Check for updates

Investigating the metabolic reprogramming mechanisms in diabetic nephropathy: a comprehensive analysis using bioinformatics and machine learning

Shan He[1], Yi Wei Chen[2], Jian Ye[1], Yu Wang[1]*, Qin Kai Chen[1]* and Si Yi Liu[1]*

[1]Department of Nephrology, The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang, China, [2]Department of Orthopaedics, Jiujiang University Affiliated Hospital, Jiujiang, China

Background: Diabetic nephropathy (DN) is a common complication of diabetes, characterized by damage to renal tubules and glomeruli, leading to progressive renal dysfunction. The aim of our study is to explore the key role of metabolic reprogramming (MR) in the pathogenesis of DN.

Methods: In our study, three transcriptome datasets (GSE30528, GSE30529, and GSE96804) were sourced from the Gene Expression Omnibus (GEO) database. These datasets were integrated for batch effect correction and subsequently subjected to differential expression analysis to identify differentially expressed genes (DEGs) between DN and control samples. The identified DEGs were cross-referenced with genes associated with MR to derive MR associated differentially expressed genes (MRRDEGs). These MRRDEGs underwent Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. To identify key genes and develop diagnostic models, four machine learning algorithms were employed in conjunction with weighted gene co-expression network analysis (WGCNA) and the protein interaction tool CytoHubba. Gene set enrichment analysis (GSEA) and CIBERSORT analysis were conducted on the key genes to assess immune cell infiltration in DN. Additionally, a competitive endogenous RNA (ceRNA) network was constructed using the key genes. Finally, the expression levels of core genes in human samples were validated through quantitative real-time PCR (qRT-PCR).

Results: We identified 256 MRRDEGs, highlighting metabolic and inflammatory pathways in DN. KEGG analysis linked these genes to the MAPK signaling pathway, suggesting its key role in DN. Six key genes were pinpointed using WGCNA, PPI, and machine learning, with their diagnostic value confirmed by ROC analysis. CIBERSORT revealed a strong link between these genes and immune cell infiltration, indicating the immune response's role in DN. GSEA showed these genes' involvement in inflammatory and metabolic processes. A ceRNA network was predicted to clarify gene regulation. qRT-PCR confirmed

the expression patterns of *CXCR2, NAMPT,* and *CUEDC2*, aligning with bioinformatics results.

Conclusion: Through bioinformatics analysis, a total of six potential MRRDEGs were identified, among which *CUEDC2, NAMPT, CXCR2* could serve as potential biomarkers.

1 Introduction

Diabetic nephropathy (DN) is a severe complication associated with diabetes mellitus (Thipsawat, 2021), characterized by progressive damage to the renal tubules and glomeruli, ultimately leading to a decline in kidney function (Levin et al., 2020). It is estimated that approximately 30%–40% of patients with diabetes will develop DN (Thomas et al., 2015), making it a significant contributor to end-stage renal disease (ESRD) and associated mortality (Gilbertson et al., 2005; Cross et al., 2021; Dw et al., 2025). Despite the implementation of blood glucose regulation and therapeutic measures aimed at diminishing urinary albumin excretion, the fundamental progression of DN remains largely unaltered (Thomas et al., 2015). Recently, researchers have been exploring new treatments like glucose stabilizers, kidney-protective agents, and therapies targeting inflammation and fibrosis (Ghose et al., 2024). SGLT2 inhibitors, for instance, help regulate blood glucose and may reduce kidney workload by decreasing glucose reabsorption (Neuen et al., 2024). Despite these advances, current treatments have limitations in slowing disease progression and cannot fully reverse DN. Understanding the molecular mechanisms behind DN could help identify new therapeutic targets and improve prevention and management strategies.

The pathogenesis of DN is intricate, with recent research highlighting the pivotal role of metabolic reprogramming (MR) in various diseases, such as tumors and metabolic disorders (Li and Liao, 2021; Huang et al., 2023). Metabolic reprogramming refers to the cellular adaptation to environmental changes through the modification of metabolic pathways under specific pathological conditions (Li et al., 2021). Recent studies have underscored the significance of MR in DN, demonstrating its close association with the onset and progression of the disease. Notably, significant alterations have been observed in energy and lipid metabolism (Wang et al., 2022a; Yu et al., 2025). A fundamental component of metabolic reprogramming in DN is the modification of mitochondrial function. Mitochondria are central to cellular energy metabolism, and their impaired function leads to increased oxidative stress, apoptosis, and dysregulated autophagy, thereby facilitating the advancement of renal fibrosis and dysfunction (Fan et al., 2024; Sharma et al., 2013). This dysfunction not only accelerates the progression of DN but also involves the reprogramming of lipid metabolism. Under diabetic conditions, renal cells exhibit significant lipid metabolic abnormalities, including increased lipid uptake, impaired fatty acid oxidation, disrupted cholesterol efflux, and enhanced lipid catabolism (Yu et al., 2025). These alterations result in the accumulation of lipids such as free fatty acids, diacylglycerol, and ceramides, which subsequently induce lipotoxicity, inflammation, and fibrosis (Yu et al., 2025). Additionally, the interplay between MR and DN is critical, as metabolic dysregulation not only exacerbates renal injury but also promotes the progression of the disease through mechanisms such as oxidative stress and inflammation (Hou et al., 2025).

Existing research has highlighted the critical role of MR in the pathological progression of DN. However, the regulatory mechanisms at the molecular level and the specific role of MR in DN remain insufficiently understood. Consequently, this study aims to elucidate the key characteristics and functional importance of MR in DN through bioinformatics analysis. We employed bioinformatics and machine learning methodologies to identify biomarkers associated with metabolic reprogramming in DN. Initially, we acquired the DN dataset from the GEO database and conducted differential gene expression analysis. This was followed by an intersection with genes related to metabolic reprogramming. We further integrated weighted gene co-expression network analysis (WGCNA), four distinct machine learning algorithms, and protein-protein interaction (PPI) network construction to pinpoint key differentially expressed genes associated with MR. Subsequently, a diagnostic model for DN was developed, and the diagnostic efficacy of the model and key genes was validated. Additionally, immune cell infiltration analysis was conducted to investigate the relationship between key genes and immune cell populations. The pivotal genes were employed to construct competitive endogenous RNA (ceRNA) networks. Finally, the expression levels of these key genes were experimentally validated. These findings have the potential to offer novel insights into the early diagnosis and therapeutic strategies for DN.

2 Materials and methods

2.1 Data acquisition

We procured gene expression data pertinent to diabetic nephropathy from the GEO database (Barrett et al., 2013), specifically datasets GSE30528 (Woroniecka et al., 2011), GSE30529 (Woroniecka et al., 2011), and GSE96804 (Pan et al., 2018; Shi et al., 2018), utilizing the GEO query package. The GSE30528 and GSE30529 dataset is derived from the GPL571 platform, while the GSE96804 dataset is based on the GPL17586 platform. The GSE30528 dataset comprises 13 control samples and 9 DN samples, whereas the GSE96804 dataset includes 20 healthy human renal tubular samples and 41 tubular samples from DN patients. Additionally, the GSE30529 dataset consists of 10 control samples and 12 DN samples. We employed the R package 'sva' (v3.50.0)

TABLE 1 GEO Dataset Information list.

Accession	GSE30528	GSE96804	GSE30529
Platform	GPL571	GPL17586	GPL571
Experiment type	Expression profiling by array	Expression profiling by array	Expression profiling by array
Species	*Homo sapiens*	*Homo sapiens*	*Homo sapiens*
Tissue	glomeruli	glomeruli	glomeruli
Samples in Control group	Control (13)	Control (20)	Control (10)
Samples in Disease group	DN (9)	DN (41)	DN (12)
Reference	Transcriptome analysis of human diabetic kidney disease	Dissection of Glomerular Transcriptional Profile in Patients With Diabetic Nephropathy: SRGAP2a Protects Podocyte Structure and Function	Transcriptome analysis of human diabetic kidney disease

DN, diabetic nephropathy.

to adjust for batch effects across the diabetic nephropathy datasets GSE30528, GSE30529, and GSE96804, resulting in a consolidated dataset, DN_Datasets, which encompasses 62 DN samples and 43 control samples. The dataset specifics are outlined in Table 1. Furthermore, we identified 1,468 Metabolic Reprogramming-Related Genes (MRRGs) from the GeneCards database (Stelzer et al., 2016) (http://www.genecards.org) by searching for protein-coding genes related to "metabolic reprogramming" with a relevance score >4.

2.2 Differentially expressed gene analysis

To identify the differentially expressed genes (DEGs) associated with DN, we employed the 'limma' (v3.58.1) package to conduct a differential analysis of the expression profile data from the DN_Datasets. The criteria for selecting DEGs were set at $|logFC| > 0.25$ and p. value <0.01. To further identify metabolic reprogramming-related differentially expressed genes (MRRDEGs), we intersected the DEGs derived from the DN_Datasets analysis with MRRGs. This intersection was visualized using a Venn diagram, and the resultant MRRDEGs were used for subsequent analyses. Additionally, a volcano plot was generated using the R package ggplot2 (v3.5.1) to illustrate the differential analysis results. For functional annotation, Gene Ontology (Mi et al., 2019) (GO) and Kyoto Encyclopedia of Genes and Genomes (Kanehisa and Goto, 2000) (KEGG) pathway analyses were conducted using the 'clusterProfiler' (v4.10.1) package (Yu et al., 2012), with an enrichment significance threshold set at p. adjust <0.05, corrected using the Benjamini–Hochberg (BH) method.

2.3 Weighted gene Co-Expression network analysis (WGCNA)

WGCNA is a computational algorithm designed to cluster genes into distinct modules and elucidate the relationships between

these modules and disease characteristics. To thoroughly investigate the genetic mechanisms underlying the pathogenesis of DN, a co-expression network was constructed utilizing the 'WGCNA' (v1.72-5) package (Langfelder and Horvath, 2008) in the R. This network was developed using the top 40% of genes with the highest variance from the DN_Datasets dataset. A dynamic tree cut method was employed to merge modules, applying a threshold of 0.15. Additional criteria for constructing the co-expression network included the use of the 'pickSoftThreshold' function, which selects powers of soft thresholds (β) based on a scale-free topology criterion (independence index $R^2 = 0.8$) (Tanimura et al., 2006), and a minimum module size of 100 genes. Spearman correlation analysis was conducted to identify potential associations between the modules and DN. The intersection of MRRDEGs and modular genes was obtained by venn diagram to identify hub MRRDEGs.

