

FN1 and TGFBI are key biomarkers of macrophage immune injury in diabetic kidney disease

Fulin Dou, MD^a, Qingzhen Liu, MD, PhD^a, Shasha Lv, MD, PhD^a, Qiaoying Xu, MD^a, Xueling Wang, MD, PhD^a, Shanshan Liu, MD, PhD^{a,*}, Gang Liu, MD, PhD^{a,b}

Abstract

The pathogenesis of diabetic kidney disease (DKD) is complex, and the existing treatment methods cannot control disease progression well. Macrophages play an important role in the development of DKD. This study aimed to search for biomarkers involved in immune injury induced by macrophages in DKD. The GSE96804 dataset was downloaded and analyzed by the CIBERSORT algorithm to understand the differential infiltration of macrophages between DKD and normal controls. Weighted gene co-expression network analysis was used to explore the correlation between gene expression modules and macrophages in renal tissue of DKD patients. Protein-protein interaction network and machine learning algorithm were used to screen the hub genes in the key modules. Subsequently, the GSE30528 dataset was used to further validate the expression of hub genes and analyze the diagnostic effect by the receiver operating characteristic curve. The clinical data were applied to explore the prognostic significance of hub genes. CIBERSORT analysis showed that macrophages increased significantly in DKD renal tissue samples. A total of ten modules were generated by weighted gene co-expression network analysis, of which the blue module was closely associated with macrophages. The blue module mainly played an important role in biological processes such as immune response and fibrosis. Fibronectin 1 (FN1) and transforming growth factor beta induced (TGFBI) were identified as hub genes of DKD patients. Receiver operating characteristic curve analysis was performed in the test cohort: FN1 and TGFBI had larger area under the curve values (0.99 and 0.88, respectively). Clinical validation showed that 2 hub genes were negatively correlated with the estimated glomerular filtration rate in DKD patients. In addition, FN1 and TGFBI showed a strong positive correlation with macrophage alternative activation. FN1 and TGFBI are promising biomarkers for the diagnosis and treatment of DKD patients, which may participate in immune response and fibrosis induced by macrophages.

Abbreviations: AUC = area under the curve, DEGs = differentially expressed genes, DKD = diabetic kidney disease, ECM = extracellular matrix, eGFR = estimated glomerular filtration rate, FN1 = fibronectin 1, GEO = gene expression omnibus, GO = gene ontology, KEGG = Kyoto encyclopedia of genes and genomes, LASSO = least absolute shrinkage and selection operator, M1 = macrophage classical activation, M2 = macrophage alternative activation, PPI = protein-protein interaction, RF = random forest, RFE = recursive feature elimination, ROC = receiver operating characteristic, SVM = support vector machines, TGFBI = transforming growth factor beta induced, WGCNA = weighted gene co-expression network analysis.

Keywords: bioinformatics, diabetic kidney disease, immune injury, machine learning algorithm, macrophages, WGCNA

1. Introduction

Diabetic kidney disease (DKD) is the major cause of end-stage renal disease,^[1] which brings a heavy economic burden to patients and society. At present, there is no specific therapy for DKD. The existing treatment strategies aim to control blood glucose and blood pressure levels and inhibit the renin-angiotensin system. However, it has been proved that this approach can only delay the disease progression, but cannot prevent or reverse the disease.^[2] Furthermore, due to the genetic heterogeneity and complexity of the disease, not all patients benefit

from these drugs.^[3] Therefore, it is urgent to improve the understanding of the pathogenesis of DKD to identify new potential diagnostic and therapeutic targets.

The immune response is involved in the pathogenesis and progress of DKD.^[4] In renal biopsy specimens of DKD patients at various stages, the infiltration of immune cells and increased expression of inflammatory mediators have been observed, among which macrophages are the most important immune cells.^[5] The accumulation of macrophages in the kidney is closely related to urinary albumin excretion, estimated glomerular filtration rate (eGFR), and interstitial fibrosis score.

This work is supported by the Clinical Research Center of Shandong University (No. 2020 SDUCRC A004).

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are publicly available.

^a Department of Nephrology, The Second Hospital of Shandong University, Jinan, China, ^b Nephrology Research Institute of Shandong University, The Second Hospital of Shandong University, Jinan, China.

*Correspondence: Shanshan Liu, Department of Nephrology, The Second Hospital of Shandong University, Jinan, 250033, China (e-mail: liushanshan630@126.com).

Copyright © 2023 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Dou F, Liu Q, Lv S, Xu Q, Wang X, Liu S, Liu G. FN1 and TGFBI are key biomarkers of macrophage immune injury in diabetic kidney disease. *Medicine* 2023;102:45(e35794).

Received: 7 March 2023 / Received in final form: 3 October 2023 / Accepted: 4 October 2023

<http://dx.doi.org/10.1097/MD.00000000000035794>

A body of evidence supports that^[6,7] macrophages induce renal injury through interacting with resident renal cells or being activated by components of the diabetic environment, resulting in the production of a large number of proinflammatory and profibrogenic factors. In addition, due to the plasticity of macrophages, they could acquire different phenotypes, namely classical activation (M1) or alternative activation (M2). The function of M1 is to promote immune inflammation leading to tissue damage, while M2 is primarily involved in tissue healing and inflammation regression, but some M2 may become to promote fibrosis.^[8] Little is known about how M2 terminates the repair response and starts fibrosis. Depletion of macrophages significantly reduces proteinuria and glomerulopathy in diabetic mice.^[9] Although many preclinical studies have shown that^[10,11] targeted therapy of the innate immune pathway in DKD may have a renoprotective effect, nonspecific anti-inflammatory therapy for DKD may increase the susceptibility to infection and it is not suitable for the DKD population with immune dysfunction.^[12] Therefore, we should actively explore biomarkers related to macrophage infiltration activation and phenotypic regulation of DKD to provide targets for more specific and less toxic treatment in the future.

To explore the key biomarkers involved in the immune damage of DKD caused by macrophages, our study first applied the CIBERSORT algorithm to confirm that macrophage infiltration in DKD samples increased significantly. Afterward, the CIBERSORT algorithm was combined with weighted gene co-expression network analysis (WGCNA) to find the gene co-expression modules highly related to macrophages. Protein-protein interaction (PPI) analysis and machine learning algorithm were performed on the genes in the module to identify the 2 hub genes in DKD. Finally, the GSE30528 dataset from the gene expression omnibus (GEO) database was used to validate the good diagnostic efficacy and prognostic value of hub genes (Fig. 1).

In this study, we aimed to identify potential biomarkers that modulate macrophages' exacerbation of immune damage in DKD by integrating immune infiltration analysis, WGCNA and

machine learning algorithms, providing a new viewpoint for the diagnosis and treatment of DKD.

2. Materials and methods

2.1. Data source and analysis

GSE96804^[13] and GSE30528^[14] were obtained from the GEO database. GSE96804 and GSE30528 were annotated based on the GPL17586 and GPL571 platform files, respectively. Among them, GSE96804 was used as the training dataset, including 41 DKD samples and 20 normal samples, and GSE30528 was used as the validation, including 9 DKD samples and 13 normal samples. All these samples were from human glomerular tissue and were expression profiling by array. The “arrayqualitymetrics” package^[15] in R was used to evaluate the training dataset and remove outliers, followed by the “oligo” package^[16] to standardize the dataset. “Limma” package^[17] in R was used to screen differentially expressed genes (DEGs) between DKD and normal samples with $P < .05$, $|\text{Log}_2$ fold change >1 . The results were visualized as volcano plot and heatmap by “ggplot2”,^[18] “pheatmap” package.^[19]

2.2. Immune cell infiltration analysis

CIBERSORT method^[20] in R was used to further analyze the gene expression matrix of GSE96804, and to explore the differences in the composition of 22 immune cells between DKD and normal kidney tissue samples. The results were visualized by the “ggplot2”,^[18] “ggpubr” and “pheatmap” packages^[19] in R.

