



OPEN ACCESS

EDITED BY

Ferdinando Carlo Sasso,
University of Campania Luigi Vanvitelli,
Italy

REVIEWED BY

Alfredo Caturano,
University of Campania Luigi Vanvitelli,
Italy
Raffaele Galiero,
University of Campania Luigi Vanvitelli,
Italy

*CORRESPONDENCE

Han Li
✉ hanli@ccmu.edu.cn

SPECIALTY SECTION

This article was submitted to
Clinical Diabetes,
a section of the journal
Frontiers in Public Health

RECEIVED 23 January 2023

ACCEPTED 13 March 2023

PUBLISHED 30 March 2023

CITATION

Zhang C, Li H and Wang S (2023) Common
gene signatures and molecular mechanisms of
diabetic nephropathy and metabolic syndrome.
Front. Public Health 11:1150122.
doi: 10.3389/fpubh.2023.1150122

COPYRIGHT

© 2023 Zhang, Li and Wang. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Common gene signatures and molecular mechanisms of diabetic nephropathy and metabolic syndrome

Chengyu Zhang, Han Li* and Shixiang Wang

Department of Nephrology, Beijing Chao-Yang Hospital, Capital Medical University, Beijing, China

Background: Diabetic nephropathy (DN) is the leading cause of end-stage renal disease. Multiple metabolic toxicities, redox stress, and endothelial dysfunction contribute to the development of diabetic glomerulosclerosis and DN. Metabolic syndrome (MetS) is a pathological state in which the body's ability to process carbohydrates, fats, and proteins is compromised because of metabolic disorders, resulting in redox stress and renal remodeling. However, a causal relationship between MetS and DN has not been proven. This study aimed to provide valuable information for the clinical diagnosis and treatment of MetS with DN.

Methods: Here, transcriptome data of DN and MetS patients were obtained from the Gene Expression Omnibus database, and seven potential biomarkers were screened using bioinformatics analysis. In addition, the relationship between these marker genes and metabolism and immune infiltration was explored. Among the identified marker genes, the relationship between *PLEKHA1* and the cellular process, oxidative phosphorylation (OXPHOS), in DN was further investigated through single-cell analysis.

Results: We found that *PLEKHA1* may represent an important biomarker that perhaps initiates DN by activating B cells, proximal tubular cells, distal tubular cells, macrophages, and endothelial cells, thereby inducing OXPHOS in renal monocytes.

Conclusion: Overall, our findings can aid in further investigation of the effects of drug treatment on single cells of patients with diabetes to validate *PLEKHA1* as a therapeutic target and to inform the development of targeted therapies.

KEYWORDS

transcription data, single-cell analysis, type 2 diabetes, oxidative phosphorylation, diabetic nephropathy

1. Introduction

Metabolic syndrome (MetS) is a group of medical conditions, including abdominal obesity, high blood pressure, high blood sugar, high triglyceride levels, and low high-density lipoprotein cholesterol levels, that increases the risk of developing cardiovascular disease and type 2 diabetes (1). MetS, independent of other covariates, is a predictor of declining renal function and worsening of albuminuria in patients with type 2 diabetes (2). Nephropathy remains a major cause of morbidity and a key determinant of mortality in patients with type 1 or type 2 diabetes mellitus (3, 4). Mitochondrial fatty acid β -oxidation is the preferred process for generating adenosine triphosphate (ATP) in the kidney, and its dysfunction results in ATP depletion and

lipotoxicity, leading to tubular injury, inflammation, and subsequent progression of fibrosis (5).

The relationship between MetS and diabetic nephropathy (DN) is complex and bidirectional; while DN is considered a common progressive disease, MetS can be inhibited and it contributes to DN development and progression. The possible mechanisms of renal injury include insulin resistance, oxidative stress, increased production of pro-inflammatory cytokines, increased production of connective tissue growth and fibrotic factors, increased microvascular injury, and renal ischemia. MetS promotes kidney injury. However, despite the strong association between MetS and DN, a causal relationship has not yet been proven. Therefore, there is an urgent need to discover a range of novel biological markers for DN and MetS.

This study aimed to provide valuable information for the clinical diagnosis and treatment of MetS with DN. To this end, we analyzed transcriptomic RNA data from patients with DN and MetS and single-cell RNA sequencing (scRNA-seq) data from patients with DN to study their gene expression profiles and subsequently explore common biological markers of DN and MetS.

2. Materials and methods

2.1. Data sources

The keywords “diabetic nephropathy” and “metabolic syndrome” were used to search for DN and MetS gene expression profiles, respectfully, in the Gene Expression Omnibus (GEO) database (Home – GEO – NCBI [nih.gov]). Furthermore, the keywords “diabetic nephropathy” and “scRNA” were used to search for the single-cell dataset of DN. Finally, the GEO datasets GSE30529, GSE99340, and GSE98895 were selected. Four GEO datasets were selected — GSE30529, GSE98895, GSE99340, and GSE131882 (6–9). Information regarding the four datasets, including GSE numbers, detection platforms, samples, and types of RNA sources, is summarized in [Supplementary material 1](#). GSE30529 and GSE98895 were also paired as a discovery cohort for weighted gene co-expression network analysis (WGCNA) and GSE99340 and GSE131882 as validated datasets for differential gene expression analysis. The gene expression profiles were then transformed, and the probes were matched to their gene symbols according to the annotation documents of the corresponding platforms. Finally, gene matrices with row names designating sample names and column names designating gene symbols were obtained for subsequent analysis.

The DN scRNA-seq dataset GSE131882 was downloaded from the GEO database, which included data for three control groups and three patients with DN (10). The original dataset contains data for 31,286 cells. The percentages of mitochondria and rRNA were then calculated using the PercentageFeatureSet function, with the genes expressed being >200 and <2,500, respectively. The selection criteria are shown in [Supplementary Figure S1](#), while the workflow is shown in [Supplementary Figure S2](#).

2.2. Identification of differentially expressed genes

Data normalization and probe annotation were performed on datasets GSE30529 and GSE98895 using the limma and GEOquery

packages of R software version 4.1.3, with adjusted $p < 0.05$ and $|\log FC| > 1$ as the DGE screening criteria (11). Subsequently, common DEGs for DN and MetS were obtained, and network enrichment analysis was performed using Metascape.¹

2.3. Co-expression modules in DN and MetS

WGCNA is an algorithm that identifies co-expressed gene modules with high biological significance and explores the relationship between gene networks and diseases. WGCNA was used here to obtain DN- and MetS-associated modules. A total of 111,107 genes in the GEO dataset were obtained from the sequencing data and were used for WGCNA. The WGCNA package in R version 4.1.3 was used to perform the analysis. The appropriate soft-threshold powers (β , ranging from 1 to 20) were then selected using the “pickSoftThreshold” function in the WGCNA package based on the standard scale-free network. The soft-threshold power value β and gene correlation matrix among all gene pairs calculated by Pearson analysis were used to build an adjacency matrix. The topological overlap matrix and the corresponding dissimilarity were then transformed from the adjacency matrix. Subsequently, a hierarchical clustering dendrogram was built, with similar gene expression profiles divided into different modules. Finally, the expression profiles of each module were summarized by the module eigengene, and the correlation between the module eigengene and clinical features was calculated (12). Finally, modules with high correlation coefficients were targeted in terms of clinical features, and the genes in these modules were selected for subsequent analysis.

2.4. Identification of shared and unique gene signatures in DN and MetS

Least absolute shrinkage and selection operator (LASSO) was applied to identify DEGs with independent pre-diagnostic values. Based on the highest lambda value selected by 1,000 cross-validations in LASSO, a set of diagnostic genes and their LASSO coefficients were defined (13).