2.4 Construction of protein-protein interaction (PPI) networks

Protein-protein interaction (PPI) networks were developed utilizing the STRING online database (Szklarczyk et al., 2019) (https://www.string-db.org/), with a confidence interaction score threshold set at 0.4 to establish significance. The construction and analysis of the PPI network were conducted using Cytoscape (v3.9.1) software (Shannon et al., 2003). Within Cytoscape, the CytoHubba (C et al., 2014) plugin was employed to identify key genes. In particular, the Maximal Clique Centrality (MCC) algorithm was successful in identifying core genes with high centrality, whereas the Degree method enabled the identification of genes with the most connections. The Protein-Protein Interaction (PPI) network was analyzed using MRRDEGs scores, and the top 80 MRRDEGs were selected based on these scores. This approach focuses on genes with the highest scores, facilitating manageable downstream analysis and highlighting key network components. A venn diagram was employed to illustrate the overlap

of genes identified by the two algorithms, thereby pinpointing the candidate genes.

2.5 Machine learning

For the identification of key genes in the diagnosis of DN, we utilized four machine learning algorithms: Random Forest (RF) (Rigatti, 2017), Extreme Gradient Boosting (XGB) (Guo and Chang, 2022), Support Vector Machine (SVM) (Tan et al., 2014), and Generalized Linear Model (GLM) (Song et al., 2013). RF was selected for its robust capability in feature importance ranking, XGB for its proficiency in capturing non-linear interactions pertinent to metabolic reprogramming, SVM with a radial basis function (RBF) kernel for its effectiveness in high-dimensional classification tasks, and GLM as an interpretable baseline model. Using the train function from the R 'caret' (v6.0-94) package, we trained RF, SVM, XGB, and GLM models, utilizing the 'randomForest' (v4.7-1.1), 'kernlab' (v0.9-33), 'xgboost' (v1.7.8.1), and 'stats' (v4.3.3) packages, respectively. Concurrently, we utilized the 'caret' (v6.0-94) package's to tune their parameters through grid search, and evaluated their performance using fivefold cross-validation. Furthermore, to ensure the reliability of the models, we generated residual boxplots, feature importance plots, reverse cumulative distribution of residuals, and receiver operating characteristic (ROC) curves for the models.

2.6 Diagnostic model construction and assessment

To determine the feasibility of this diagnostic model as a diagnostic factor, a nomogram model was performed for six genes using the 'rms' (v6.7-1) package. The reliability of the model predictions was then assessed using ROC curves. To evaluate the predictive performance and clinical applicability of the models, calibration curves and decision curve analysis (DCA) were employed. Furthermore, based on the results of the DCA, we assessed the clinical impact curves (CIC).

2.7 Immune infiltration analysis

The immune infiltration matrix was derived from the DN gene expression dataset using the 'CIBERSORT' (v0.1.0) package in R. The CIBERSORT algorithm was then used to compare the distribution of immune cell infiltration between patients with DN and controls (Newman et al., 2015). Additionally, Spearman's correlation analysis was conducted to explore the relationship between MRRDEGs, DN, and immune infiltration. The immune infiltration matrix for each sample and group was visualized using the 'ggplot2' (v3.5.1) package.

2.8 Single-sample gene set enrichment analysis (ssGSEA)

We employed single-sample Gene Set Enrichment Analysis (ssGSEA), an advanced analytical technique that leverages molecular feature databases, to elucidate the effects of gene expression (Subramanian et al., 2005). The calculations were conducted using the 'gsva' function within the clusterProfiler package (v4.10.1). This method utilizes a precise algorithm for functional enrichment analysis, allowing us to investigate the potential biological pathways associated with key MRRDEGs.

2.9 ceRNA networks and validation of key gene expression

To predict microRNAs (miRNAs) targeting the model genes, we utilized TargetScan (http://www.targetscan.org/) with a screening criterion of pancancer num >7. For predicting lncRNA-miRNA interactions, we employed Starbase (https://starbase.sysu.edu.cn/starbase2/) using a screening criterion of pancancer num >12. Following this, we utilized Cytoscape (v3.9.1) software to construct and visualize the lncRNA-miRNA-mRNA regulatory networks. To assess the differential expression of crucial genes between disease and control groups within the DN_Datasets, we employed the Mann-Whitney U test (Wilcoxon rank-sum test). The results of this differential analysis were visualized using group comparison graphs generated with the R package ggplot2 (v3.5.1).

2.10 qRT-PCR experiments

Additionally, qRT-PCR validation was conducted. Between March to April 2025, twelve whole-blood samples were collected from the First Affiliated Hospital of Nanchang University, including six from DN patients and three from healthy controls, aged 30 to 60. All participants gave informed consent, and the study was approved by the ethics committee (Ethical number: (2024) CDYFYYLK (07–026)). Blood samples ranging from 3 to 5 mL were collected into EDTA tubes for subsequent white blood cell enrichment. Total leukocytes were isolated through the lysis of red blood cells using ACK Lysing Buffer (Thermo Fisher, USA). The whole blood was combined with 10 volumes of the lysis buffer, incubated at room temperature for 10 min, and then subjected to centrifugation at 300 g for 5 min. The resulting leukocyte pellet underwent three washes with PBS to ensure the thorough removal of erythrocyte debris. RNA was extracted from both DN patients' and healthy controls' samples, with DN samples taken before treatment. Reverse transcription was performed using the Servicebio® RT First Strand cDNA Synthesis Kit from Wuhan, China, followed by quantitative PCR. In the next step, qPCR was performed using the 2 × SYBR Green qPCR Master Mix (None ROX) from Servicebio, Wuhan, China, following the manufacturer's instructions. The thermocycling protocol included an initial phase at 95 °C for 5 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 30 s β-actin was used as the reference gene for data normalization, and gene expression was calculated using the $2^{-\Delta\Delta CT}$ method. The primers used are listed in Table 2.

2.11 Statistical analysis

All statistical analyses were performed using R software, version 4.3.3. For comparing two groups with normally distributed data,

TABLE 2 Primer sequences for quantitative real-time PCR.

Gene Names	Forward (5′-3′)	Reverse (3′-5′)
CXRC2	CTCCCTTTCATAGGTCACAG	AAACTTAAATCCTGACTGGGTC
CUEDC2	GGAACAAAGAGAACCTGCA	CTTCTTTGAGCATTTCGGG
NAMPT	GAAATGTTCTCTTCACGGTGG	GACTGAACAAGAATAGTCTCAATCC
ATF3	AGAAGGAGAAGACGGAGTG	TATGCAGGTCTTCAGGACC
GDF15	GCTGGGAAGATTCGAACAC	ACTTCTGGCGTGAGTATCC
CEBPD	AGAAGTTGGTGGAGCTGTC	GCAGCTGCTTGAAGAACTG

we applied the Student's t-test. The chi-square test was used to compare categorical and pairwise features across different groups. The Mann-Whitney U test was employed to determine statistically significant differences between two groups, while the Kruskal–Wallis test was used to assess significant differences among multiple independent groups. Pearson's correlation test was applied to evaluate correlations between normally distributed variables, and Spearman's correlation test was used for non-normally distributed variables. All statistical tests were two-sided, and a p-value of less than 0.05 was considered statistically significant unless otherwise specified.

3 Results

3.1 Calibration of data set, variance analysis, and functional enrichment analysis

We employed the R package 'sva' (v3.50.0) to address batch effects in the DN_Datasets by performing de-batching, resulting in a refined dataset. We then conducted a comparative analysis of the datasets before and after batch effect removal using Principal Component Analysis (PCA) plots (Figures 1A,B). The PCA results indicate that the batch effect in the DN_Datasets was effectively mitigated. To identify differentially expressed genes (DEGs) associated with DN across various subgroups, we utilized the R package 'limma' (v3.58.1) to analyze the expression profiling data of the DN_Datasets. This analysis revealed a total of 2,080 DEGs, comprising 1,007 upregulated and 1,073 downregulated genes in the DN group compared to the control group. Subsequently, a volcano plot (Figure 1C) was generated to visualize the differential analysis results. To identify the MRRDEGs, we determined the intersection of DEGs from the DN_Datasets, which pertains to DN, with MRRGs. This process yielded a total of 256 genes designated as MRRDEGs for further analysis. A venn diagram was constructed to illustrate these results (Figure 1D). GO functional analysis revealed that the biological processes (BPs) were predominantly enriched in pathways such as response to insulin, hexose metabolic process, monosaccharide metabolic process, carbohydrate biosynthetic process, and glucose

metabolic process. The cellular components (CC) were primarily associated with the mitochondrial matrix and chromosomal region. Regarding molecular function (MF), significant enrichment was observed in DNA-binding transcription factor binding, calmodulin binding, ubiquitin-like protein ligase binding, and transcription factor binding (Figure 1E). The KEGG pathway analysis indicated significant enrichment in the AMPK signaling pathway, insulin resistance, cell cycle, and PI3K-Akt signaling pathway (Figure 1F).