2.3. Construction of gene co-expression network module and identification of key modules

To screen the potential genes associated with DKD, we first ranked the genes by the median absolute deviation and included the top 5000 genes. Afterward, the WGCNA package^[21] in R

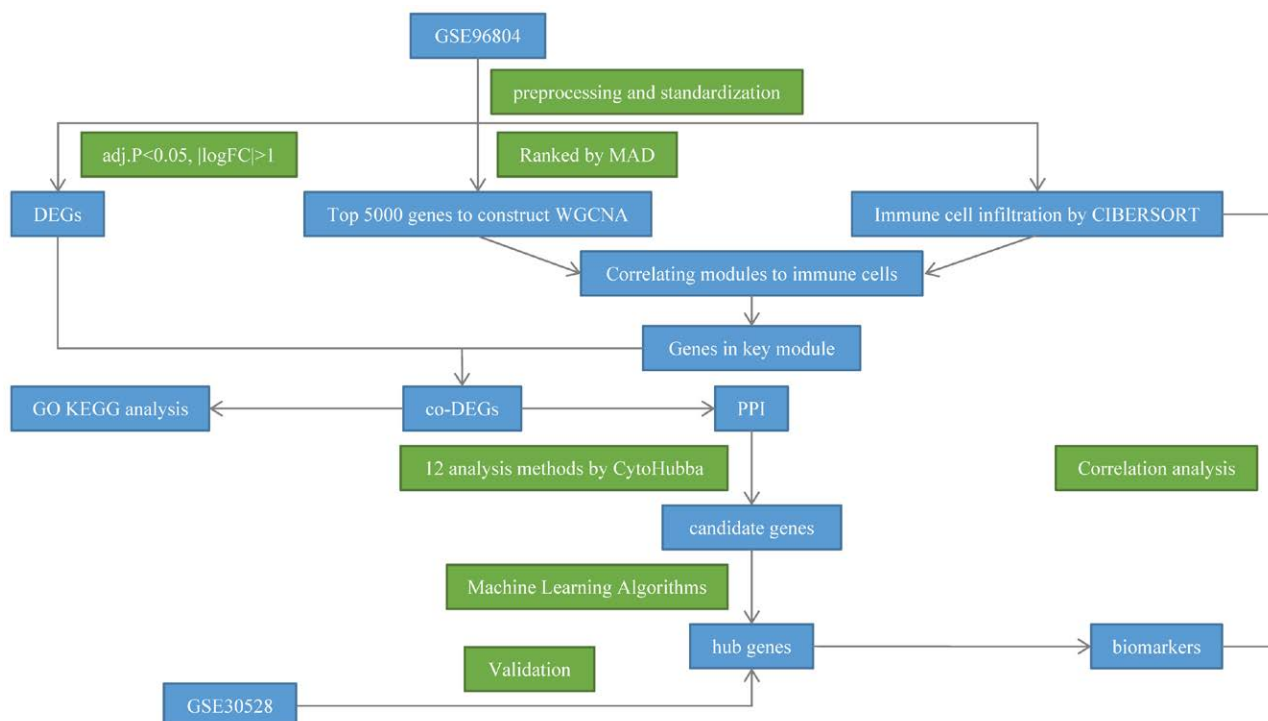


Figure 1. Study flowchart (DEGs = differentially expressed genes. MAD = median absolute deviation; WGCNA = weighted gene co-expression network analysis; GO = Gene Ontology; KEGG = Kyoto Encyclopedia of Genes and Genomes; PPI = protein-protein interaction).

was applied to construct the co-expression network modules by merging modules with the degree of dissimilarity <0.25 and setting the minimum number of genes in the module to 30. Then, we explored the correlation between gene expression modules and DKD, monocytes/macrophages. Ultimately, a key module was confirmed based on the *P* value and the correlation coefficients in macrophage. The overlapping genes of key module and DEGs were selected as target genes for subsequent analysis.

2.4. Enrichment analysis of the co-DEGs

In this study, DAVID (<https://david.ncifcrf.gov/>)^[22] tool was used for gene ontology (GO) functional enrichment analysis and Kyoto encyclopedia of genes and genomes (KEGG) pathway analysis for co-DEGs. The number of enrichment genes (count number) ≥4 and *P* value < .05 were chosen as cutoff criteria.

2.5. Construction of protein interaction networks and identification of candidate genes

The STRING software^[23] (version 10.0; <https://string-db.org/>) was applied to construct the PPI networks of co-DEGs, where the comprehensive score threshold was >0.4. Cytoscape software^[24] was used to visualize the above results and the plug-in cytohubba was applied to explore important nodes in biological networks. We ranked all nodes by the 12 topological analysis methods provided by CytoHubba. Each algorithm computed all node scores, and then 1 to 50 points were assigned based on the rank. According to all points, we selected the top 20 genes as candidate genes.

2.6. Identification of hub genes for DKD based on Machine Learning Algorithms

First, we used the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm^[25] by the “glmnet” R package^[26] to screen these candidates for potential genes. Then, we

applied the support vector machines (SVM)-recursive feature elimination (RFE) algorithm^[27] and the random forest (RF) algorithm to filter these candidates again based on the “e1071” R package^[28] and the “randomforest” R package,^[29] respectively. Finally, overlapping genes among potential genes generated via LASSO, SVM-RFE and RF algorithms were considered hub genes in the DKD.

2.7. Diagnostic efficacy and prognostic risk assessment of hub genes

First, we downloaded the GSE30528 dataset from the GEO database for cross-validation, compared the log2 transformation mRNA expression level of hub genes between the DKD and the control group, and used receiver operator characteristics (ROC) analysis to evaluate the discrimination ability of hub genes to DKD by calculating the area under the curve (AUC). The AUC value was calculated by the “pROC” package.^[30] After that, the clinical data of the validation dataset GSE30528 was obtained from the “Nephroseq” online platform (<http://v5.nephroseq.org>) to explore the Pearson correlation between the mRNA expression levels of hub genes and eGFR values. Finally, we analyzed the correlation between hub genes and 22 immune cell types.

3. Results

3.1. Identification of DEGs

The outlier sample GSM2544307 was deleted from the GSE96804 dataset and finally, 40 cases of DKD and 20 cases of normal were included. A volcano plot was used to show 490 DEGs screened by GSE96804, of which 176 up-regulated genes and 314 down-regulated genes. We used heatmap to show the DEGs of the top 20 up and down (Fig. 2A and B).

3.2. Immune cell infiltration characteristics

Our results more intuitively showed the proportion of various immune cell compositions in each sample (Fig. 3A), as well

