Article

J. Biol. Regul. Homeost. Agents. 2024; 38(9): 5831–5843
https://doi.org/10.23812/j.biol.regul.homeost.agents.20243809.465

A Predictive Model for Diabetic Kidney Disease Based on Inflammatory Gene Signatures and Its Regulatory Network

Mengyun Xiao[1], Yongping Lu[2], Ting Zhu[2], Yue Peng[1], Zigan Xu[3], Shaodong Luan[3], Lianghong Yin[1], Donge Tang[4,*], Yong Dai[5,*]

[1]Institute of Nephrology and Blood Purification, The First Affiliated Hospital of Jinan University, 510632 Guangzhou, Guangdong, China
[2]Department of Nephrology, Center of Kidney and Urology, The Seventh Affiliated Hospital of Sun Yat-sen University, 518039 Shenzhen, Guangdong, China
[3]Department of Nephrology, Shenzhen Longhua District Central Hospital, 518110 Shenzhen, Guangdong, China
[4]Guangdong Provincial Engineering Research Center of Autoimmune Disease Precision Medicine, Shenzhen People's Hospital, The Second Clinical Medical College, Jinan University, 518020 Shenzhen, Guangdong, China
[5]The First Hospital of Anhui University of Science and Technology, School of Medicine, Anhui University of Science and Technology, 232001 Huainan, Anhui, China
*Correspondence: donge66@126.com (Donge Tang); daiyong22@aliyun.com (Yong Dai)
Submitted: 1 March 2024 Revised: 19 March 2024 Accepted: 22 March 2024 Published: 30 August 2024

Background: Diabetic kidney disease (DKD) is a leading cause of end-stage renal disease (ESRD) globally, characterized by increased albuminuria and reduced glomerular filtration rate. Recent evidence points to inflammation as a vital contributor to the development and progression of DKD, involving interactions among immune cells, cytokines, and chemokines. Our study focused on uncovering inflammation-related genes in DKD to understand its mechanisms and developed an inflammation-centric predictive model. We aimed to bridge molecular insights with immune interactions, paving the way for innovative treatments.
Methods: This study involves comprehensive data collection from gene expression omnibus (GEO) datasets (GSE1009 and GSE30528) to identify differentially expressed genes (DEGs) between patients with DKD and healthy controls (HC). Using the ComBat method for batch effect removal, R package Limma for DEGs identification, and Metascape for enrichment analysis, we focused on the interplay between inflammation-associated genes and immune cell infiltration. We developed a predictive model for DKD using the least absolute shrinkage and selection operator (LASSO) regression, centered on six potential candidate genes: chitinase-3-like protein 1 (*CHI3L1*), coagulation factor V (*F5*), decay-accelerating factor (*CD55*), insulin-like growth factor 1 (*IGF1*), vascular endothelial growth factor A (*VEGFA*), and 15-hydroxyprostaglandin dehydrogenase (*HPGD*), within a training cohort. This model was subsequently validated in a test cohort utilizing data extracted from the GEO dataset GSE96804. Immune cell infiltration was determined using CIBERSORT, followed by Pearson correlation analysis to elucidate the interactions between hub genes, immune cells, and chemokines.
Results: We identified 349 DEGs, including 99 upregulated and 250 downregulated genes, highlighting the significant role of inflammation in DKD. Through weighted gene co-expression network analysis (WGCNA), a module consisting of 784 genes strongly associated with DKD was identified. Within this module, six inflammatory-related genes were identified as crucial for the predictive model, achieving an area under the receiver operating characteristic curve (AUC) of 1 in training and 0.76 in validation. Analysis of immune cells revealed significant differences between DKD patients and controls, while Pearson correlation analysis highlighted key associations with immune infiltration and regulation.
Conclusions: Our study provides novel insights into the genetic and inflammatory landscape of DKD, establishing a predictive model with high accuracy compared to existing models. We pinpoint significant correlations between hub genes and immune cell dynamics, potentially opening avenues for new therapeutic strategies. Our findings underscore the promise of precision medicine in diagnosing and treating DKD.

Keywords: diabetic kidney disease (DKD); inflammation; predictive model; immune cells infiltration; gamma delta T cells; cytokines; chemokines

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

Introduction

Diabetic kidney disease (DKD) is the primary cause of end-stage renal disease (ESRD) worldwide and a significant microvascular complication of diabetes mellitus, with multifactorial mechanisms [1]. Recent studies have highlighted inflammation as a pivotal factor in the onset and progression of DKD [2–4]. Growing evidence suggests that DKD-related inflammation is complex and multifaceted, involving a web of interactions among immune cells, cytokines, chemokines, and both the innate and adaptive immune systems [2,4–6]. The persistent, mild inflammation common in diabetes accelerates endothelial dysfunction, significantly contributing to the pathology of DKD [2]. Elevated levels of complement factors, such as complement component 3 (C3) and complement component 5 (C5), along with their activated forms C3a and C5a, have been observed in plasma, urine, and kidney tissues of diabetic patients [2,5]. These factors are associated with an increased risk of developing DKD and poorer kidney outcomes in individuals affected by the disease [2,5,7–9]. Targeting the receptors for C3a and C5a has demonstrated a reduction in inflammation-related gene activity, a decrease in albuminuria, and an improvement in kidney fibrosis, both *in vivo* and *in vitro* [10,11].

Immune cell infiltration by monocytes, macrophages, neutrophils, natural killer (NK) cells, and dendritic cells significantly contributes to the progression of DKD [12–14]. Strategies that reduce or block signaling molecules that attract macrophages have been particularly effective in kidney protection [15]. Furthermore, high levels of pro-inflammatory cytokines such as Tumor Necrosis Factor-α (TNF-α), interleukin-1β (IL-1β), and interleukin-6 (IL-6) in the kidneys of DKD patients are closely linked to more severe albuminuria and kidney damage [16].

Based on these findings, inflammation seems to be prevalent and potentially crucial in the progression of DKD. Our research focuses on identifying genes that undergo changes contributing to DKD-associated inflammation and constructing a predictive model for these genes. Additionally, we investigated the interaction between these altered genes and infiltrating immune cells, along with their potential underlying mechanisms.

Materials and Methods

Data Collection, Processing, Differentially Expressed Genes (DEGs) Screening, and Enrichment Analysis

Early intervention is essential in DKD to prevent and slow down the progression of the disease. To elucidate the inflammatory characteristics in glomerular tissue during early-stage DKD, we analyzed three datasets: GSE1009, GSE30528, and GSE96804. These datasets contain gene expression profiles of glomeruli from patients with DKD and are publicly available in the gene expression omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). The original clinical trials included patients with type 2 diabetes mellitus (T2DM) classified as stages 1 to 3 DKD, representing early disease phases with varying degrees of kidney damage ranging from minor to mild. Detailed dataset information is provided in Table 1 (Ref. [17–19]).

We integrated the data from GSE1009 and GSE30528, employing the ComBat algorithm to address batch effects. DEGs between DKD patients and healthy controls (HC) were identified using the Limma package in R, based on criteria of an adjusted *p*-value (adj. *p*) below 0.05 and an absolute log2 fold change (|log2 FC|) exceeding 1. Heatmaps generated with the Pheatmap package in R facilitated the visualization of these DEGs, while enrichment analysis was conducted using the Metascape database (https://metascape.org). Probes corresponding to the same gene with missing values were excluded, and expression levels were averaged to ensure accuracy. This rigorous methodology enables a deeper understanding of the molecular mechanisms underlying early DKD, highlighting the importance of prompt therapeutic interventions.

Weighted Gene Co-Expression Network Analysis (WGCNA)

The WGCNA package was used to build co-expression networks for genes in the dataset, facilitating further analysis with a soft-thresholding power of 9. This approach converts the weighted adjacency matrix into a topological overlap matrix (TOM), which estimates network connectivity. Subsequently, hierarchical clustering was applied to TOM to create a dendrogram, where different branches represent distinct gene modules, and different colors indicate distinct modules. Genes were classified based on their weighted correlation coefficients and expression patterns. Genes exhibiting similar patterns were grouped into a single module, effectively organizing thousands of genes into multiple modules according to their gene expression patterns.