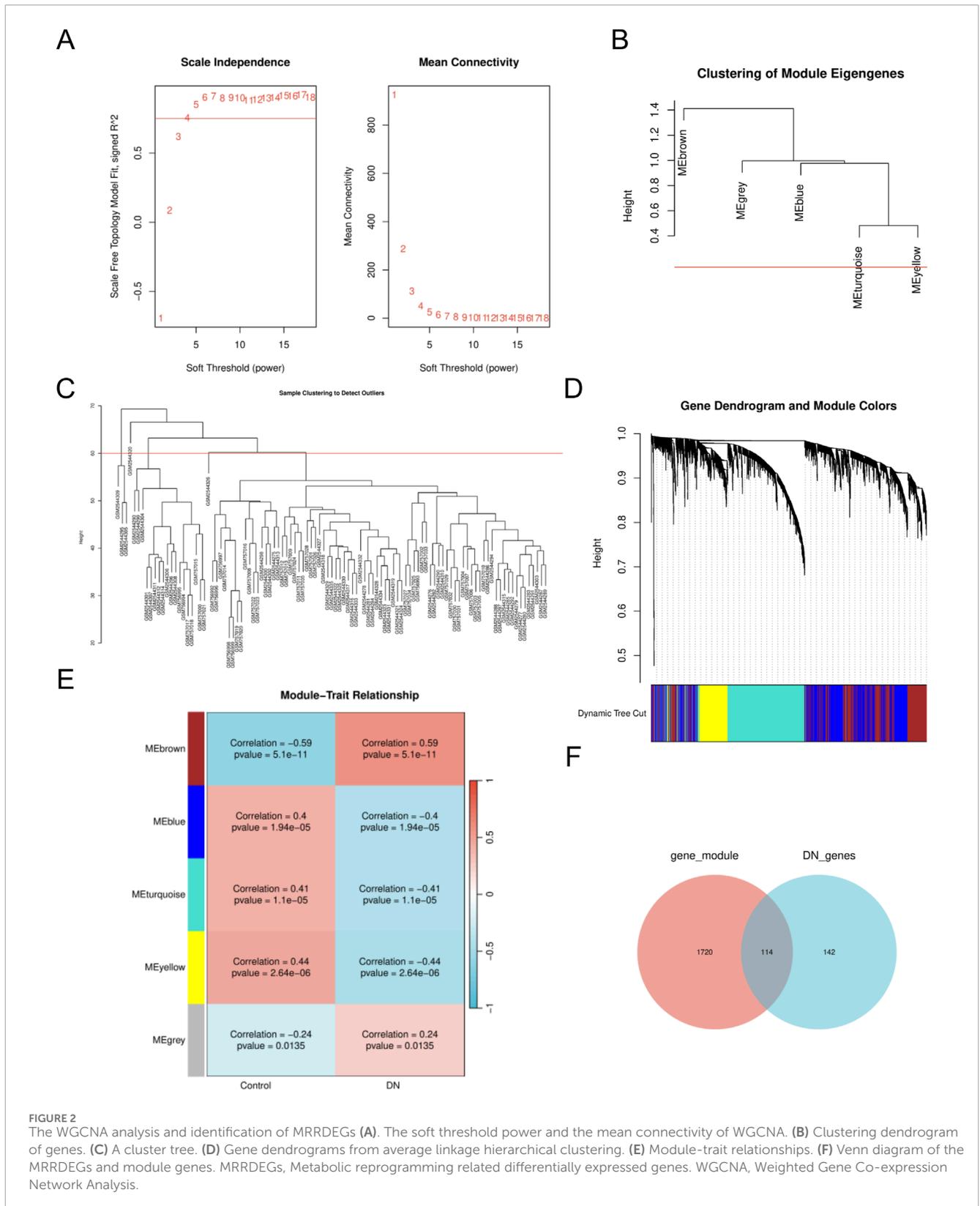## 3.2 Weighted gene Co-expression network analysis (WGCNA)

To identify key gene modules associated with DN, we employed the WGCNA algorithm to construct co-expression networks and modules for both DN and control groups. Initially, we calculated the expression variance of each gene within the DN datasets and selected the top 40% of genes exhibiting the highest variance for further analysis. A threshold of $\beta = 10$ (scale-free R = 0.8) was utilized to construct a scale-free network, facilitating the identification of co-expression gene modules (Figure 2A). Through hierarchical clustering of the samples, eight distinct co-expression modules, each represented by a unique color, were identified using the dynamic tree cut algorithm (Figures 2B–D). Subsequently, we assessed the co-expression similarity and adjacency of these modules in relation to clinical characteristics of the control and DN groups. Our analysis revealed that the blue and yellow modules were strongly associated with DN, encompassing a total of 1,834 genes (Figure 2E). Additionally, we examined the extent of overlap between module genes and MRRDEGs, identifying 114 genes in common (Figure 2F).

3.3 Protein-protein interaction (PPI) networks

PPI networks (Figure 3A) were developed utilizing the STRING database (https://www.string-db.org/) for the 118 genes obtained in the previous step. The construction and analysis of the PPI network were conducted using Cytoscape software. Within Cytoscape
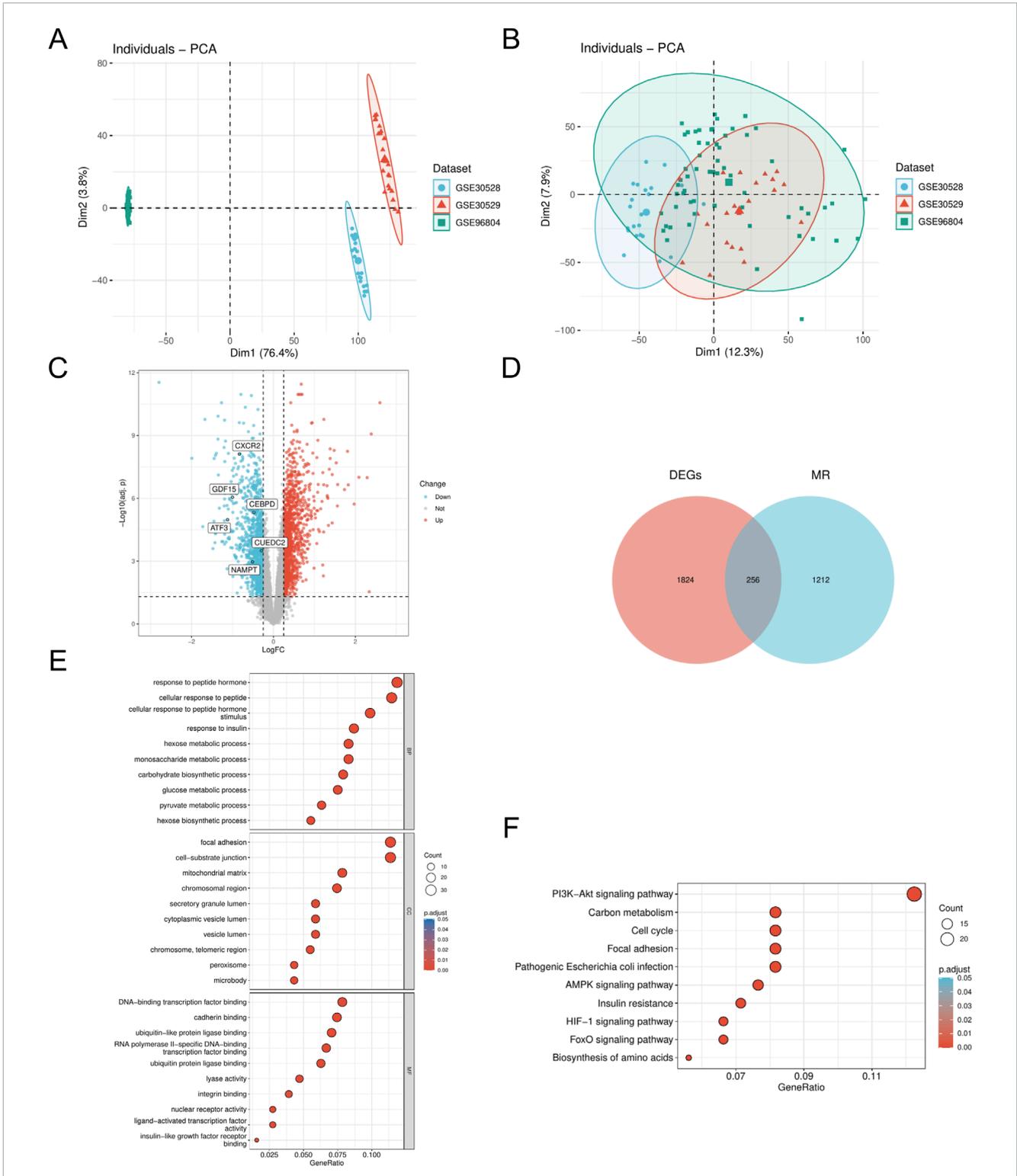
FIGURE 1
Dataset correction and analysis of differentially expressed genes **(A)**. PCA plot of DN- Datasets before correction. **(B)** PCA plot of the corrected DN-Datasets. **(C)** Volcano plot of differential analysis results between DN and Control groups in the DN- Datasets dataset. **(D)** Venn diagram of the DEGs and MR. **(E)** GO enrichment analyses of MRRDEGs. **(F)** KEGG enrichment analyses of MRRDEGs. DN, Diabetic nephropathy. MR, Metabolic reprogramming. GO, Gene ontology. KEGG, Kyoto encyclopedia of genes and genomes. BP, Biological process. CC, Cellular component. MF, Molecular function. The screening criteria for GO/KEGG enrichment items were p. Adj <0.05 and FDR value (q. value) < 0.25, and the p value correction method was Benjamini–Hochberg (BH).

FIGURE 2
The WGCNA analysis and identification of MRRDEGs **(A)**. The soft threshold power and the mean connectivity of WGCNA. **(B)** Clustering dendrogram of genes. **(C)** A cluster tree. **(D)** Gene dendrograms from average linkage hierarchical clustering. **(E)** Module-trait relationships. **(F)** Venn diagram of the MRRDEGs and module genes. MRRDEGs, Metabolic reprogramming related differentially expressed genes. WGCNA, Weighted Gene Co-expression Network Analysis.

(v3.9.1), the CytoHubba plugin was employed to apply two distinct computational methods, namely, Maximal Clique Centrality (MCC) and Degree, for the identification of key genes. The scores are detailed in Supplementary Table S1, S2. Subsequently, a total of 78 key genes (Supplementary Table S3), referred to as hub MRRDEGs, were identified through the use of a Venn diagram (Figure 3B).