2.5. Construction of the XGBoost model

XGBoost was used to select a few key genomic features to build prediction models. XGBoost comprises an ensemble of K regression trees $[T_1(X, Y) \dots T_k(X, Y)]$, where X is the feature vector and Y is the corresponding risk. Assuming that the dataset contains n examples and p features $D = \{(x_i, y_i) \mid |D| = n, x_i \in X, y_i \in Y\}$, the ensemble XGBoost model uses K trees to predict patient risk:

$$\hat{y}_i = \mathcal{O}(x_i) = \sum_{k=1}^K f_k(x_i), f_k \in \mathcal{F}$$

¹ <http://metascape.org>

where f represents the space of the regression tree, q is the structure of the tree, T is the number of leaves in each tree, and f_k represents the regression tree's structure q with weight w . This method was implemented using the XGBoost package in R. All other parameters in our study used the default values in the R package XGBoost (14). The genes selected by XGBoost were considered candidate genes related to DN and MetS.

2.6. Analysis of immune cell infiltration and metabolic pathway

The CIBERSORT (<https://cibersortx.stanford.edu/>) deconvolution algorithm was used to evaluate differential immune cell infiltration. CIBERSORT is an analysis tool that uses gene expression data to estimate the abundance of member cell types in a mixed cell population (15). The LM22 gene file provided by CIBERSORT was used to define and infer the relative proportions of 22 types of immune-infiltrating cells in both disease and normal gene expression data (16). The default signature matrix of 100 permutations was used in the algorithm for this study (17). To ensure confidence in the results, CIBERSORT uses Monte Carlo sampling to derive the deconvolution p -value for each sample, and, in this study, only data with p -values <0.05 were retained. The results were then visualized using the ggplot2 package in R, before performing a correlation analysis between the 22 immune cells and key genes using Spearman's rank correlation test. Both MetS and DN are related to metabolism, and the set of metabolic signatures in the Molecular Signatures Database was used as a reference for gene set enrichment analysis, where $p < 0.05$ and a false discovery rate $q < 0.05$ were considered to indicate significant enrichment (18).

2.7. scRNA-seq data clustering dimension reduction

The merged data were first normalized using log normalization. Simultaneously, all genes were scaled using the ScaleData function, while the RunPCA function was used to reduce the principal component analysis (PCA) dimension for the first 2000 highly variable genes screened above. $\text{Dim} = 20$ was chosen before clustering the cells through the "FindNeighbors" and "FindClusters" functions (resolution = 0.8; Supplementary Figure S3) to find the cell clusters (19, 20). The top 50 principal components were then selected to further reduce dimensionality using the Uniform Manifold Approximation and Projection (UMAP) method. UMAP is a method of data dimensionality reduction, which assumes that the available data samples are uniformly distributed in the topological space and that these limited data samples can be approximated and mapped to a low-dimensional space. The "FindAllMarkers" function was subsequently used to screen the marker genes of 22 subgroups with $|\log\text{FC}| = 0.5$ and $\text{min.pct} = 0.35$ (21). Finally, a corrected $p < 0.05$ was used to screen the marker genes.

2.8. Cell-type identification by estimating relative subsets of oxidative phosphorylation

OXPHOS is an electron transfer chain driven by substrate oxidation coupled to ATP synthesis through an electrochemical

transmembrane gradient. Hallmark_oxidative_phosphorylation is a pathway based on the input matrix of gene expression file that is used to accurately estimate metabolism in tissues. This approach was used to compare metabolic differences between different cells in different groups. Spearman correlation analysis was performed to explore the relationship between metabolism and various cells in the kidneys of DN patients. The ggplot2 software package was then used to visualize the differences and results of related analyses (14, 22–26).

2.9. Statistical analysis

R version 4.1.3 was used for statistical analysis, while Student's t -test was performed to assess significant differences among distinct groups. In addition, the glmnet R package was used for the LASSO and Cox regression analyses. p -values <0.05 indicated statistical significance ($*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$).

2.10. Ethics approval and informed consent

GEO is a public database, and ethical approval was obtained for the studies that collected the various data in the database. Users can freely download relevant data for use in research and publication of relevant articles. Our study was based on open-source data; therefore, there are no ethical issues or other conflicts of interest.

3. Results

3.1. Co-expression modules in DN and MetS

Three datasets (GSE30529, GSE98895, and GSE99340) were integrated here, before removing batch effects. A soft threshold of $\beta = 4$ was chosen for consistency with the scale-free network. A total of 18 modules were identified in GSE30529 and GSE98895 using WGCNA, with each color representing a different module. A module-trait heatmap was also constructed according to the Spearman correlation coefficient to evaluate the association between each module and disease (Figures 1A,B). In the heatmap, the pink module showed the strongest association with DN ($\text{cor} = 0.3$, $p = 5.6 \times 10^{-7}$; Figure 1C) and MetS ($\text{cor} = 0.48$, $p = 7.5 \times 10^{-17}$; Figure 1D), which included 268 genes (Supplementary material 2). Enriched clusters up to 100 showed that the functionally enriched WGCNA pink and turquoise modules were primarily involved in the following processes: metabolic, immune system, and cellular processes, as well as the regulation of biological processes and response to stimuli. The molecular pathway was related to cellular macromolecule catabolic process, RNA metabolism, signaling by Rho and mitochondrial Rho GTPases and RHOBTB3, and protein phosphorylation (Figures 1E,F). Figure 1G further validates the correlation of comorbid WGCNA modules with metabolism and immunity.

3.2. Common gene signatures in DN and MetS

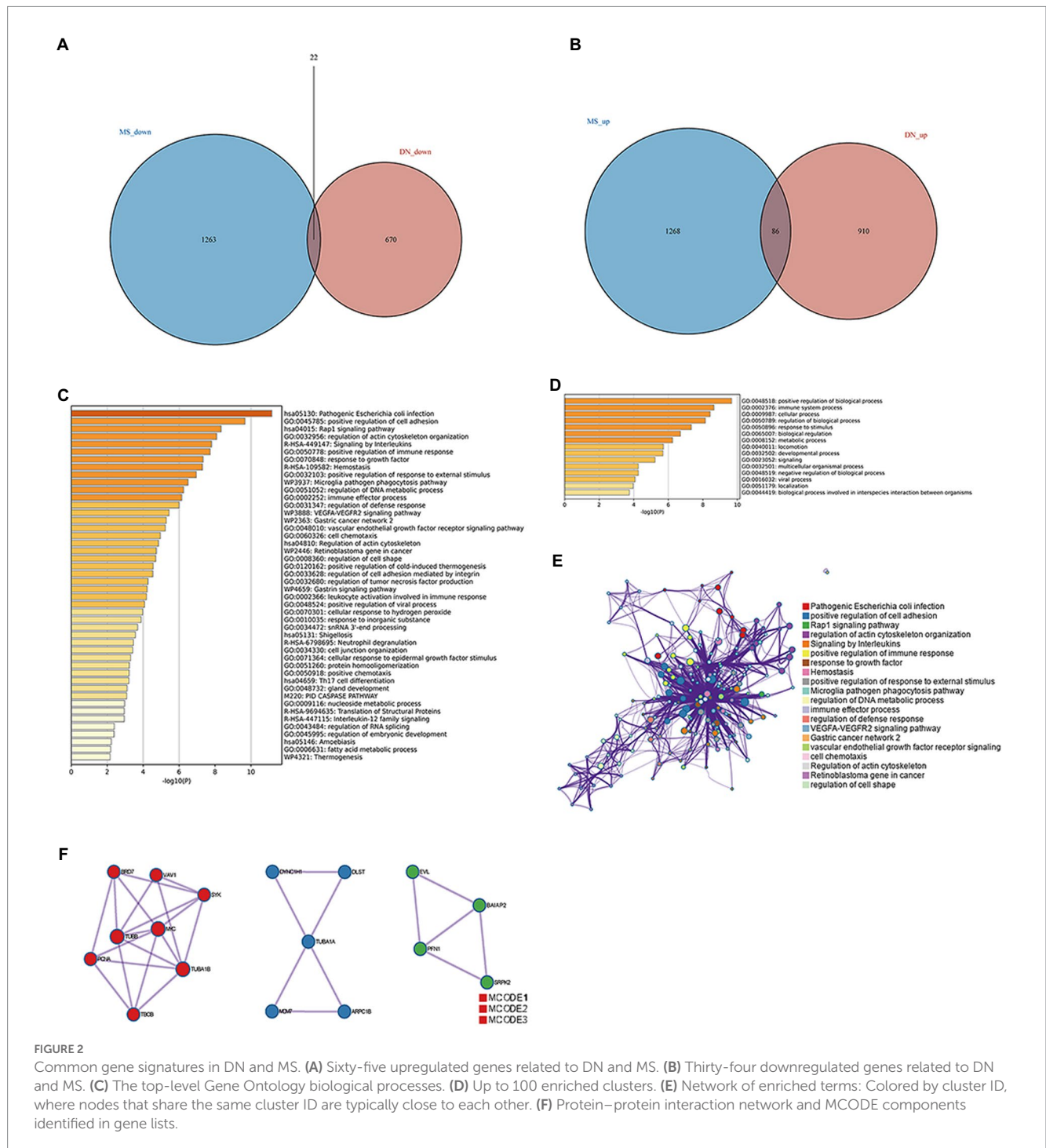
Up-and-down regulated DEGs were obtained for DN and MetS by setting the cut-off values at an adjusted p -value of 0.05 and