Model Construction

Datasets GSE1009 and GSE30528 were combined to create the training cohort, with dataset GSE96804 serving as the test cohort. By intersecting 28 inflammatory genes with 784 genes associated with DKD as determined by WGCNA, we identified 12 candidate genes. These genes were used to build a predictive model for DKD through least absolute shrinkage and selection operator (LASSO) regression, which highlighted six key genes as significant indicators of DKD. Patient risk scores were calculated by combining gene expression values with their respective LASSO regression coefficients. Based on the median risk score, patients were divided into low or high-risk groups. The model's predictive performance was assessed by analyzing the area under the receiver operating characteristic curve (AUC).

Table 1. The information of GEO datasets in this study.

	Tissue	HC	DKD	Isolation of glomeruli	Experiment type	Attribute	Reference
GSE1009-GPL8300	Glomeruli (Homo sapiens)	Cadaveric donor kidneys (n = 3)	Stage 1-stage 3 (patients with T2DM) (n = 3)	Dissected by Sieving techniques	by array	Test	[17]
GSE30528-GPL571	Glomeruli (Homo sapiens)	Living allograft donors, surgical nephrectomies, and leftover portions of diagnostic kidney biopsies (n = 13)	Stage 3-stage 4 (not specified the type of diabetes) (n = 9)	Microdissection	by array	Test	[18]
GSE96804-GPL17586	Glomeruli (Homo sapiens)	Surgical nephrectomies (n = 20)	Stage 1-stage 3 (patients with T2DM) (n = 41)	Microdissection	by array	Validation	[19]

GEO, gene expression omnibus; HC, healthy controls; DKD, diabetic kidney disease; T2DM, type 2 diabetes mellitus.

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

Immune Cell Infiltration and Correlation Analysis

CIBERSORT (https://cibersortx.stanford.edu/index.php), a method employing support vector regression, was used to deconvolve the expression matrix into 22 different immune cell phenotypes. This algorithm estimated the composition of immune cells in patient samples. Spearman correlation analysis was conducted to investigate the relationship between gene expression and immune cell prevalence.

Pearson's correlation analysis was used to examine the relationship between immune cells, hub genes, as well as between hub genes and chemokines along with their receptors.

Single-Gene Gene Set Enrichment Analysis (GSEA)

We ranked genes from the DKD samples across the combined datasets based on their relative expression levels of chitinase-3-like protein 1 (*CHI3L1*), coagulation factor V (*F5*), decay-accelerating factor (*CD55*), insulin-like growth factor 1 (*IGF1*), vascular endothelial growth factor A (*VEGFA*), and 15-hydroxyprostaglandin dehydrogenase (*HPGD*), categorizing them into either the top 10th percentile (indicative of high gene expression) or the bottom 10th percentile (indicative of low gene expression) for GSEA. Subsequently, we compared the differential enrichment of the kyoto encyclopedia of genes and genomes (KEGG) signaling pathways between the high and low-expression groups. The analysis involved 1000 permutations, with phenotypes determining the permutation type.

Regulatory Network Analysis and Inverse Prediction of microRNAs (miRNAs) for Hub Genes

We employed the RcisTarget package for predicting transcription factor (TF) regulation through motif enrichment analysis. The significance of each motif's overrepresentation in a gene set was assessed by calculating the AUC of the recovery curve in comparison to motif ranking. The normalized enrichment score (NES) for each motif was derived from the AUC for all motifs within a set. Gene-motif rankings were obtained from the DRIMust database (https://drimust.technion.ac.il). To predict upstream miRNAs for hub genes, we employed the miRcode database (http://www.mircode.org). Visualization of the miRNA-mRNA regulatory network was accomplished using Cytoscape (version 3.7.2: https://github.com/cytoscape/cytoscape/releases/3.7.2/).

Statistical Analysis

In this analysis, the R language (version 4.0: https://cran.r-project.org/src/base/R-4/) was used, with statistical significance set at $p < 0.05$.

Results

Identifying Inflammatory DEGs in Glomeruli between Patients with DKD and HC

Fig. 1 outlines the study workflow, while Table 1 provides details on the datasets employed. We acquired the expression profile datasets for DKD, specifically GSE1009 and GSE30528, from the GEO database, comprising 16 healthy individuals and 12 patients with DKD. The data were integrated, and batch effects were corrected using the ComBat algorithm. Principal component analysis (PCA) revealed a significant reduction in batch effects after ComBat adjustment, as depicted in **Supplementary Fig. 1A,B**. Using the limma package with criteria of $p < 0.05$ and $|\log2FC| > 1$ for screening, we identified 349 DEGs in contrast to healthy controls and DKD patients. The volcano plot and heatmap of DEGs revealed that 99 genes were upregulated, whereas 250 were downregulated in patients with DKD compared to controls (Fig. 2A, **Supplementary Fig. 1C**). Subsequent analysis of these DEGs using gene ontology (GO) and KEGG highlighted key biological processes and signaling pathways. Enriched GO terms included 'regulation of cell-cell adhesion' and 'positive regulation of leukocyte activation', with signaling pathways such as 'cytokine-cytokine receptor interaction' and 'complement and coagulation cascades'. These results indicate a significant role of inflammation in the progression of DKD (Fig. 2B,C). The involvement of inflammation in diabetic DKD has increasingly gained attention in recent years [2,5,6].

Hyperglycemia, the presence of advanced glycation end products (AGEs), and the activation of the Renin-Angiotensin-Aldosterone System (RAAS) contribute to damage in renal cells, leading to chronic kidney inflammation [20]. To further explore the influence of inflammation on DKD, we retrieved a list of 11,449 inflammation-associated genes from GeneCards, which we refined to 528 based on a relevance score cutoff exceeding 5. A Venn diagram showed that 28 of these genes overlapped with previously identified DEGs (Fig. 2D).

A Predictive Model for DKD was Robustly Constructed Using Key Inflammatory Genes

The WGCNA algorithm was employed to determine DKD-related genes by integrating expression data from the GSE1009 and GSE3052 datasets. Utilizing the WGCNA package in R software, a scale-free co-expression network was generated from a tom matrix with a threshold of $\beta = 9$, achieving a scale-free R^2 of 0.9 (Fig. 3A,B). The DKD expression profile revealed 13 distinct gene modules, each represented by a different color (Fig. 3C). Through Pearson's correlation analysis applied to genes within each module across different groups, the turquoise module (comprising 784 genes) emerged with the strongest correlation coefficient of –0.93, $p = 4 \times 10^{-13}$ with DKD (Fig. 3D).