FIGURE 3
Identification of mitochondria-related DEGs (A). PPI networks of overlapping MRRDEGs constructed in the STRING database. (B) Venn diagram of hub MRRDEGs identified by two different algorithms in Cytohubba (MCC and Degree).
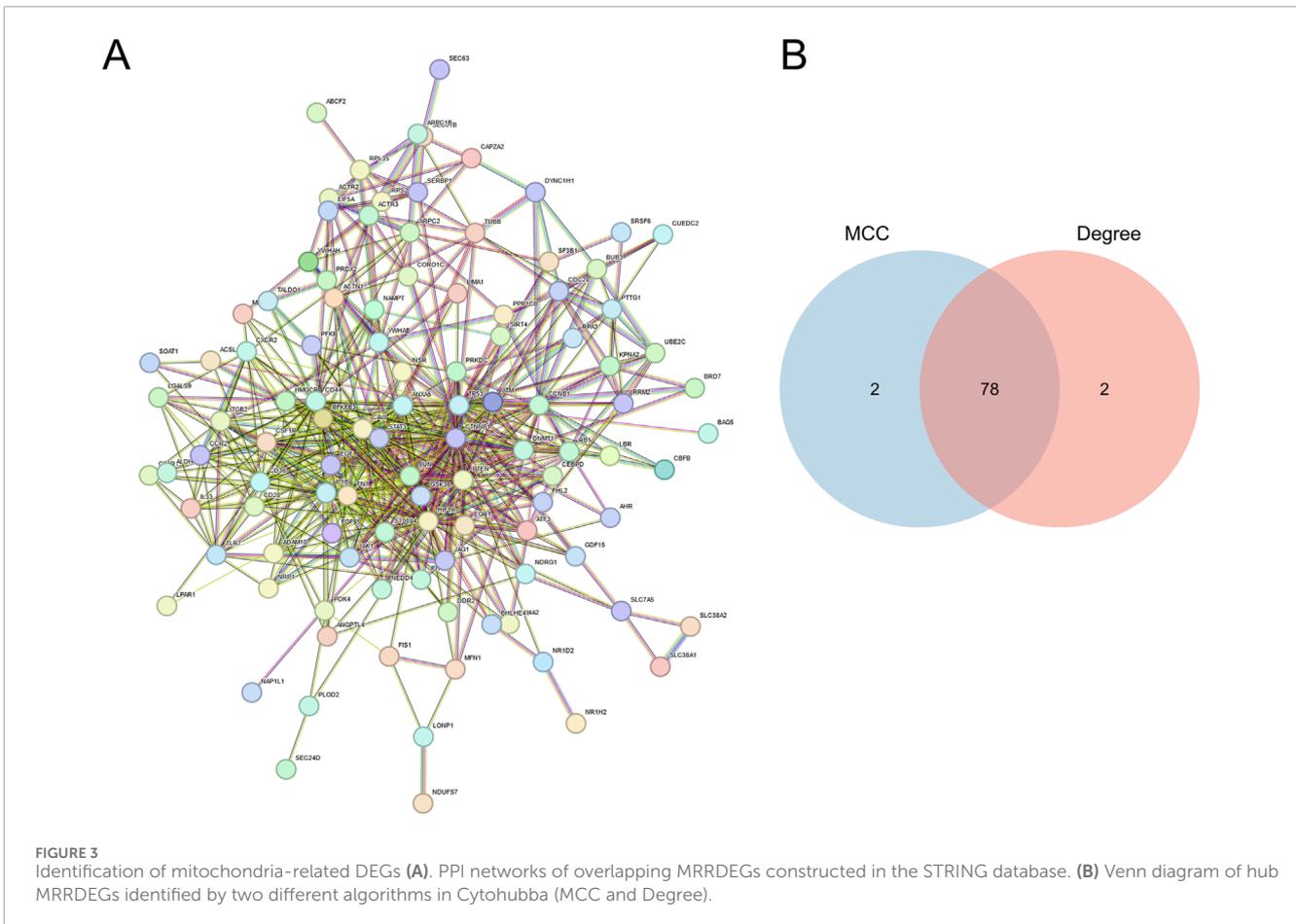
3.4 Machine learning

In this study, we utilized four different machine learning algorithms: RF, SVM, XGBoost, and GLM. The RF algorithm was utilized to develop the model, leading to the identification of 40 genes (Figures 4A,B). Subsequently, the SVM algorithm was applied to select the 10 genes with the highest accuracy (AUC = 0.949) (Figures 4C–F). Furthermore, the XGBoost algorithm was used to construct the model, resulting in the selection of 10 genes (Figures 5A,B). Finally, the GLM algorithm was used to screen the key genes, and six genes were identified as *UEDC2*, *CXCR2*, *ATF3*, *GDF15*, *CEBPD* and *NAMPT* (Figures 5C–F).

3.5 Diagnostic model construction and evaluation

We created a nomogram model (Figure 6A) and assessed the diagnostic performance of genes using a six-gene ROC analysis. *CXCR2* showed the highest AUC of 0.846 (Figure 6B), while the other genes had AUCs of 0.704, 0.8, 0.689, 0.77, and 0.704 (Figures 6B,C). Calibration curve analysis confirmed that predicted rates matched observed rates well (Figure 6D). Decision curve analysis (DCA) indicated that the logistic regression model with the six diagnostic factors was stable at high-risk thresholds (Figure 6E).

In the high-risk threshold range of 0.4–1, clinical impact curves showed strong predictive ability, as the "number of people at high risk" closely matched the "number of people at high risk of having an event" (Figure 6F).

