



Article ID 1007-1202(2025)06-0613-16 DOI <https://doi.org/10.1051/wujns/2025306613>

Cite this article: HE Zhanqi, CHEN Xinxin, HUANG Wenwen, *et al.* CPPED1 Links Neutrophil Extracellular Traps to Diabetic Nephropathy: Bioinformatics-Driven Expression Validation[J]. *Wuhan Univ J of Nat Sci*, 2025, 30(6): 613-628.

CPPED1 Links Neutrophil Extracellular Traps to Diabetic Nephropathy: Bioinformatics-Driven Expression Validation

□ HE Zhanqi, CHEN Xinxin, HUANG Wenwen, ZHAO Yanling[†]

Department of Nephrology, The Second Affiliated Hospital of Wenzhou Medical University, Wenzhou 325000, Zhejiang, China

Abstract: Diabetic nephropathy (DN) begins with diabetes-related disruptions in glucose metabolism, with oxidative stress playing a crucial role. Neutrophil extracellular traps (NETs) are extensive web-like formations composed of cytosolic and granule proteins, which are dependent on oxidative stress for their formation and function. This study aimed to identify potential targets for DN progression, focusing on NETs, using bioinformatic analysis and quantitative Real-Time PCR(qRT-PCR). We performed differential gene expression (DEG) analysis on two DN-related RNA-seq datasets (GSE142025 and GSE163603) and NETs-related genes. Subsequent analysis included gene set enrichment analysis (GSEA), gene set variation analysis (GSVA), GeneMANIA, and receiver operating characteristic (ROC) curves. Immune cell infiltration levels were assessed via single-sample GSEA, and a regulatory network involving RNA-binding proteins (RBPs) and their associated target mRNA was constructed. qRT-PCR was conducted on high glucose (HG)-treated and control HK-2 cells. Our analysis identified a set of 22 hub genes through the intersection of differentially expressed genes (DEGs) with NETs-related genes. Immune infiltration assessments revealed significant differences across 23 immune cell types among the analyzed groups. Hub genes including calcineurin-like phosphoesterase domain-containing 1 (CPPED1) showed high diagnostic values (AUC over 0.6). qRT-PCR indicated reduced gene expression levels. In summary, this study identified 22 significantly upregulated DEGs that play a vital role in DN by using gene expression omnibus (GEO) database, GSEA, GSVA, immune infiltration analysis and ROC. The expression levels of CPPED1 may serve as novel diagnostic biomarkers and therapeutic targets.

Key words: gene expression omnibus (GEO) database; diabetic nephropathy (DN); neutrophil extracellular traps (NETs); bioinformatics; quantitative Real-Time PCR (qRT-PCR); calcineurin-like phosphoesterase domain-containing 1 (CPPED1)

CLC number: R587.2

0 Introduction

Diabetic nephropathy (DN) stands as a severe complication linked with diabetes, posing a considerable health risk to millions worldwide. This disorder affects

approximately 30%-40% of diabetes patients globally, with epidemiological estimates suggesting over 840 million individuals exhibit subclinical manifestations^[1]. As the primary contributor to chronic kidney disease (CKD), DN leads to a deterioration in renal function and

Received date: 2025-05-07 © Wuhan University 2025

Foundation item: Supported by the Wenzhou Basic Scientific Research Project (Y2023103) and the Basic Public Welfare Research Program of Zhejiang Province (LGD19H070003)

Biography: HE Zhanqi, female, Master candidate, research direction: diabetic nephropathy. E-mail: hmpaxl@outlook.com

[†] Corresponding author. E-mail: zy1131@163.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

an increased risk of cardiovascular diseases^[2]. Common pharmacological options for DN such as sodium-glucose cotransporter 2 (SGLT2) inhibitors not only aid in glyce-mic control but also for their proven cardiorenal protec-tive effects^[3]. However, their inability to reverse estab-lished renal pathology (e.g., renal tubulointerstitial fibro-sis) underscores the urgent need for mechanistically novel interventions^[4].

Neutrophil extracellular traps (NETs), which are complex, net-like formations made up of DNA and pro-teins secreted by neutrophils, initially function as a de-fense mechanism against pathogens^[5]. Recent research has illuminated the pivotal role of NETs in various in-flammatory and autoimmune disorders. For instance, re-searchers have demonstrated that NETs facilitate osteo-clastogenesis in rheumatoid arthritis, where they in-crease inflammation and bone erosion^[6-7]. Emerging evi-dence implicates NETs in the pathogenesis of DN. Clini-cal and preclinical studies demonstrate elevated NETosis biomarkers (e. g., citrullinated histone H3, cell-free DNA) in both diabetic patients and murine DN models, suggesting active NETs involvement in disease progres-sion^[7]. Mechanistically, NETs promote the activation of the NLRP3 inflammasome under elevated glucose condi-tions^[8]. These findings position NETs inhibition as a po-tential therapeutic strategy, though the precise signaling pathways governing NET-mediated renal injury remain incompletely characterized. Bioinformatics analysis con-ducted by Zhang *et al*^[7] identified calcineurin-like phos-phoesterase domain-containing 1 (CPPED1) as a NET-related gene exhibiting increased expression. This phos-phoesterase modulates cellular signaling via substrate de-phosphorylation, notably inhibiting AKT1 activity through Ser473 residue dephosphorylation^[9]. While CPPED1 is established as a negative regulator of insulin-stimulated glucose uptake in adipocytes, its functional role in DN-associated NET formation warrants system-atic investigation^[10].

In this study, we employed bioinformatics methods such as differential gene expression and gene set enrich-

ment analysis (GSEA) to explore the differential genes related to DN and NETs, with functional annotation re-vealing enrichment in biological processes, including in-flammatory response, apoptosis, and metabolism. We constructed a protein-protein interaction (PPI) network and performed receiver operating characteristic curve analysis to assess the diagnostic capabilities of candidate genes. Concurrent immune infiltration profiling comple-mented these analyses, enabling multidimensional char-acterization of NETs-mediated pathogenic mechanisms to DN pathogenesis. Additionally, we conducted experi-mental validation of the significantly differential hub gene CPPED1, aiming to identify novel targets for the diagnosis and therapeutic intervention of DN.

1 Materials and Methods

1.1 Data Acquisition

The datasets utilized in the investigation were pro-cured from the GEO database, a well-established plat-form that provides public access to a diverse array of gene expression data from multiple areas. Specifically, the series accession codes for the datasets that were thor-oughly analyzed in this study are GSE142025 (includes samples from DN with a total of 21 samples confirmed by biopsy, and a control group consisting of 9 subjects from the non-involved regions of tumor nephrectomies) and GSE163603 (comprises 6 subjects with DN and 9 subjects in the control group). The genome-wide gene expression profiles from datasets GSE142025 and GSE163603 were efficiently retrieved through the GEO-query R package, enabling streamlined data extraction. Batch effects caused by non-biotechnical deviations were corrected using the ComBat method in R package "sva" based on the "empirical Bayes" framework^[11]. Then principal component analysis (PCA) was used to ensure result reliability and analytical accuracy. Addi-tionally, the NETs related genes were listed in Table 1^[12].

Table 1 The NETs related genes^[12]

PMID	Gene symbol
35688556	ALPL BST1 CD93 CEACAM3 CREB5 CRISPLD2 CSF3R CYP4F3 DYSF FCAR FCGR3B CPPED1 FPR1 FPR2 G0S2 HIST1H2BC HPSE CXCR1 CXCR2 KCNJ15 LILRB2 MGAM MME PDE4B S100A12 SIGLEC5 SLC22A4 SLC25A37
	TECPR2 TNFRSF10C VNN3 AKT1 AKT2 ATG7 CLEC6A CSF3 CTSG CYBB DNASE1 ELANE ENTPD4 F3 HMGB1
	IL17A IL1B IL6 IL8 IRAK4 ITGAM ITGB2 KCNN3 MAPK1 MAPK3 MMP9 MPO MTOR PADI4 PTAFR PIK3CA RIPK1 RIPK3 SELP SELPLG SIGLEC14 TLR2 TLR4 TLR7 TLR8 TNF

1.2 Differential Gene Expression Analysis

To discern differentially expressed genes (DEGs) between the control group ($n=18$) and the DN group ($n=27$), the R package "limma"^[13] was used. The selection criteria for DEGs encompassed a $|\log_2\text{Fold Change}| > 0.5$ and $p\text{-value} < 0.05$. Moreover, we utilized the pheatmap package to visualize the results of this analysis, which allow us to generate a heatmap that was based on Euclidean distance and hierarchical clustering, effectively illustrating the DEGs between different groups.

1.3 Gene Set Enrichment Analysis (GSEA)

We performed GSEA to evaluate whether a specified set of genes demonstrates statistically significant differences under two distinct biological conditions^[14]. This analysis used the R package "clusterProfiler". In this process, genes were meticulously ranked according to their $\log_2\text{Fold Change}$ values. We executed 1 000 permutations of gene sets to ensure robust results. The reference gene set utilized for the analysis was derived from the "c2.cp.kegg.v7.5.1.symbols" gene set from Molecular Signatures Database (MSigDB)^[15-16]. We defined significant enrichment as a $p\text{-value}$ below 0.05, which indicates meaningful biological relevance.

1.4 Gene Set Variation Analysis (GSVA)

GSVA is a nonparametric approach utilized for assessing the relationships between pathways and gene signatures to examine the functional disparities between the control and DN groups. It reveals a variety of pathways with differential expression that were effectively validated in a comprehensive heatmap. We utilized the "c2.cp.kegg.v7.5.1.symbols" gene set, executing the analysis using the R package GSVA. The outcomes of the GSVA were illustrated through the "pheatmap" package show the data clearly. GSVA scores for gene sets were performed utilizing the single-sample GSEA method. And by using the "limma" package, we compared the differences in GSVA scores between the control and DN groups. By converting gene expression data into a pathway-centric view, GSVA enhances the comprehension of the complex biological processes underlying genomic data in DN.