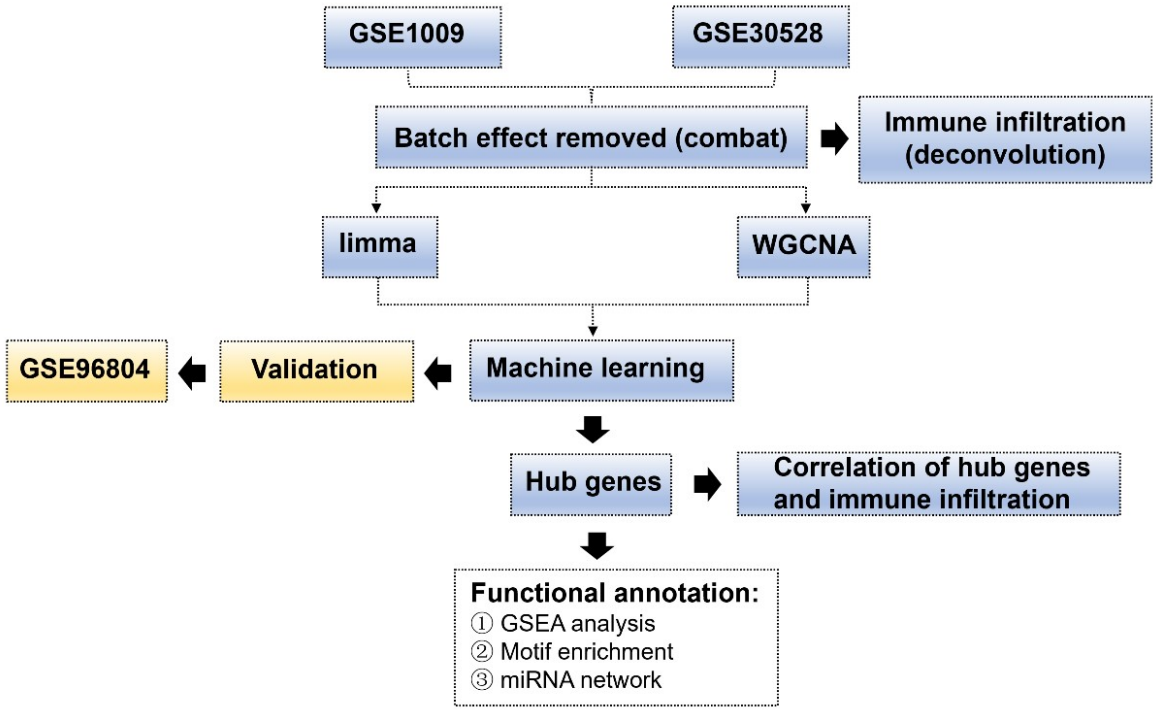

Fig. 1. Workflow of the current study. This figure presents the integrative process used to identify inflammatory gene signatures linked to diabetic kidney disease (DKD) progression. Initially, data from the gene expression omnibus (GEO) datasets GSE1009 and GSE30538 were combined and subsequently cleaned to eliminate batch effects. The Limma package was utilized to identify differentially expressed genes (DEGs) by comparing DKD patients with healthy controls. By intersecting these DEGs with a predefined list of inflammatory genes, a subset specifically related to inflammatory responses was identified. Through weighted gene co-expression network analysis (WGCNA), genes implicated in DKD progression were determined and further refined by their overlap with inflammatory DEGs. A predictive model was developed using the least absolute shrinkage and selection operator regression on a selection of 12 candidate genes, ultimately isolating six key genes as central hubs of inflammation in DKD within a training cohort, validated by an area under the receiver operating characteristic curve (AUC) of 1. In the test cohort, the model's validity was confirmed with an AUC of 0.76. CIBERSORT was employed for immune cell infiltration profiling, and Pearson's correlation analysis was conducted to elucidate the complex interplay among hub genes, immune cell dynamics, and chemokine signaling. Additionally, single-gene gene set enrichment analysis offered insights into the active signaling pathways associated with each hub gene. Finally, a detailed examination was performed on transcription factors and microRNAs that might regulate the activity of these central hub genes.

Twelve candidate genes (*CHI3L1*, lipoprotein lipase (*LPL*), *F5*, *VEGFA*, phospholipase C gamma 2 (*PLCG2*), *IGF1*, tachykinin precursor 1 (*TAC1*), serine peptidase inhibitor kazal type 1 (*SPINK1*), *HPGD*, coagulation factor III (*F3*), *CD55*, and interleukin-33 (*IL-33*)) were identified by analyzing 784 turquoise module genes with 28 inflammation-related DEGs for further analysis (Fig. 4A). The datasets GSE1009 and GSE30528 were used as training sets, with GSE96804 serving as the validation set. LASSO regression analysis was applied to the 12 inflammation-related DEGs to identify key genes for constructing the DKD predictive model. Through iterative analysis using a 10-fold cross-validation method, an optimal model with the minimum number of variables was obtained at $\lambda = 0.063$ ($\log \lambda = -3.62$) (Fig. 4B). The regression coefficient was then calculated (Fig. 4C) to aid in predicting DKD.

Six genes (*HPGD*, *IGF1*, *CD55*, *VEGFA*, *F5*, and *CHI3L1*) were identified as hub genes for predicting DKD (Fig. 4D). After training and model optimization, the model demonstrated a strong diagnostic accuracy for DKD, achieving an AUC of 1 in the training dataset (Fig. 4E). In the validation dataset GSE96804, the AUC for DKD was 0.76 (Fig. 4F).

Hub Genes Correlate with Immune Cell Infiltration in DKD

The immune microenvironment is a complex assembly primarily consisting of immune-related fibroblasts, immune cells, the extracellular matrix, a variety of growth and inflammatory factors, unique physical and chemical characteristics, and diseased cells [21,22]. This immune microenvironment significantly affects the diagnosis, survival rates, and therapeutic responses of numerous prevalent diseases

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

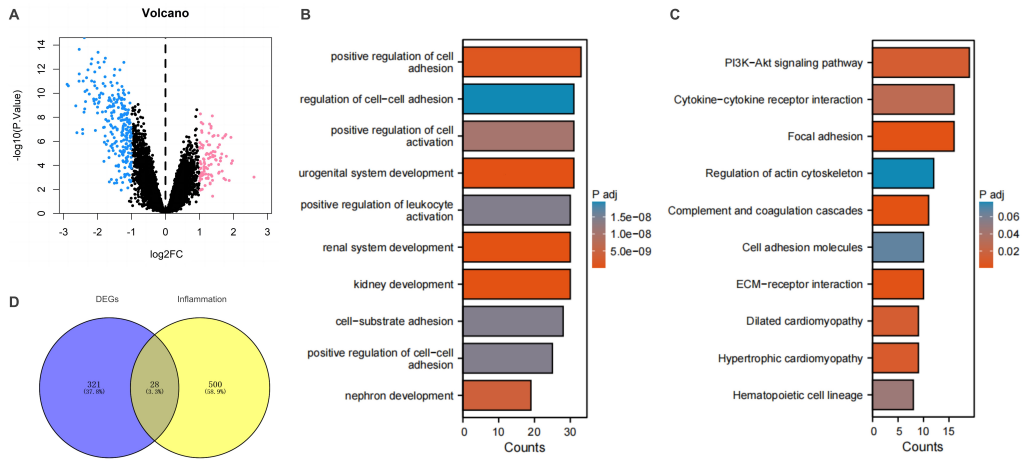

Fig. 2. Identification of inflammatory differentially expressed genes (DEGs) in diabetic kidney disease (DKD). (A) The volcano plot displays the differential expression between patients with DKD and healthy controls, highlighting 349 DEGs with 99 upregulated and 250 downregulated genes. (B,C) The bar plots provide a summary of the DEGs enrichment analysis illustrated through gene ontology terms (B) and kyoto encyclopedia of genes and genomes pathways (C). (D) The Venn diagram identifies a subset of 28 genes common to both the identified DEGs and the inflammation-related genes, which were selected from GeneCards with a relevance score exceeding 5.

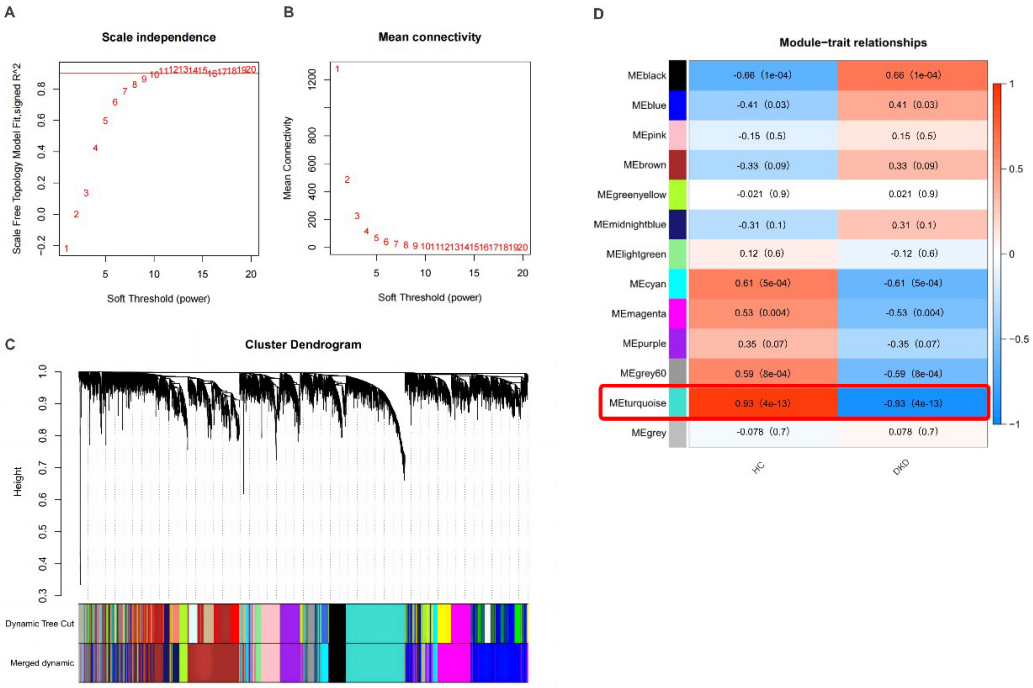

Fig. 3. Identification of genes associated with diabetic kidney disease (DKD) through weighted gene co-expression network analysis. (A,B) Determination of network threshold parameters. The plot of the scale-free fit index against the soft-thresholding power demonstrated the selection of the power threshold that achieves a scale-free network (A). Additionally, the graph of average node connectivity against soft-thresholding power aids in determining network stability (B). (C) Gene clustering and module identification. The dendrogram represents gene clustering based on the topological overlap, with gene modules displayed in different colors alongside the tree. (D) Analysis of module-trait relationship. Each row represents the eigengene of a module, while each column corresponds to a phenotype trait. The cells indicate the correlation and significance between module eigengenes and clinical traits, with colors indicating the strength and direction of the correlation. This section of the figure also shows the distribution of mean gene significance and error within modules linked to DKD. The turquoise module (784 genes) exhibited the highest correlation coefficient of –0.93 with DKD (marked by the red box).