## 3.6 Immune infiltration analysis

To investigate potential changes in the immune system between DN patients and controls, we conducted immune cell characterization utilizing the CIBERSORT algorithm. Our analysis was designed to elucidate variations in the composition of the immune cell population. The findings revealed significant disparities in the proportions of various immune cell types between the DN and control (Figures 7A,B). Notably, the disease group exhibited an elevated presence of M1 macrophages, M2 macrophages, and gamma delta T cells relative to the control group, indicating that alterations in the immune system might be critically involved in the development of DN. Furthermore, interactions among different immune cell types were evident (Figure 7C). The microenvironment of DN was evaluated using the CIBERSORT algorithm, and the relationship between key genes and immune cells was investigated using Spearman's correlation analysis (p < 0.05, |correlation coefficient|>0.3). Significant associations were identified between most key genes and immune cell components.
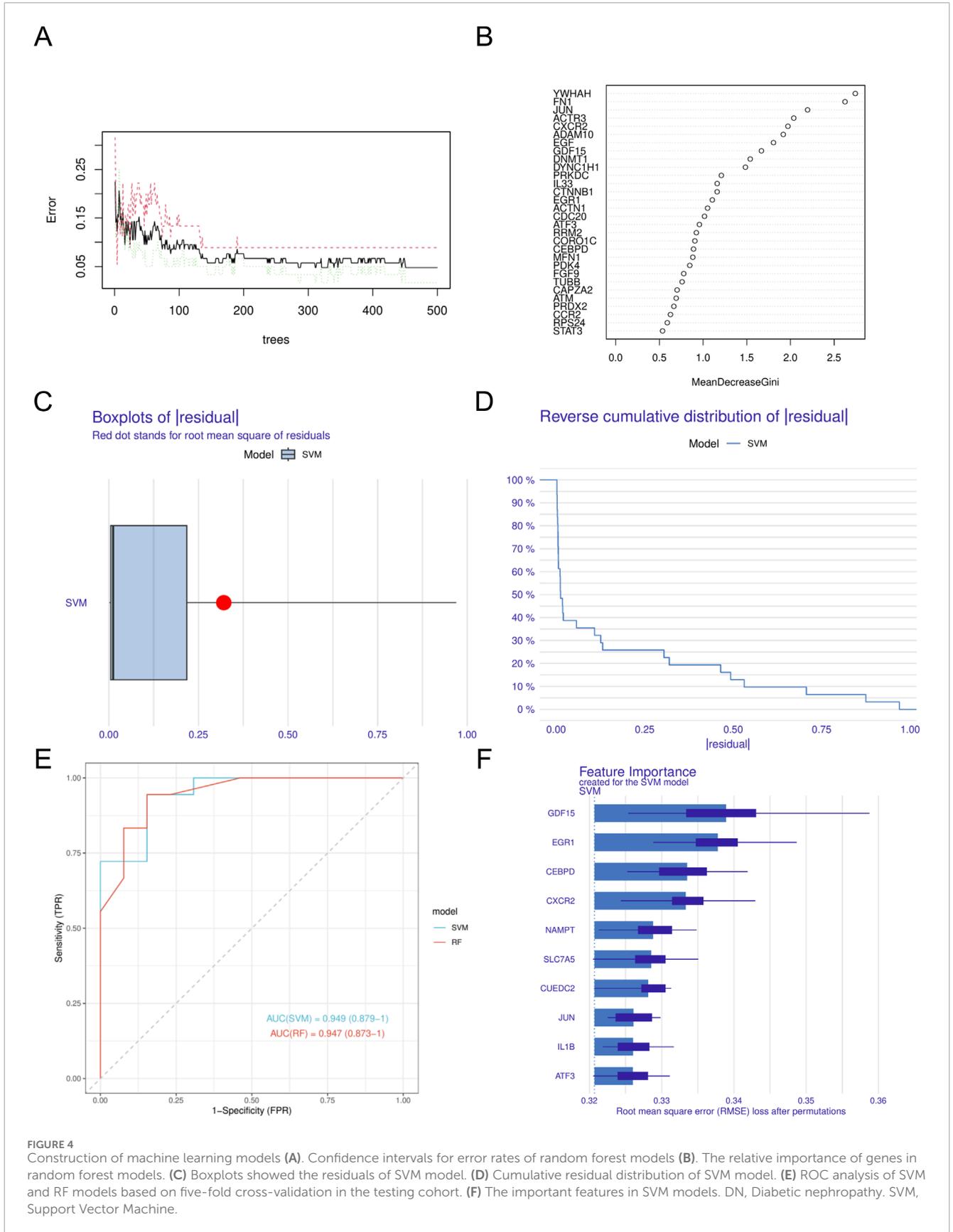
FIGURE 4
Construction of machine learning models **(A)**. Confidence intervals for error rates of random forest models **(B)**. The relative importance of genes in random forest models. **(C)** Boxplots showed the residuals of SVM model. **(D)** Cumulative residual distribution of SVM model. **(E)** ROC analysis of SVM and RF models based on five-fold cross-validation in the testing cohort. **(F)** The important features in SVM models. DN, Diabetic nephropathy. SVM, Support Vector Machine.
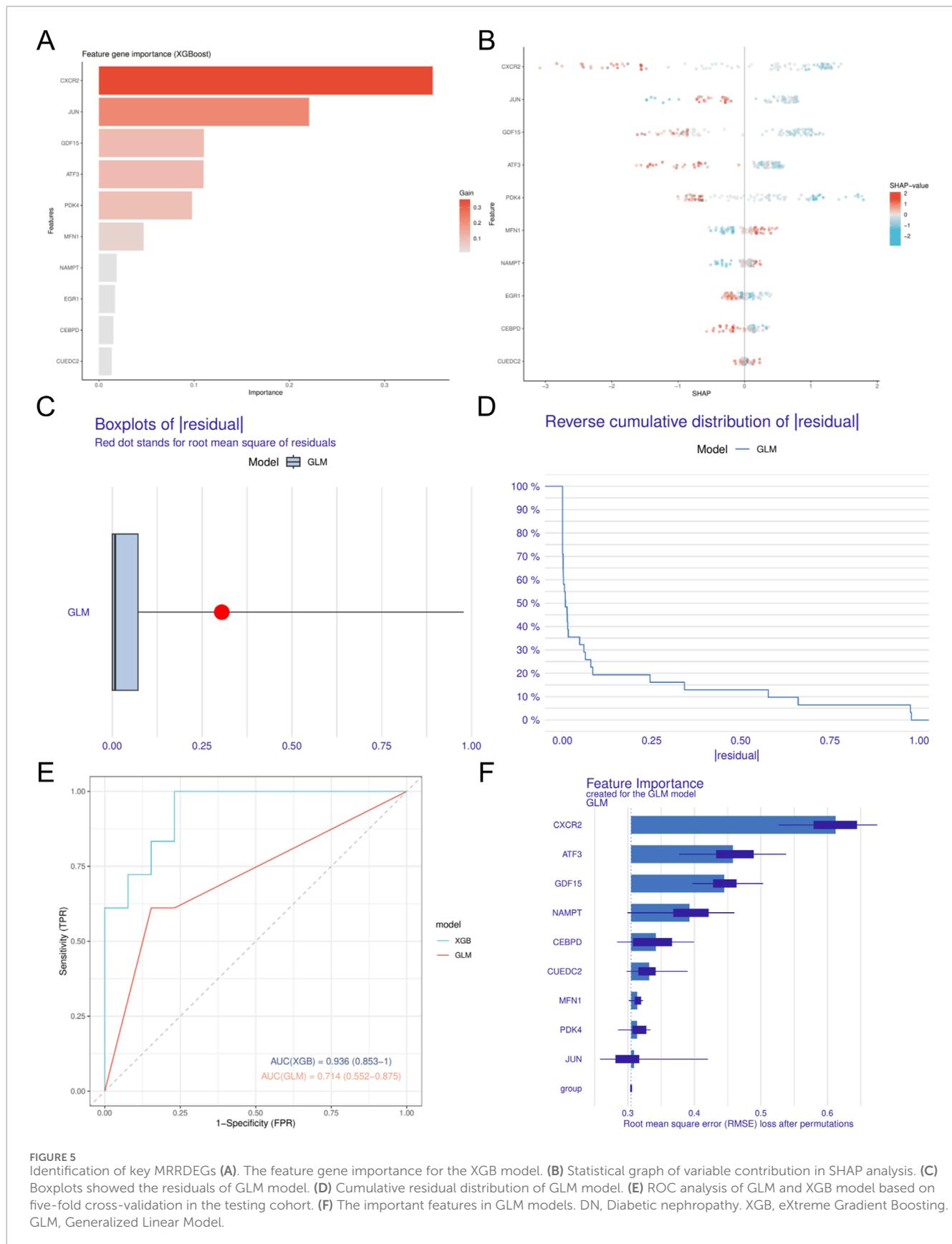
FIGURE 5
Identification of key MRRDEGs **(A)**. The feature gene importance for the XGB model. **(B)** Statistical graph of variable contribution in SHAP analysis. **(C)** Boxplots showed the residuals of GLM model. **(D)** Cumulative residual distribution of GLM model. **(E)** ROC analysis of GLM and XGB model based on five-fold cross-validation in the testing cohort. **(F)** The important features in GLM models. DN, Diabetic nephropathy. XGB, eXtreme Gradient Boosting. GLM, Generalized Linear Model.
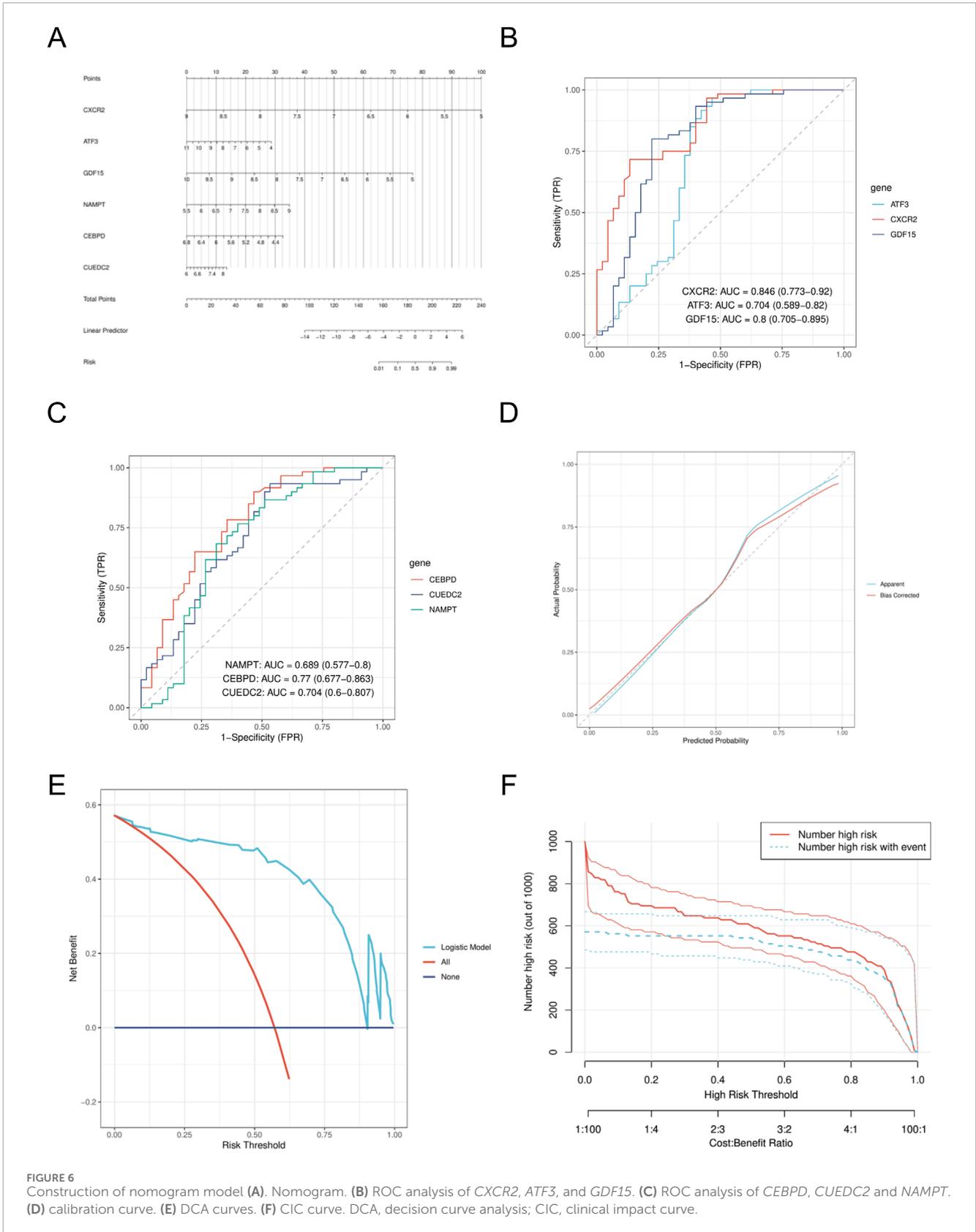
FIGURE 6
Construction of nomogram model **(A)**. Nomogram. **(B)** ROC analysis of *CXCR2, ATF3,* and *GDF15*. **(C)** ROC analysis of *CEBPD, CUEDC2* and *NAMPT*. **(D)** calibration curve. **(E)** DCA curves. **(F)** CIC curve. DCA, decision curve analysis; CIC, clinical impact curve.

A notable observation was the positive correlation identified between *CUEDC2* and M1 macrophages. Additionally, *CXCR2* demonstrated a significant positive correlation with neutrophils.

In contrast, a significant negative correlation was observed between *NAMPT* expression and the M0 macrophage population (Figures 7D–F).

3.7 Single-sample gene set enrichment analysis (ssGSEA)

To elucidate the enrichment pathways associated with the characterized genes, we conducted a ssGSEA. Target genes were ranked, and groups with high and low expression were delineated based on expression level differences. The enrichment degree was evaluated by calculating the cumulative score for the target gene set within the ranked list. The GSEA results indicated significant enrichment in pathways such as WP_INFLAMMATORY_ RESPONSE_PATHWAY, REACTOME_SIGNALING_BY_ INTERLEUKINS, and cellular metabolism-related pathways including KEGG_OXIDATIVE_PHOSPHORYLATION, along with other biologically relevant functions and signaling pathways (Enrichment score >0.5, p < 0.001) (Figures 8A–F). Normalized enrichment scores (NES), false discovery rates (FDR qvalues), and P-values for all pathways are included in Supplementary Table S1, S4.

3.8 ceRNA networks and expression verification

The miRNA-mRNA interactions were obtained from TargetScan and subsequently screened to identify miRNAs interacting with five specific genes, resulting in 36 overlapping targeting relationships. Additionally, the StarBase database was employed to predict 82 miRNA-targeted lncRNAs. The datasets were subsequently integrated into Cytoscape software to generate a ceRNA regulatory network (Figure 9A). This network offers insights into the potential regulatory mechanisms driving metabolic reprogramming in DN. We analyzed the expression differences of six key genes in the DN_ Datasets dataset between the DN and control groups using the WilCoxon rank-sum test, and the results of the expression difference analysis were presented by the subgroup comparison graph (Figure 9B). The results showed that in the DN_Datasets dataset, the expression of the six key genes was statistically significantly different between the DN and Control groups (P < 0.05). Fold-change and effect size are included in Supplementary Table S5.