1.5 Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Enrichment Analysis

GO analysis, which encompasses the three categories of biological processes (BP), molecular functions (MF), and cellular components (CC), was performed to

conduct large-scale functional enrichment studies aimed at advancing our comprehension of disease mechanisms and pinpointing potential therapeutic targets. The KEGG database, esteemed for its exhaustive collection of information related to genomes, biological pathways, drugs, served as a resource to investigate the significantly enriched metabolic pathways^[17]. In this article, the "clusterProfiler" package facilitated GO and KEGG analysis of DEGs associated with the NETs and DN, aiming to systematically elucidate gene contributions to complex biological processes in DN, with a $p\text{-value} < 0.05$ ^[18].

1.6 GeneMANIA

The GeneMANIA website (<http://genemania.org>) stands as a comprehensive platform for construction and analysis of gene networks, encompassing a diverse spectrum of interactions. By harnessing the extensive capabilities of the GeneMANIA platform, we successfully established a detailed PPI network for these hub genes selected in DEGs. This network has provided an in-depth understanding of the complex interrelations that govern the functionality of hub genes, offering insights into the complex interactions within the networks involved in NETs in DN.

1.7 Receiver Operating Characteristic (ROC) Curve Analysis

The ROC curve analysis is a valuable tool for assessing diagnostic test efficacy. The ROC curve effectively represents the relationship between sensitivity and specificity, with the area under curve (AUC) serving as a widely used metric derived from it. It represents the operating characteristics associated with the sensitivity and specificity of the subjects under investigation. ROC curves were generated by "pROC" package aimed at calculating the AUC to identify the signature genes, thereby assessing the diagnostic significance^[19]. Typically, AUC values range from 0.5 to 1, with values approaching 1 signifying superior diagnostic performance.

1.8 Immune Infiltration Analysis

Immune Infiltration analysis is a technique employed to evaluate the extent of immune cell infiltration in various diseases, for uncovering the presence and distribution of immune cells within tissue or disease micro-environments through bioinformatic methods^[20]. This approach utilizes high-throughput transcriptomics technology, gene expression profile data, and a range of algorithms to conduct both quantitative and localization analyses of immune cells. Furthermore, data from the

Tumor Immune System Interaction Database (TISIDB) encompass a wide range of immune cell types^[21]. And the relative enrichment scores for each immune cell type were meticulously calculated using gene expression profiles, providing valuable insights into the immune landscape associated with the studied conditions. The levels of immune infiltration in DN and control groups were graphically represented utilizing the R package "ggplot2"^[22]. This analysis not only elucidates the immunological mechanisms associated with DN but also establishes a foundation for the identification of potential diagnostic biomarkers.

1.9 RNA Binding Protein (RBP)-mRNA Network

The construction of the RBP-mRNA network was undertaken with the aid of StarBase platform, which facilitated a comprehensive analysis of the interactions between non-coding RNAs and mRNAs. This analysis involved a detailed investigation of the relationships between mRNAs and RBPs through the utilization of various data types, including degradome-seq and RNA-RNA interaction data. Significant mRNA-RBP pairs were identified based on stringent cutoff criteria of p -value < 0.05, clusterNum $\geq$ 5, and clipExpNum $\geq$ 5. The RBP-mRNA network was constructed successfully through Cytoscape software, renowned for its capability to manage complex network data.

1.10 Cell Culture and Experimental Models

The HK-2 cell line, derived from human renal tubular epithelial cells, was sourced from Procell Life Science & Technology Co., Ltd., situated in Wuhan, China. Prior to experimentation, cell line underwent routine short tandem repeat profiling and testing for mycoplasma contamination to ensure its integrity and reliability. The cells were cultured in Procell full medium (CM-0109) at a controlled temperature of 37 °C within a CO₂ incubator, providing an optimal environment for growth. This specialized medium was composed of minimum essential medium (MEM) supplemented with 10% fetal bovine serum (FBS), nonessential amino acids and 1% penicillin-streptomycin (P/S). And three distinct experimental conditions were applied to the cell cultures over a period of 72 h to investigate the effects of varying glucose concentrations: one group was subjected to an exposure of 5.5×10^{-3} mol/L glucose, representing normal glucose (NG) levels, another group was subjected to 25×10^{-3} mol/L glucose, simulating high glucose (HG) conditions, and a third group received 5.5×10^{-3} mol/L glucose com-

bined with 19.5×10^{-3} mol/L mannitol, serving as an osmolar control (OSM).

1.11 Detection of Reactive Oxygen Species (ROS)

In this study, intracellular levels of ROS were evaluated using the fluorescent probe 2',7'-dichlorodihydrofluorescein diacetate (DCFH-DA) for a precise measurement of oxidative stress within the cells. In summary, cells seeded in two 6-well plates were divided into three groups (NG, HG, and OSM) with three wells per group and incubated for 72 h. Following the careful removal of the culture medium, the cells were meticulously treated with 2 mL of DCFH-DA, a compound prepared at a concentration of 10 μ mol/L in Dulbecco's modified eagle medium (DMEM), and then incubated at a controlled temperature of 37 °C for a duration of 20 min, allowing for optimal interaction and reaction between the cells and the treatment. Fluorescence intensity was meticulously visualized under a fluorescence microscope (200 $\times$). Uniform exposure time and gain values were set, and five non-overlapping visual fields were randomly selected for each well, covering both central and peripheral areas. Image acquisition was followed by background subtraction and signal correction using Image-Pro Plus; the "Measure" function was then used to quantify fluorescence intensity, which was normalized to the cell count. All experiments were carried out in triplicate, incorporating a minimum of three biological replicates.

1.12 Quantitative Real-Time PCR (qRT-PCR) and RNA Extraction

Total RNA was meticulously isolated from samples which were treated under three distinct conditions: NG, HG, and OSM, utilizing the well-established TRIzol® reagent from Invitrogen. Using the PrimeScript™ RT reagent Kit from Takara, RNA was reverse transcribed into complementary DNA (cDNA). qRT-PCR was performed using TB Green Premix Ex Taq II, initiating process at 95 °C for 30 s, 40 cycles of 95 °C for 5 s and 60 °C for 30 s in 25 μ L reaction volume. The primer sequences were: CPPED1 forward: 5'-GGGTGTTTTCCA CAGAGCCA-3', CPPED1 reverse: 5'-GATCTCCTGTT CCCATTCGTCA-3', GAPDH forward: 5'-GGAAGCTT GTCATCAATGGAAATC-3', GAPDH reverse: 5'-TGAT GACCCTTTTGGCTCCC-3'.

1.13 Statistical Analysis

Statistical evaluations used R4.1.2; Spearman's cor-

relation assessed parameter relationships, Wilcoxon test compared two groups, and Kruskal-Wallis test analyzed three or more groups, with $p < 0.05$ as significant. Statistical analysis of the ROS data was performed using an independent-samples t test and data visualization was conducted using the "ggplot2" package.

2 Results

2.1 Differential Gene Expression Related to DN

In this investigation, a total of 4 222 DEGs were discerned by contrasting samples obtained from DN patients with those from healthy controls. These identified genes exhibited statistically significant variations between two cohorts. Among the samples collected from DN patients, a substantial number of 2 178 genes were

found to be upregulated, indicating an increase in their expression levels, while 2 044 genes were downregulated, reflecting a decrease in their expression. The DEGs were visually represented through a volcano plot (Fig. 1(a)), providing a clear illustration of the expression changes, and a heatmap was employed to illustrate the five most significant upregulated genes (DKK3, C1orf216, RNASE1, MFAP4, SIT1) and the five most downregulated genes (NR4A1, DUSP1, NR4A2, KBTBD11, VEGFA) (Fig. 1(b)). A rank-sum test was conducted to validate the notable variations in the expression levels of the ten most prominent genes when comparing the DN group to the control group (Fig. 1(c)). Furthermore, we identified 22 NETs-related DEGs, which were designated as hub genes through the identification of the overlap between the DEGs and those genes linked to NETs.