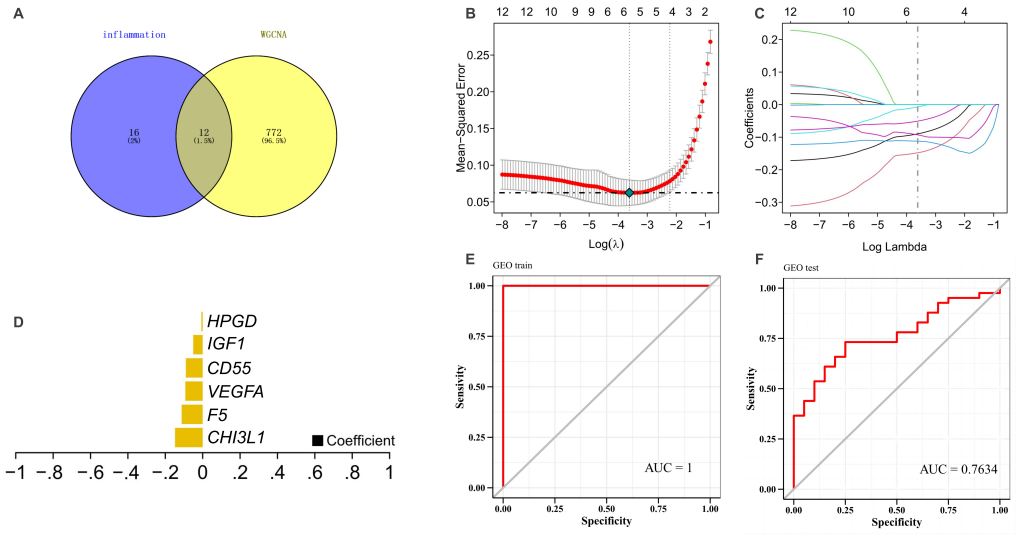

Fig. 4. Construction of a predictive model with inflammatory differentially expressed genes (DEGs) for diabetic kidney disease (DKD). (A) The Venn diagram shows the identification of 12 candidate genes by intersecting 28 inflammatory DEGs with 784 turquoise module genes. (B,C) Screening of distinctive genes and construction of the DKD predictive model using the least absolute shrinkage and selection operator regression. The range between two dotted lines represents the positive and negative standard deviations of log (λ). The left dotted line indicates the value of the harmonic parameter log (λ) when the error of the model is minimized. Six variables were selected at log (λ) = –3.62 (B). A vertical line marks the value chosen through 10-fold cross-validation. Decreasing λ enhances model compression and variable selection (C). (D) Development of a predictive model for DKD based on the expression levels of six pivotal genes. For each patient in the study cohort, a risk score was computed, and the median of these scores was established as a cut-off point. A distribution chart of the risk scores illustrates this bifurcation, with a vertical demarcation representing the median, effectively categorizing patients into low-risk or high-risk classifications. (E,F) The area under the receiver operating characteristic curves demonstrate the robust diagnostic accuracy of the DKD predictive model, which is based on the six-gene signature. These curves demonstrate the model's effectiveness in distinguishing between DKD cases and controls in both (E) the training cohort and (F) validation cohort.

[23]. To gain more insight into the role of inflammation in the progression of DKD, we investigated the correlation between hub genes and the infiltration of immune cells.

The number of immune cells within the glomeruli of each individual is shown in **Supplementary Fig. 2**. The immune cell distribution varied significantly between HC and DKD patients, as indicated by the combined expression profiles of the GSE1009 and GSE30528 datasets (**Supplementary Fig. 3**). Specifically, there was a notable increase in CD4 naive T-cells, gamma delta T cells, and resting dendritic cells, while the counts of resting NK cells, M2 macrophages, and activated dendritic cells were significantly reduced in DKD patients compared to their healthy counterparts (**Supplementary Fig. 3**).

Gamma delta T cells have been reported to produce abundant cytokines among the immune cells showing notable changes. A KEGG enrichment analysis indicated that 16 genes were enriched in cytokine-related signaling pathways (**Supplementary Table 1**). These findings align with the increased expression of gamma delta T cells, as 11 out of the 16 DEGs were upregulated (**Supplementary Table 1**).

We conducted Pearson correlation analysis to examine the relationship among different immune cells showing significant alterations in the glomeruli of patients with DKD. Our findings revealed a positive correlation factor of 0.41 between gamma delta T-cells and CD4 naive T-cells (**Supplementary Fig. 4**). Additionally, M2 macrophages and activated dendritic cells displayed a positive correlation coefficient factor of 0.5 (**Supplementary Fig. 4**). However, CD4 naive T-cells and resting NK cells were negatively correlated, with a coefficient factor of –0.46 (**Supplementary Fig. 4**).

Correlation between Hub Genes and Chemokines and their Receptors in Regulating Immune Cell Infiltration in DKD Glomerular Tissue

Investigating the impact of hub genes on the regulatory mechanisms of immune infiltration in DKD patients, we conducted a Pearson correlation analysis between hub genes and immune cells. Our analysis revealed various correlations between hub genes and significantly altered immune cells in DKD patients. Notably, all hub genes exhibited positive correlations with resting NK cells, while

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

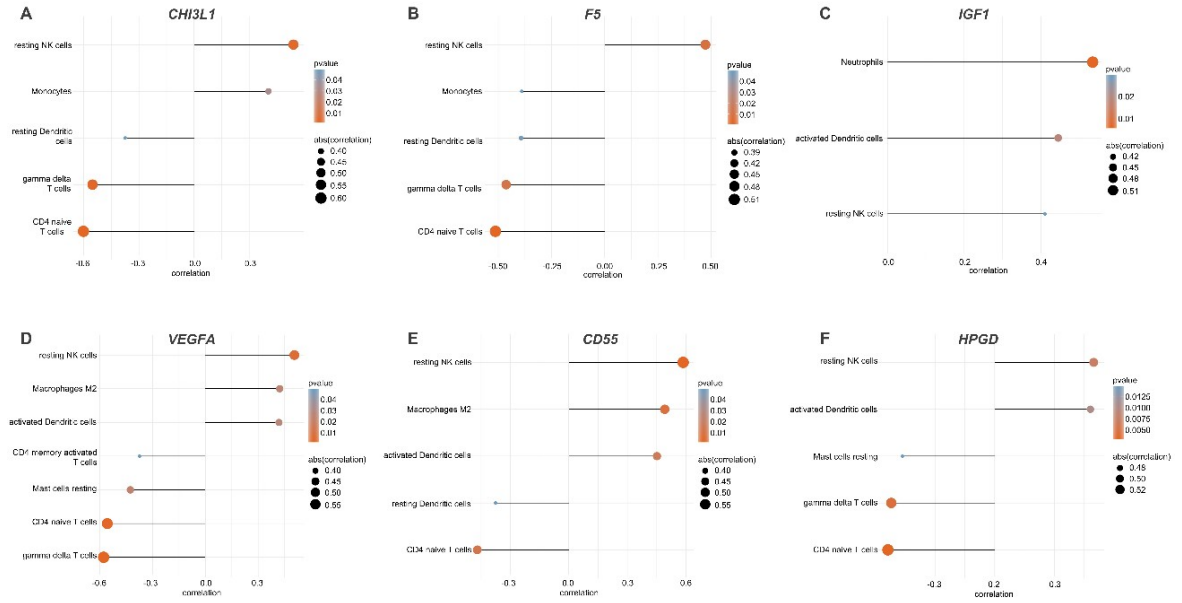

Fig. 5. Correlations between the hub genes and infiltrated immune cells in glomeruli in diabetic kidney disease (DKD). Lollyplots illustrate the correlations between different immune cells infiltrating the glomerular tissue in DKD patients and six hub genes: (A) chitinase-3-like protein 1 (*CHI3L1*), (B) coagulation factor V (*F5*), (C) insulin-like growth factor 1 (*IGF1*), (D) vascular endothelial growth factor A (*VEGFA*), (E) decay-accelerating factor (*CD55*), and (F) 15-hydroxyprostaglandin dehydrogenase (*HPGD*). Each panel corresponds to one of the hub genes and displays the strength of their positive (red) or negative (blue) association with various immune cell types in patients with DKD.

