Journal of Urology and Renal Diseases
Li Y, et al. J Urol Ren Dis 07: 1294.
www.doi.org/10.29011/2575-7903.001294
www.gavinpublishers.com
Research Article

Bioinformatics Facilitates the use of Microarrays to Identify Potential Markers in Diabetic Nephropathy

Yao Li[1], Daoxun Wu[1], Yun Zhou[1], Mei Wu[1], Lizhi Gao[1], and Dan Zhao[2*]

[1]West Yunnan University of Applied Sciences, Dali 671006,China

[2]Da Li Agriculture and Forestry Career Technical College, Dali 671003,China

[*]**Corresponding author:** Dan Zhao, Da Li Agriculture and Forestry Career Technical College, Dali 671003,China

Citation: Li Y, Wu D, Zhou Y, Wu M, Hao L, Zhao D (2022) Bioinformatics Facilitates the use of Microarrays to Identify Potential Markers in Diabetic Nephropathy. J Urol Ren Dis 07: 1294. DOI: 10.29011/2575-7903.001294.

Received Date: 09 November, 2022**; Accepted Date:** 15 November, 2022**; Published Date:** 18 November 2022

Abstract

Background: Diabetic Nephropathy (DN) is the leading cause of end-stage renal disease worldwide. Extensive studies have been performed to elucidate the underlying mechanisms of DN, which still need to be clarified. The identification of key biomarkers using integrated bioinformatics could provide a certain theoretical foundation for future research and provide experimental direction for subsequent experimental verification.

Methods: GSE1009, GSE30528, and GSE96804 were downloaded from the Gene Expression Omnibus (GEO) database to screen Differentially Expressed Genes (DEGs) between normal renal tissue and DN renal tissue by using the limma package. Then, the Robust Rank Aggreg (RRA) method was used to integrate and analyze the three datasets to obtain integrated DEGs. Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed to determine the molecular mechanisms of integrated DEGs involved in the progression of DN. A Protein-Protein Interaction (PPI) network of integrated DEGs was constructed via the STRING database, PPI network visualization and module analyses were performed by using Cytoscape software, and the hub genes in the PPI network were selected by topological analysis. Finally, the Nephroseq v5 online platform was utilized to explore the correlation between hub genes and clinical features of DN

Results: In total, 249 integrated DEGs, including 191 upregulated genes and 58 downregulated genes, were identified and enriched in pathways involved in several functions and expression pathways, such as extracellular matrix,complement and coagulation cascades, focal adhesion, ECM-receptor interactions, cytokine-cytokine receptor interaction, the renin-angiotensin system, and chemokine signaling pathways. The top 10 hub genes identified from the PPI network were ALB, FN1, VEGFA, IGF1, JUN, FOS, CTGF, C3, COL1A2, and CLU. In addition, a KEGG pathway analysis of the top 2 modules identified from the PPI network revealed that Module 1 was mainly involved in focal adhesion and ECM-receptor interactions, while Module 2 was mainly involved in cytokine-cytokine receptor interactions, the chemokine signaling pathway, the Toll-like receptor signaling pathway, and the TGF-beta signaling pathway.Correlation and subgroup analyses of 10 hub genes and the clinical characteristics of DN indicated that ALB, FN1, VEGFA, IGF1, JUN, FOS, CTGF, C3, COL1A2, and CLU may participate in the development of DN.

Conclusions: The identification of hub genes may be a key biomarker for early DN diagnosis and targeted treatment.

Keywords: Biological pathways; Biomarkers; Diabetic nephropathy; Differentially expressed genes; Hub genes

Introduction

Diabetic Nephropathy (DN) is one of the most serious microvascular complications and the leading cause of Chronic Kidney Disease (CKD) and End-Stage Renal Disease (ESRD) [1,2]. DN is characterized pathologically by progressive accumulation of extracellular matrix in the glomerular mesangium, basement membrane thickening, and glomerular hypertrophy, conditions that lead to glomerulosclerosis [3-7]. The International Diabetes Federation (IDF) estimated that in 2021, 537 million adults (20-79 years) were living with diabetes worldwide. This number is expected to increase to 643 million by 2030 and 783 million by 2045 [8], of whom 40% will develop diabetic kidney disease [9]. The irreversible deterioration of DN eventually leads to glomerulosclerosis and renal failure. However, long-term dialysis or kidney transplantation treatment will result in large global economic and social burdens.Hemodynamic and metabolic factors are the main cause of the onset of DN [10,11]. Other risk factors, such as congregation of advanced glycation end-products (AGEs), Oxidative Stress (OS), and activation of Protein Kinase C (PKC), are also believed to be involved in the pathogenesis of diabetes and its related complications [12,13]. Additionally, various inflammatory factors, endoplasmic reticulum (ER) stress, cellular autophagy, and microRNAs can activate the process of DN [14-18]. Although strict control of blood sugar and blood pressure levels and inhibiting the RAS system to reduce albuminuria can delay the progression of DN [19,20], this treatment is not sufficient. DN is the result of multiple gene interactions, and the molecular mechanisms of DN remain poorly understood; its prognosis is poor, and therapy is very difficult [21,22]. Therefore, identifying key biomarkersfor the early detection, diagnosis, and treatment of DN is urgently needed.