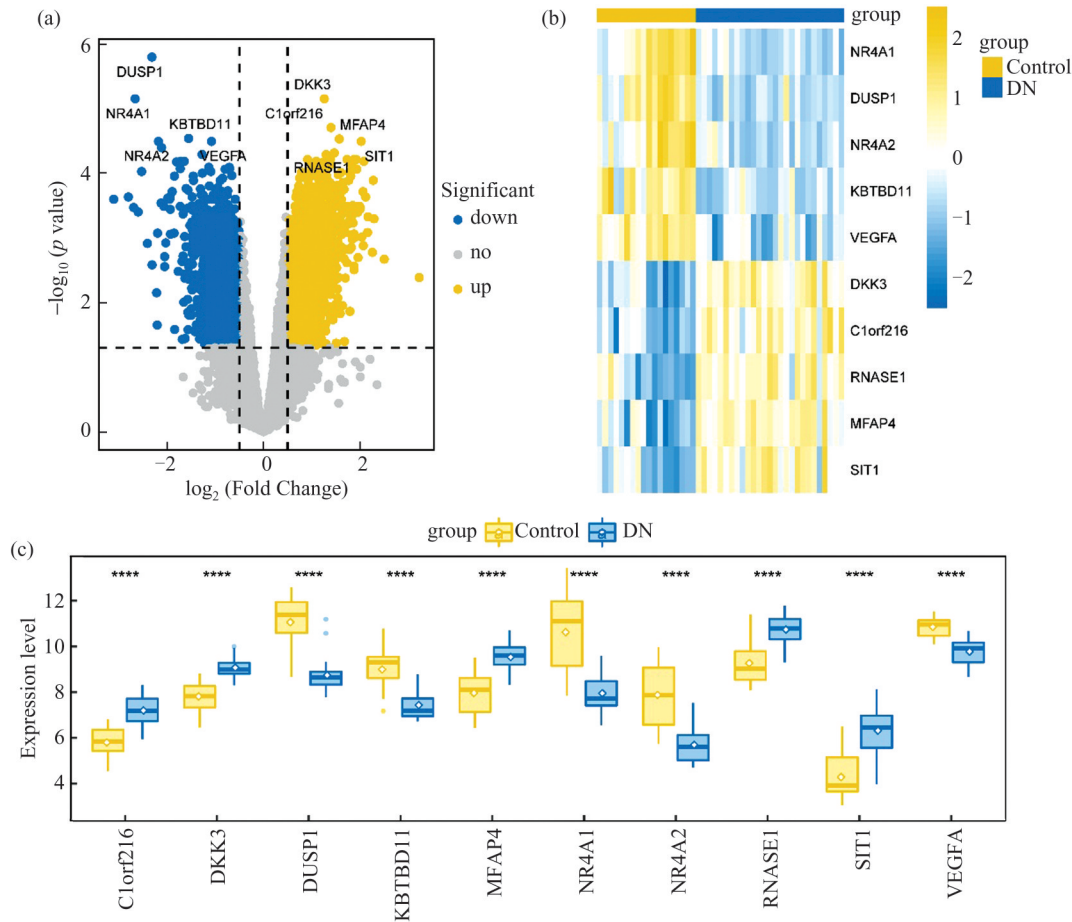


Fig. 1 The DEGs in DN vs. control

(a) Volcano plot for DEG distribution; (b) Heatmap for top DEGs; (c) Box plot for expression differences
**** $p < 0.0001$.

2.2 Analysis Results of GSEA

The GSEA aimed to reveal the mechanisms linked to the identified DEGs. We pinpointed the signaling

pathways that exhibit the highest levels of enrichment based on their normalized enrichment scores (NES) utilizing the MSigDB database (see Table 2). GSEA indi-

Table 2 Pathways and functions in GSEA analysis

Description	Set-size	Enrichment score	NES	p
KEGG_RIBOSOME	84	0.779 915 232	3.270	0.001 760 563
KEGG_CHEMOKINE_SIGNALING_PATHWAY	158	0.557 120 183	2.621	0.001 594 896
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	180	0.556 108 13	2.653	0.001 605 14
KEGG_PROXIMAL_TUBULE_BICARBONATE_RECLAMATION	29	-0.714 102 76	-2.291	0.002 066 116
KEGG_CITRATE_CYCLE/TCA_CYCLE	16	-0.621 900 279	-2.196	0.002 222 222
KEGG_VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	42	-0.599 894 236	-2.303	0.002 237 136

cated substantial enrichment in the RIBOSOME pathway, the CYTOKINE-CYTOKINE RECEPTOR INTERACTION pathway (CCRIP), the CHEMOKINE SIGNALING pathway (CSP), etc. Additionally, pathways such as the citrate cycle (also known as TCA cycle), and proximal tubule bicarbonate reclamation (PTBR) were also found to be significantly enriched, highlighting the complex biological processes involved (Fig. 2(a-f)).

2.3 Analysis Results of GSVA

The GSVA method further assessed the comparative expression differences of various biological pathways between the DN and control groups, revealing a multitude of pathways with differentially expressed that were represented in heatmap. In the DN cohort, the expression levels of the CITRATE CYCLE TCA CYCLE and BUTANOATE METABOLISM were markedly diminished, indicating a significant downregulation, while the RIBOSOME and PATHOGENIC_ESCHERICHIA_COLI_INFECTION pathways exhibited a significant elevation in contrast to the control groups (Fig. 2(g)).

2.4 Enrichment Analyses of GO and KEGG

GO and KEGG enrichment analyses were performed to further explore the biological functions linked to NETs. The GO analysis uncovered significant enrichment in genes related to neutrophil-mediated immunity and cellular response to abiotic stimulus within the BP category, highlighting the critical roles these genes play in immune responses. Within the CC category, an enrichment of genes was identified in tertiary granule, tertiary granule membrane, and secretory granule membrane, indicating their localization and functional significance. For MF, enrichment was observed in activities such as pattern recognition receptor activity, NAD+ nucleosidase activity (Fig. 3(a), (c-e)). The KEGG pathway analysis further revealed notable enrichment in pathways including neutrophil extracellular trap formation, Leish-

maniasis, and the Toll-like receptor signaling pathway, among others (Fig. 3(b),(f)).

2.5 Analysis of Hub Gene Interactions

We constructed a PPI network for 22 hub genes, utilizing the GeneMANIA database (Fig. 4(a)). In order to investigate the roles of the hub genes, we performed GO and KEGG analyses on a comprehensive set of 42 genes (encompassing 22 hub genes and 20 associated genes). The GO enrichment analysis highlighted significant pathways, including the enhancement of the ERK1 and ERK2 cascade (GO: 0070374), endoderm development (GO: 0007492), multi-multicellular organism process (GO:0044706) (BP), rough endoplasmic reticulum (GO: 0005791), caveola (GO: 0005901), vacuolar membrane (GO: 0005774) (CC), monosaccharide binding (GO: 0048029) (MF) (Fig. 4(b)). Besides, the KEGG analysis indicated noteworthy enrichment in Neutrophil extracellular trap formation (hsa04613), Leishmaniasis (hsa05140), Leukocyte transendothelial migration (hsa04670), Phagosome (hsa04145), Lipid and atherosclerosis (hsa05417), Malaria (hsa05144), Tuberculosis (hsa05152), and Coronavirus disease - COVID-19 (hsa05171) (Fig. 4(c)).

2.6 Validation of Gene Sets in Key Pathways

The enrichment of gene sets tumor necrosis factor (TNF) superfamily cytokine production and the TLR signaling pathway were conducted through the assessment of hub genes. As illustrated in Fig. 5(a-b), the findings reveal that the expression levels of most critical genes are markedly increased in the DN group.

2.7 Diagnostic Value of Hub Genes

To ascertain their diagnostic capabilities, ROC curves were constructed to systematically evaluate their performance. The AUC values for several pivotal hub genes, including MMP9 with an impressive AUC of 0.85, MAPK3 at 0.84, PTAFR also at 0.84, and CPPED1 with a value of 0.757, were all notably greater than the

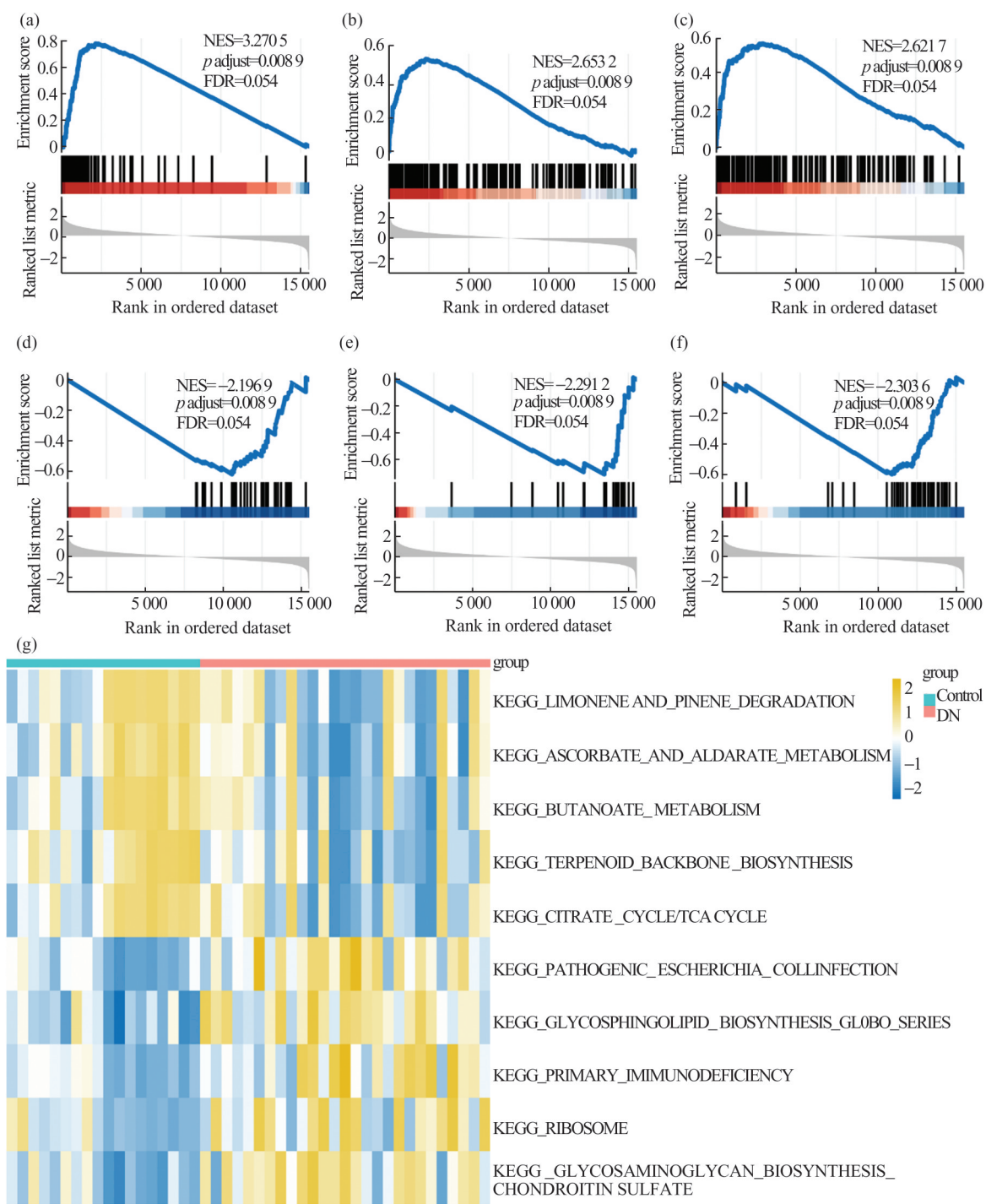


Fig. 2 Significantly enriched pathways

(a) RIBOSOME; (b) CYTOKINE CYTOKINE RECEPTOR INTERACTION; (c) CHEMOKINE SIGNALING PATHWAY; (d) CITRATE CYCLE TCA CYCLE; (e) PROXIMAL TUBULE BICARBONATE RECLAMATION; (f) VALINE LEUCINE AND ISOLEUCINE DEGRADATION; (g) Visualized through a heatmap of GSVA; FDR: false discovery rate.

threshold of 0.6. This finding indicates their potential utility as biomarkers for DN (Fig. 6(a-f)).