3.9 qRT-PCR

The expression levels of the key genes in DNA and control blood samples were examined using qRT-PCR. The findings indicated significant differences in the expression levels of three genes (*CUEDC2*, *CXCR2*, and *NAMPT*), and the differences were statistically significant (P < 0.05). Specifically, the expression of *CUEDC2*, *CXCR2*, and *NAMPT* exhibited lower expression in the DN group relative to the control group. In contrast, no significant difference in expression was observed for *ATF3*, *GDF15*, or *CEBPD* between the DN and control groups (Figures 10A–F).
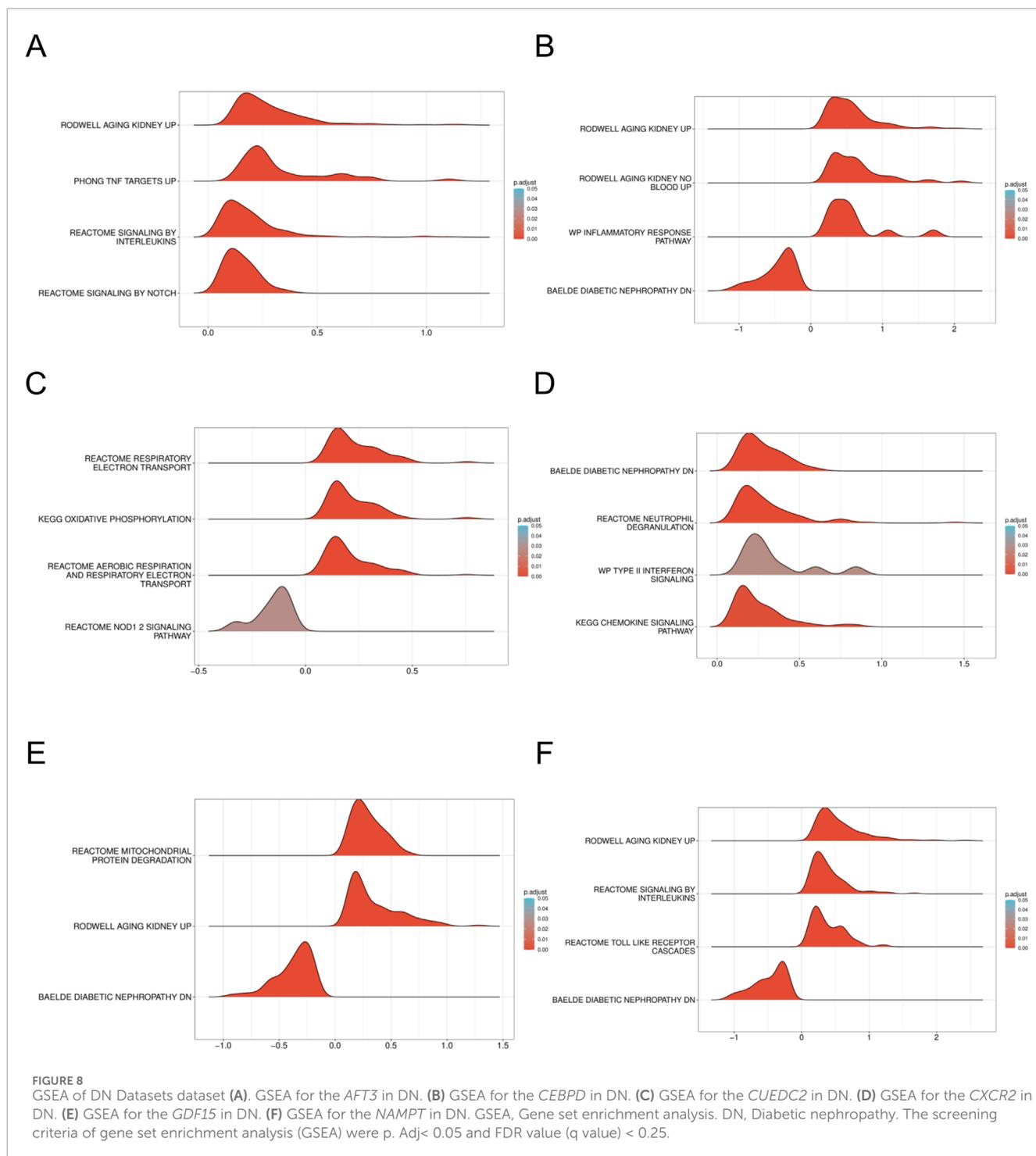
4 Discussion

In the context of DN, renal cells experience substantial metabolic reprogramming, characterized by a transition from mitochondrial oxidative phosphorylation to glycolysis. This metabolic shift is believed to be instrumental in the development and progression of DN (Wang et al., 2022b). The rising prevalence of DN necessitates an urgent investigation into its underlying mechanisms and potential therapeutic targets. In this study, we aimed to clarify the crucial function of metabolic reprogramming in the development of DN using a bioinformatics analysis. We identified 256 MRRDEGs, and GO analysis indicated that these genes were significantly enriched in biological processes related to insulin signaling, carbohydrate catabolism, and glucose control. These observations are consistent with previous studies demonstrating a metabolic disorder in the pathogenesis of DN (Khoshjou and Dadras, 2014). KEGG pathway analysis showed that these genes are mainly linked to the MAPK signaling pathway, which is crucial in DN progression. Hyperglycemia triggers MAPK cascades (ERK/JNK/p38), leading to renal inflammation, fibrosis, apoptosis, and oxidative stress (Thongrung et al., 2025). This pathway connects metabolic signals to tissue damage, making it a key mechanism in DN, aligning with prior research (Wang et al., 2021; Han et al., 2020). We combined WGCNA, Cytohubba (Degree/MCC), and four machine learning algorithms (XGBoost, SVM-RFE, LASSO, Random Forest) to overcome individual model limitations. WGCNA finds co-expression modules but needs hub gene refinement, Cytohubba identifies network hubs but relies on PPI completeness, and machine learning algorithms provide strong feature selection but have biases. We prioritized genes identified by at least two Cytohubba methods and 3 ML algorithms to ensure robust biomarker selection. This approach led to the identification of six key genes: *CXCR2*, *ATF3*, *GDF15*, *NAMPT*, *CEBPD*, and *CUEDC2*. To ascertain the diagnostic utility of the core genes, we assessed their diagnostic performance by constructing the nomogram. Further investigation involved examining immune infiltration in DN using the CIBERSORT algorithm, and evaluating the correlation between key genes and infiltrating immune cells to reveal relevant immune mechanisms. Additionally, we predicted the ceRNA regulatory network to further clarify the regulation of these core genes. While bioinformatics models provide valuable insights, experimental validation is essential for confirming the therapeutic potential of the identified biomarkers. Lastly, qRT-PCR was performed to confirm that *CXCR2*, *NAMPT*, and *CUEDC2* could potentially serve as biomarkers for the clinical diagnosis and risk evaluation of DN patients. *CXCR2* and *CUEDC2* predominantly regulate inflammatory pathways, including NF-κB signaling and macrophage polarization, whereas *NAMPT* plays a critical role in linking metabolic dysfunction (NAD + depletion and mitochondrial impairment) to inflammatory responses and oxidative stress, crucial for DN progression.

CXCR2, a receptor that mediates the effects of specific chemokines, functions as a signaling molecule and exerts a key influence on inflammation and tissue damage (Leslie et al., 2022). The deletion of *CXCR2* has been demonstrated to significantly enhance renal function in mice with DN, while concurrently inhibiting the activation of the NF-κB signaling pathway. This pathway is known to regulate inflammation, restore the endothelial glycocalyx, and mitigate DN, as evidenced by a study utilizing a mouse model with a specific knockout of the *CXCR2* gene (Cui et al., 2024). Furthermore, another study suggests that the IL-8-CXCR1/2 axis may contribute to DN by inducing podocyte injury (Loretelli et al., 2021). Inhibition of CXCR1/2 resulted in
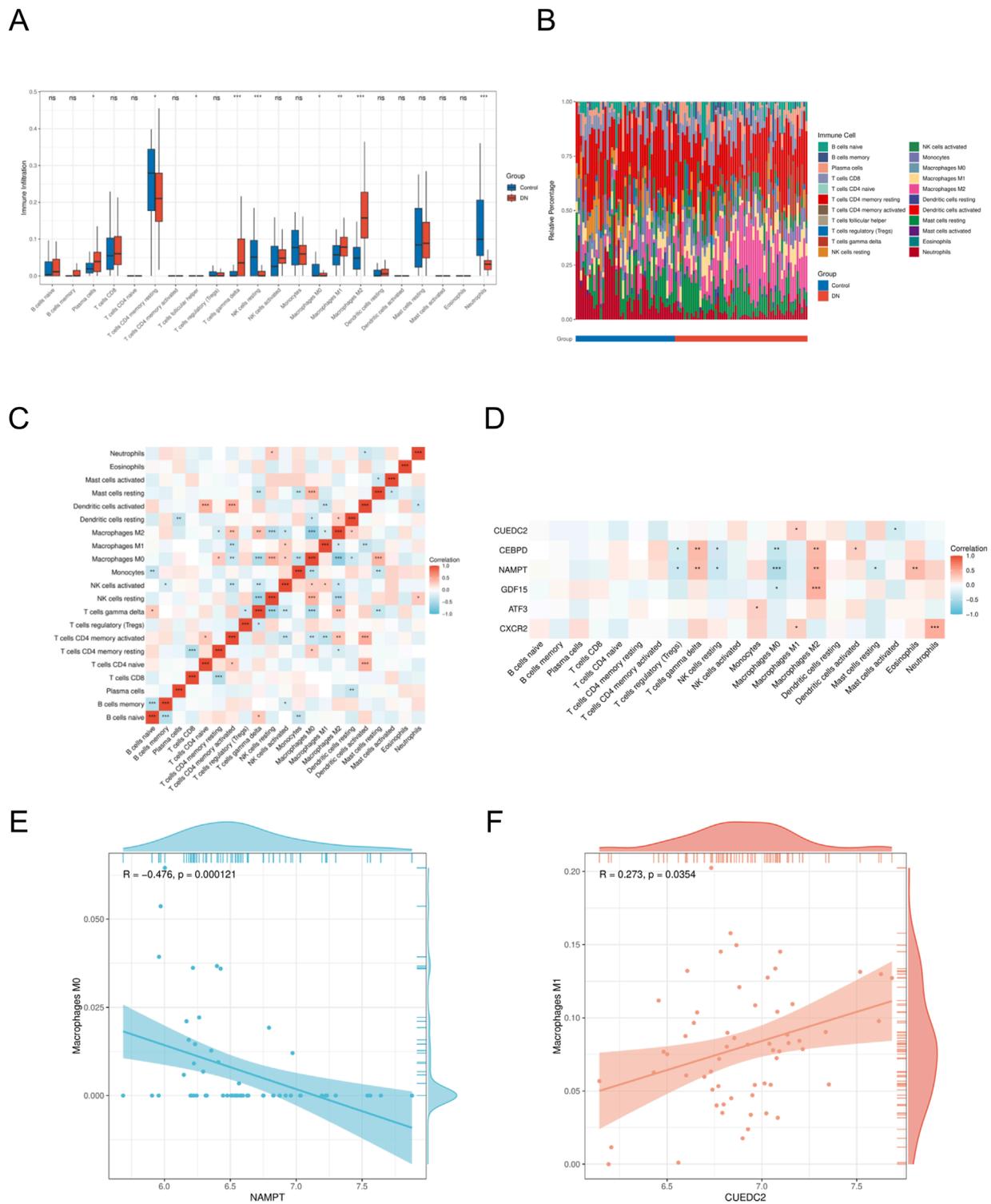
FIGURE 7
Immune cell infiltration analyses **(A)**. Immune cell distribution map in DN. **(B)** boxplot showing the comparison of 22 kinds of immune cells between DN and the control group. **(C)** heatmap representing the associations of the differentially infiltrated immune cells with immune cells. **(D)** heatmap representing the associations of the differentially infiltrated immune cells with the six hub genes. **(E)** Scatter plot of correlation between *NAMPT* and immune cell. **(F)** Scatter plot of correlation between *CUEDC2* and immune cell. DN, Diabetic nephropathy. The symbol * is equivalent to P < 0.05, which is statistically significant. The symbol ** is equivalent to P < 0.01, which is highly statistically significant. The symbol *** is equivalent to P < 0.001 and highly statistically significant.