$|\log_2FC| > 1$ (27). Ultimately, 86 upregulated and 22 downregulated genes associated with DN and MetS were discovered using the intersection of the differential genes of the two disorders (Figures 2A,B; Supplementary materials 3, 4). These genes were subjected to functional enrichment and module analyses using Metascape. Based on enrichment clusters up to 100, the common gene functions of DN and MetS were primarily enriched as follows: immune system function, cellular function, and regulation of biological processes (Figures 2C–E). The common DN and MetS related to metabolism and immunity were further investigated using MCODE, with three core modules subsequently obtained: module 1 included 8 genes: SYK,

VAV1, BRD7, TUBB, MYC, TUBA1B, PCNA, and TBCB; module 2 included 5 genes: DYNC1H1, DLST, TUBA1A, MCM7, and ARPC1B; and module 3 included 4 genes: EVL, BAIAP2, SRPK2, and PFN (Figure 2F).

3.3. Screening for common biological markers of DN and MetS

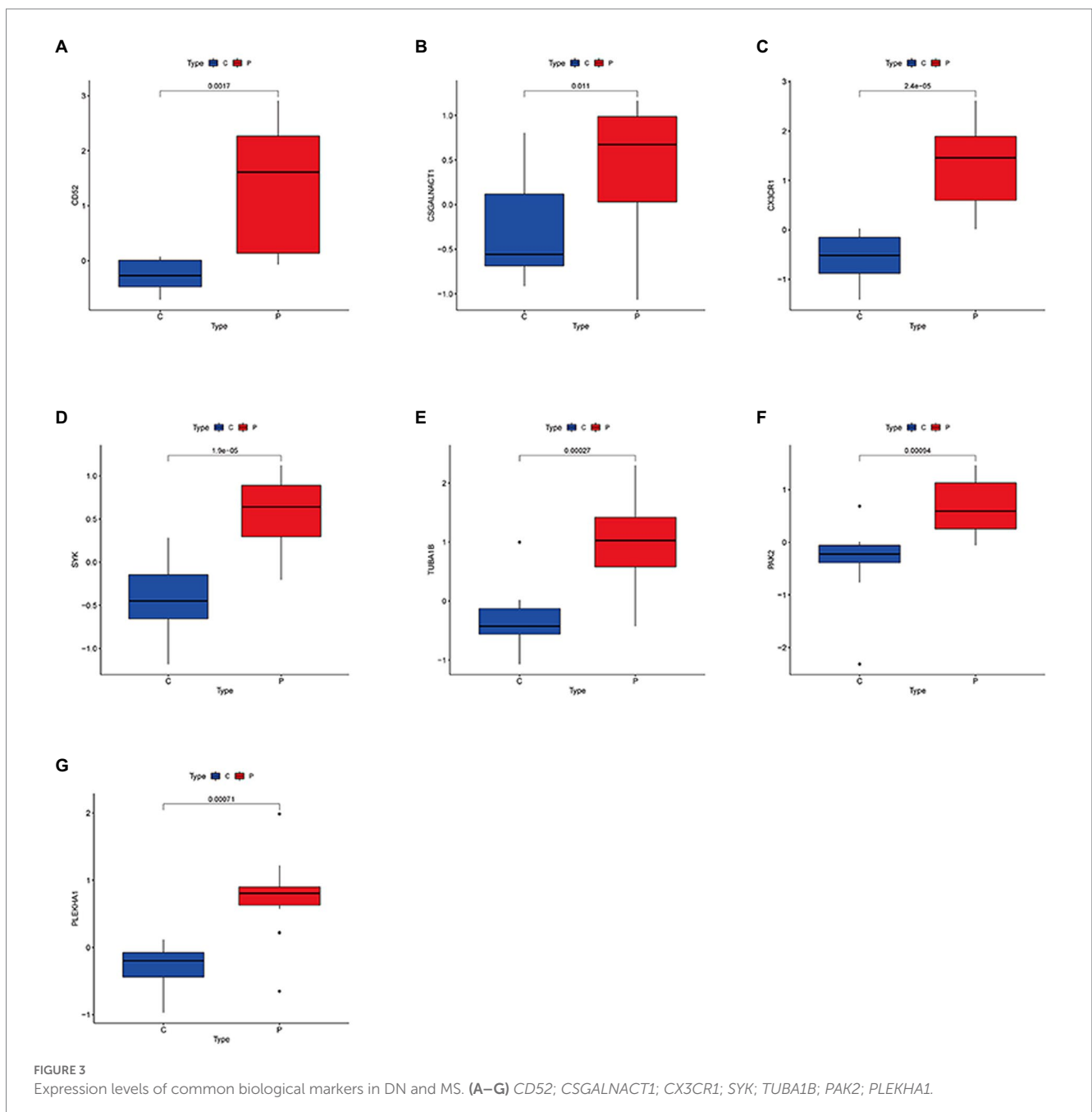
To further identify potential biomarkers for DN and MetS from the 108 target genes, Cox regression analysis of the LASSO data



was used to reduce the number of candidate genes. The change in trajectory of each gene is shown in [Supplementary Figure S4](#). Nine potential biological markers were obtained by LASSO regression, while seven potential biomarkers were obtained using the DN external validation dataset GSE99340. We also found that in GSE30529, seven genes were highly upregulated in DN samples ([Figures 3A–G](#)). As shown in [Supplementary Figure S5](#), receiver operating characteristic (ROC) curves of the seven potential biological markers were well adapted in both the training and validation datasets (28). The Coef index for the seven identified genes is presented in [Supplementary material 5](#). Ultimately, the seven genes were considered potential biomarkers for DN and MetS.

3.4. Building risk prediction models using XGBoost

The XGBoost method was used to construct a lightweight model of the seven potential biological markers. Both the areas under the ROC curve for the DN training set GSE30529 and the precision-recall (PR) curve were 1 ([Figures 4A,B](#)). Furthermore, the area under the ROC curve in the validation set GSE99340 was 0.778, while that under the PR curve was 0.955 ([Figures 4C,D](#)). Simultaneously, XGBoost curve fitting was performed on the GSE988985 MetS dataset, yielding solid curve-fitting results (ROC=0.995, PR=0.995; [Figures 4E,F](#)). These results suggest that the seven potential biological markers exhibited favorable diagnostic efficacy for both MetS and DN.