2.8 Immune Infiltration Analysis

Immune infiltration analysis demonstrated notable variations in 23 of the 28 immune cell types when com-

paring DN patients to healthy controls ($p < 0.05$) (Fig. 7 (a)). Specifically, a marked increase in infiltration levels was observed for various immune cells, including regulatory T cells and activated CD8 T cell, within DN groups relative to the control group (Fig. 7(b)). Correlation analysis further indicated a robust association between

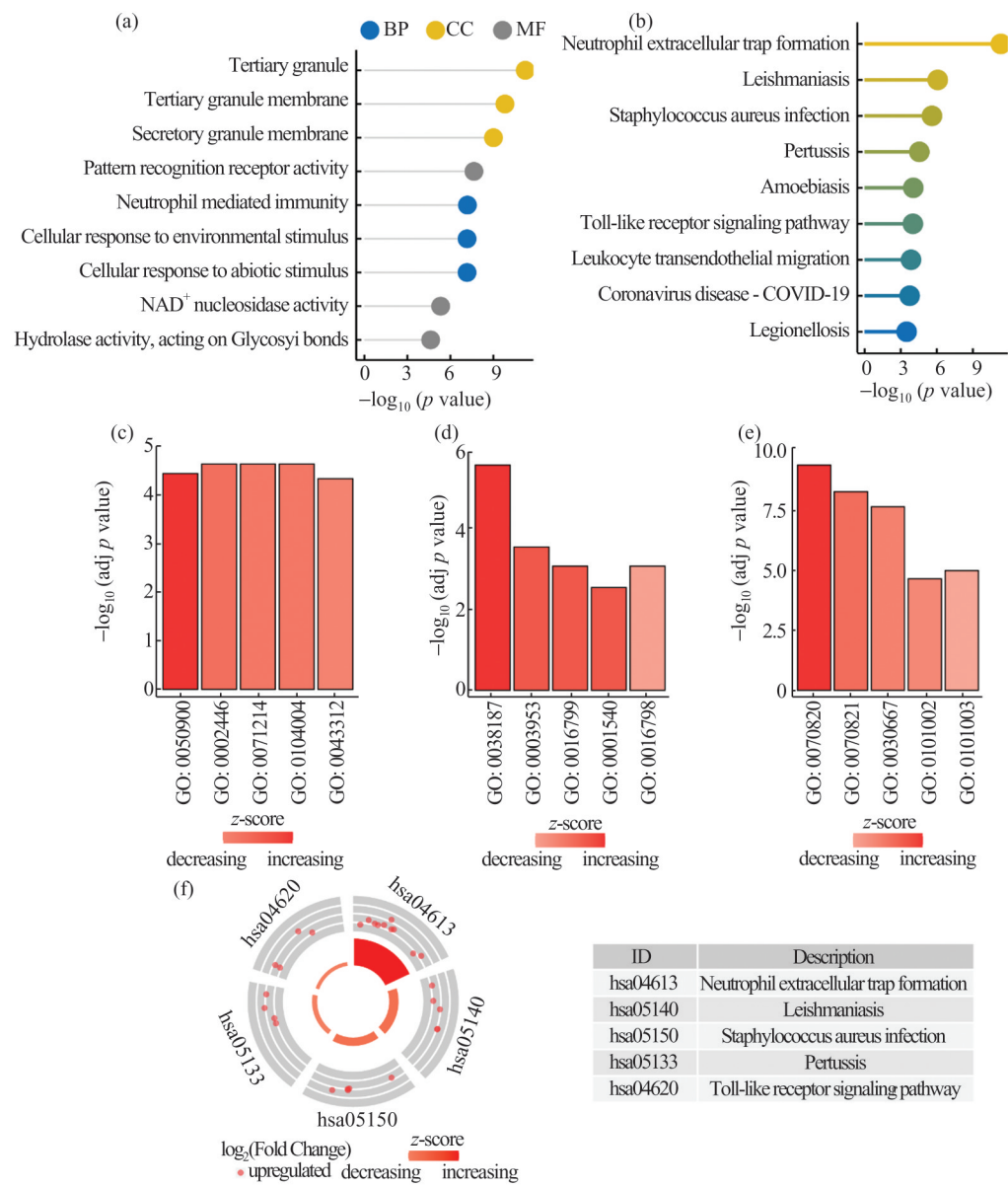


Fig. 3 Enrichment analysis of DEGs pertaining to NETs

(a) Lollipop chart of GO terms; (b) Lollipop chart of KEGG pathways; (c) Bar chart of BP; (d) Bar chart of MF; (e) Bar chart of CC; (f) Circular diagram of KEGG pathways.

regulatory T cells and ITGAM ($R=0.921$, $p<0.001$), and a significant association between myeloid-derived suppressor cells and SELPLG ($R=0.932$, $p<0.001$) (Fig. 7 (c-d)).

2.9 Signaling Pathways Associated with Hub Genes

We conducted an investigation into the association between the five hub genes exhibiting the most notable differential expression and the 50 Hallmark signaling pathways (Fig. 8(a)). Furthermore, we also investigated the differences between the DN and control group within the same pathways utilizing GSVA. In DN pa-

tients, 23 Hallmark signaling pathways exhibited significant upregulation, specifically HALLMARK_ANGIOGENESIS, HALLMARK_APICAL_JUNCTION, HALLMARK_APOPTOSIS. Additionally, there were five significantly downregulated pathways, including HALLMARK_HEME_METABOLISM, HALLMARK_KRAS_SIGNALING_DN, HALLMARK_PANCREAS_BETA_CELLS, HALLMARK_SPERMATOGENESIS, HALLMARK_XENOBIOTIC_METABOLISM (Fig. 8(b)).