FIGURE 8
GSEA of DN Datasets dataset **(A)**. GSEA for the *AFT3* in DN. **(B)** GSEA for the *CEBPD* in DN. **(C)** GSEA for the *CUEDC2* in DN. **(D)** GSEA for the *CXCR2* in DN. **(E)** GSEA for the *GDF15* in DN. **(F)** GSEA for the *NAMPT* in DN. GSEA, Gene set enrichment analysis. DN, Diabetic nephropathy. The screening criteria of gene set enrichment analysis (GSEA) were p. Adj< 0.05 and FDR value (q value) < 0.25.

reduced proteinuria, decreased thylakoid dilatation, and diminished podocyte apoptosis and DNA damage in diabetic mice. Given the observed significant downregulation of *CXCR2* in the DN group, it is suggested that *CXCR2* may not only be involved in the progression of the disease but may also represent the body's attempt to ameliorate the condition by reducing the activity of the *CXCR2* signaling axis. This complex association underscores the multifaceted role of CXCR2 in maintaining kidney health. Nicotinamide phosphoribosyltransferase (*NAMPT*), an essential

enzyme, is pivotal in the biosynthetic pathway of nicotinamide adenine dinucleotide (NAD+), a molecule integral to renal physiology and central to cellular energy metabolism and redox balance. Research indicates that the downregulation of *NAMPT* results in reduced NAD + levels, while *NAMPT* deficiency leads to an overload of mitochondrial ribosomes. This overload impairs the translation of proteins associated with the mitochondrial inner membrane's oxidative phosphorylation complexes I (CI), III (CIII), IV (CIV), and V (CV), culminating in mitochondrial dysfunction.

FIGURE 9
ceRNA network and expression verification. **(A)** ceRNA network of key genes. **(B)** Expresion level of key genes in DN_datasets. ceRNA, competing endogenous RNA. DN, Diabetic nephropathy. The symbol * is equivalent to P < 0.05, which is statistically significant. The symbol ** is equivalent to P < 0.01, which is highly statistically significant. The symbol *** is equivalent to P < 0.001 and highly statistically significant.

FIGURE 10
qRT-qPCR. **(A-F)** RT-qPCR validation of relative expression of key genes. The symbol * is equivalent to P < 0.05, which is statistically significant. The symbol ** is equivalent to P < 0.01, which is highly statistically significant. The symbol *** is equivalent to P < 0.001 and highly statistically significant.

Furthermore, *NAMPT* downregulation is linked to diminished expression of the transcriptional repressor HIC1, which exacerbates mitochondrial ribosome overload and contributes to diabetic albuminuria and type IV collagen deposition (Hasegawa et al., 2024). These findings imply a protective role of *NAMPT* in DN, aligning with our study's conclusions. *CUEDC2* is a protein characterized by the presence of the CUE structural domain, which is integral to various biological processes, including the cell cycle, inflammation, and tumorigenesis. Nonetheless, the involvement of *CUEDC2* in the onset and progression of DN remains to be elucidated. Previous studies have indicated that *CUEDC2* may influence bone formation and regeneration through the regulation of the SOCS3-STAT3 pathway (Kim et al., 2020). Additionally, *CUEDC2* has been associated with the oxidative capacity of cardiomyocytes, impacting their oxidative stress response by modulating the stability of GPX1 (Jian et al., 2016). In the context of DN, oxidative stress and inflammation are critical components of its pathophysiological process. Therefore, *CUEDC2* may contribute to the pathogenesis of DN by modulating oxidative stress and the inflammatory response.

The polarization state of macrophages, specifically the equilibrium between M1 pro-inflammatory macrophages and M2 anti-inflammatory macrophages, plays a crucial role in the pathogenesis of DN. Research indicates that excessive activation of M1 macrophages intensifies renal inflammation and fibrosis, while M2 macrophages aid in mitigating these pathological alterations (Yuan et al., 2020; Xie et al., 2020). Our immunological analyses further elucidated distinct patterns of immune infiltration within DN tissues, characterized by an increased presence of M1 macrophages, M2 macrophages, and T cells gamma delta. This altered immune landscape indicates a dynamic interplay of immune responses within DN, particularly marked by the heightened infiltration of T cells gamma delta. T cells gamma delta represent a unique subset of T cell populations with specialized functions, playing a crucial role in immunosurveillance and tissue repair (Ribot et al., 2021). Previous studies have demonstrated that T cells gamma delta can modulate the activity of other immune cells through cytokine secretion and direct cell-cell interactions (Ziegler, 2004). In DN tissues, increased T cells gamma delta likely regulate local immune responses and aid tissue repair. *CUEDC2* positively correlates with M1 macrophages, indicating a role in macrophage polarization and inflammation. *CXCR2* is significantly positively correlated with neutrophils, which is consistent with previous studies. *CXCR2* can recruit neutrophils (Holloman et al., 2024). *NAMPT* negatively correlates with M0 macrophages, suggesting its role in their regulation. M0 macrophages' differentiation into M1 or M2 influences disease progression (Chen et al., 2018), and *NAMPT* downregulation may boost M0 activation.

Our GSEA analysis confirms that metabolic reprogramming drives DN progression through coordinated activation of inflammatory and metabolism. The dysregulation of *CXCR2*, *NAMPT*, and *CUEDC2*, as identified and validated in this study, represents key molecular intersections within this pathogenic network. Further study of these genes, including their immune connections and interactions with miRNAs and lnRNAs, could guide future DN targeting and immunotherapy focused on correcting metabolic-immune dysregulation. We plan to investigate their mechanisms in DN using molecular biology experiments.

Our study identified *CUEDC2*, *NAMPT*, and *CXCR2* as potential MRRDEGs crucial in DN development. These genes showed high diagnostic value in ROC analysis, suggesting their potential as biomarkers for early DN diagnosis. Their roles in immune cell infiltration and metabolic reprogramming highlight their diagnostic and therapeutic potential. Targeting these genes may offer new strategies to slow or reverse DN progression. However, the study has limitations (Thipsawat, 2021): The current analysis of specific markers in the plasma of DN patients is still in its early stages. To improve the rigor and generalizability of our results, we intend to increase our sample size in future research. Moreover, we plan to utilize additional validation techniques, including Western blotting and ELISA, to confirm the RT-qPCR data and ensure that the observed alterations in gene expression precisely mirror the actual plasma levels of these markers. These further experiments will yield more substantial evidence to substantiate our findings regarding DN biomarkers (Levin et al., 2020). Bias in dataset selection can arise from sample variability, which future research could mitigate by using more uniform datasets or cross-validation. Even with batch-effect correction methods like ComBat, differences between datasets (such as platforms or cohorts) may still cause bias. Future studies should involve larger prospective cohorts (Thomas et al., 2015). Theoretical ceRNA regulatory networks need experimental validation to confirm their therapeutic potential. Both predicted ceRNA networks and CIBERSORT-inferred immune proportions are computational and require experiments, such as luciferase assays for miRNA interactions and flow cytometry/IHC for immune cells, to verify their biological and therapeutic relevance.

5 Conclusion

In conclusion, *CXRC2*, *NAMPT*, and *CUEDC2* were identified as key genes involved in DN metabolic reprogramming through GEO database analysis and confirmed by qRT-PCR. These insights improve our understanding of DN's molecular mechanisms and support future research into new therapeutic strategies targeting these genes. More experiments and integrated multi-omics data are required to confirm these findings.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving humans were approved by the Medical Research Ethics Committee of the First Affiliated Hospital of Nanchang University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

SH: Conceptualization, Methodology, Visualization, Writing – original draft, Formal Analysis, Data curation, Investigation, Software. YC: Methodology, Data curation, Visualization, Software, Writing – original draft, Supervision, Investigation. JY: Validation, Formal Analysis, Investigation, Conceptualization, Writing – original draft, Methodology. YW: Data curation, Validation, Funding acquisition, Resources, Writing – review and editing, Supervision, Formal Analysis, Project administration. QC: Supervision, Conceptualization, Project administration, Writing – review and editing. SL: Funding acquisition, Supervision, Resources, Writing – review and editing.

Funding

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

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/fcell.2025.1630708/full#supplementary-material

References

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets–update. *Nucleic Acids Res.* 41 (Database issue), D991–D995. doi:10.1093/nar/gks1193

Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T., and Lin, C.-Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* 8, S11. doi:10.1186/1752-0509-8-S4-S11

Chen, W., Sandoval, H., Kubiak, J. Z., Li, X. C., Ghobrial, R. M., and Kloc, M. (2018). The phenotype of peritoneal mouse macrophages depends on the mitochondria and ATP/ADP homeostasis. *Cell Immunol.* 324, 1–7. doi:10.1016/j.cellimm.2017.11.003

Cross, S. H., Lakin, J. R., Mendu, M., Mandel, E. I., and Warraich, H. J. (2021). Trends in place of death for individuals with deaths attributed to advanced chronic or end-stage kidney disease in the United States. *J. Pain Symptom Manage* 61 (1), 112–120.e1. doi:10.1016/j.jpainsymman.2020.08.001

Cui, S., Chen, X., Li, J., Wang, W., Meng, D., Zhu, S., et al. (2024). Endothelial CXCR2 deficiency attenuates renal inflammation and glycocalyx shedding through NF-κB signaling in diabetic kidney disease. *Cell Commun. Signal* 22 (1), 191. doi:10.1186/s12964-024-01565-2

Dwivedi, S., and Sikarwar, M. S. (2025). Diabetic nephropathy: pathogenesis, mechanisms, and therapeutic strategies. *Hormone metabolic Res. = Hormon- und Stoffwechselforschung = Hormones metabolisme* 57 (1), 7–17. doi:10.1055/a-2435-8264

Fan, X., Yang, M., Lang, Y., Lu, S., Kong, Z., Gao, Y., et al. (2024). Mitochondrial metabolic reprogramming in diabetic kidney disease. *Cell Death Dis.* 15 (6), 442. doi:10.1038/s41419-024-06833-0

Ghose, S., Satariano, M., Korada, S., Cahill, T., Shah, R., and Raina, R. (2024). Advancements in diabetic kidney disease management: integrating innovative therapies and targeted drug development. *Am. J. Physiol. Endocrinol. Metab.* 326 (6), E791–E806. doi:10.1152/ajpendo.00026.2024