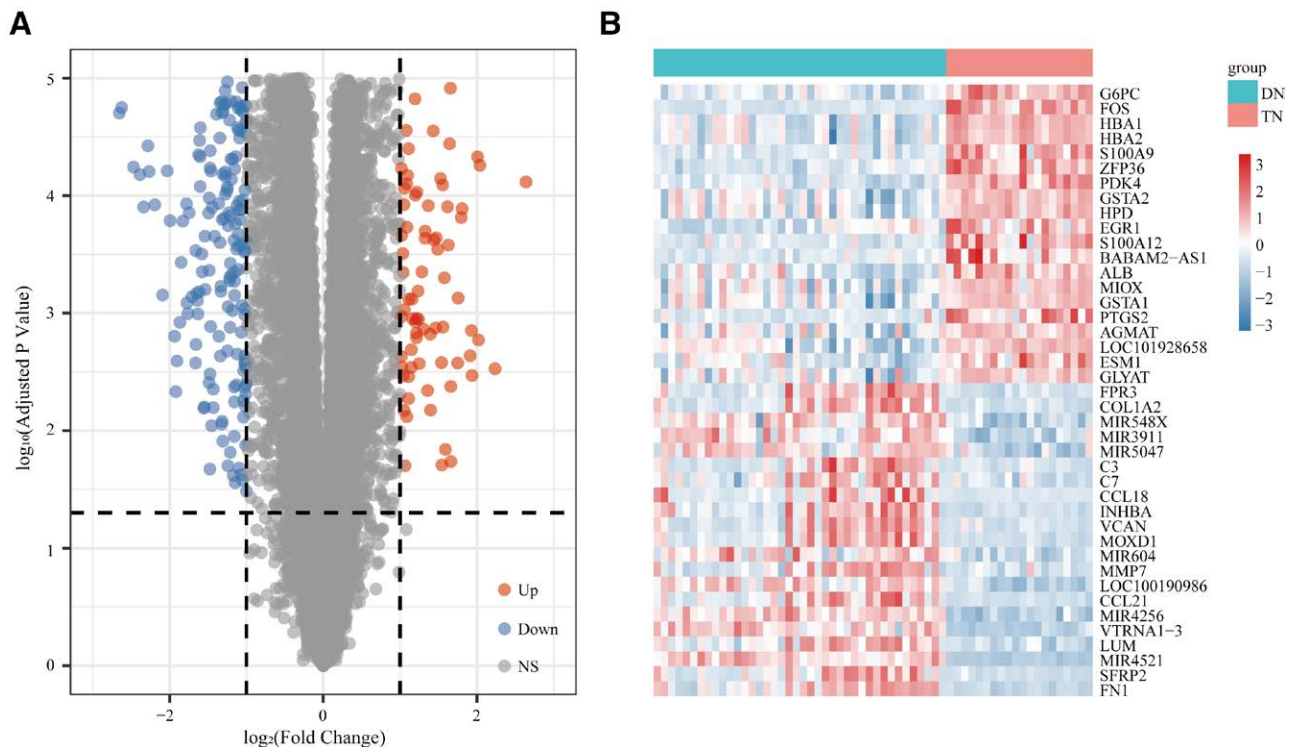


Figure 2. Identification of differentially expressed genes (DEGs) in the GSE96804 dataset. (A) Volcano plot and (B) heatmap of DEGs in GSE96804.

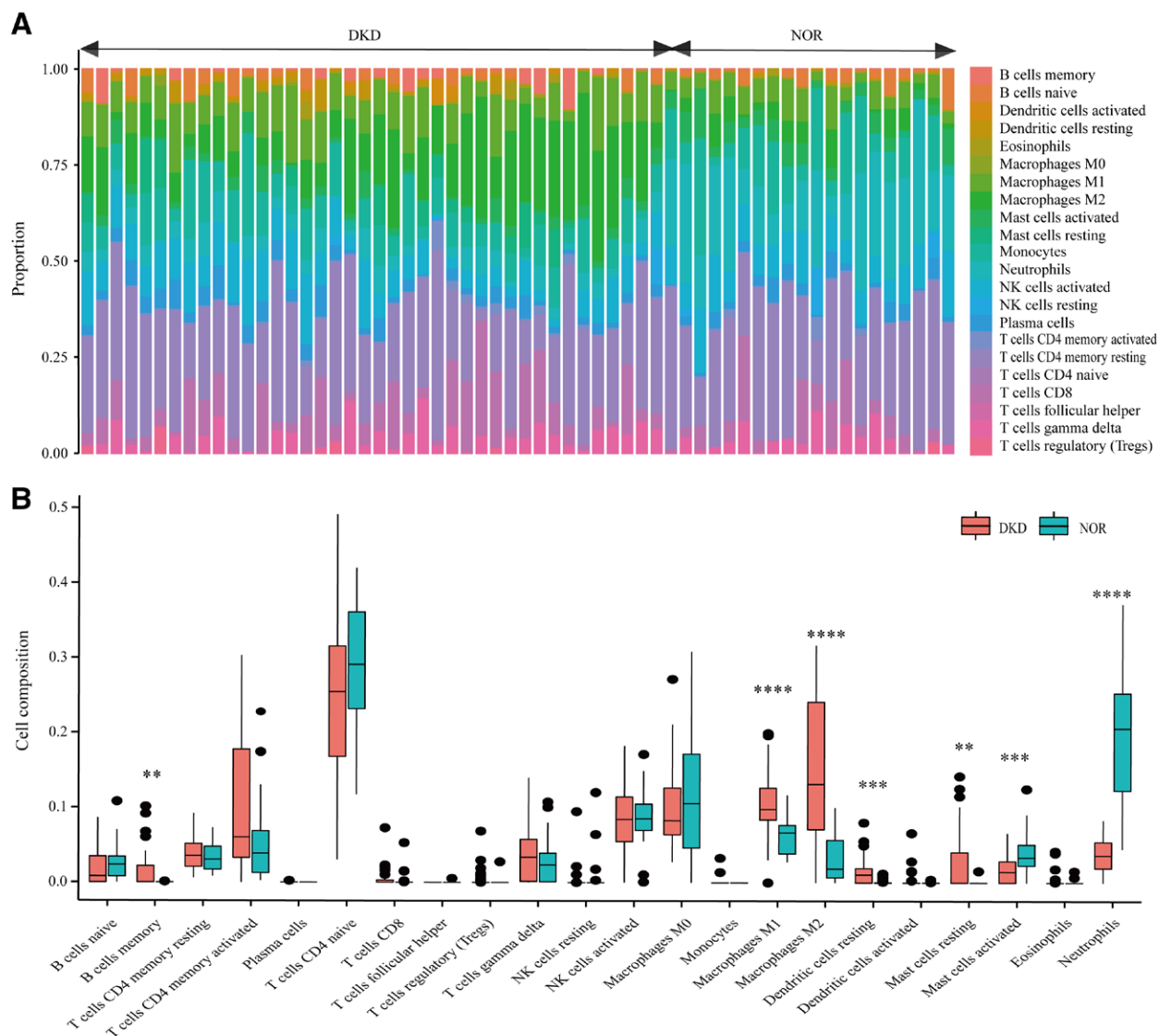


Figure 3. Immune cell infiltration differs between DKD and normal tissues. (A) Proportions of immune cell subsets in DKD and normal groups. (B) Statistical differences in immune cell subsets in DKD and normal groups. * $P < .05$, ** $P < .01$, *** $P < .001$, **** $P < .0001$. DKD = diabetic kidney disease.

as the immune cell infiltration with a significant difference between the 2 groups ($P < .05$). M1, M2, memory B cells, and resting mast cells increased in the DKD group, while activating mast cells and neutrophils increased in the normal group (Fig. 3B).

3.3. Co-expression network and key modules

The clustering analysis of all samples showed that GSM2544295 and GSM2544309 samples were poorly clustered, so the above outlier samples were removed in the subsequent WGCNA analysis (Fig. 4A). Afterward, the network topology of 1 to 30 threshold weights was analyzed, and the soft threshold parameter $\beta = 12$ ($R_2 > 0.8$) was selected to ensure a scale-free network (Fig. 4B). Subsequently, we screened 10 co-expression network modules, of which the gray module was composed of genes excluded from the other 9 modules (Fig. 4C). We analyzed the correlation of 10 modules with DKD and monocytes/macrophages, and then plotted the network heatmap. The result showed that the blue

module was considered a key module, indicating a close correlation with DKD and macrophages M2 (Fig. 4D). 79 DEGs in the blue module, namely co-DEGs, were regarded as target genes for subsequent analysis (Fig. 4E).

3.4. Biofunctional and disease enrichment of genes in key modules

The biological function of co-DEGs in the blue module was studied by GO and KEGG enrichment analysis. The biological process of GO analysis showed that (Fig. 5A) target genes were mainly involved in cell adhesion, extractable matrix organization, collagen fiber organization, immune response, monocyte chemotaxis, and cellular response to tumor necrosis factor. In addition, KEGG analysis showed that (Fig. 5B) target genes were mainly enriched in immune response and fibrosis, such as completion and coagulation cascades, extracellular matrix (ECM)-receptor interaction, PI3K-Akt, and TGF- β signaling pathways.