2.10 Interaction Network of Hub mRNAs and Their Associated RBPs

We also constructed an interaction network of hub

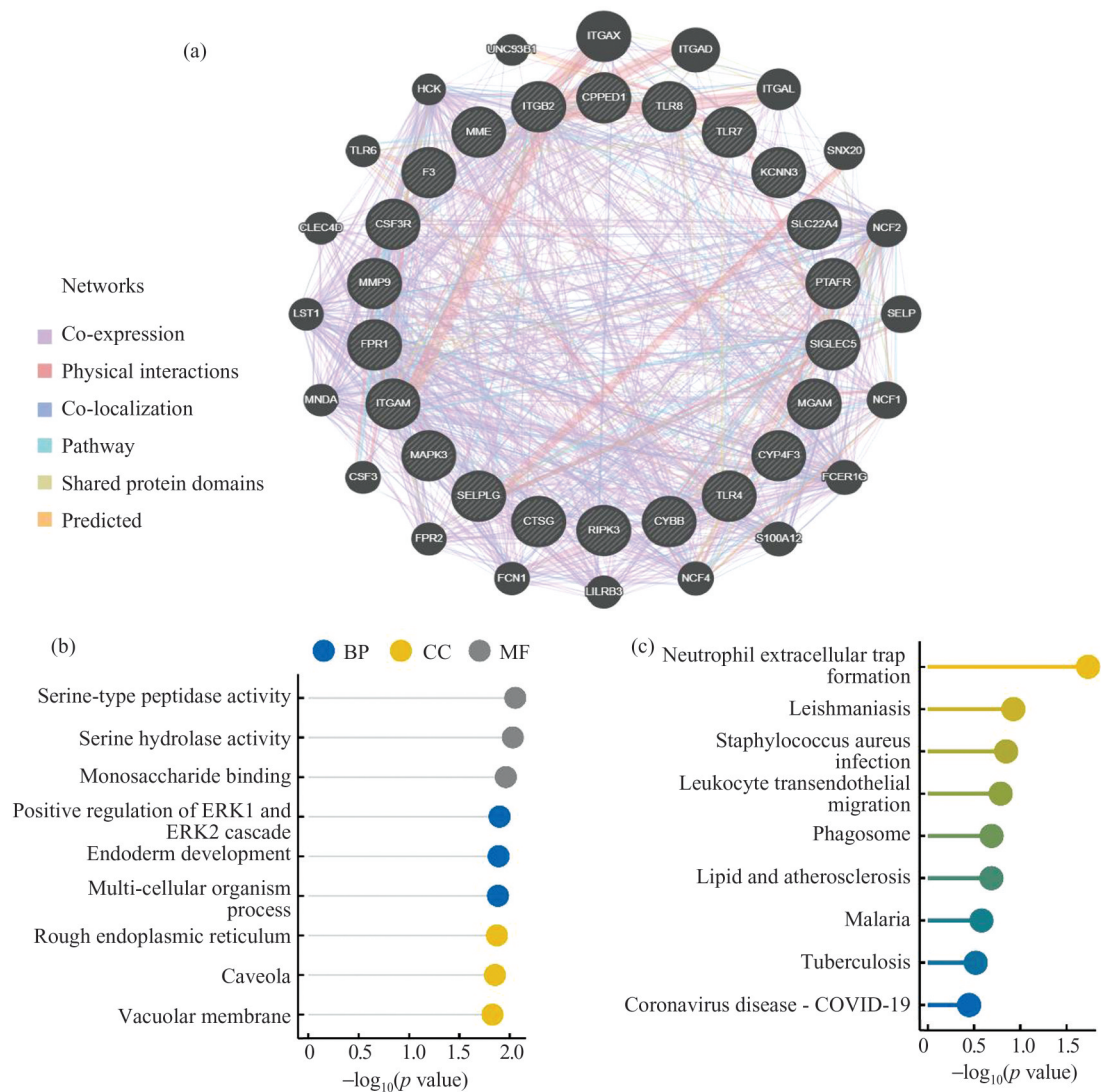


Fig. 4 Interaction analysis of hub genes

(a) PPI network, (b) GO enrichment results, (c) KEGG pathway enrichment results.

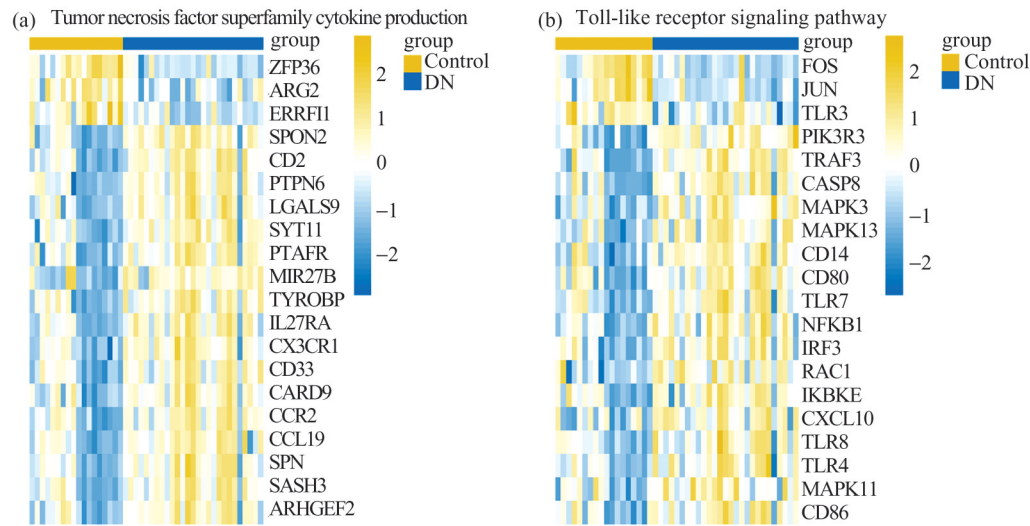


Fig. 5 Heatmaps of key gene expression profiles in the gene sets for (a) TNF superfamily cytokine production and (b) TLR signaling pathway in DN and control

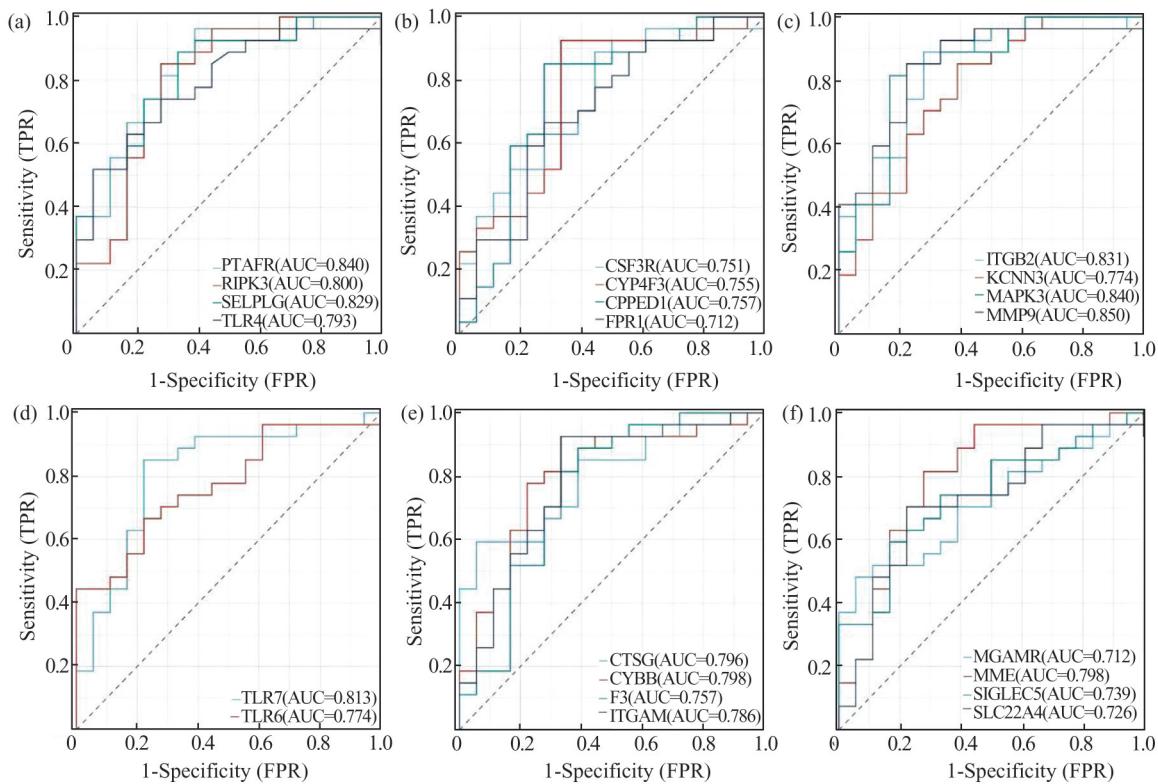


Fig. 6 The ROC curve of identified hub genes (a-f)

mRNAs and their associated RBPs using the StarBase online database, identifying 22 hub mRNAs. Subsequently, we obtained the relevant mRNA/RBP pairs for 18 of these mRNAs, leading to the construction of an RBP-mRNA network based on interactions sourced from an online dataset. As shown in Fig. 9, the network consists of 58 nodes, including 40 RBPs and 18 mRNAs, along with a total of 145 edges.

2.11 Preliminary Experimental Verification

Measurement of intracellular ROS levels in HK-2 cells using DCFH-DA demonstrated significantly stronger fluorescence intensity in HG group than the NG and OSM groups, indicating elevated ROS production under hyperglycemic conditions. Therefore, high glucose concentrations have the potential to trigger oxidative stress in HK-2 cells, leading to a marked increase in intracellular ROS levels. Notably, no statistically significant difference in fluorescence intensity was detected when comparing the NG group to the OSM group, implying comparable ROS levels and further confirming that osmotic pressure variations did not substantially alter the oxidative stress status in these cells. To investigate the hub genes associated with DN within the biological network, we conducted qRT-PCR on CPPED1 utilizing cDNA derived from HG, OSM, and control cell samples.

Figure 10 shows that CPPED1 expression in HG groups was significantly lower than that in the control group.

3 Discussion

The study of DN holds considerable clinical significance, given that it constitutes a major complication of diabetes mellitus that exerts a significant impact on patients' quality of life and survival outcomes. NETs, which are intricate web-like formations produced by neutrophils and composed of DNA, histones, and granule proteins, were initially believed to function solely as a defense mechanism against pathogens^[23]. Nevertheless, emerging evidence has begun to illuminate a strong association between NETs and the inflammatory response observed in DN. Targeting the dysregulation of neutrophils, particularly concerning the development of NETs, may significantly modulate neutrophil-mediated inflammation of patients suffering from diabetes. In this study, NETs contribute to renal inflammation and damage, highlighting their role in both host defense and the pathophysiology of DN^[24]. Therefore, elucidating the complex and multifaceted role of NETs in DN could provide invaluable insights that may inform clinical interventions and specific inhibitors of NETosis (e.g., DNase