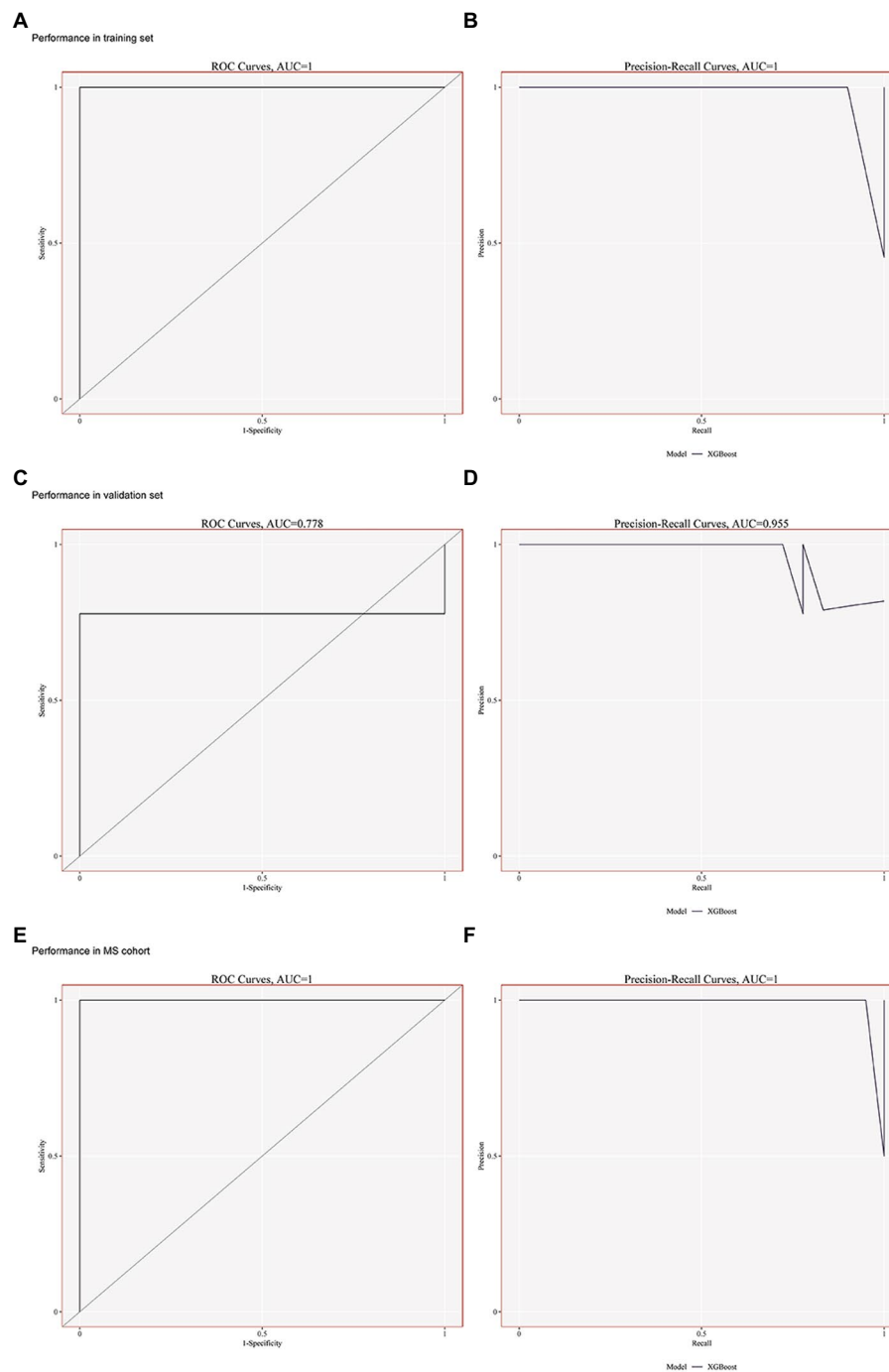


FIGURE 4

Building risk prediction models by XGBoost. (A–B) Performance of the ROC and PR curves in GSE30529 (training set). (C–D) Performance of the ROC and PR curves in GSE99340 (validation set). (E–F) Performance of ROC and PR curves in GSE98895 (MS cohort).

3.5. Immune cell infiltration and metabolic pathway analysis

There was a difference in the composition of immune cells. Tumor microenvironment (TME) analysis results indicated the presence of immune infiltration in DN GSE30529, primarily memory B cells (MBCs), CD8⁺ T cells, $\gamma\delta$ T cells, resting memory CD4⁺ T cells, and M₀ macrophages (Supplementary Figure S6A). TME was grouped according to DN and control groups, with the results showing immune

infiltration in DN GSE30529, mainly natural killer (NK) cells, monocytes, M₁ macrophages, dendritic cells, and neutrophils (Supplementary Figure S6B). The immune-infiltrating cells in DN GSE30529 were predominantly macrophages, T cells, and MBCs (Supplementary Figures S6C,D). There was also a correlation between immune-infiltrating cells, with the results showing that $\gamma\delta$ T cells and M₁ macrophages were both highly correlated ($cor=0.85$; Supplementary Figure S6E). The seven genes screened above were analyzed for their correlation with immune cells in the GSE30529 DN

dataset, with the results showing that *SYK* highly correlated with $\gamma\delta$ T cells and NK cells. *PLEKHA1* was also highly correlated with macrophages and resting memory CD4⁺ T cells (Supplementary Figure S6F). Furthermore, nine metabolic pathways were analyzed in the GSE30529 DN dataset. The upregulation of *SYK*, *CD52*, *CX3CR1*, *TUBA1B*, and *CSGALNACT1* led to the most noticeable metabolic abnormalities, which were reflected primarily in heme metabolism, bile acid metabolism, adipogenesis, OXPHOS, xenobiotic metabolism, and fatty acid metabolism (Supplementary Figure S7G). TME analysis also indicated the presence of immune infiltration in the GSE98895 MetS dataset, mainly monocytes, naïve CD4⁺ T cells, and CD8⁺ T cells (Supplementary Figure S7A). TME was grouped according to MetS and control groups, and the results showed immune infiltration in the GSE98895 MetS dataset, primarily monocytes, CD8⁺ T cells, and naïve CD4⁺ T cells (Supplementary Figure S7B). Monocytes and T cells constituted the major proportion of immune-infiltrating cells in the GSE98895 MetS dataset (Supplementary Figures S7C,D). GSE98895 immune-infiltrating cells were correlated, and the results showed that M₂ macrophages and dendritic cells were highly correlated ($\text{cor}=0.84$; Supplementary Figure S7E). In the GSE98895 MetS dataset, *CX3CR1* was associated with immune infiltration, with elevated *CX3CR1* levels leading to NK cell and monocyte activation. Furthermore, elevated expression levels of *PAK2* activated monocytes (Supplementary Figure S7F). *PLEKHA1*, *TUBA1B*, *CX3CR1*, *PAK2*, and *CD52* were the most highly correlated genes in the GSE98895 MetS dataset. *TUBA1B* expression was also elevated, resulting in OXPHOS, glycolysis, adipogenesis, xenobiotic metabolism, and fatty acid metabolism (Supplementary Figure S7G). Furthermore, *CX3CR1* expression was elevated, leading to increased OXPHOS, glycolysis, bile acid metabolism, xenobiotic metabolism, and fatty acid metabolism (Supplementary Figure S4G). *PAK2* expression was also elevated, leading to increased OXPHOS, xenobiotic metabolism, and fatty acid metabolism; moreover, *CD52* expression was elevated, leading to increased OXPHOS (Supplementary Figure S7G). Finally, *TUBA1B* expression was elevated, which led to increased OXPHOS, glycolysis, adipogenesis, xenobiotic metabolism, and fatty acid metabolism (Supplementary Figure S7G).