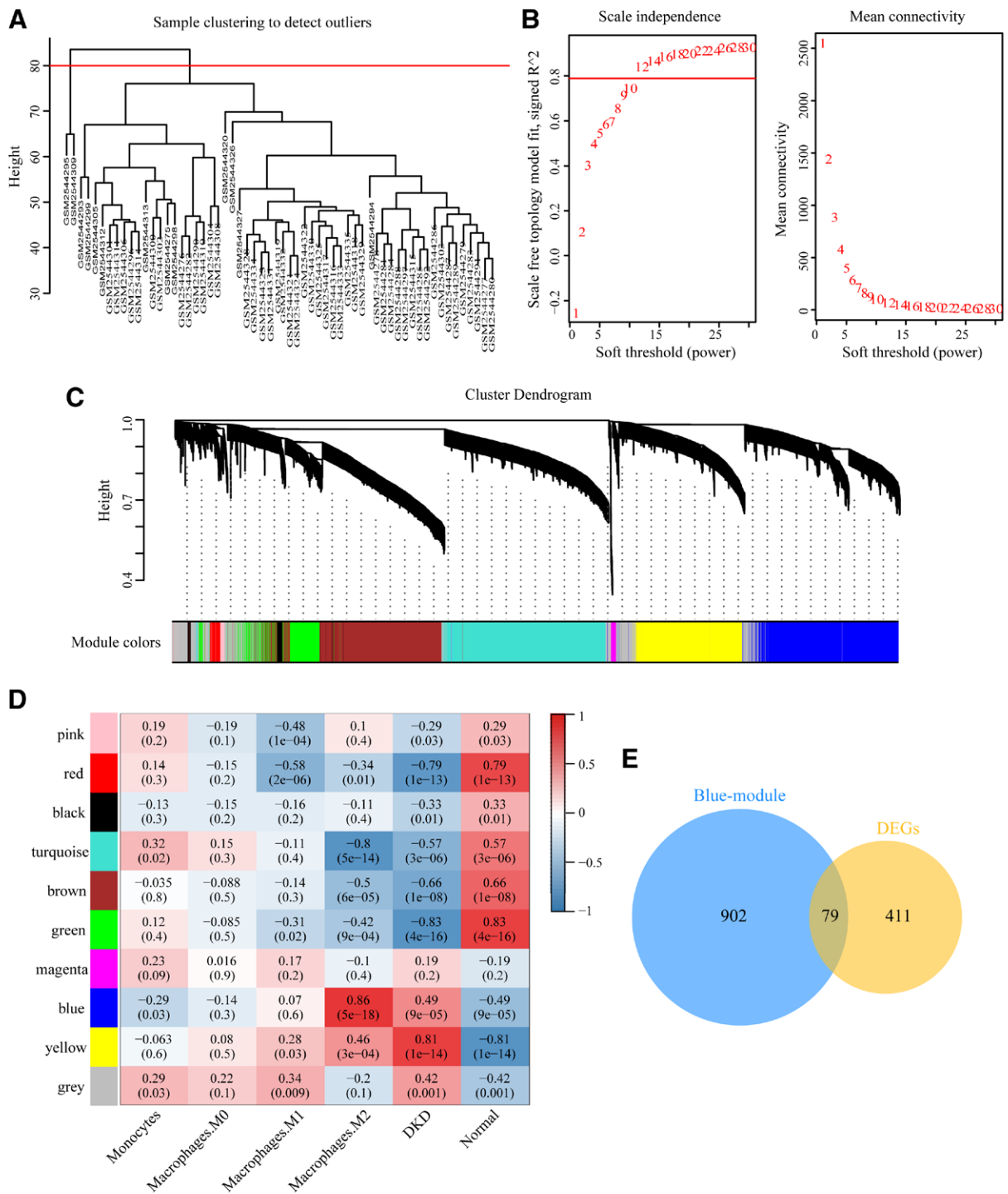


Figure 4. Construction of the WGCNA and identification of key module. (A) Sample clustering diagram (set height to 80 and remove 2 outliers). (B) Determine the optimal soft threshold powers according to the scale free topology model fit (R^2) and mean connectivity, $\beta = 12$. (C) Clustering tree of coexpressive gene modules. Each branch represents 1 gene, and each color below represents 1 co-expression module. (D) Correlation between modules and DKD, monocytes/macrophages. Correlation coefficient and P value are marked in each grid. (E) Venn diagram shows the overlapping genes of differentially expressed genes (DEGs) and blue module, co-DEGs. WGCNA = weighted gene co-expression network analysis.

3.5. PPI network construction and candidate genes screening

We constructed a PPI network to study the protein interaction of co-DEGs in the blue module. As shown in Figure 6A, the PPI network consisted of 52 nodes and 223 edges. The scores of all nodes for each algorithm were calculated by 12 topological analysis methods in CytoHubba and ranked from large to small, and then 1 to 50 points were allocated based on the ranking.

According to the total score of each node, the top 20 nodes were selected as candidate genes, as shown in Figure 6B.

3.6. Hub-genes identification

We chose to apply machine learning algorithms to identify potential biomarkers of DKD from 20 candidates. First, the

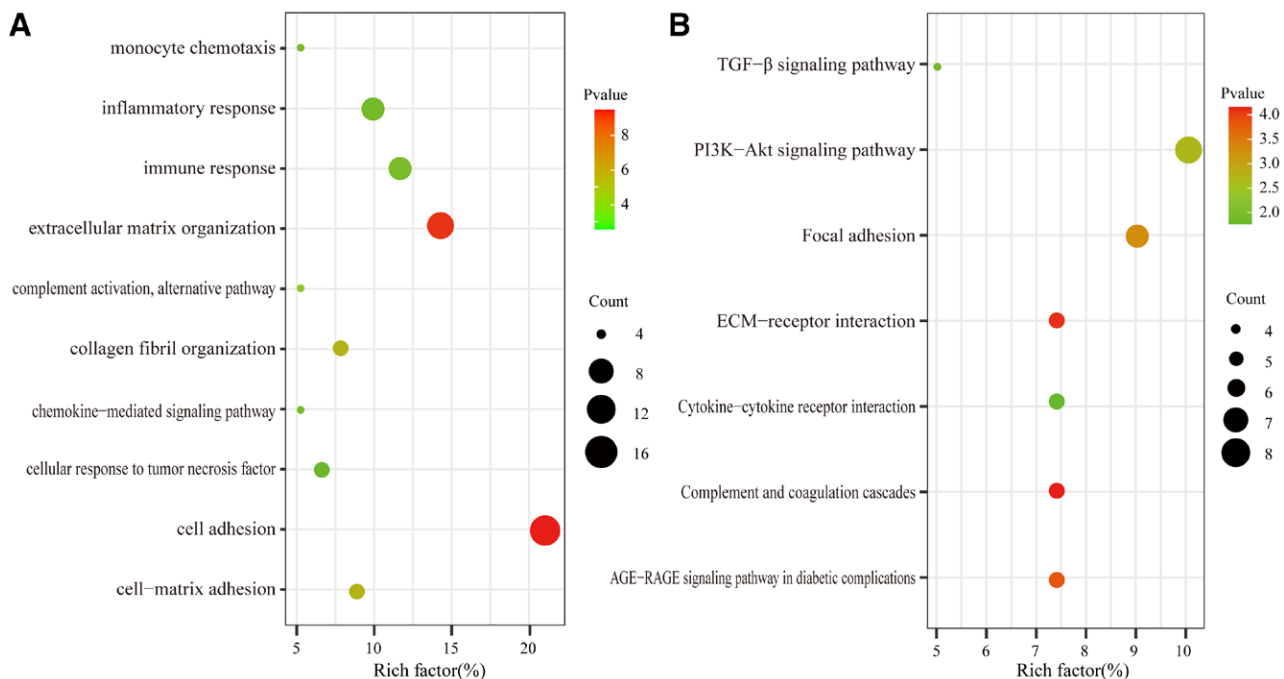


Figure 5. GO and KEGG enrichment analysis of co-DEGs. The color depth of the nodes refers to the *P* value. The size of the nodes refers to the number of genes (A) GO biological process (B) KEGG pathway. DEGs = differentially expressed genes, GO = gene ontology, KEGG = Kyoto encyclopedia of genes and genomes.