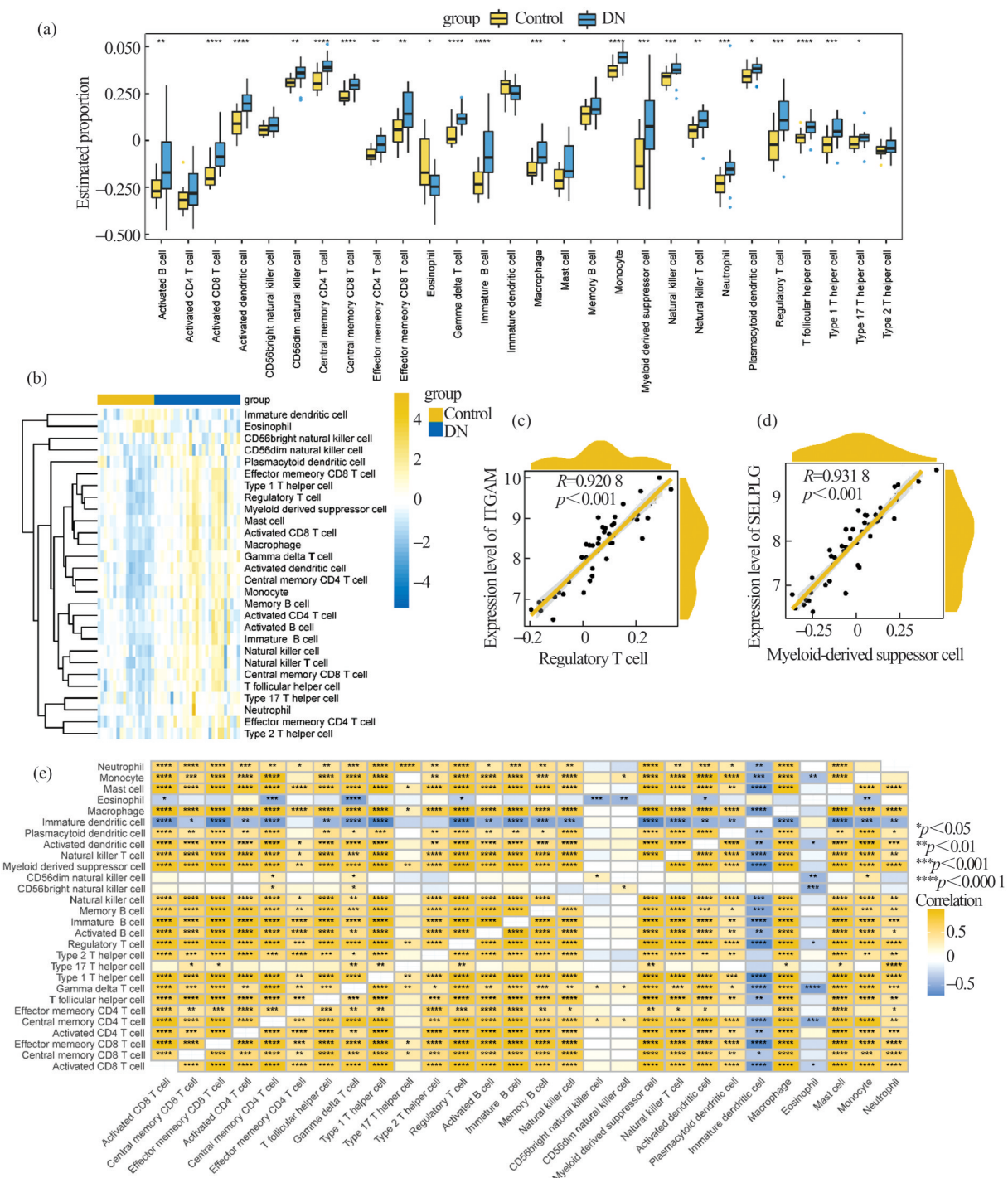


Fig. 7 Immune infiltration differences in DN vs control

(a) Bar chart showing immune infiltration abundance; (b) Heatmap of immune infiltration levels; (c) Scatter plot of correlation between ITGAM and regulatory T cells; (d) Scatter plot of correlation between SELPLG and myeloid-derived suppressor cells; (e) Correlation matrix of immune cells.

I) may attenuate renal damage.

To investigate this hypothesis, we employed a systems biology approach combined with multiple databases and basic experimental validation. Additionally, we integrated DEGs, GSEA, and immune infiltration analysis, and validated NETs-related DEGs using qRT-

PCR. Notably, our DGE analysis revealed CPPED1, a NETs-related gene was significantly downregulated in DN, suggesting a potential regulatory link to NETosis, which provides an important foundation for subsequent research and clinical applications.

This study using GSVA and GSEA found signifi-

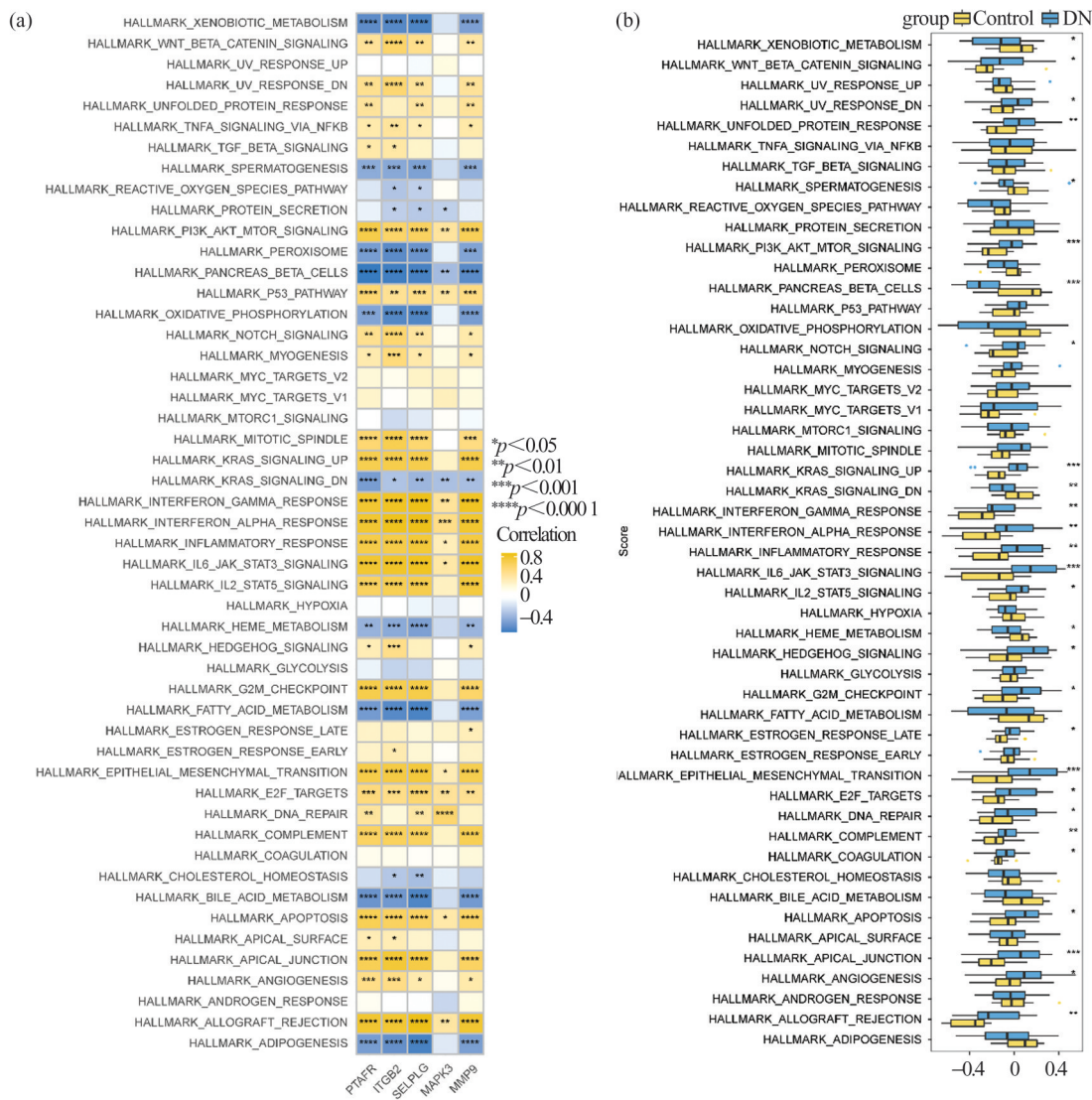


Fig. 8 Correlation between hub genes (a) and 50 hallmark signaling pathways (b)

cant enrichment of the CCRIP and CSP in DN, linked to inflammatory and immune responses^[25]. Cytokines, encompassing TNF, interleukins and interferons, can facilitate the formation of dimeric or multimeric complexes that activate intracellular signaling cascades, notably the JAK/STAT pathway, thereby modulating inflammatory progression and immune cell activity^[26-28]. Experimental studies have shown that inhibiting TNF- α using soluble TNF- α receptor fusion proteins can delay disease progression in DN^[29]. Chemokines are a class of small molecular cytokines with chemotactic effects that guide directed cell migration through concentration gradients, primarily secreted by immune cells, endothelial cells, or tissue cells^[30-31]. In DN, monocyte chemoattractant protein-1 (MCP-1) can recruit monocytes and macrophages, thereby inducing inflammatory responses in the kid-

neys^[32]. In summary, prior research has validated the significant roles of these pathways in inflammatory mechanisms. The enrichment of CCRIP and CS found in the DN group further supports that chronic low-grade inflammation and immune cell infiltration are characteristics of DN.