CHI3L1, *F5*, *VEGFA*, and *HPGD* showed negative correlations with gamma delta T-cells (Fig. 5A–F). These results suggest that hub genes may play a role in regulating immune infiltration and sculpting the immune environment in DKD.

Chemokines and their receptors play critical roles in mediating immune cell infiltration in DKD [24]. Specifically, C-C motif chemokine ligand 19 (*CCL19*) and its receptor C-C motif chemokine receptor 7 (*CCR7*) are key factors in immune inflammation, with previous reports indicating the upregulation of *CCL19* in DKD [22,25]. The interaction between C-C motif chemokine ligand 5 (*CCL5*) and C-C motif chemokine receptor 5 (*CCR5*) is known to regulate the migration of monocytes and macrophages, exacerbating inflammation-induced damage to endothelial cells [21]. To better understand the mechanisms by which hub genes regulate immune cell infiltration, we examined the correlation between these hub genes, chemokines, and their receptors using the TISIDB database (http://cis.hku.hk/TISIDB/). The analysis revealed that chemokines such as *CCL5* and *CC19*, along with their corresponding receptors *CCR5* and *CCR7* were negatively correlated with all six hub genes (Fig. 6A,B).

Enrichment of Signaling Pathways and Regulatory Networks Involving Hub Genes in DKD

We applied single-gene GSEA to investigate the molecular mechanisms contributing to DKD progression by examining the enrichment of specific signaling pathways associated with each hub gene. The mRNA expression matrix of patients with DKD was categorized based on the expression levels of *CHI3L1*, *F5*, *IGF1*, *CD55*, *VEGFA*, and *HPGD*. The GSEA analysis revealed the ubiquitin-mediated proteolytic pathway as a common signaling pathway within the ranked gene matrix for DKD (Fig. 7A–F). This result indicates that inflammation-related hub genes in DKD may be controlled by a shared regulatory network.

Subsequently, we examined the TFs and miRNAs that potentially regulate these hub genes. As expected, these hub genes appeared to be regulated by multiple TFs. We then conducted additional enrichment analyses for these TFs, as demonstrated by the cumulative recovery curves (Fig. 8A). The results detailing all enriched motifs and their corresponding TFs for the hub genes are shown in **Supplementary Table 2**. The cumulative recovery curves for the top three motifs with the highest NES are presented in Fig. 8B–D. Notably, among these motifs, cisbp_M6431 exhibited the highest NES of 6.22, with *F5*, *IGF1*, *VEGFA*, and *HPGD* being enriched within this motif (Fig. 8B, **Supplementary Table 2**). We also performed a reverse prediction of the six hub genes using the Mircode database, result-

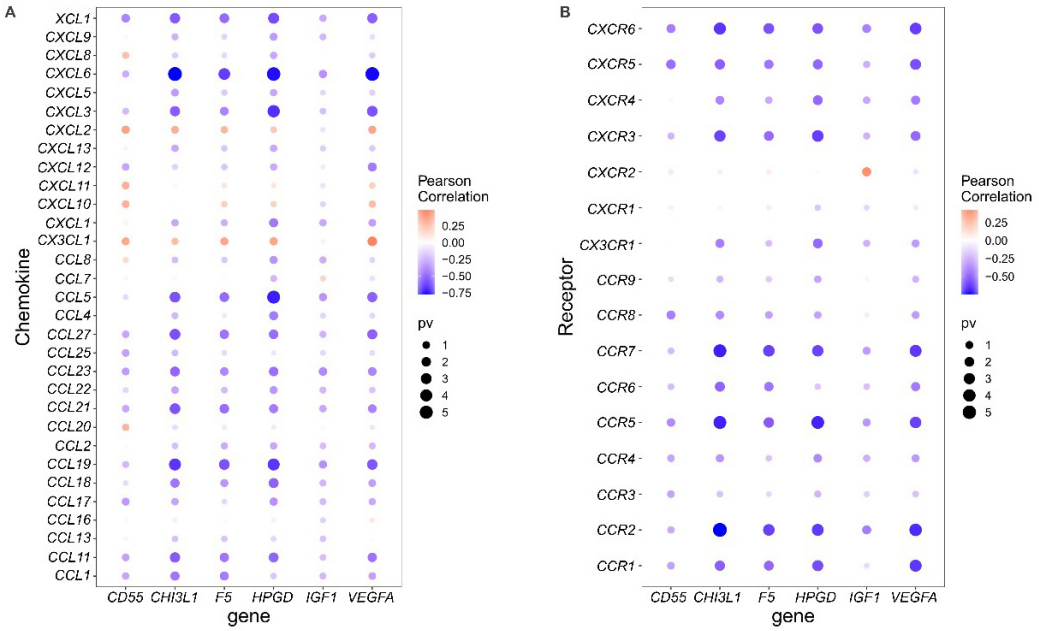

Fig. 6. Correlations between the hub genes and chemokines and receptors of chemokines. Bubble charts illustrate the correlations between hub genes and an array of (A) chemokines and (B) their receptors in diabetic kidney disease. The size and placement of each bubble indicate the magnitude and direction, either positive or negative, of the correlation between the respective hub gene and chemokine or receptor. Additionally, the color coding of the bubbles distinguishes between positive (red) and negative (blue) correlations. *XCL1*, X-C motif chemokine ligand 1; *CXCL*, C-X-C motif chemokine ligand; *CCL*, C-C motif chemokine ligand; *CXCR*, C-X-C motif chemokine receptor; *CCR*, C-C motif chemokine receptor; *CHI3L1*, chitinase-3-like protein 1; *F5*, coagulation factor V; *IGF1*, insulin-like growth factor 1; *VEGFA*, vascular endothelial growth factor A; *CD55*, decay-accelerating factor; *HPGD*, 15-hydroxyprostaglandin dehydrogenase.