3.6. Definition of clusters and dimensionality reduction for visual representation of cells

The “ScaleData” function was implemented to scale all genes extracted from the scRNA-seq dataset GSE131882 and PCA dimensionality reduction was performed to identify anchor points. A total of 21 clusters were identified (Figure 5A; Supplementary material 6). These identified clusters were then labeled as various cell types using the top five differential marker genes. Cell marker genes were downloaded from PanglaoDB² and the Human Protein Atlas.³ Ultimately, 11 cell types were defined (Figure 5B), and after grouping into DN and blank controls, a single-cell analysis was

performed (Figure 5C). The five differential genes with the most prominent contributions were screened, as shown in Figure 5D.

3.7. Common biological markers for single-cell analysis

The relationship between potential DN genes and single cells of DN after screening by proportional values subsequently showed that *CSGALNACT1*, *PAK2*, *PLEKHA1*, and *SYK* were all related to DN. Figure 6A shows the proportional relationship between the seven potential biological markers and single kidney cells. *SYK*, *PLEKHA1*, *PKA2*, and *CSGALNACT1* were found to have the highest associations with DN OXPHOS. Figure 6B shows the expression relationship between the seven potential biomarkers and individual kidney cells, demonstrating that the above four genes were the most relevant for OXPHOS in DN.

3.8. Association between *PLEKHA1* and single-cell OXPHOS in DN

The relationship between upregulated *PLEKHA1* and the single-cell OXPHOS score for DN was presented in a violin plot; the results in Figures 7A–E show that upregulated *PLEKHA1* induced OXPHOS in B cells, proximal tubular cells, distal tubular cells, macrophages, and endothelial cells. Figure 7F also shows that the upregulation of *PLEKHA1* led to elevated levels of OXPHOS in DN. Furthermore, Supplementary material 7 presents single-cell OXPHOS fractions for various types of DN.

4. Discussion

Diabetes mellitus is the most common cause of chronic kidney disease worldwide and can lead to multiple complications, including end-stage renal disease, cardiovascular disease, infection, and ultimately death (29, 30). MetS is a common metabolic disorder arising from the increasing prevalence of obesity (1). In this study, differential genes associated with DN and MetS were found to be primarily enriched in immune system function, cellular function, and biological process regulation. One of the potential biological markers identified in this study, *PLEKHA1*, induces OXPHOS in B cells, proximal tubular cells, distal tubule cells, macrophages, and endothelial cells.

The induction of oxidative stress, inflammation, fibrosis, and apoptosis by the accumulation of metabolites is known as lipotoxicity (31). Recent evidence has indicated that both the quantity and quality of lipids are involved in renal damage associated with lipotoxicity by activating inflammation, oxidative stress, mitochondrial dysfunction, and cell death. The pathological effects of lipotoxicity have also been observed in renal cells and promote podocyte injury, tubular damage, mesangial proliferation, endothelial activation, and macrophage-derived foam cell formation (32). Metabolic abnormalities and OXPHOS are the predominant contributors to DN.

DN is also associated with elevated expression levels of OXPHOS-related genes and pathways (33). Mitochondria are the site of cellular respiration and generate energy as ATP through OXPHOS; therefore,

2 <https://panglaodb.se/index.html>

3 <https://www.proteinatlas.org/>

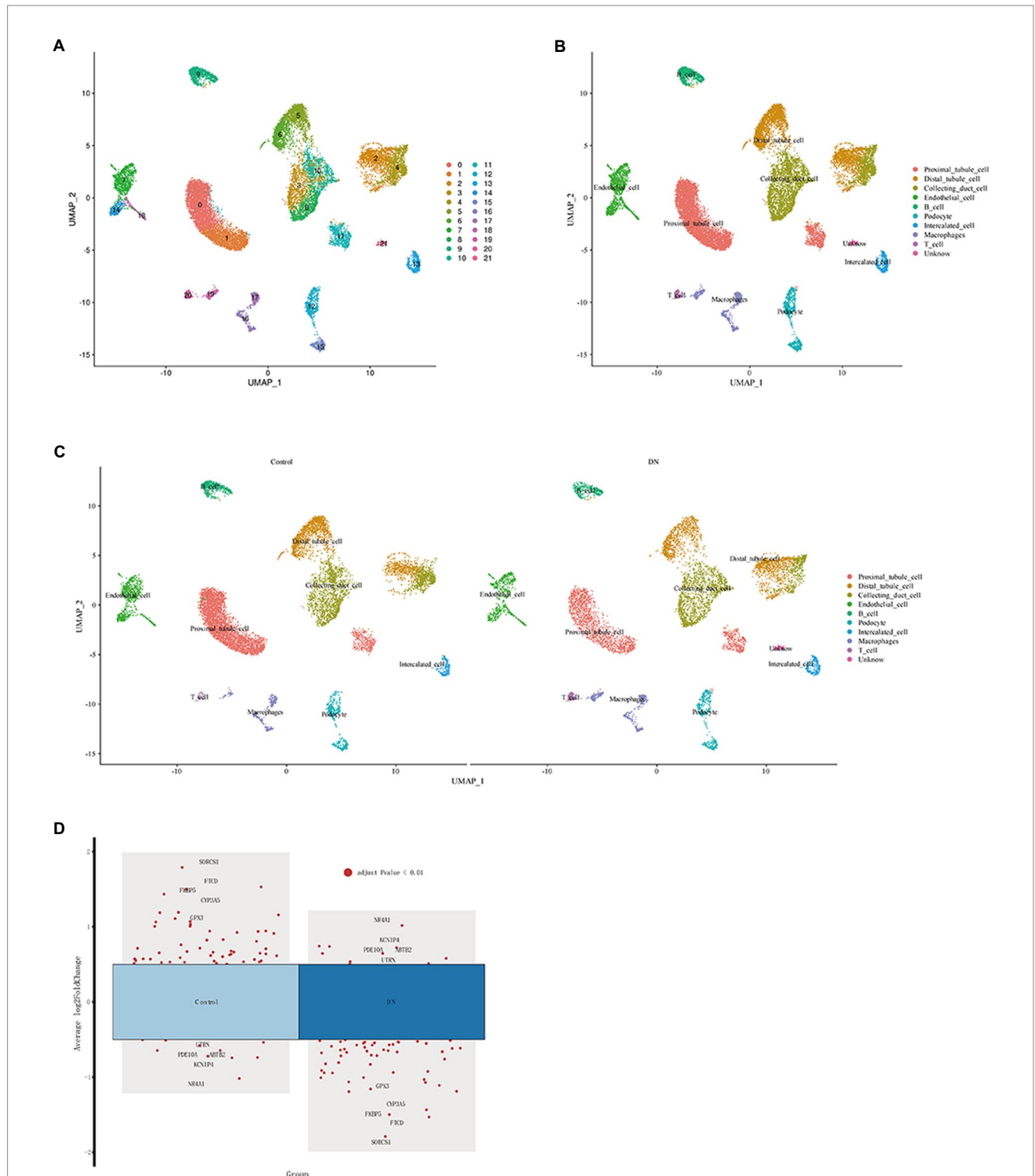


FIGURE 5 Definition of clusters and dimensionality reduction for visual representation of the cells. (A) UMAP of single cells from DN samples and normal samples before comments. (B) UMAP of single cells from DN samples and normal samples after the comment. (C) UMAP of single cells from DN samples and normal samples between DN and the control. (D) Single-cell analysis of differential genes in DN and control.