Through the analysis of immune cell infiltration in DN and control groups, it was discerned that among 28 immune cell subtypes, 23 revealed notable variations in immune infiltration abundance. The level of neutrophil infiltration in the DN group significantly increased, suggesting the participation of neutrophils in the inflammatory processes linked to DN. Neutrophils serve as the initial line of defense for the host organism and represent the predominant type of white blood cells found in circulation, constituting approximately 50% to 70% of the to-

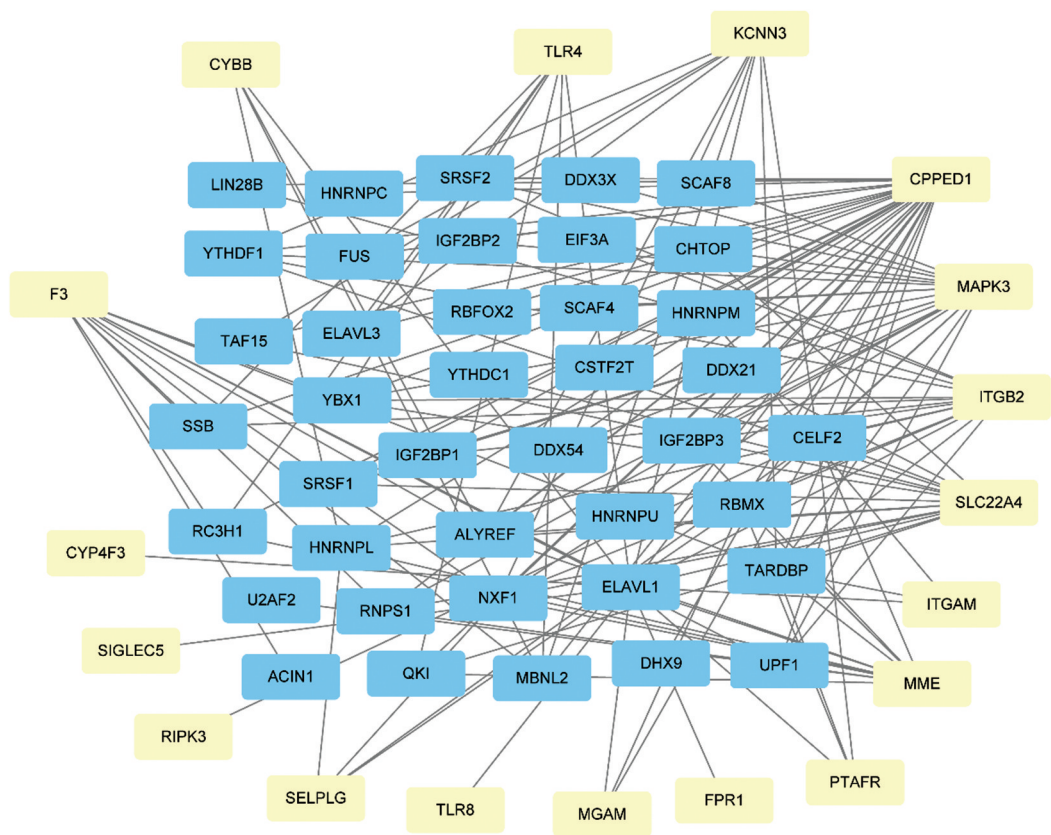


Fig. 9 RBP-mRNA regulatory network
Yellow symbolizes mRNAs and blue symbolizes RBPs.

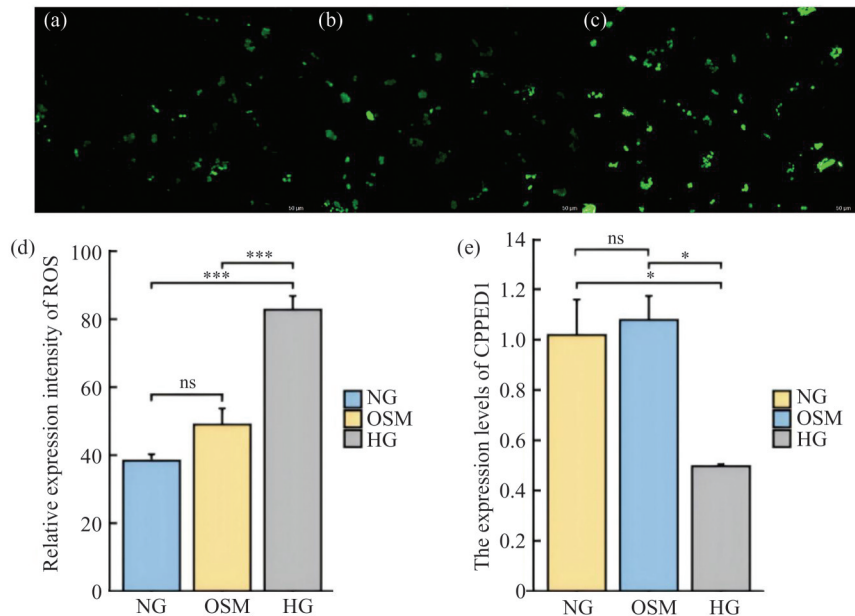


Fig. 10 Comparison of ROS fluorescence levels within NG, OSM and HG (a-c), fluorescence results from NG, OSM and HG groups (d), qRT-PCR experiment of CPPED1 (e)

* $p < 0.05$; *** $p < 0.001$; $n=3$; ns: $p > 0.05$.

tal white blood cell population in the bloodstream. They primarily defend against infections through methods

such as phagocytosis, degranulation, and NETs formation^[33-34]. This further proved the association of NETs

with DN. Neutrophils are crucial contributors to inflammatory responses, largely through the secretion of cytokines and chemokines that influence the activity of diverse immune cells. This includes the secretion of B cell activating factor (BAFF) and various chemokine ligands 2/3/19/20 (CCL2/3/19/20) to promote the accumulation of monocytes at the site of inflammation^[35-37]. This finding aligns with the outcomes observed in GSEA and GSVA, where inflammatory pathways were significantly upregulated. Additionally, the level of ROS in the DN cell model was detected using DCFH-DA in this study, revealing a significant increase in oxidative stress levels within the DN group, corroborating the previously mentioned findings.

To ascertain the diagnostic capabilities of hub genes, we employed ROC, revealing that all 22 hub genes exhibited AUC values exceeding 0.6, indicating their potential utility as biomarkers. Furthermore, GSVA indicated significant variations across multiple Hallmark pathways among the hub genes, and we constructed interaction networks with RBPs for these hub genes, indicating a model to improve understanding of underlying mechanisms. In order to experimentally validate these findings, we conducted qRT-PCR analysis and discovered that CPPED1 expression was markedly reduced in DN samples compared to the control group. CPPED1, encoding a phosphatase-domain protein, functions as a critical regulator of cellular signaling pathways through dephosphorylation of key substrates. Notably, it suppresses Protein Kinase B alpha (also known as AKT1) activity via dephosphorylating the Ser473 residue of AKT1, a central node in metabolic and survival signaling^[9]. Emerging evidence implicates CPPED1 in oncogenesis and glucose metabolism. In non-invasive bladder cancer, CPPED1 expression is significantly downregulated, while its overexpression correlates with tumor regression^[9]. Moreover, in adipocytes, CPPED1 negatively regulates insulin sensitivity, as its knockdown enhances insulin-stimulated glucose uptake^[10]. In our DN cell model under hyperglycemic conditions, CPPED1 exhibited marked downregulation in renal tubular epithelial cells. This observation aligns with prior studies linking CPPED1 to glucose metabolism but reveals an inverse regulatory pattern in renal pathophysiology. The reduced expression of CPPED1 in DN models suggests its potential role in disease pathogenesis, which may promote NETosis by dysregulating AKT1-dependent survival signaling, thereby exacerbating renal inflamma-

tion. However, current research on the relationship between CPPED1 and the kidneys is limited, which needs further validation. We hypothesize that CPPED1, identified as a gene related to NETs, plays a significant role in the inflammatory cascades driven by NETs, which may contribute to the progression and worsening of DN. The findings from our research suggest that CPPED1 could serve as a promising therapeutic target, with the potential to restore glucose homeostasis and mitigate inflammation associated with DN, thereby offering a new avenue for treatment strategies aimed at improving patient outcomes.

In conclusion, this study aims to explore the regulation of NETs by investigating the hub genes and pathways involved in DN. Although we conducted thorough analyses and laboratory experiments, the lack of *in vivo* validation (e.g., CPPED1 knockout mouse models) and clinical cohort analyses limits the translatability of our findings. Furthermore, while using multiple datasets enhances the robustness of our analysis, this approach may also introduce batch effects that are not fully corrected by Combat. Moreover, the GSE163603 dataset is relatively small, which may affect the reliability of the differentially expressed genes analysis. In the future, we will increase animal model experiments and clinical samples for further validation.

This study employed bioinformatics analyses and experimental validation to elucidate NET-related mechanisms in DN. Differential expression analysis of DN and NETs-associated genes identified 22 hub genes. Multidimensional approaches—including GSEA, GSVA, GeneMANIA interaction networks, and ROC curve assessments—were integrated to delineate NET-related gene expression profiles. Single-sample GSEA quantified immune cell infiltration, while PPI networks mapped RBP-mRNA regulatory axes. Among hub genes, CPPED1 demonstrated high diagnostic value (AUC = 0.757), corroborated by qRT-PCR showing significant downregulation in DN samples ($p < 0.01$). However, the underlying mechanisms remain unclear, warranting further investigation into stage-specific molecular dynamics.