In recent decades, microarray technology and bioinformatics have been widely used to screen biomarkers involved in disease progression, to obtain further insights into disease pathogenesis, and to elucidate the underlying molecular mechanisms of disease [23,24]. With the widespread application of genome transcriptome analysis, large-scale microarray data have been generated and can be acquired from public databases, such as Gene Expression Omnibus (GEO) [25]. Reanalyzing available public data can provide valuable clues for new research. To date, many studies have screened numerous Differentially Expressed Genes (DEGs) involved in DN [26-31]. However, the heterogeneity of experimental samples in independent studies and the application of different detection platforms and different processing methods can lead to inconsistent results.The Robust Rank Aggreg (RRA) method is suitable for comparing several sequenced gene lists based on the assumption that each gene in each dataset is randomly arranged [32]. If a gene with a higher ranking in all datasets has a lower p value, it is more likely to become a DEG. The RRA is a robust and easy-to-implement method, not strictly requiring the use of a certain subset of problems or requiring all data to be of high quality. The RRA algorithm can also handle variable gene content from different microarray platforms. Therefore, the integration of gene expression datasets from multiple databases using RRA can offer a better understanding of the molecular mechanisms of disease genes [33].

In this study, we reanalyzed three original microarray datasets, GSE1009, GSE30528, and GSE96804, from the GEO database. First, the DEGs between DN and normal glomerular samples were identified, and the RRA method was performed to integrate the results and obtain the integrated DEGs. Second, the integrated DEGs were subjected to functional enrichment analysis by using the DAVID database to explore the molecular mechanisms involved in DN. Finally, the hub genes were screened by PPI network analysis, and modules were mined from the PPI network. Furthermore, the correlations between hub genes and clinical features of DN were analyzed by using the Nephroseq v5 online platform to further explore the pathogenesis and pathophysiological and molecular mechanisms involved in DN. Our research aimed to identify key biomarkers that may contribute to the early diagnosis and therapy of DN. Our flowchart is shown in Figure 1.