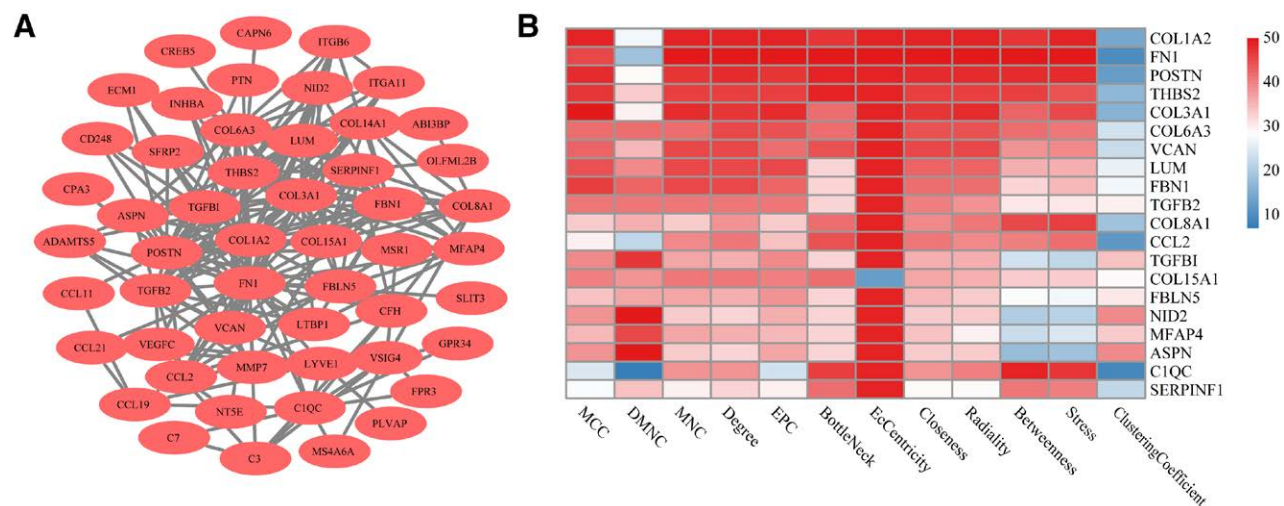


Figure 6. PPI network and identification of candidate genes. (A) PPI network of co-DEGs. (B) Heatmap of the CytoHubba analysis score. DEGs = differentially expressed genes, PPI = protein-protein interaction.

LASSO model was built based on DKD and control samples, with $\lambda = 0.016$. Thus, ASPN, C1QC, COL15A1, COL6A3, fibronectin 1 (FN1), LUM, POSTN, SERPINF1, transforming growth factor beta induced (TGFB1), and THBS2 were identified as potential genes for building the LASSO module (Fig. 7A and B). On the other hand, SVM-RFE analysis revealed that the SVM model based on 6 characteristic genes showed an optimum error rate. Therefore, FBLN5, TGFB1, NID2, LUM, FBN1, and FN1 were identified as potential genes. Meanwhile, the RF algorithm identified the top 5 genes from the 20 candidates, including FBN1, FN1, TGFB1, NID2, and COL6A3 (Fig. 7C). Finally, combining the results of the above 3 algorithms (Fig. 7D), FN1 and TGFB1 were regarded as hub genes for DKD patients.

3.7. Hub-genes validation

To further evaluate the consistent changes of hub genes in DKD, we analyzed the validation dataset GSE30528. Compared with the controls, we found that FN1 and TGFB1 mRNA levels in DKD patients were significantly higher (Fig. 8A and B). Furthermore, ROC analysis showed that the AUC values of FN1, TGFB1 were >0.85 in both the GSE96804 and GSE30528 datasets (Fig. 8C and D), suggesting that they could be used as valid indicators for definitive diagnosis in DKD patients.

In order to verify the correlation between hub genes and renal function progression, the clinical data of GSE30528

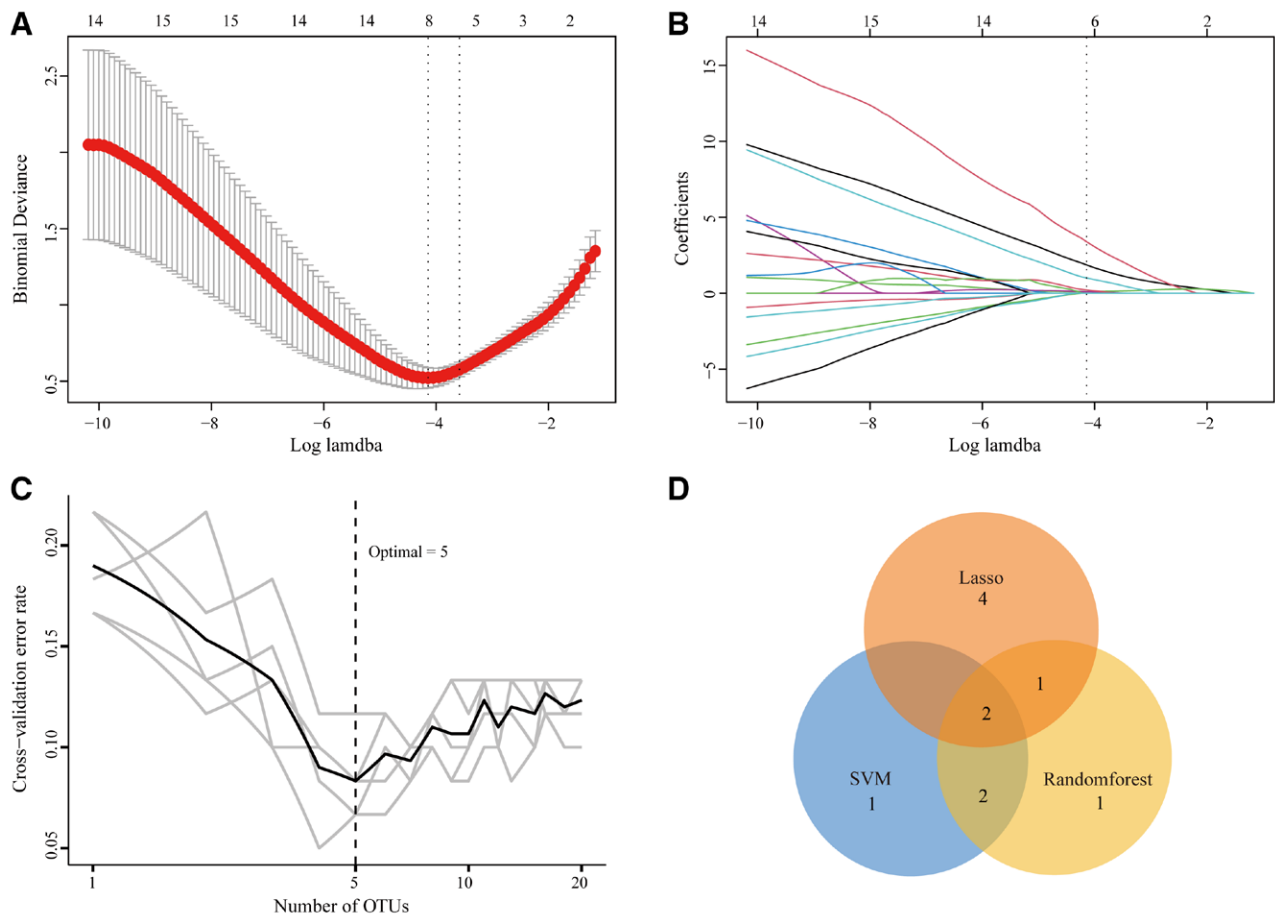


Figure 7. Identification of hub genes in DKD based on machine learning algorithms. (A) Optimal lambda of candidates selection in least absolute shrinkage and selection operator (LASSO) regression. (B) LASSO regression coefficient profiles of ten risk factors. (C) The optimum error rate of RF model based on 5 characteristic genes. (D) Venn diagram shows the overlapping genes in LASSO, support vector machines (SVM) and RF modules, namely hub genes. DKD = diabetic kidney disease, RF = randomforest.

were obtained from Nephroseq V5. Pearson correlation analysis showed that the expression levels of FN1 and TGFBI were negatively correlated with eGFR ($r = -0.68, P = 4.94 \times 10^{-4}$, $r = -0.744, P = 7.25 \times 10^{-5}$) (Fig. 8E and F).

Correlation analysis showed that a strong positive correlation between FN1, TGFBI and Macrophages M2 ($R > 0.7, P < .001$) (Fig. 9).

4. Discussion

DKD is one of the microvascular complications of diabetes and is currently considered a chronic immune-inflammatory disease.^[31] With the depth of research, the role of macrophages in the pathogenesis of DKD has been recognized. In our study, CIBERSORTx was used to identify the infiltration of immune cells in the glomerular tissue of DKD. The results suggested that macrophage infiltration increased and accounted for a high proportion in the immune cells (M1 and M2), which was consistent with previous reports.

