2 diabetic kidney disease. *Clin J Am Soc Nephrol.* (2017) 12:1366–73. doi: 10.2215/cjn.11111016
32. Papadopoulou-Marketou N, Chrousos GP, Kanaka-Gantenbein C. Diabetic nephropathy in type 1 diabetes: a review of early natural history, pathogenesis, and diagnosis. *Diabetes Metab Res Rev.* (2017) 33:e2841. doi: 10.1002/dmrr.2841
33. Xu Q, Li B, Wang Y, Wang C, Feng S, Xue L, et al. Identification of VCAN as hub gene for diabetic kidney disease immune injury using integrated bioinformatics analysis. *Front Physiol.* (2021) 12:651690. doi: 10.3389/fphys.2021.651690
34. Deng YJ, Ren EH, Yuan WH, Zhang GZ, Wu ZL, Xie QQ. GRB10 and E2F3 as diagnostic markers of osteoarthritis and their correlation with immune infiltration. *Diagnostics.* (2020) 10:171. doi: 10.3390/diagnostics10030171
35. He D, Zhang X, Tu J. Diagnostic significance and carcinogenic mechanism of pan-cancer gene POU5F1 in liver hepatocellular carcinoma. *Cancer Med.* (2020) 9:8782–800. doi: 10.1002/cam4.3486
36. Tao Y, Wei X, Yue Y, Wang J, Li J, Shen L, et al. Extracellular vesicle-derived AEBP1 mRNA as a novel candidate biomarker for diabetic kidney disease. *J Transl Med.* (2021) 19:326. doi: 10.1186/s12967-021-03000-3
37. Cui C, Cui Y, Fu Y, Ma S, Zhang S. Microarray analysis reveals gene and microRNA signatures in diabetic kidney disease. *Mol Med Rep.* (2018) 17:2161–8. doi: 10.3892/mmr.2017.8177
38. Garcia-Magariños M, Antoniadis A, Cao R, González-Manteiga W. Lasso logistic regression, GSoft and the cyclic coordinate descent algorithm: application to gene expression data. *Stat Appl Genet Mol Biol.* (2010) 9:Article30. doi: 10.2202/1544-6115.1536
39. Nan S, Sun L, Chen B, Lin Z, Toh KA. Density-dependent quantized least squares support vector machine for large data sets. *IEEE Trans Neural Netw Learn Syst.* (2017) 28:94–106. doi: 10.1109/tnnls.2015.2504382
40. Alderden J, Pepper GA, Wilson A, Whitney JD, Richardson S, Butcher R, et al. Predicting pressure injury in critical care patients: a machine-learning model. *Am J Crit Care.* (2018) 27:461–8. doi: 10.4037/ajcc2018525
41. Caunt CJ, Keyse SM. Dual-specificity MAP kinase phosphatases (MKPs) Shaping the outcome of MAP kinase signalling. *FEBS J.* (2013) 280:489–504. doi: 10.1111/j.1742-4658.2012.08716.x
42. Ge Y, Wang J, Wu D, Zhou Y, Qiu S, Chen J, et al. lncRNA NR_038323 suppresses renal fibrosis in diabetic nephropathy by targeting the miR-324-3p/DUSP1 axis. *Mol Ther Nucleic Acids.* (2019) 17:741–53. doi: 10.1016/j.omtn.2019.07.007
43. Qin XY, Zhang YL, Chi YF, Yan B, Zeng XJ, Li HH, et al. Angiotensin II regulates Th1 T cell differentiation through angiotensin II Type 1 Receptor-PKA-Mediated activation of proteasome. *Cell Physiol Biochem.* (2018) 45:1366–76. doi: 10.1159/000487562
44. Zhang Y, Reynolds JM, Chang SH, Martin-Orozco N, Chung Y, Nurieva RI, et al. MKP-1 is necessary for T cell activation and function. *J Biol Chem.* (2009) 284:30815–24. doi: 10.1074/jbc.M109.052472
45. Enns LC, Morton JF, Treuting PR, Emond MJ, Wolf NS, Dai DF, et al. Disruption of protein kinase A in mice enhances healthy aging. *PLoS One.* (2009) 4:e5963. doi: 10.1371/journal.pone.0005963
46. Taylor SS, Ilouz R, Zhang P, Kornev AP. Assembly of allosteric macromolecular switches: lessons from PKA. *Nat Rev Mol Cell Biol.* (2012) 13:646–58. doi: 10.1038/nrm3432
47. Lucia K, Wu Y, Garcia JM, Barlier A, Buchfelder M, Saeger W, et al. Hypoxia and the hypoxia inducible factor 1 α activate protein kinase A by repressing RII beta subunit transcription. *Oncogene.* (2020) 39:3367–80. doi: 10.1038/s41388-020-1223-6
48. Xia L, Sun J, Xie S, Chi C, Zhu Y, Pan J, et al. PRKAR2B-HIF-1 α loop promotes aerobic glycolysis and tumour growth in prostate cancer. *Cell Prolif.* (2020) 53:e12918. doi: 10.1111/cpr.12918
49. Lin YC, Chang YH, Yang SY, Wu KD, Chu TS. Update of pathophysiology and management of diabetic kidney disease. *J Formos Med Assoc.* (2018) 117:662–75. doi: 10.1016/j.jfma.2018.02.007
50. Goel S, Saheb Sharif-Askari F, Saheb Sharif Askari N, Madkhana B, Alwaa AM, Mahboub B, et al. SARS-CoV-2 Switches ‘on’ MAPK and NF κ B signaling via the reduction of nuclear DUSP1 and DUSP5 expression. *Front Pharmacol.* (2021) 12:631879. doi: 10.3389/fphar.2021.631879