mitochondrial dysfunction has been implicated in DN. Mitochondria are also an important cellular source of reactive oxygen species (ROS) through the OXPHOS pathway (34, 35). Oxidative stress plays an important role in diabetic vascular complications (36), while hyperglycemia induces intracellular ROS production in interstitial

cells and diabetic kidneys (37, 38). Based on indirect evidence, it has been hypothesized that increased oxidative stress contributes to the development of DN. *Smad4* promotes DN by regulating glycolysis and OXPHOS (39). Furthermore, OXPHOS, electron transport system complex III, the citric acid cycle, propionate metabolism, and

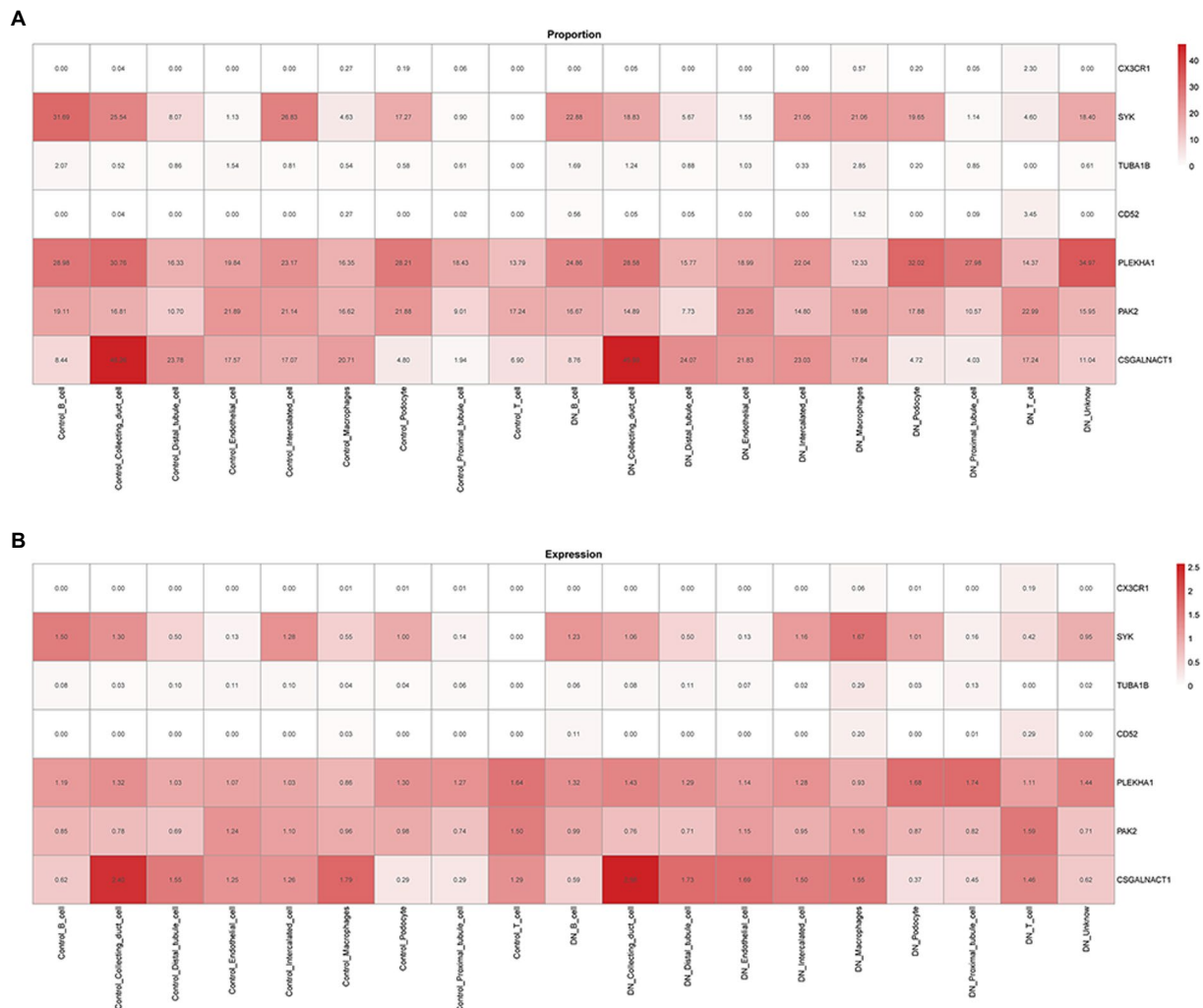


FIGURE 6 Relationship between biological markers. (A) The relationship between PLEKHA1 and CSGALNACT1. (B) The relationship between PLEKHA1 and PAK2.

transcription factors are all key contributors to metabolic abnormalities in DN (33). The paradigm that high glucose drives the overproduction of superoxide in mitochondria, as a unifying theory to explain end-organ damage in complications resulting from diabetes, has been upheld for more than a decade (40).

Considering its widespread prevalence and massive toll on health and finances, the diagnosis and management of DN are of great clinical and social relevance. Correlation and subgroup analyses of pivotal genes associated with the clinical features of DN showed that *ALB*, *ANXA1*, *APOH*, *C3*, *CCL19*, *COL1A2*, *COL3A1*, *COL4A1*, *COL6A3*, *CXCL6*, *DCN*, *EGF*, *HRG*, *KNG1*, *LUM*, *SERPINA3*, *SPARC*, *SRGN*, and *TIMP1* may all be involved in diabetic renal tubular interstitial injury (41). A urine transcriptome test also showed that urinary sediment *CCL5* and *CXCL1* mRNAs were upregulated in patients with DN, while being associated with a decline in renal function and the degree of renal interstitial fibrosis. Therefore, urinary sediment *CCL5* mRNA could be a potential prognostic biomarker of DN (42). Transcriptome analysis of osteoporosis in DN patients treated with traditional Chinese medicine showed that miR-574 may play an important role in DN-related osteoporosis, with the

therapeutic effects of kaempferol and quercetin on Leri-Weill dyschondrosteosis in DN-related osteoporosis potentially mediated by miR-574 by targeting *MAPK1* (43).

In this study, we first identified potential biological gene markers of MetS and DN before further analyzing the metabolic and immune pathways involved in these conditions. The role of *PLEKHA1* in OXPPOS, which is involved in DN, was investigated *via* single-cell analysis, and *PLEKHA1* was identified as a potential biological marker of DN worthy of further study.

SGLT2 inhibitors have potential as an antifibrotic therapeutic intervention by regulating inflammation, oxidative stress, mitochondrial function, and autophagy. GLP-1 agonists reduce inflammation and oxidative stress, which can damage the kidneys in people with diabetes. CCR2/5 inhibitors block the chemokine receptors CCR2/CCR5, reducing immune cell infiltration, cytokine production, and slowing DN progression (44, 45). However, more research is needed to verify their safety and efficacy in the treatment of fibrotic diseases (46). Our research highlighted *PLEKHA1* as a key gene in the development of DN. Future studies should investigate the effects of drug treatment on single cells of patients with diabetes to