Gilbertson, D. T., Liu, J., Xue, J. L., Louis, T. A., Solid, C. A., Ebben, J. P., et al. (2005). Projecting the number of patients with end-stage renal disease in the United States to the year 2015. *J. Am. Soc. Nephrol.* 16 (12), 3736–3741. doi:10.1681/ASN.2005010112

Guo, C.-Y., and Chang, K.-H. (2022). A novel algorithm to estimate the significance level of a feature interaction using the extreme gradient boosting machine. *Int. J. Environ. Res. Public Health* 19 (4), 2338. doi:10.3390/ijerph19042338

Han, J., Pang, X., Zhang, Y., Peng, Z., Shi, X., and Xing, Y. (2020). Hirudin protects against kidney damage in Streptozotocin-Induced diabetic nephropathy rats by inhibiting inflammation *via* P38 MAPK/NF-κB pathway. *Drug Des. Devel Ther.* 14, 3223–3234. doi:10.2147/DDDT.S257613

Hasegawa, K., Tamaki, M., Sakamaki, Y., and Wakino, S. (2024). Nmnat1 deficiency causes mitoribosome excess in diabetic nephropathy mediated by transcriptional repressor HIC1. *Int. J. Mol. Sci.* 25 (12), 6384. doi:10.3390/ijms25126384

Holloman, B. L., Wilson, K., Cannon, A., Nagarkatti, M., and Nagarkatti, P. S. (2024). Indole-3-carbinol attenuates lipopolysaccharide-induced acute respiratory distress syndrome through activation of AhR: role of CCR2+ monocyte activation and recruitment in the regulation of CXCR2+ neutrophils in the lungs. *Front. Immunol.* 15, 1330373. doi:10.3389/fimmu.2024.1330373

Hou, G., Dong, Y., Jiang, Y., Zhao, W., Zhou, L., Cao, S., et al. (2025). Immune inflammation and metabolic interactions in the pathogenesis of diabetic nephropathy. *Front. Endocrinol. (Lausanne)* 16, 1602594. doi:10.3389/fendo.2025.1602594

Huang, P., Fan, X., Yu, H., Zhang, K., Li, H., Wang, Y., et al. (2023). Glucose metabolic reprogramming and its therapeutic potential in obesity-associated endometrial cancer. *J. Transl. Med.* 21 (1), 94. doi:10.1186/s12967-022-03851-4

Jian, Z., Liang, B., Pan, X., Xu, G., Guo, S.-S., Li, T., et al. (2016). CUEDC2 modulates cardiomyocyte oxidative capacity by regulating GPX1 stability. *EMBO Mol. Med.* 8 (7), 813–829. doi:10.15252/emmm.201506010

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27

Khoshjou, F., and Dadras, F. (2014). Mitochondrion and its role in diabetic nephropathy. *Iran. J. Kidney Dis.* 8 (5), 355–358.

Kim, J.-W., Oh, S.-H., Lee, M. N., Song, J. H., Jeong, B.-C., Yang, J.-W., et al. (2020). CUEDC2 controls osteoblast differentiation and bone formation *via* SOCS3-STAT3 pathway. *Cell Death Dis.* 11 (5), 344. doi:10.1038/s41419-020-2562-5

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559

Leslie, J., Mackey, J. B. G., Jamieson, T., Ramon-Gil, E., Drake, T. M., Fercoq, F., et al. (2022). CXCR2 inhibition enables NASH-HCC immunotherapy. *Gut* 71 (10), 2093–2106. doi:10.1136/gutjnl-2021-326259

Levin, A., Reznichenko, A., Witasp, A., Liu, P., Greasley, P. J., Sorrentino, A., et al. (2020). Novel insights into the disease transcriptome of human diabetic glomeruli and tubulointerstitium. *Nephrol. Dial. Transpl.* 35 (12), 2059–2072. doi:10.1093/ndt/gfaa121

Li, C.-H., and Liao, C.-C. (2021). The metabolism reprogramming of microRNA Let-7-Mediated glycolysis contributes to autophagy and tumor progression. *Int. J. Mol. Sci.* 23 (1), 113. doi:10.3390/ijms23010113

Li, Y., Sha, Z., and Peng, H. (2021). Metabolic reprogramming in kidney diseases: evidence and therapeutic opportunities. *Int. J. Nephrol.* 2021, 5497346. doi:10.1155/2021/5497346

Loretelli, C., Rocchio, F., D'Addio, F., Ben Nasr, M., Castillo-Leon, E., Dellepiane, S., et al. (2021). The IL-8-CXCR1/2 axis contributes to diabetic kidney disease. *Metabolism* 121, 154804. doi:10.1016/j.metabol.2021.154804

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P. D. (2019). PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* 47 (D1), D419-D426–D426. doi:10.1093/nar/gky1038

Neuen, B. L., Heerspink, H. J. L., Vart, P., Claggett, B. L., Fletcher, R. A., Arnott, C., et al. (2024). Estimated lifetime cardiovascular, kidney, and mortality benefits of combination treatment with SGLT2 inhibitors, GLP-1 receptor agonists, and nonsteroidal MRA compared with conventional care in patients with type 2 diabetes and albuminuria. *Circulation* 149 (6), 450–462. doi:10.1161/CIRCULATIONAHA.123.067584

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Pan, Y., Jiang, S., Hou, Q., Qiu, D., Shi, J., Wang, L., et al. (2018). Dissection of glomerular transcriptional profile in patients with diabetic nephropathy: SRGAP2a protects podocyte structure and function. *Diabetes* 67 (4), 717–730. doi:10.2337/db17-0755

Ribot, J. C., Lopes, N., and Silva-Santos, B. (2021). γδ T cells in tissue physiology and surveillance. *Nat. Rev. Immunol.* 21 (4), 221–232. doi:10.1038/s41577-020-00452-4

Rigatti, S. J. (2017). Random Forest. *J. Insur Med.* 47 (1), 31–39. doi:10.17849/insm-47-01-31-39.1

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Sharma, K., Karl, B., Mathew, A. V., Gangoiti, J. A., Wassel, C. L., Saito, R., et al. (2013). Metabolomics reveals signature of mitochondrial dysfunction in diabetic kidney disease. *J. Am. Soc. Nephrol.* 24 (11), 1901–1912. doi:10.1681/asn.2013020126

Shi, J.-S., Qiu, D.-D., Le, W.-B., Wang, H., Li, S., Lu, Y.-H., et al. (2018). Identification of transcription regulatory relationships in diabetic nephropathy. *Chin. Med. J.* 131 (23), 2886–2890. doi:10.4103/0366-6999.246063

Song, L., Langfelder, P., and Horvath, S. (2013). Random generalized linear model: a highly accurate and interpretable ensemble predictor. *BMC Bioinforma.* 14, 5. doi:10.1186/1471-2105-14-5

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., et al. (2016). The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinforma.* 54, 1. doi:10.1002/cpbi.5

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47 (D1), D607-D613–D613. doi:10.1093/nar/gky1131

Tan, M., Pu, J., and Zheng, B. (2014). Optimization of breast mass classification using sequential forward floating selection (SFFS) and a support vector machine (SVM) model. *Int. J. Comput. Assist. Radiol. Surg.* 9 (6), 1005–1020. doi:10.1007/s11548-014-0992-1

Tanimura, S., Kuroiwa, C., and Mizota, T. (2006). Proportional symbol mapping in R. *J. Stat. Softw.* 15 (5), 1–7. doi:10.18637/jss.v015.i05

Thipsawat, S. (2021). Early detection of diabetic nephropathy in patient with type 2 diabetes mellitus: a review of the literature. *Diabetes and Vasc. Dis. Res.* 18 (6), 14791641211058856. doi:10.1177/14791641211058856

Thomas, M. C., Brownlee, M., Susztak, K., Sharma, K., Jandeleit-Dahm, K. A. M., Zoungas, S., et al. (2015). Diabetic kidney disease. *Nat. Rev. Dis. Prim.* 1, 15018. doi:10.1038/nrdp.2015.18

Thongrung, R., Lapmanee, S., Bray, P. T., Suthamwong, P., Deandee, S., Pangjit, K., et al. (2025). Gambogic acid mitigates nephropathy by inhibiting oxidative stress and inflammation in diabetic rats. *Int. J. Mol. Cell. Med.* 14 (1), 448–461. doi:10.22088/ijmcm.Bums.14.1.448

Wang, Y., Song, S., Qiu, D., Wu, G., Zheng, R., Zhao, L., et al. (2021). Effects of MiR-23b/MAPK on renal fibrosis in rats with diabetic nephropathy. *Minerva Med*. doi:10.23736/S0026-4806.20.07250-X

Wang, M., Pang, Y., Guo, Y., Tian, L., Liu, Y., Shen, C., et al. (2022a). Metabolic reprogramming: a novel therapeutic target in diabetic kidney disease. *Front. Pharmacol.* 13, 970601.

Wang, M., Pang, Y., Guo, Y., Tian, L., Liu, Y., Shen, C., et al. (2022b). Metabolic reprogramming: a novel therapeutic target in diabetic kidney disease. *Front. Pharmacol.* 13, 970601. doi:10.3389/fphar.2022.970601

Woroniecka, K. I., Park, A. S. D., Mohtat, D., Thomas, D. B., Pullman, J. M., and Susztak, K. (2011). Transcriptome analysis of human diabetic kidney disease. *Diabetes* 60 (9), 2354–2369. doi:10.2337/db10-1181

Xie, F., Lei, J., Ran, M., Li, Y., Deng, L., Feng, J., et al. (2020). Attenuation of diabetic nephropathy in diabetic mice by fasudil through regulation of macrophage polarization. *J. Diabetes Res.* 2020, 4126913. doi:10.1155/2020/4126913

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yu, W., Haoyu, Y., Ling, Z., Xing, H., Pengfei, X., Anzhu, W., et al. (2025). Targeting lipid metabolic reprogramming to alleviate diabetic kidney disease: molecular insights and therapeutic strategies. *Front. Immunol.* 16, 1549484. doi:10.3389/fimmu.2025.1549484

Yuan, Y., Li, L., Zhu, L., Liu, F., Tang, X., Liao, G., et al. (2020). Mesenchymal stem cells elicit macrophages into M2 phenotype *via* improving transcription factor EB-mediated autophagy to alleviate diabetic nephropathy. *Stem cells Dayt. Ohio* 38 (5), 639–652. doi:10.1002/stem.3144

Ziegler, H. K. (2004). The role of gamma/delta T cells in immunity to infection and regulation of inflammation. *Immunol. Res.* 29 (1-3), 293–302. doi:10.1385/ir:29:1-3:293