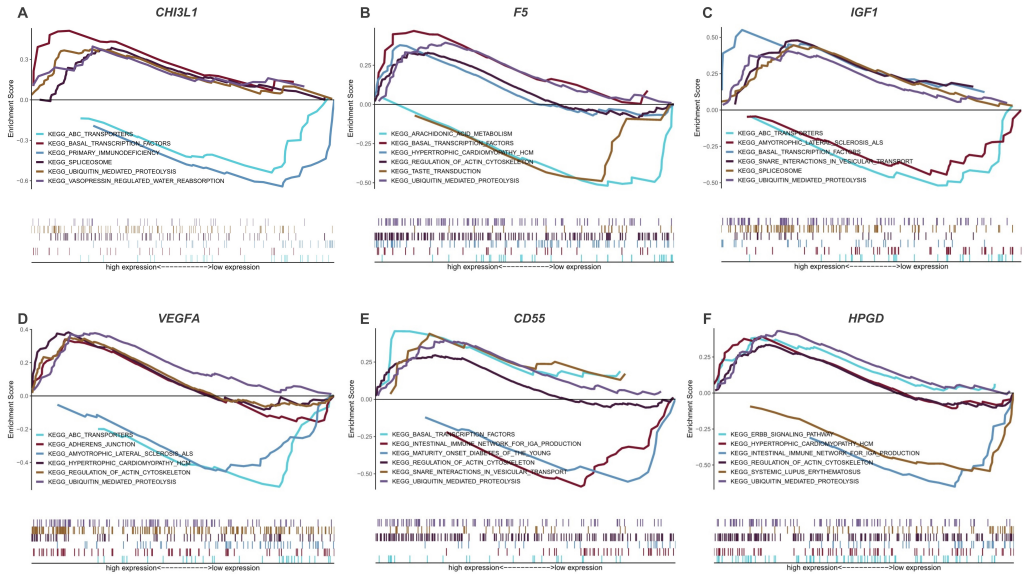

Fig. 7. Single-gene gene set enrichment analysis (GSEA) elucidating molecular mechanisms in diabetic kidney disease (DKD). Single-gene GSEA analysis was performed separately on (A) chitinase-3-like protein 1 (*CHI3L1*), (B) coagulation factor V (*F5*), (C) insulin-like growth factor 1 (*IGF1*), (D) vascular endothelial growth factor A (*VEGFA*), (E) decay-accelerating factor (*CD55*), and (F) 15-hydroxyprostaglandin dehydrogenase (*HPGD*) to explore their respective roles in the progression of DKD. The figure was constructed to display the enrichment scores and pathways for each gene, providing insights into the potential mechanistic influences of these genes on DKD pathophysiology. The ubiquitin-mediated proteolytic pathway was enriched as a common pathway among all hub genes.

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

ing in 80 miRNAs and 271 mRNA-miRNA pairs. These associations were visualized in Cytoscape (**Supplementary Fig. 5**). Among the miRNAs analyzed, *Let-7* was identified to regulate *CD55*, *VEGFA*, *IGF1*, and miR-146a, influencing the expressions of *F5*, *VEGFA*, and *HPGD* (**Supplementary Fig. 5**).

Discussion

Our study conducts a detailed exploration of the inflammatory landscape within DKD, uncovering potential molecular drivers and therapeutic avenues. We identified 349 DEGs, including 99 upregulated and 250 downregulated genes, which distinguish patients with DKD from HC. Through GO and KEGG analyses, we have uncovered significant enrichment in inflammation-related processes such as 'regulation of cell-cell adhesion' and 'positive regulation of leukocyte activation'. These processes are crucial for immune cell activation and infiltration, with pathways like 'cytokine-cytokine receptor interaction' and 'complement and coagulation cascades' implicated. The leukocyte infiltration—including neutrophils, macrophages, dendritic cells, T and B lymphocytes, and mast cells—has been shown to contribute to kidney damage in diabetes [3,13]. Given the role of cytokines in mediating both innate and adaptive immune responses and the involvement of the complement system in DKD progression [2,6,8], these findings underscore the complex inflammatory pathophysiology of DKD.

DKD is the leading cause of ESRD globally, characterized by its multifaceted and intricate mechanisms [26]. The significant emphasis on the role of inflammation in DKD is supported by evidence of altered immune cells and elevated levels of pro-inflammatory markers [3,12–14]. The increase in inflammatory cytokines in the initial stages suggests the potential of inflammation-related markers for predicting DKD [6,27,28].

In our pursuit of precision medicine, we have developed a predictive model using an inflammation-related gene signature. Previous models have incorporated a wide array of predictors, ranging from patient demographics to detailed metabolic data, including factors such as age, ethnicity, antidiabetic medication, hypertension, diabetic retinopathy, eGFR, albuminuria, as well as parameters like prothrombin time, platelet large cell ratio, fibrinogen levels, and red blood cell distribution width [29–34]. Our approach integrates six inflammation-related DEGs identified through RNA sequencing analysis of glomerular tissue obtained from T2DM patients in stages 1 to 3. Despite the study's limited sample size, our model demonstrated exceptional predictive accuracy, achieving an AUC of 1 in the training set and a notable AUC of 0.76 in the validation set. This innovative model, the first to utilize inflammation-related DEGs for DKD prediction, competes well with existing models primarily designed for patients with T2DM. Other models reported AUC values ranging from 0.74 to 0.86 (with an average of 0.81) [29–34]. However, further research is needed to expand these predictive models to Type 1 diabetics and validate our findings across larger and more diverse patient cohorts. Additionally, exploring the expression of these hub genes in more accessible samples such as blood and urine could offer broader clinical utility.

Our analysis has shed light on immune cell distribution differences between DKD patients and healthy controls, enhancing our understanding of DKD's immune response nuances [21,35,36]. The correlations we identified between DEGs and specific immune cell types, such as gamma delta T cells and CD4 naive T cells, reveal potential mechanistic pathways that could influence the progression of DKD. Therapeutic interventions aimed at restraining the activation or mitigating the impacts of these cells may help reduce inflammation-induced kidney damage [37]. Consequently, exploring the therapeutic impacts of antibodies or small molecules that specifically target signals critical for the activation or functional roles of gamma delta T cells in DKD holds considerable interest. Potential strategies include utilizing antibodies against T-cell receptors or *CD27*, as well as employing small molecule inhibitors like phosphoinositide 3-kinase inhibitors or protein kinase C inhibitors to modulate gamma delta T cell activity [38]. Additionally, our findings highlight an intriguing trend in macrophage population dynamics, providing insights into the mechanisms underlying early-stage renal damage response and fibrosis development in DKD.

The study further elaborates on the complex regulatory network of immune infiltration in DKD, highlighting a notable negative correlation between hub genes and the *CCL5-CCR5* axis. These findings, combined with the identification of shared signaling pathways such as ubiquitin-mediated proteolysis, suggest novel molecular dysregulation factors pivotal to the progression of DKD [39,40].

The elucidation of a comprehensive regulatory network comprising both transcriptional and post-transcriptional elements, including TFs and miRNAs, provides profound insights into the complex nature of gene regulation in DKD. This intricate network underscores the importance of developing highly precise and nuanced therapeutic strategies capable of effectively navigating the labyrinth of gene expression dynamics characteristic of DKD.

The identification of specific motifs through motif enrichment analysis, revealing TFs collaboratively modulating key hub genes, unveils potential therapeutic targets. The discovery of highly enriched motifs, particularly implicating genes like *F5*, *IGF1*, and *VEGFA*, suggests these motifs as likely binding sites for critical TFs such as *PPARA* and *PPARG* [41,42]. Given the significant role of these TFs in mediating inflammation, a hallmark of diabetes pathology, their involvement accentuates the potential for targeting these transcription factors and their associated motifs to alleviate the inflammatory processes in DKD [41,42].