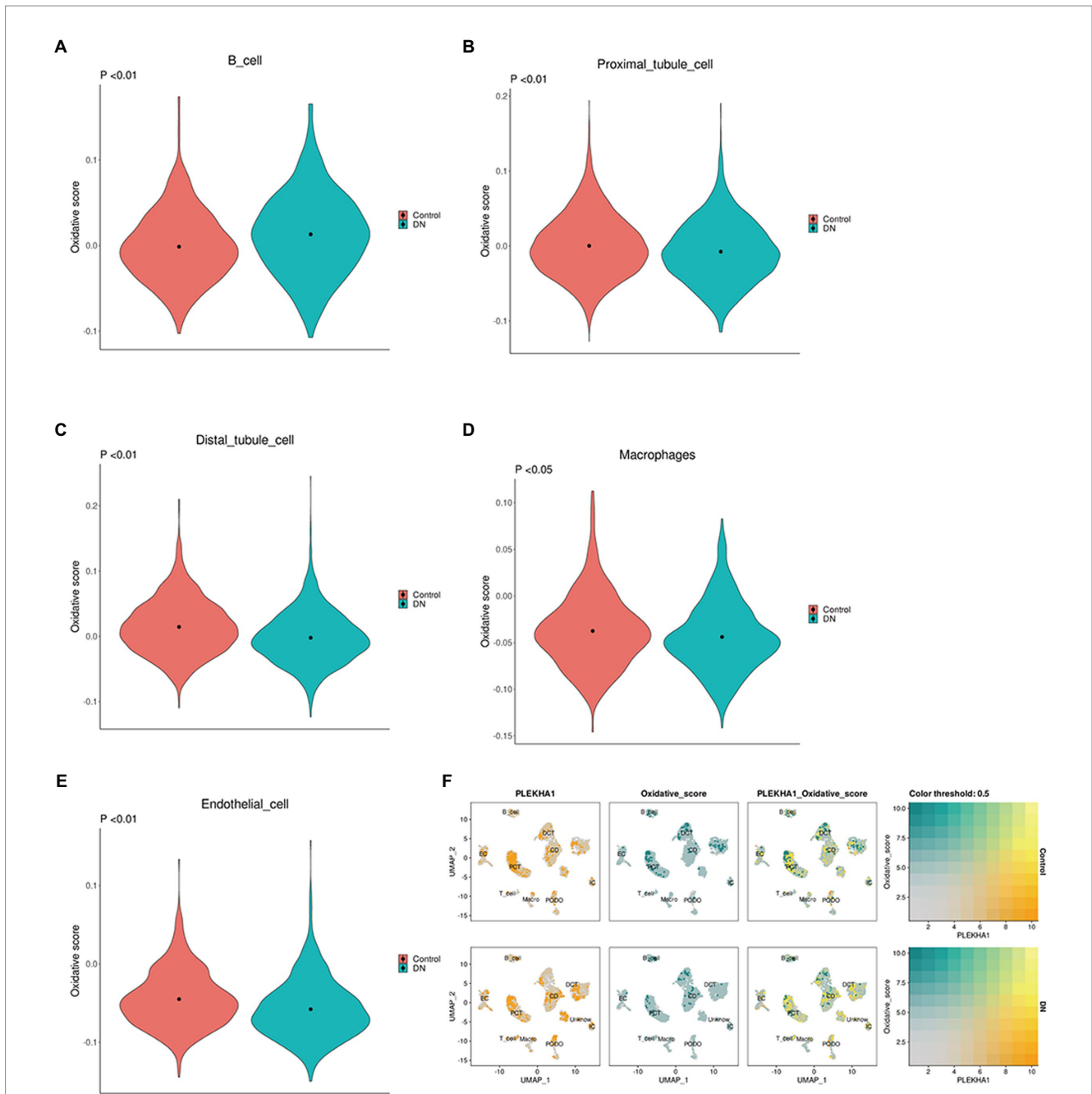


FIGURE 7 Association between the PLEKHA1 gene and single-cell oxidative phosphorylation. (A–E) Violin plot of potential biological markers and single-cell oxidative stress in DN: B cell. Proximal tubular cell. Distal tubular cell. Macrophages. Endothelial cell. (F) The relationship between PLEKHA1 and single-cell oxidative stress in DN was demonstrated by UMAP diagrams.

validate PLEKHA1 as a therapeutic target and inform the development of targeted therapies.

Our study of the shared biological markers of DN and MetS is critical to understanding the comorbidities between these conditions. To increase our sample size, we used transcriptome datasets of DN from the GEO database and adjusted for batch effects. We incorporated both transcriptome and single-cell analyses, allowing us to draw a definite conclusion. However, our study is not without limitations. We could not collect clinical data, which limited our ability to conduct further analysis on clinical prognosis. Furthermore, our study focused

solely on the genetic aspect and lacked a control treatment group. These limitations underscore the need for future research to focus on clinical applications and treatments to enhance the practical implications of our findings.

In conclusion, a functional enrichment analysis was performed based on the common patterns of differential gene expression between DN and MetS, and these differential genes were found to be related to immunity and metabolism. Further screening of potential biological markers resulted in the identification of seven well-fitting genes. Subsequently, the relationship between *PLEKHA1* and OXPHOS in DN

was further investigated by single-cell analysis. This study shows that overexpression of *PLEKHA1* in B cells, proximal tubular cells, distal tubular cells, macrophages, and endothelial cells caused oxidative stress in the kidneys. The role of *PLEKHA1* in OXPHOS, which is involved in DN, was investigated *via* single-cell analysis, and *PLEKHA1* was identified as a potential biological marker of DN worthy of further study.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: GSE30529, GSE98895, GSE99340, and GSE131882 GEO Accession viewer <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30529>. GEO Accession viewer <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98895>. GEO Accession viewer <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99340>. GEO Accession viewer <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE131882>.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

CZ performed the study, collected the data, and wrote the paper. HL conceived and designed the study and revised the paper. SW

curated the data and administered the project, and revised the paper. All authors contributed to the final manuscript and approved the submitted version.

Acknowledgments

We would like to thank the authors of the studies that generated the GSE30529, GSE98895, GSE99340, and GSE131882 datasets for their contributions.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1150122/full#supplementary-material>

References

- Fernandez ML. The metabolic syndrome. *Nutr Rev.* (2007) 65:S30–4. doi: 10.1111/j.1753-4887.2007.tb00325.x
- Chuang S-M, Shih H-M, Chien M-N, Liu S-C, Wang C-H, Lee C-C. Risk factors in metabolic syndrome predict the progression of diabetic nephropathy in patients with type 2 diabetes. *Diabetes Res Clin Pract.* (2019) 153:6–13. doi: 10.1016/j.diabres.2019.04.022
- Fineberg D, Jandeleit-Dahm KAM, Cooper ME. Diabetic nephropathy: diagnosis and treatment. *Nat Rev Endocrinol.* (2013) 9:713–23. doi: 10.1038/nrendo.2013.184
- Thorn LM, Forsblom C, Wadén J, Saraheimo M, Tolonen N, Hietala K, et al. Metabolic syndrome as a risk factor for cardiovascular disease, mortality, and progression of diabetic nephropathy in type 1 diabetes. *Diabetes Care.* (2009) 32:950–2. doi: 10.2337/dc08-2022
- Wilson PC, Wu H, Kirita Y, Uchimura K, Ledru N, Rennke HG, et al. The single-cell transcriptomic landscape of early human diabetic nephropathy. *Proc Natl Acad Sci U S A.* (2019) 116:19619–25. doi: 10.1073/pnas.1908706116
- GEO Accession viewer. Available at: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30529> (Accessed December 6, 2018).
- GEO Accession viewer. Available at: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98895> (Accessed July 25, 2021).
- GEO Accession viewer. Available at: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE99340> (Accessed July 25, 2021).
- GEO Accession viewer. Available at: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE131882> (Accessed May 6, 2021).
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* (2015) 43:e47. doi: 10.1093/nar/gkv007
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* (2008) 9:559. doi: 10.1186/1471-2105-9-559
- Alhamzawi R, Ali HTM. The Bayesian adaptive lasso regression. *Math Biosci.* (2018) 303:75–82. doi: 10.1016/j.mbs.2018.06.004
- Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol.* (2017) 13:e1005752. doi: 10.1371/journal.pcbi.1005752
- Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* (2015) 12:453–7. doi: 10.1038/nmeth.3337
- Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. "Profiling tumor infiltrating immune cells with CIBERSORT." In: Stechow L Von, editor. *Cancer systems biology*. New York, NY: Springer New York (2018). p. 243–259.
- Zhao E, Xie H, Zhang Y. Predicting diagnostic gene biomarkers associated with immune infiltration in patients with acute myocardial infarction. *Front Cardiovasc Med.* (2020) 7:586871. doi: 10.3389/fcvm.2020.586871
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. *Cell Syst.* (2015) 1:417–25. doi: 10.1016/j.cels.2015.12.004
- Der E, Ranabothu S, Suryawanshi H, Akat KM, Clancy R, Morozov P, et al. Single cell RNA sequencing to dissect the molecular heterogeneity in lupus nephritis. *JCI Insight.* (2017) 2:e93009. doi: 10.1172/jci.insight.93009
- He H, Suryawanshi H, Morozov P, Gay-Mimbrera J, Del Duca E, Kim HJ, et al. Single-cell transcriptome analysis of human skin identifies novel fibroblast subpopulation and enrichment of immune subsets in atopic dermatitis. *J Allergy Clin Immunol.* (2020) 145:1615–28. doi: 10.1016/j.jaci.2020.01.042
- The Accelerating Medicines Partnership Rheumatoid Arthritis and Systemic Lupus Erythematosus (AMP RA/SLE) Consortium Der E, Suryawanshi H, Morozov P, Kustagi M, Goilav B, et al. Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. *Nat Immunol.* (2019) 20:915–27. doi: 10.1038/s41590-019-0386-1