51. Chuang HC, Tan TH. MAP4K family kinases and DUSP family phosphatases in T-cell signaling and systemic lupus erythematosus. *Cells.* (2019) 8:1433. doi: 10.3390/cells8111433

52. Elliott MR, Tolnay M, Tsokos GC, Kammer GM. Protein kinase A regulatory subunit type II beta directly interacts with and suppresses CREB transcriptional activity in activated T cells. *J Immunol.* (2003) 171:3636–44. doi: 10.4049/jimmunol.171.7.3636
53. Kelly KJ, Dominguez JH. Rapid progression of diabetic nephropathy is linked to inflammation and episodes of acute renal failure. *Am J Nephrol.* (2010) 32:469–75. doi: 10.1159/000320749
54. Wong SL, Demers M, Martinod K, Gallant M, Wang Y, Goldfine AB, et al. Diabetes primes neutrophils to undergo NETosis, which impairs wound healing. *Nat Med.* (2015) 21:815–9. doi: 10.1038/nm.3887
55. Wierusz-Wysocka B, Wykretowicz A, Byks H, Sadurska K, Wysocki H. Polymorphonuclear neutrophils adherence, superoxide anion (O₂⁻) production and HBA1 level in diabetic patients. *Diabetes Res Clin Pract.* (1993) 21:109–14. doi: 10.1016/0168-8227(93)90057-c
56. Huang J, Xiao Y, Xu A, Zhou Z. Neutrophils in type 1 diabetes. *J Diabetes Investig.* (2016) 7:652–63. doi: 10.1111/jdi.12469
57. Wang Y, Zhao M, Zhang Y. Identification of fibronectin 1 (FN1) and complement component 3 (C3) as immune infiltration-related biomarkers for diabetic nephropathy using integrated bioinformatic analysis. *Bioengineered.* (2021) 12:5386–401. doi: 10.1080/21655979.2021.1960766
58. Xiao X, Ma B, Dong B, Zhao P, Tai N, Chen L, et al. Cellular and humoral immune responses in the early stages of diabetic nephropathy in NOD mice. *J Autoimmun.* (2009) 32:85–93. doi: 10.1016/j.jaut.2008.12.003
59. Chow FY, Nikolic-Paterson DJ, Ozols E, Atkins RC, Tesch GH. Intercellular adhesion molecule-1 deficiency is protective against nephropathy in type 2 diabetic db/db mice. *J Am Soc Nephrol.* (2005) 16:1711–22. doi: 10.1681/asn.2004070612
60. Chow FY, Nikolic-Paterson DJ, Ma FY, Ozols E, Rollins BJ, Tesch GH. Monocyte chemoattractant protein-1-induced tissue inflammation is critical for the development of renal injury but not type 2 diabetes in obese db/db mice. *Diabetologia.* (2007) 50:471–80. doi: 10.1007/s00125-006-0497-8
61. Webster L, Abordo EA, Thornalley PJ, Limb GA. Induction of TNF alpha and IL-1 beta mRNA in monocytes by methylglyoxal- and advanced glycated endproduct-modified human serum albumin. *Biochem Soc Trans.* (1997) 25:250S. doi: 10.1042/bst025250s
62. Cipollone F, Iezzi A, Fazio M, Zucchelli M, Pini B, Cuccurullo C, et al. The receptor RAGE as a progression factor amplifying arachidonate-dependent inflammatory and proteolytic response in human atherosclerotic plaques: role of glycemic control. *Circulation.* (2003) 108:1070–7. doi: 10.1161/01.Cir.0000086014.80477.0d
63. Eller K, Kirsch A, Wolf AM, Sopper S, Tagwerker A, Stanzl U, et al. Potential role of regulatory T cells in reversing obesity-linked insulin resistance and diabetic nephropathy. *Diabetes.* (2011) 60:2954–62. doi: 10.2337/db11-0358
64. Kim SM, Lee SH, Lee A, Kim DJ, Kim YG, Kim SY, et al. Targeting T helper 17 by mycophenolate mofetil attenuates diabetic nephropathy progression. *Transl Res.* (2015) 166:375–83. doi: 10.1016/j.trsl.2015.04.013
65. Okoń K, Stachura J. Increased mast cell density in renal interstitium is correlated with relative interstitial volume, serum creatinine and urea especially in diabetic nephropathy but also in primary glomerulonephritis. *Pol J Pathol.* (2007) 58:193–7.
66. Zheng JM, Yao GH, Cheng Z, Wang R, Liu ZH. Pathogenic role of mast cells in the development of diabetic nephropathy: a study of patients at different stages of the disease. *Diabetologia.* (2012) 55:801–11. doi: 10.1007/s00125-011-2391-2
67. Zhang X, Yang Y, Zhao Y. Macrophage phenotype and its relationship with renal function in human diabetic nephropathy. *PLoS One.* (2019) 14:e0221991. doi: 10.1371/journal.pone.0221991
68. Huang H, Liu H, Tang J, Xu W, Gan H, Fan Q, et al. M2 macrophage-derived exosomal miR-25-3p improves high glucose-induced podocytes injury through activation autophagy via inhibiting DUSP1 expression. *IUBMB Life.* (2020) 72:2651–62. doi: 10.1002/iub.2393