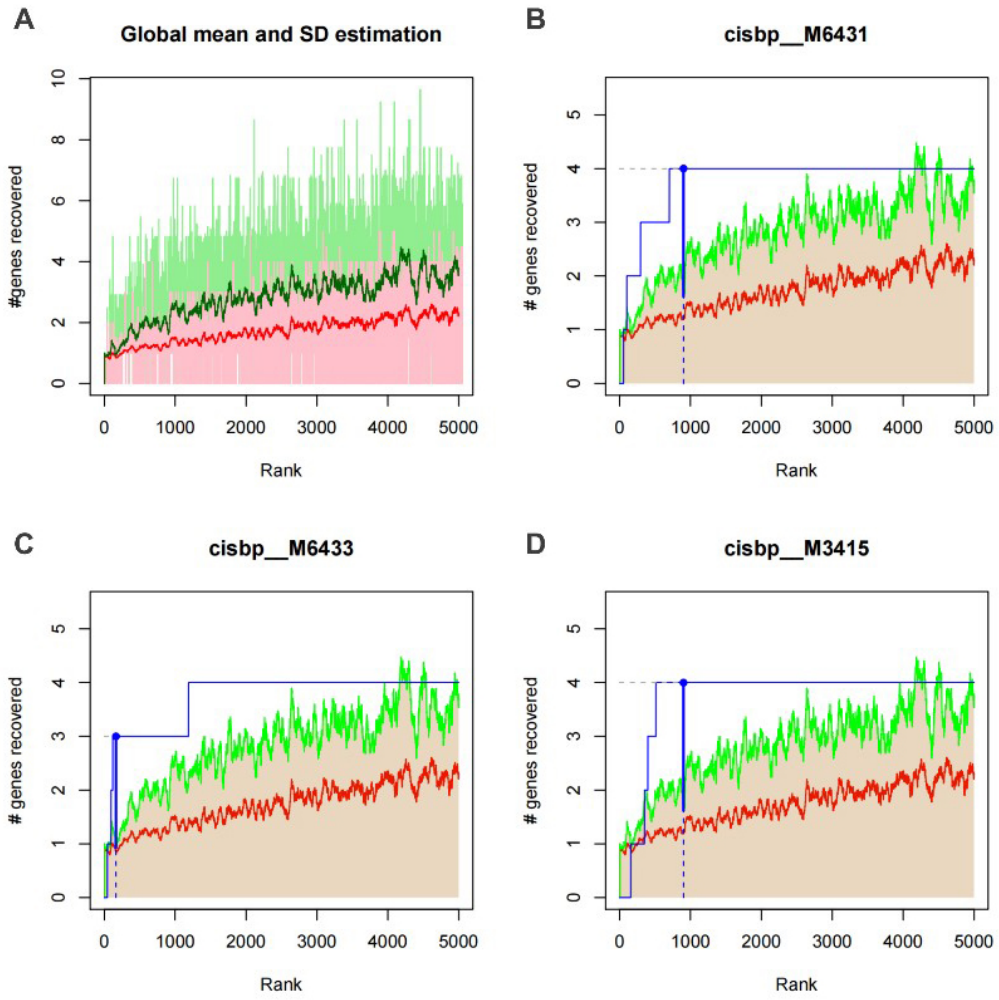

Fig. 8. The regulatory network analysis of hub genes. (A) The enrichment patterns of transcription factors (TFs) for six hub genes central to the pathophysiology of diabetic kidney disease (DKD) were presented. These patterns were illustrated using cumulative recovery curves, a statistical tool that provides insight into how frequently a certain TF is associated with the selected gene set compared to a random distribution. The steeper the slope at these points, the stronger the association between the motif and the hub genes. (B–D) The top three motifs with the highest normalized enrichment scores were (B) cisbp_M6431, (C) cisbp_M6433, and (D) cisbp_M3415, respectively. These motifs indicate potential TF binding sites that could play a central role in the gene regulation processes pertinent to the progression of DKD.

Moreover, the significant role of miRNAs in governing renal fibrosis, a critical aspect of DKD progression, highlights an additional layer of regulatory complexity. The observed downregulation of *Let-7* and its impact on genes such as *CD55*, *VEGFA*, and *IGF1* introduce a targetable pathway for potentially mitigating TGF-β-induced renal fibrosis in DKD [43,44]. Conversely, the decrease in miR-146 within the context of DKD and its predicted regulatory impact on genes like *F5*, *VEGFA*, and *HPGD* further elucidate a nuanced pathway where therapeutic interventions could potentially suppress inflammation and oxidative stress, cornerstones in the progression of DKD [45].

These findings suggest a potential shift towards precision medicine in DKD management. This approach involves customizing therapies to target specific elements within this regulatory network. By focusing on the modulation of specific TFs and miRNAs implicated in DKD pathogenesis, particularly those influencing inflammation and renal fibrosis, precision medicine approaches could provide more effective and personalized treatment modalities. The ability to address the disease at a molecular level, targeting the precise genetic and epigenetic alterations driving DKD, could significantly improve therapeutic outcomes by attenuating disease progression and improving patient quality of life. This shift towards precision medicine, supported by a comprehensive understanding of the regulatory networks in DKD, marks a new era in managing this complex condition, offering promising prospects for more effective, targeted, and individualized therapies.

Journal of
BIOLOGICAL REGULATORS
and Homeostatic Agents

While this study has made valuable contributions, it is important to note its limitations. These include the relatively small cohort size and the exclusive focus on type 2 DKD datasets, which may limit its generalizability. Moreover, the use of computational methods to estimate immune cell infiltration instead of direct experimental validation, along with the logistical challenges in obtaining renal tissues, highlight the need for additional *in vivo* studies to thoroughly validate these findings. This is particularly crucial in understanding the involvement of gamma delta T cells.

Conclusions

This study offers an extensive overview of the genetic changes in DKD, emphasizing the pivotal role of inflammation in its progression. By establishing a solid foundation for future research, the interaction of genetic insights and immune microenvironment deserves further exploration. The clinical application of our findings depends on continuous refinements to our predictive models and the validation of suggested biomarkers and molecular targets. Ultimately, integrating bioinformatics advancements with clinical evaluation will determine the influence of our findings on improving care for DKD patients.

Availability of Data and Materials

The R codes and other detailed methods used in this article can be obtained by contacting the corresponding authors if needed.

Author Contributions

MX, DT, and YD designed the study; MX, YL, LY, and TZ collected and analyzed the data; YP, ZX, and SL visualized the data; MX drafted the original manuscript; DT and LY requested fundings for this work; DT, LY, and YD supervised this work; YL, TZ, YP, ZX, SL, DT, LY, and YD reviewed and edited the paper. All authors contributed to important editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

The layout of the figures by Mr. Hao Tian is gratefully acknowledged.

Funding

This work was supported by the Shenzhen Science and Technology Program (NO. JCYJ20210324113214039 and NO. GJHZ20220913142614027), Guangzhou Entrepreneurship Leading Team (No. 202009030005).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.23812/j.biol.regul.homeost.agents.20243809.465.

References

[1] Tommerdahl KL, Shapiro ALB, Nehus EJ, Bjornstad P. Early microvascular complications in type 1 and type 2 diabetes: recent developments and updates. Pediatric Nephrology (Berlin, Germany). 2022; 37: 79–93.

[2] Tang SCW, Yiu WH. Innate immunity in diabetic kidney disease. Nature Reviews. Nephrology. 2020; 16: 206–222.

[3] Zhou W, Liu Y, Hu Q, Zhou J, Lin H. The landscape of immune cell infiltration in the glomerulus of diabetic nephropathy: evidence based on bioinformatics. BMC Nephrology. 2022; 23: 303.

[4] Pérez-Morales RE, Del Pino MD, Valdivielso JM, Ortiz A, Mora-Fernández C, Navarro-González JF. Inflammation in Diabetic Kidney Disease. Nephron. 2019; 143: 12–16.

[5] Xiao M, Tang D, Luan S, Hu B, Gong W, Pommer W, *et al*. Dysregulated coagulation system links to inflammation in diabetic kidney disease. Frontiers in Clinical Diabetes and Healthcare. 2023; 4: 1270028.

[6] Hofherr A, Williams J, Gan LM, Söderberg M, Hansen PBL, Woollard KJ. Targeting inflammation for the treatment of Diabetic Kidney Disease: a five-compartment mechanistic model. BMC Nephrology. 2022; 23: 208.

[7] Jiao Y, Jiang S, Wang Y, Yu T, Zou G, Zhuo L, *et al*. Activation of complement C1q and C3 in glomeruli might accelerate the progression of diabetic nephropathy: Evidence from transcriptomic data and renal histopathology. Journal of Diabetes Investigation. 2022; 13: 839–849.

[8] Tan SM, Snelson M, Østergaard JA, Coughlan MT. The Complement Pathway: New Insights into Immunometabolic Signaling in Diabetic Kidney Disease. Antioxidants & Redox Signaling. 2022; 37: 781–801.

[9] Li L, Chen L, Zang J, Tang X, Liu Y, Zhang J, *et al*. C3a and C5a receptor antagonists ameliorate endothelial-myofibroblast transition via the Wnt/β-catenin signaling pathway in diabetic kidney disease. Metabolism: Clinical and Experimental. 2015; 64: 597–610.

[10] Kong L, Andrikopoulos S, MacIsaac RJ, Mackay LK, Nikolic-Paterson DJ, Torkamani N, *et al*. Role of the adaptive immune system in diabetic kidney disease. Journal of Diabetes Investigation. 2022; 13: 213–226.

[11] Morigi M, Perico L, Corna D, Locatelli M, Cassis P, Carminati CE, *et al*. C3a receptor blockade protects podocytes from injury in diabetic nephropathy. JCI Insight. 2020; 5: e131849.

[12] Chen J, Liu Q, He J, Li Y. Immune responses in diabetic nephropathy: Pathogenic mechanisms and therapeutic target. Frontiers in Immunology. 2022; 13: 958790.