WGCNA is a comprehensive analysis technology based on a biological network, which divides potentially relevant genes into multiple modules and explores the correlation between key modules and clinical features to identify potential biomarkers or therapeutic targets.^[21] In this study, we found the blue module strongly associated with macrophage M2 infiltration in DKD samples by WGCNA, and identified a total of 79 target genes by crossing with DEGs. Interestingly, these 79 target genes were related to an extractable matrix organization, collagen fiber

organization, immune response, monocyte chemotaxis and so on. Therefore, we speculate that these genes are involved in the immune response and fibrosis process induced by macrophages in the DKD state.

Numerous studies have shown that,^[32,33] macrophage accumulation could cause the progression of renal function and the deterioration of histopathological manifestations in DKD. After macrophage recruitment, it releases inflammatory mediators and stimulates the production of reactive oxygen species and proteases, resulting in tissue damage and fibrosis. Tumor necrosis factor- α in the kidney is mainly produced by macrophages.^[34] However, how macrophages of different phenotypes participate in the progression of DKD has not been fully defined.^[35] Traditionally, it is believed that macrophage M1 is involved in the initial stage of inflammation and is associated with tissue damage and pro-inflammatory function, whereas the M2 phenotype mediates repair and wound healing, mainly playing an anti-inflammatory and pro-angiogenic role. Both usually play opposite roles and coexist in DKD kidney tissue. Although some studies^[36,37] have reduced renal injury in DKD by inhibiting M1 activation and promoting M2 transformation, some studies have found that^[38] not all M2 therapies are beneficial and effective, and some M2 therapies can become pro-fibrotic. Notably, as revealed by Anders H.J. et al,^[39] the infiltration of macrophage M2 is positively correlated with the progression of renal fibrosis, and the possible mechanism is macrophage M2 to myofibroblast transition in the process, resulting in excessive deposition of ECM.^[40] However, it remains unclear how the M2 phenotype becomes pro-fibrotic and what the influence of the environment

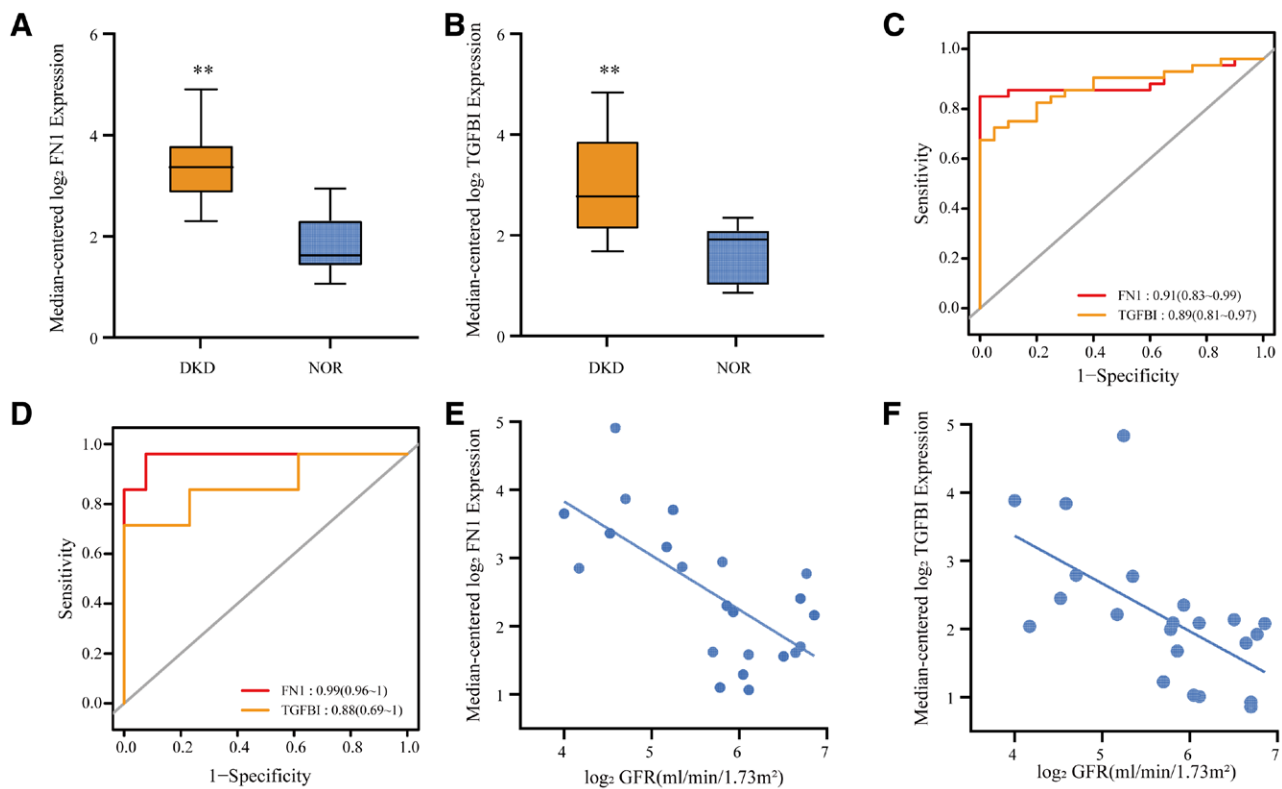


Figure 8. Validation of hub genes. (A and B) Gene expression levels of fibronectin 1 (FN1) and transforming growth factor beta induced (TGFBI) between DKD and healthy controls in GSE30528, respectively. ** $P < .01$. (C and D) Receiver operator characteristics (ROC) curves show hub genes in GSE96804 and validation in GSE30528, respectively. (E and F) Pearson correlation analysis of GFR and hub genes. FN1 and TGFBI are negatively related to GFR in DKD glomerulus samples ($r = -0.68, P = 4.94 \times 10^{-4}$, $r = -0.744, P = 7.25 \times 10^{-5}$). DKD = diabetic kidney disease.

is. Therefore, macrophages are potential immune cell targets with research prospects, and regulation of the number of macrophages in the kidney and M1/M2 phenotypic transformation may play an important role in disease progression. At this stage, several studies have been carried out,^[41,42] with therapeutic strategies aimed at reducing macrophage infiltration or antibodies against downstream inflammatory mediators, thus reducing the progression of renal injury, but most remain at the stage of animal experiments.^[43] Therefore, actively exploring more specific biomarkers mediated macrophage induced immune injury in DKD remains an important issue that we urgently need to address.

Our study constructed a PPI network for 79 target genes, and subsequently used machine learning algorithms, namely LASSO logistic regression, SVM-RFE and RF algorithms to identify FN1, and TGFBI as hub genes in DKD. Next, we verified the diagnostic efficacy and prognostic risk of hub genes in the test cohort and clinical database, all showing statistical significance. Finally, the correlation results suggested that hub genes were strongly positively correlated with macrophage M2. Therefore, we hypothesize that FN1 and TGFBI are the diagnostic and prognostic markers of DKD, and may play a role in immune injury and fibrosis caused by macrophages.

FN1 is an ECM protein involved in various biological processes, such as cell adhesion, migration, differentiation and so on. It is well known that FN1, as a fibrogenic cytokine, plays an important role in the occurrence and progression of fibrosis in DKD.^[44,45] Recently, overexpression of FN1 has been found to promote disease progression in a variety of diseases, including heart failure^[46] and neoplastic diseases,^[47] and is associated with macrophage M2 polarization. It has been reported^[48] that FN1 promotes the secretion of macrophage M2 markers (IL-10) and

reduces the production of M1 markers tumor necrosis factor- α in a concentration-dependent manner. In conjunction with our findings, it is speculated that in the development of DKD, FN1 may not only promote fibrosis, but also exert damaging effects by participating in macrophage polarization. It is a promising target for immunotherapy.