21. Ito K, Murphy D. Application of ggplot2 to Pharmacometric graphics. *Pharmacometrics Syst Pharmacol.* (2013) 2:e79. doi: 10.1038/psp.2013.56
22. Jia A, Xu L, Wang Y. Venn diagrams in bioinformatics. *Brief Bioinform.* (2021) 22:bbab108. doi: 10.1093/bib/bbab108
23. Download R-4.1.3 for Windows. The R-project for statistical computing. Available at: <https://cran.microsoft.com/snapshot/2022-03-25/bin/windows/base/> (Accessed March 25, 2022).
24. Friedman J, Hastie T, Tibshirani R, Narasimhan B, Tay K, Simon N, et al. Glmnet: Lasso and elastic-net regularized generalized linear models. (2022). Available at: <https://CRAN.R-project.org/package=glmnet> (Accessed February 21, 2023).
25. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, et al. xgboost: Extreme Gradient Boosting (2023). Available at: <https://CRAN.R-project.org/package=xgboost> (Accessed February 21, 2023).
26. Wickham H, Chang W, Henry L, Pedersen TL, Takahashi K, Wilke C, et al. ggplot2: Create elegant data Visualisations using the grammar of graphics (2023). Available at: <https://CRAN.R-project.org/package=ggplot2> (Accessed February 21, 2023).
27. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *Bioinformatics.* (2011) 12:77. doi: 10.1186/1471-2105-12-77
28. Afkarian M, Zelnick LR, Hall YN, Heagerty PJ, Tuttle K, Weiss NS, et al. Clinical manifestations of kidney disease among US adults with diabetes, 1988-2014. *JAMA.* (2016) 316:602. doi: 10.1001/jama.2016.10924
29. Gregg EW, Li Y, Wang J, Rios Burrows N, Ali MK, Rolka D, et al. Changes in diabetes-related complications in the United States, 1990-2010. *N Engl J Med.* (2014) 370:1514-23. doi: 10.1056/NEJMoal310799
30. De Boer IH. Temporal trends in the prevalence of diabetic kidney disease in the United States. *JAMA.* (2011) 305:2532. doi: 10.1001/jama.2011.861
31. Thongnak L, Pongchaidecha A, Lungkaphin A. Renal lipid metabolism and lipotoxicity in diabetes. *Am J Med Sci.* (2020) 359:84-99. doi: 10.1016/j.amjms.2019.11.004
32. Opazo-Ríos L, Mas S, Marín-Royo G, Mezzano S, Gómez-Guerrero C, Moreno JA, et al. Lipotoxicity and diabetic nephropathy: novel mechanistic insights and therapeutic opportunities. *IJMS.* (2020) 21:2632. doi: 10.3390/ijms21072632
33. Huang C, Kim Y, Caramori ML, Moore JH, Rich SS, Mychaleckyj JC, et al. Diabetic nephropathy is associated with gene expression levels of oxidative phosphorylation and related pathways. *Diabetes.* (2006) 55:1826-31. doi: 10.2337/db05-1438
34. Coughlan MT, Higgins GC, Nguyen T-V, Penfold SA, Thallas-Bonke V, Tan SM, et al. Deficiency in apoptosis-inducing factor recapitulates chronic kidney disease via aberrant mitochondrial homeostasis. *Diabetes.* (2016) 65:1085-98. doi: 10.2337/db15-0864
35. Suzuki S, Hinokio Y, Komatu K, Ohtomo M, Onoda M, Hirai S, et al. Oxidative damage to mitochondrial DNA and its relationship to diabetic complications. *Diabetes Res Clin Pract.* (1999) 45:161-8. doi: 10.1016/S0168-8227(99)00046-7
36. Baynes JW. Role of oxidative stress in development of complications in diabetes. *Diabetes.* (1991) 40:405-12. doi: 10.2337/diab.40.4.405
37. Han HJ, Lee YJ, Park SH, Lee JH, Taub M. High glucose-induced oxidative stress inhibits Na⁺/glucose cotransporter activity in renal proximal tubule cells. *Am J Physiol Renal Physiol.* (2005) 288:F988-96. doi: 10.1152/ajprenal.00327.2004
38. Kashihara N, Haruna Y, Kondeti VK, Kanwar YS. Oxidative stress in diabetic nephropathy. *Curr Med Chem.* (2010) 17:4256-69. doi: 10.2174/092986710793348581
39. Li J, Sun YBY, Chen W, Fan J, Li S, Qu X, et al. Smad4 promotes diabetic nephropathy by modulating glycolysis and OXPHOS. *EMBO Rep.* (2020) 21:e48781. doi: 10.15252/embr.201948781
40. Coughlan MT, Sharma K. Challenging the dogma of mitochondrial reactive oxygen species overproduction in diabetic kidney disease. *Kidney Int.* (2016) 90:272-9. doi: 10.1016/j.kint.2016.02.043
41. Zeng M, Liu J, Yang W, Zhang S, Liu F, Dong Z, et al. Multiple-microarray analysis for identification of hub genes involved in tubulointerstitial injury in diabetic nephropathy. *J Cell Physiol.* (2019) 234:16447-62. doi: 10.1002/jcp.28313
42. Feng S-T, Yang Y, Yang J-F, Gao Y-M, Cao J-Y, Li Z-L, et al. Urinary sediment CCL5 messenger RNA as a potential prognostic biomarker of diabetic nephropathy. *Clin Kidney J.* (2022) 15:534-44. doi: 10.1093/ckj/sfab186
43. Liu MM, Lv NN, Geng R, Hua Z, Ma Y, Huang GC, et al. Uncovering the key miRNAs and targets of the Liuwei Dihuang pill in diabetic nephropathy-related osteoporosis based on weighted gene co-expression network and network pharmacology analysis. *EMIDDT.* (2022) 22:274-89. doi: 10.2174/1871530321666210215161921
44. Rambacher KM, Moniri NH. The β 2-adrenergic receptor-ROS signaling axis: an overlooked component of β 2AR function? *Biochem Pharmacol.* (2020) 171:113690. doi: 10.1016/j.bcp.2019.113690
45. Hu Y, Zeng N, Ge Y, Wang D, Qin X, Zhang W, et al. Identification of the shared gene signatures and biological mechanism in type 2 diabetes and pancreatic cancer. *Front Endocrinol.* (2022) 13. doi: 10.3389/fendo.2022.847760
46. Salvatore T, Galiero R, Caturano A, Rinaldi L, Di Martino A, Albanese G, et al. An overview of the cardiorenal protective mechanisms of SGLT2 inhibitors. *IJMS.* (2022) 23:3651. doi: 10.3390/ijms23073651