References

- [1] Duan J Y, Wang C J, Liu D W, *et al.* Prevalence and risk factors of chronic kidney disease and diabetic kidney disease in Chinese rural residents: A cross-sectional survey[J]. *Scien-*

- tific Reports*, 2019, **9**: 10408.
- [2] Tuttle K R, Agarwal R, Alpers C E, *et al*. Molecular mechanisms and therapeutic targets for diabetic kidney disease[J]. *Kidney International*, 2022, **102**(2): 248-260.
 - [3] Neuen B L, Heerspink H J L, Vart P, *et al*. 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[J]. *Circulation*, 2024, **149**(6): 450-462.
 - [4] Ishibashi T, Morita S, Furuta H, *et al*. Renoprotective potential of concomitant medications with SGLT2 inhibitors and renin-angiotensin system inhibitors in diabetic nephropathy without albuminuria: A retrospective cohort study[J]. *Scientific Reports*, 2023, **13**: 16373.
 - [5] Hidalgo A, Libby P, Soehnlein O, *et al*. Neutrophil extracellular traps: From physiology to pathology[J]. *Cardiovascular Research*, 2022, **118**(13): 2737-2753.
 - [6] O'Neil L J, Oliveira C B, Wang X H, *et al*. Neutrophil extracellular trap-associated carbamylation and histones trigger osteoclast formation in rheumatoid arthritis[J]. *Annals of the Rheumatic Diseases*, 2023, **82**(5): 630-638.
 - [7] Zhang H, Wang Y, Qu M D, *et al*. Neutrophil, neutrophil extracellular traps and endothelial cell dysfunction in sepsis[J]. *Clinical and Translational Medicine*, 2023, **13**(1): e1170.
 - [8] Gupta A, Singh K, Fatima S, *et al*. Neutrophil extracellular traps promote NLRP3 inflammasome activation and glomerular endothelial dysfunction in diabetic kidney disease[J]. *Nutrients*, 2022, **14**(14): 2965.
 - [9] Haapalainen A M, Daddali R, Hallman M, *et al*. Human CPPED1 belongs to calcineurin-like metallophosphoesterase superfamily and dephosphorylates PI3K-AKT pathway component PAK4[J]. *Journal of Cellular and Molecular Medicine*, 2021, **25**(13): 6304-6317.
 - [10] Vaitinen M, Kaminska D, Käkälä P, *et al*. Downregulation of CPPED1 expression improves glucose metabolism *in vitro* in adipocytes[J]. *Diabetes*, 2013, **62**(11): 3747-3750.
 - [11] Leek J T, Johnson W E, Parker H S, *et al*. The sva package for removing batch effects and other unwanted variation in high-throughput experimentsFree[J]. *Bioinformatics*, 2012, **28**(6): 882-883.
 - [12] Zhang Y, Guo L P, Dai Q C, *et al*. A signature for pan-cancer prognosis based on neutrophil extracellular traps[J]. *Journal for Immunotherapy of Cancer*, 2022, **10**(6): e004210.
 - [13] Ritchie M E, Phipson B, Wu D, *et al*. Limma powers differential expression analyses for RNA-sequencing and microarray studies[J]. *Nucleic Acids Research*, 2015, **43**(7): e47.
 - [14] Subramanian A, Tamayo P, Mootha V K, *et al*. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**(43): 15545-15550.
 - [15] Liberzon A, Birger C, Thorvaldsdóttir H, *et al*. The Molecular Signatures Database (MSigDB) hallmark gene set collection[J]. *Cell Systems*, 2015, **1**(6): 417-425.
 - [16] Liberzon A, Subramanian A, Pinchback R, *et al*. Molecular Signatures Database (MSigDB) 3.0[J]. *Bioinformatics*, 2011, **27**(12): 1739-1740.
 - [17] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes[J]. *Nucleic Acids Research*, 2000, **28**(1): 27-30.
 - [18] Yu G C, Wang L G, Han Y Y, *et al*. clusterProfiler: An R package for comparing biological themes among gene clusters[J]. *Omics*, 2012, **16**(5): 284-287.
 - [19] Robin X, Turck N, Hainard A, *et al*. pROC: An open-source package for R and S+ to analyze and compare ROC curves[J]. *BMC Bioinformatics*, 2011, **12**(1): 77.
 - [20] Wu S, Lv X K, Li Y L, *et al*. Integrated machine learning and single-sample gene set enrichment analysis identifies a TGF-beta signaling pathway derived score in headneck squamous cell carcinoma[J]. *Journal of Oncology*, 2022, **2022**: 3140263.
 - [21] Ru B B, Wong C N, Tong Y, *et al*. TISIDB: An integrated repository portal for tumor-immune system interactionsFree [J]. *Bioinformatics*, 2019, **35**(20): 4200-4202.
 - [22] Ito K, Murphy D. Application of ggplot2 to pharmacometric graphics[J]. *CPT: Pharmacometrics & Systems Pharmacology*, 2013, **2**(10): 79.
 - [23] Brinkmann V, Reichard U, Goosmann C, *et al*. Neutrophil extracellular traps kill bacteria[J]. *Science*, 2004, **303**(5663): 1532-1535.
 - [24] Dowe R, Iqbal A, Heller S R, *et al*. A bittersweet response to infection in diabetes; targeting neutrophils to modify inflammation and improve host immunity[J]. *Frontiers in Immunology*, 2021, **12**: 678771.
 - [25] Bhat A A, Nisar S, Singh M, *et al*. Cytokine- and chemokine-induced inflammatory colorectal tumor microenvironment: Emerging avenue for targeted therapy[J]. *Cancer Communications*, 2022, **42**(8): 689-715.
 - [26] Galozzi P, Bindoli S, Doria A, *et al*. The revisited role of interleukin-1 alpha and beta in autoimmune and inflammatory disorders and in comorbidities[J]. *Autoimmunity Reviews*, 2021, **20**(4): 102785.
 - [27] Xue X, Liu X H, Lu C L, *et al*. Chinese patent herbal medicine Huaiqihuang for Henoch-Schonlein purpura nephritis in children: A systematic review of randomized controlled trials [J]. *BMC Complementary Medicine and Therapies*, 2021, **21**

- (1): 278.
- [28] Zhao H K, Wu L, Yan G F, *et al.* Inflammation and tumor progression: Signaling pathways and targeted intervention [J]. *Signal Transduction and Targeted Therapy*, 2021, **6**: 263.
- [29] Guan Y K, Peiffer B, Feng D C, *et al.* IL-8+ neutrophils drive inexorable inflammation in severe alcohol-associated hepatitis[J]. *The Journal of Clinical Investigation*, 2024, **134** (9): e178616.
- [30] Hughes C E, Nibbs R J B. A guide to chemokines and their receptors[J]. *The FEBS Journal*, 2018, **285**(16): 2944-2971.
- [31] Serra H M, Baena-Cagnani C E, Eberhard Y. Is secondary lymphoid-organ chemokine (SLC/CCL21) much more than a constitutive chemokine? [J]. *Allergy*, 2004, **59**(11): 1219-1223.
- [32] Liu Y S, Li Y, Liu Y X, *et al.* Investigation of the shared biomarkers in heterotopic ossification between ossification of the ligamentum flavum and ankylosing spondylitis[J]. *Global Spine Journal*, 2025, **15**(1): 161-174.
- [33] Gierlikowska B, Stachura A, Gierlikowski W, *et al.* Phagocytosis, degranulation and extracellular traps release by neutrophils-the current knowledge, pharmacological modulation and future prospects[J]. *Frontiers in Pharmacology*, 2021, **12**: 666732.
- [34] Metzemaekers M, Gouwy M, Proost P. Neutrophil chemoattractant receptors in health and disease: Double-edged swords[J]. *Cellular & Molecular Immunology*, 2020, **17**(5): 433-450.
- [35] Mayadas T N, Cullere X, Lowell C A. The multifaceted functions of neutrophils[J]. *Annual Review of Pathology*, 2014, **9**: 181-218.
- [36] Yarchoan M, Ho W J, Mohan A, *et al.* Effects of B cell-activating factor on tumor immunity[J]. *JCI Insight*, 2020, **5** (10): e136417.
- [37] Yeo S C, Barratt J. The contribution of a proliferation-inducing ligand (APRIL) and other TNF superfamily members in pathogenesis and progression of IgA nephropathy[J]. *Clinical Kidney Journal*, 2023, **16**: 9-18.

CPPED1 关联中性粒细胞胞外诱捕网与糖尿病肾病：生物信息学驱动的表达验证

何湛淇, 陈欣欣, 黄文文, 赵艳玲[†]

温州医科大学 第二附属医院肾内科, 浙江 温州 3250000

摘要: 糖尿病肾病 (Diabetic nephropathy, DN) 始于与糖尿病相关的葡萄糖代谢紊乱, 氧化应激在其中发挥着关键作用。中性粒细胞胞外陷阱 (Neutrophil extracellular traps, NETs) 是由细胞质和颗粒蛋白组成的松散的网状结构, 其形成和功能依赖于氧化应激。本文旨在通过生物信息学分析和实时荧光定量PCR (quantitative Real-Time PCR, qRT-PCR), 识别DN进展的潜在靶点, 重点关注NETs。对两个与DN相关的RNA测序数据集 (GSE142025 和 GSE163603) 及 NETs 相关基因进行了差异表达基因分析。随后, 进行了基因集富集分析 (gene set enrichment analysis, GSEA)、基因集变异分析 (gene set variation analysis, GSVA)、GeneMANIA 和受试者工作特征曲线 (receiver operating characteristic, ROC) 分析。我们用单样本GSEA评估了免疫细胞浸润水平, 还建立了RNA结合蛋白 (RNA-binding proteins, RBP) 及其相关靶 mRNA 的相互作用网络, 并对高糖 (high glucose, HG) 处理和对照的 HK-2 细胞进行了 qRT-PCR 实验。通过差异表达基因 (differentially expressed genes, DEGs) 与 NETs 相关基因的交集, 本研究识别出 22 个核心基因。免疫细胞浸润评估显示在分析组之间 23 种免疫细胞类型存在显著差异, 包括 CPPED1 在内的核心基因显示出较高诊断价值, AUC 超过 0.6。qRT-PCR 结果显示, CPPED1 基因表达水平显著降低, 提示其有望成为新的诊断生物标志物与治疗靶点。综上, 本研究基于 GEO 数据库数据, 整合 GSEA、GSVA、免疫细胞浸润分析及 ROC 曲线分析等方法, 筛选出 22 个在 DN 中显著上调且具有关键调控作用的差异表达基因, 为 DN 的分子机制研究与临床诊疗提供了新的参考依据。

关键词: 基因表达综合数据库 (GEO 数据库); 糖尿病肾病; 中性粒细胞胞外陷阱; 生物信息学; 实时荧光定量 PCR (qRT-PCR); 含钙调磷酸酶样磷酸酯酶结构域蛋白 1 (CPPED1)

□