The results of our study showed that the expression of TGFBI was elevated in DKD patients, and the expression level was negatively correlated with eGFR, and significantly positively correlated with macrophage M2 cell infiltration. Combined with the research results of Robert J. Moritz et al,^[49] the hyperglycemic environment promotes macrophage infiltration, and macrophage derived TGF- β 1 inducing TGFBI gene expression. The above-mentioned suggests that macrophage M2 infiltration may promote TGFBI gene expression in DKD patients. After the activation of the TGFBI gene, it encodes an ECM protein called BIGH3. The C-terminal portion of this protein is cleaved to derive an integrin-ligand peptide that induces apoptosis.^[50] The increase in the number of apoptotic cells further promotes macrophage infiltration, and macrophage uptake of apoptosis induces macrophage release of more TGF- β 1. In addition, it can also stimulate the synthesis and secretion of BIGH3 protein by macrophages themselves, forming a potential feedback regulatory mechanism and promoting the progress of DKD.^[49,51] In addition, BIGH3 may also contribute to the progression of DKD by participating in the pro-fibrotic pathway driven by TGF- β 1.^[52] In recent years, clinical studies have also reported that urinary BIGH3 levels are positively correlated with TGF- β and albumin excretion rate in patients with DKD,^[53] and are also elevated in diabetic patients with normal albuminuria.^[54] Urinary BIGH3 may become a promising early renal injury marker in DKD patients. In summary, we concluded that increased

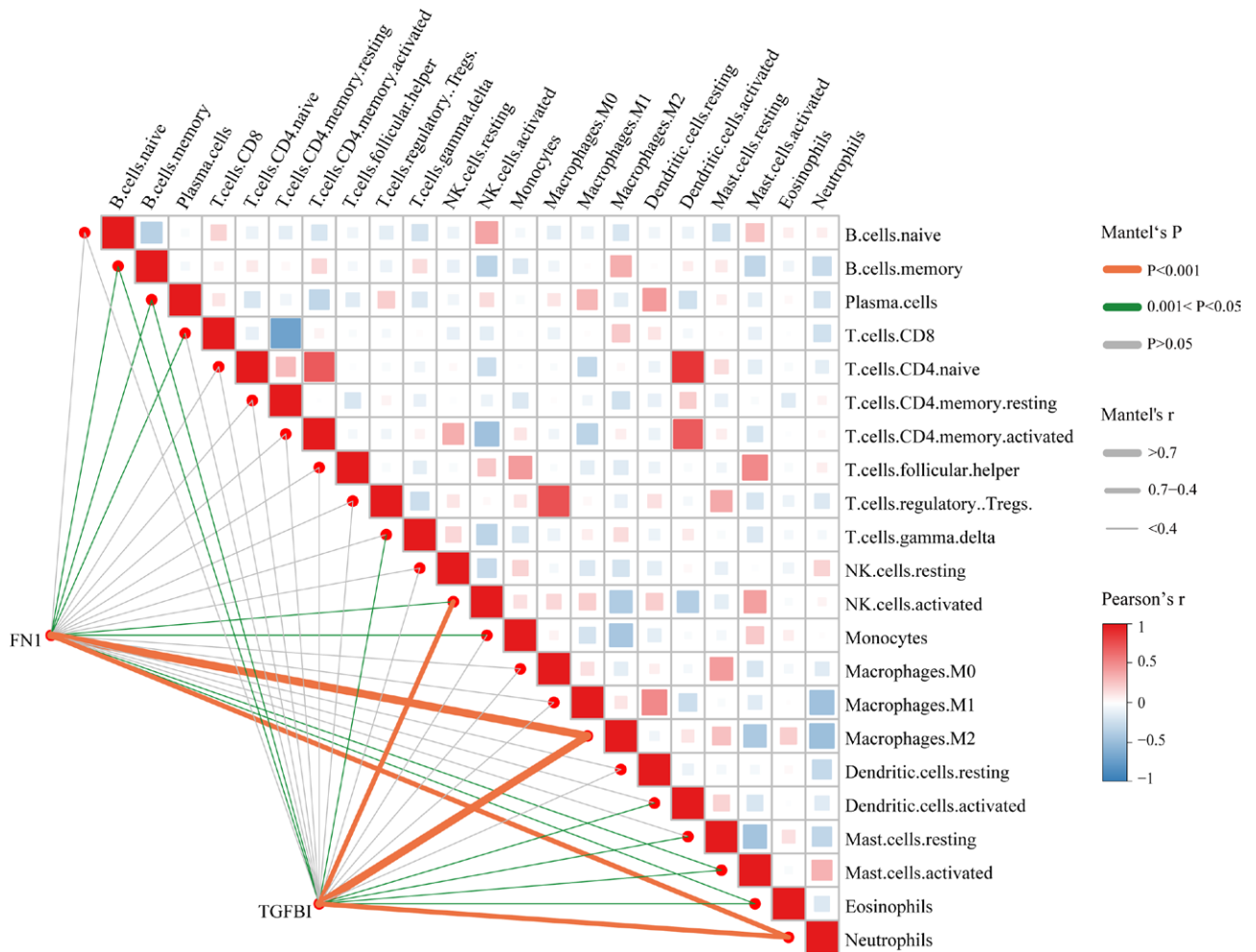


Figure 9. Correlation analysis between fibronectin 1 (FN1), transforming growth factor beta induced (TGFB1) and 22 kinds of immune cells. The color and the thickness of the line represents the *P* value and the correlation, respectively.

Macrophage M2 infiltration in DKD patients may promote the expression of TGFB1 gene and the synthesis of BIGH3 protein, inducing renal cell apoptosis through the formation of a feedback regulatory mechanism, while promoting fibrosis, leading to the progression of DKD. It suggests that TGFB1 may be involved in macrophage-induced immune injury and fibrosis.

However, there are some limitations to this study. First, the sample size from the GSE96804 dataset was limited and did not cover all patients with diabetic nephropathy. In addition, the microarray data were not classified according to the pathological stages of DKD, and the expression levels of certain genes may not be identical in different pathological stages. Finally, the specific molecular mechanisms and biological functions of these biomarkers remain to be validated by further experimental studies. Therefore, in the follow-up study, we will include more cases of DKD, including patients with different pathological stages, to verify the expression of hub genes in DKD patients, as well as the correlation between gene expression levels and clinical pathological stages. In addition, we will demonstrate the expression levels of FN1 and TGFB1 at the cellular and animal experimental levels, further investigate the correlation between genes in DKD and macrophage expression, and further validate the actual regulatory function of genes and their impact on functional phenotype through subsequent experiments. This will provide new directions and therapeutic targets for the future treatment of DKD from the perspective of macrophages and immune inflammation.

5. Conclusion

In this study, we applied immune infiltration analysis, WGCNA combined with machine learning algorithms to reveal the biological function of macrophages in the progression of DKD, that is, they participate in the immune response and fibrosis process. We identified FN1 and TGFB1 as promising biomarkers, and validated their good diagnostic effect and prognostic value. It will help to provide potential new therapeutic targets for DKD from the perspective of macrophages and immune inflammation in the future.

Acknowledgments

We thank the researchers and patients who contributed data to the GEO database.

Author contributions

Conceptualization: Fulin Dou.
Formal analysis: Fulin Dou.
Data curation: Qingzhen Liu, Shasha Lv.
Writing – original draft: Qiaoying Xu, Xueling Wang.
Writing – review & editing: Shanshan Liu, Gang Liu.

References

[1] Lu Y, Stamm C, Nobre D, et al. Changing trends in end-stage renal disease patients with diabetes. *Swiss Med Wkly.* 2017;147:w14458.