[13] Fu J, Sun Z, Wang X, Zhang T, Yuan W, Salem F, *et al*. The single-cell landscape of kidney immune cells reveals transcriptional heterogeneity in early diabetic kidney disease. Kidney International. 2022; 102: 1291–1304.

[14] Li HD, You YK, Shao BY, Wu WF, Wang YF, Guo JB, *et al*. Roles and crosstalks of macrophages in diabetic nephropathy. Frontiers in Immunology. 2022; 13: 1015142.
[15] Cantero-Navarro E, Rayego-Mateos S, Orejudo M, Tejedor-Santamaria L, Tejera-Muñoz A, Sanz AB, *et al*. Role of Macrophages and Related Cytokines in Kidney Disease. Frontiers in Medicine. 2021; 8: 688060.
[16] Bermejo-Martin JF, Martín-Fernandez M, López-Mestanza C, Duque P, Almansa R. Shared Features of Endothelial Dysfunction between Sepsis and Its Preceding Risk Factors (Aging and Chronic Disease). Journal of Clinical Medicine. 2018; 7: 400.
[17] Baelde HJ, Eikmans M, Doran PP, Lappin DWP, de Heer E, Bruijn JA. Gene expression profiling in glomeruli from human kidneys with diabetic nephropathy. American Journal of Kidney Diseases: the Official Journal of the National Kidney Foundation. 2004; 43: 636–650.
[18] Woroniecka KI, Park ASD, Mohtat D, Thomas DB, Pullman JM, Susztak K. Transcriptome analysis of human diabetic kidney disease. Diabetes. 2011; 60: 2354–2369.
[19] Pan Y, Jiang S, Hou Q, Qiu D, Shi J, Wang L, *et al*. Dissection of Glomerular Transcriptional Profile in Patients With Diabetic Nephropathy: SRGAP2a Protects Podocyte Structure and Function. Diabetes. 2018; 67: 717–730.
[20] Rayego-Mateos S, Rodrigues-Diez RR, Fernandez-Fernandez B, Mora-Fernández C, Marchant V, Donate-Correa J, *et al*. Targeting inflammation to treat diabetic kidney disease: the road to 2030. Kidney International. 2023; 103: 282–296.
[21] Zhao N, Dang H, Ma L, Martin SP, Forgues M, Ylaya K, *et al*. Intratumoral γδ T-Cell Infiltrates, Chemokine (C-C Motif) Ligand 4/Chemokine (C-C Motif) Ligand 5 Protein Expression and Survival in Patients With Hepatocellular Carcinoma. Hepatology (Baltimore, Md.). 2021; 73: 1045–1060.
[22] Chen H, Zhang Z, Zhou L, Cai T, Liu B, Wang L, *et al*. Identification of CCL19 as a Novel Immune-Related Biomarker in Diabetic Nephropathy. Frontiers in Genetics. 2022; 13: 830437.
[23] Liu Y, Guo Q, Yang H, Zhang XW, Feng N, Wang JK, *et al*. Allosteric Regulation of IGF2BP1 as a Novel Strategy for the Activation of Tumor Immune Microenvironment. ACS Central Science. 2022; 8: 1102–1115.
[24] Jung SW, Moon JY. The role of inflammation in diabetic kidney disease. The Korean Journal of Internal Medicine. 2021; 36: 753–766.
[25] Raju R, Gadakh S, Gopal P, George B, Advani J, Soman S, *et al*. Differential ligand-signaling network of CCL19/CCL21-CCR7 system. Database: the Journal of Biological Databases and Curation. 2015; 2015: bav106.
[26] Hoogeveen EK. The epidemiology of diabetic kidney disease. Kidney and Dialysis. 2022; 2: 433–442.
[27] Milas O, Gadalean F, Vlad A, Dumitrascu V, Velciov S, Gluhovschi C, *et al*. Pro-inflammatory cytokines are associated with podocyte damage and proximal tubular dysfunction in the early stage of diabetic kidney disease in type 2 diabetes mellitus patients. Journal of Diabetes and its Complications. 2020; 34: 107479.
[28] Liu GW, Zeng JE, Li LF. Correlation analysis of serum IGF-1 and IL-6 and urinary albumin/creatinine ratio in patients with type 2 diabetic kidney disease. Frontiers in Endocrinology. 2022; 13: 1082492.
[29] Sabanayagam C, He F, Nusinovici S, Li J, Lim C, Tan G, *et al*. Prediction of diabetic kidney disease risk using machine learning models: A population-based cohort study of Asian adults. eLife. 2023; 12: e81878.
[30] Mu X, Wu A, Hu H, Zhou H, Yang M. Prediction of Diabetic Kidney Disease in Newly Diagnosed Type 2 Diabetes Mellitus. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy. 2023; 16: 2061–2075.
[31] Liu XZ, Duan M, Huang HD, Zhang Y, Xiang TY, Niu WC, *et al*. Predicting diabetic kidney disease for type 2 diabetes mellitus by machine learning in the real world: a multicenter retrospective study. Frontiers in Endocrinology. 2023; 14: 1184190.
[32] Pan L, Wo M, Xu C, Wu Y, Ye Y, Han F, *et al*. Predictive significance of joint plasma fibrinogen and urinary alpha-1 microglobulin-creatinine ratio in patients with diabetic kidney disease. PloS One. 2022; 17: e0271181.
[33] Allen A, Iqbal Z, Green-Saxena A, Hurtado M, Hoffman J, Mao Q, *et al*. Prediction of diabetic kidney disease with machine learning algorithms, upon the initial diagnosis of type 2 diabetes mellitus. BMJ Open Diabetes Research & Care. 2022; 10: e002560.
[34] Dong Z, Wang Q, Ke Y, Zhang W, Hong Q, Liu C, *et al*. Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. Journal of Translational Medicine. 2022; 20: 143.
[35] Liu Y, Lv Y, Zhang T, Huang T, Lang Y, Sheng Q, *et al*. T cells and their products in diabetic kidney disease. Frontiers in Immunology. 2023; 14: 1084448.
[36] Chang TT, Chen JW. The Role of Chemokines and Chemokine Receptors in Diabetic Nephropathy. International Journal of Molecular Sciences. 2020; 21: 3172.
[37] Markle JGM, Mortin-Toth S, Wong ASL, Geng L, Hayday A, Danska JS. γδ T cells are essential effectors of type 1 diabetes in the nonobese diabetic mouse model. Journal of Immunology (Baltimore, Md.: 1950). 2013; 190: 5392–5401.
[38] Zhao Y, Niu C, Cui J. Gamma-delta (γδ) T cells: friend or foe in cancer development? Journal of Translational Medicine. 2018; 16: 3–16.
[39] Yang Y, Xiao H, Lin Z, Chen R, Li S, Li C, *et al*. The ubiquitination of CKIP-1 mediated by Src aggravates diabetic renal fibrosis (original article). Biochemical Pharmacology. 2022; 206: 115339.
[40] Guo F, Song Y, Wu L, Zhao Y, Ma X, Wang J, *et al*. SUMO specific peptidase 6 regulates the crosstalk between podocytes and glomerular endothelial cells in diabetic kidney disease. Biochimica et Biophysica Acta. Molecular Basis of Disease. 2023; 1869: 166685.
[41] Sharma V, Patial V. Peroxisome proliferator-activated receptor gamma and its natural agonists in the treatment of kidney diseases. Frontiers in Pharmacology. 2022; 13: 991059.
[42] Zhou L, Li Y, He W, Zhou D, Tan RJ, Nie J, *et al*. Mutual antagonism of Wilms' tumor 1 and β-catenin dictates podocyte health and disease. Journal of the American Society of Nephrology: JASN. 2015; 26: 677–691.
[43] Tang J, Yao D, Yan H, Chen X, Wang L, Zhan H. The Role of MicroRNAs in the Pathogenesis of Diabetic Nephropathy. International Journal of Endocrinology. 2019; 2019: 8719060.
[44] Wang T, Zhu H, Yang S, Fei X. Let 7a 5p may participate in the pathogenesis of diabetic nephropathy through targeting HMGA2. Molecular Medicine Reports. 2019; 19: 4229–4237.
[45] Wan RJ, Li YH. MicroRNA 146a/NAPDH oxidase4 decreases reactive oxygen species generation and inflammation in a diabetic nephropathy model. Molecular Medicine Reports. 2018; 17: 4759–4766.