- [2] Arora M, Singh U. Molecular mechanisms in the pathogenesis of diabetic nephropathy: an update. *Vascul Pharmacol*. 2013;58:259–71.
- [3] Tuomi T, Santoro N, Caprio S, et al. The many faces of diabetes: a disease with increasing heterogeneity. *Lancet*. 2014;383:1084–94.
- [4] Tang S, Yiu W. Innate immunity in diabetic kidney disease. *Nat Rev Nephrol*. 2020;16:206–22.
- [5] Klessens C, Zandbergen M, Wolterbeek R, et al. Macrophages in diabetic nephropathy in patients with type 2 diabetes. *Nephrol Dial Transplant*. 2017;32:1322–9.
- [6] Tesch G. Role of macrophages in complications of type 2 diabetes. *Clin Exp Pharmacol Physiol*. 2007;34:1016–9.
- [7] Lim A, Tesch G. Inflammation in diabetic nephropathy. *Mediators Inflamm*. 2012;2012:146154.
- [8] Guiteras R, Sola A, Flaquer M, et al. Macrophage overexpressing NGAL ameliorated kidney fibrosis in the UUO mice model. *Cell Physiol Biochem*. 2017;42:1945–60.
- [9] You H, Gao T, Cooper T, et al. Macrophages directly mediate diabetic renal injury. *Am J Physiol Renal Physiol*. 2013;305:F1719–1727.
- [10] Heerspink H, Parving HH, Andress DL, et al. Atrasentan and renal events in patients with type 2 diabetes and chronic kidney disease (SONAR): a double-blind, randomised, placebo-controlled trial. *Lancet*. 2019;393:1937–47.
- [11] Dick, de, Zeeuw, P, Bekker, E, et al. The effect of CCR2 inhibitor CCX140-B on residual albuminuria in patients with type 2 diabetes and nephropathy: a randomised trial. *Lancet Diabetes Endocrinol*. 2015;3:687–96.
- [12] Tuttle KR, Brosius FC, Adler SG, et al. JAK1/JAK2 inhibition by baricitinib in diabetic kidney disease: results from a Phase 2 randomized controlled clinical trial. *Nephrol Dial Transplant*. 2018;33:1950–59.
- [13] Shi JS, Qiu DD, Le WB, et al. Identification of transcription regulatory relationships in diabetic nephropathy. *Chin Med J (Engl)*. 2018;131:2886–90.
- [14] Woroniecka KI, Park AS, Mohtat D, et al. Transcriptome analysis of human diabetic kidney disease. *Diabetes*. 2011;60:2354–69.
- [15] Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. *Bioinformatics*. 2009;25:415–6.
- [16] Carvalho B, Irizarry R. A framework for oligonucleotide microarray preprocessing. *Bioinformatics*. 2010;26:2363–7.
- [17] Ritchie M, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
- [18] Ito K, Murphy D. Application of ggplot2 to Pharmacometric Graphics. *CPT Pharmacometrics Syst Pharmacol*. 2013;2:e79.
- [19] Kolde R. Pheatmap: pretty heatmaps. 2015.
- [20] Newman A, Liu C, Green M, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12:453–7.
- [21] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. 2008;9:559.
- [22] Huang W, Sherman B, Lempicki R. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protocols*. 2009;4:44–57.
- [23] Szklarczyk D, Morris J, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45:D362–8.
- [24] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–504.
- [25] Robert T. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B (Methodol)*. 1996;58:267–88.
- [26] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat*. 2010;33:1–22.
- [27] Suykens JAK, Vandewalle V. Least squares support vector machine classifiers. *Neural Process Lett*. 1999;9:293–300.
- [28] Huang ML, Yung-Hsiang H, Lee WM, et al. SVM-RFE based feature selection and taguchi parameters optimization for multiclass SVM classifier. *ScientificWorldJournal*. 2014;2014:795624.
- [29] Liaw A, Wiener M. Classification and regression by randomForest. *R News*. 2002;23:18–22.
- [30] Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. 2011;12:77.
- [31] Tesch GH. Diabetic nephropathy - is this an immune disorder? *Clin Sci (Colch)*. 2017;131:2183–99.
- [32] Nguyen D, Ping FU, Wei MU, et al. Macrophage accumulation in human progressive diabetic nephropathy. *Nephrology*. 2006;11:226–31.
- [33] Chow F, Ozols E, Nikolic-Paterson DJ, et al. Macrophages in mouse type 2 diabetic nephropathy: correlation with diabetic state and progressive renal injury. *Kidney Int*. 2004;65:116–28.
- [34] Awad AS, You H, Gao T, et al. Macrophage-derived tumor necrosis factor- α mediates diabetic renal injury. *Kidney Int*. 2015;88:722–33.
- [35] Calle P, Hotter G. Macrophage phenotype and fibrosis in diabetic nephropathy. *Int J Mol Sci*. 2020;21:2806.
- [36] Zhang XL, Guo YF, Song ZX, et al. Vitamin D prevents podocyte injury via regulation of macrophage M1/M2 phenotype in diabetic nephropathy rats. *Endocrinology*. 2014;155:4939–50.
- [37] Ji L, Chen Y, Wang H, et al. Overexpression of Sirt6 promotes M2 macrophage transformation, alleviating renal injury in diabetic nephropathy. *Int J Oncol*. 2019;55:103–15.
- [38] Guiteras R, Sola A, Flaquer M, et al. Macrophage overexpressing NGAL ameliorated kidney fibrosis in the UUO mice model. *Cell Physiol Biochem*. 2017;42:1945–60.
- [39] Anders HJ, Ryu M. Renal microenvironments and macrophage phenotypes determine progression or resolution of renal inflammation and fibrosis. *Kidney Int*. 2011;80:915–25.
- [40] Meng XM, Wang S, Huang XR, et al. Inflammatory macrophages can transdifferentiate into myofibroblasts during renal fibrosis. *Cell Death Dis*. 2016;7:e2495.
- [41] You H, Gao T, Cooper TK, et al. Macrophages directly mediate diabetic renal injury. *Am J Physiol Renal Physiol*. 2013;305:F1719–27.
- [42] Jan M, Dirk E, Diana B, et al. C-C motif-ligand 2 inhibition with emapticap pegol (NOX-E36) in type 2 diabetic patients with albuminuria. *Nephrol Dial Transplant*. 2017;32:307–15.
- [43] Zhu Y, Ling Y, Wang X. Alantolactone mitigates renal injury induced by diabetes via inhibition of high glucose-mediated inflammatory response and macrophage infiltration. *Immunopharmacol Immunotoxicol*. 2020;42:84–92.
- [44] Schena FP, Geusaldo L. Pathogenetic mechanisms of diabetic nephropathy. *J Am Soc Nephrol*. 2005;16 Suppl 1:S30–3.
- [45] Kretzler M, Brosius FC, Qian Y, et al. Mechanisms of glomerulosclerosis in diabetic nephropathy.
- [46] Zheng Y, Lang Y, Qi Z, et al. Macrophage-related genes biomarkers in left ventricular remodeling induced by heart failure. *Rev Cardiovasc Med*. 2022;23:109.
- [47] Geng Q, Huang T, Li L, et al. Over-expression and prognostic significance of FN1, correlating with immune infiltrates in thyroid cancer. *Front Med (Lausanne)*. 2021;8:812278.
- [48] Yoshida Y, Kang K, Chen G, et al. Cellular fibronectin is induced in ultraviolet-exposed human skin and induces IL-10 production by monocytes/macrophages. *J Invest Dermatol*. 1999;113:49–55.
- [49] Moritz R, LeBaron R, Phelix C, et al. β Macrophage TGF-1 and the proapoptotic extracellular matrix protein BIGH3 induce renal cell apoptosis in prediabetic and diabetic conditions. *Int J Clin Med*. 2016;7:496–510.
- [50] Kim J, Kim S, Jeong H, et al. RGD peptides released from beta ig-h3, a TGF-beta-induced cell-adhesive molecule, mediate apoptosis. *Oncogene*. 2003;22:2045–53.
- [51] Nacu N, Luzina I, Highsmith K, et al. Macrophages produce TGF-beta-induced (beta-ig-h3) following ingestion of apoptotic cells and regulate MMP14 levels and collagen turnover in fibroblasts. 2008;180:5036–44.
- [52] Brennan E, Morine M, Walsh D, et al. Next-generation sequencing identifies TGF- β 1-associated gene expression profiles in renal epithelial cells reiterated in human diabetic nephropathy. *Biochim Biophys Acta*. 2012;1822:589–99.
- [53] Ha S, Kim H, Bae J, et al. Elevation of urinary betaig-h3, transforming growth factor-beta-induced protein in patients with type 2 diabetes and nephropathy. *Diabetes Res Clin Pract*. 2004;65:167–73.
- [54] Cha D, Kim I, Kang Y, et al. Urinary concentration of transforming growth factor-beta-inducible gene-h3(beta ig-h3) in patients with Type 2 diabetes mellitus. *Diabet Med*. 2005;22:14–20.