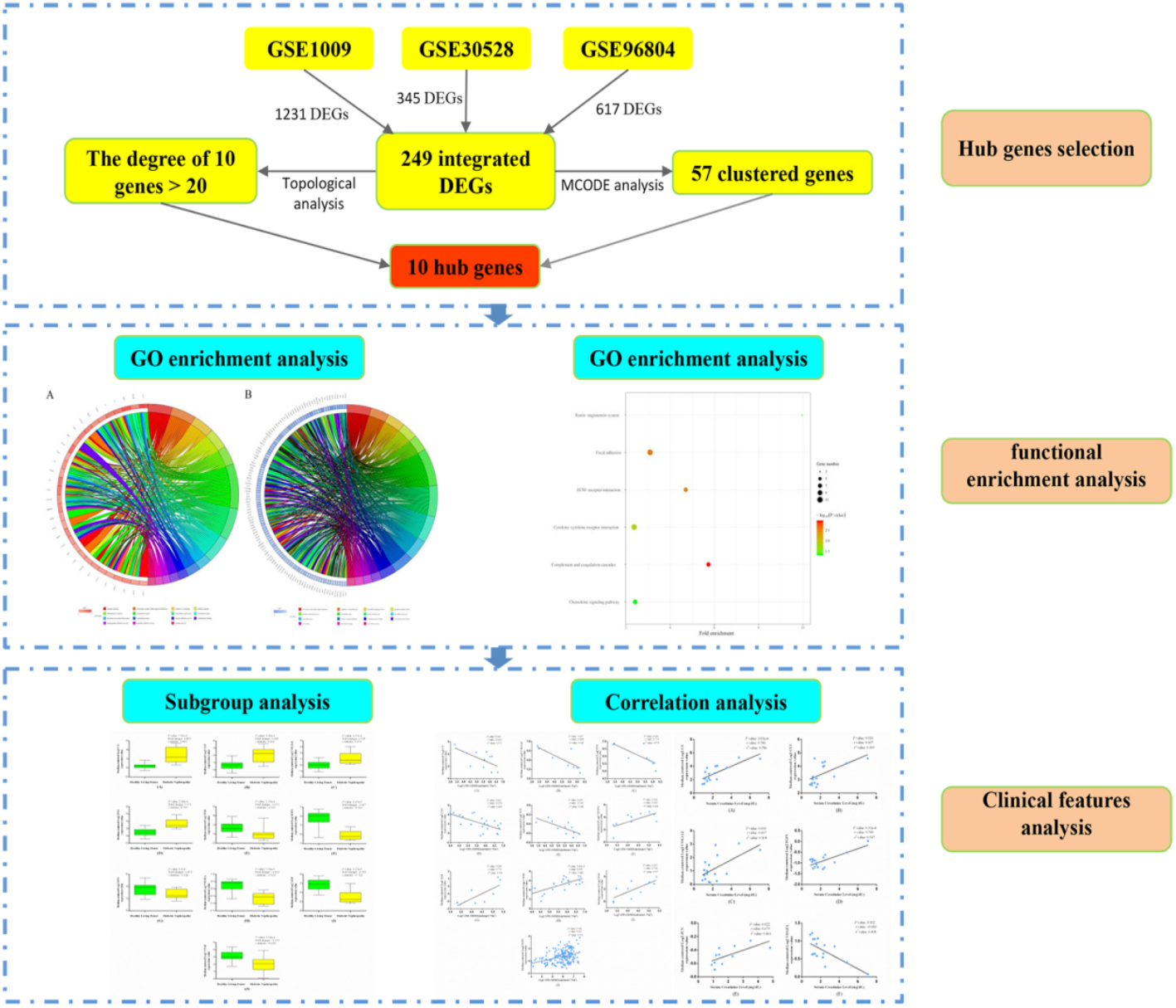

Figure 1: The whole framework based on an integration strategy of bioinformatics analysis.

Materials and Methods

Gene Expression Data Preprocessing

The gene expression profiles ofGSE1009, GSE30528, and GSE96804 were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) . The three datasets were tissue samples from glomeruli. GSE1009 contains 3 samples from controls and 3 samples from DN patients. GSE30528 includes 13 samples from controls and 9 samples from DN patients.GSE30528 consists of 20 control samples and 41 DN patient samples.

Degs Screening and Microarray Data Integration

We downloaded the series matrix TXT files and platform TXT files and used the limma package in R (v.3.6.0) language to process the gene expression data from the three datasets by performing quartile data normalization. The gene expression data of GSE1009 were

log2 transformed. DEGs were screened via the limma R package in each dataset. The settings of the thresholds were P value< 0.05 and |log-Fold Change (FC)| > 1. the log FC of each gene in each dataset was calculated and sorted. Three genes listswere integrated using the RRA package (v.1.1),and the integrated upregulated and downregulated DEG lists were used for subsequent analysis.

Functional and Pathway Analysis of Degs

GO annotation and KEGG pathway enrichment analyses were performed using the DAVID 6.7 database to determine the potential functions of the integrated DEGs. The significant screening threshold was set as a P value< 0.05 for screening correlated GO terms and KEGG pathways.The DEG-KEGG pathway network was constructed and displayed by using Cytoscape (v.3.6.1).

PPI Network Construction and Analysis of Modules

To understand the relationships between different integrated DEGs, we analyzed the PPI network of integrated DEGs via the STRING 11.0 database. The hub genes in the PPI network were screened based on node degree by using Cytoscape. Module partitioning for the PPI network was performed using the Cytoscape MCODE plugin. The default parameters were as follows: degree cutoff ≥ 2, node score cutoff ≥ 0 2, K-core ≥ 2. The module KEGG pathway enrichment analyses were performed using DAVID 6.7.

Association Between Hub Genes and Clinical Features of DN and Statistical Analysis

Correlation and subgroup analyses between hub genes and clinical features of DN were performed using Nephroseq v5 to evaluate the potential effects of hub genes on DN. Pearson correlation analysis between hub genes and Glomerular Filtration Rate (GFR) and Serum Creatinine (Scr) in patients with DN was carried out. Unpaired Students'st test was used to compare the two groups. All tests were two-tailed, with a P value < 0.05 considered statistically significant. GraphPad Prism v 7.0 was used to perform statistical analyses. The insignificant results are not displayed.

Results

Data Preprocessing and DEG Screening

The datasets from GSE1009, GSE30528, and GSE96804were normalized, and the results are shown in Figure 2. The DEGs were screened using the limma R package (P < 0.05 and |log Fold Change (FC)| > 1). A total of 1231 DEGs were obtained from the GSE1009 dataset, including 581 upregulated and 650 downregulated DEGs. A total of 345 DEGs were obtained from the GSE30528 dataset, including 99 upregulated and 246 downregulated DEGs. The GSE96804 dataset contained 617 DEGs, including 336 upregulated and 281 downregulated DEGs. TheDEGs of GSE1009, GSE30528, and GSE96804 are shown in Figure 3. The cluster heatmap of the top 100 DEGs from 3 different datasets is shown in Figure 4.

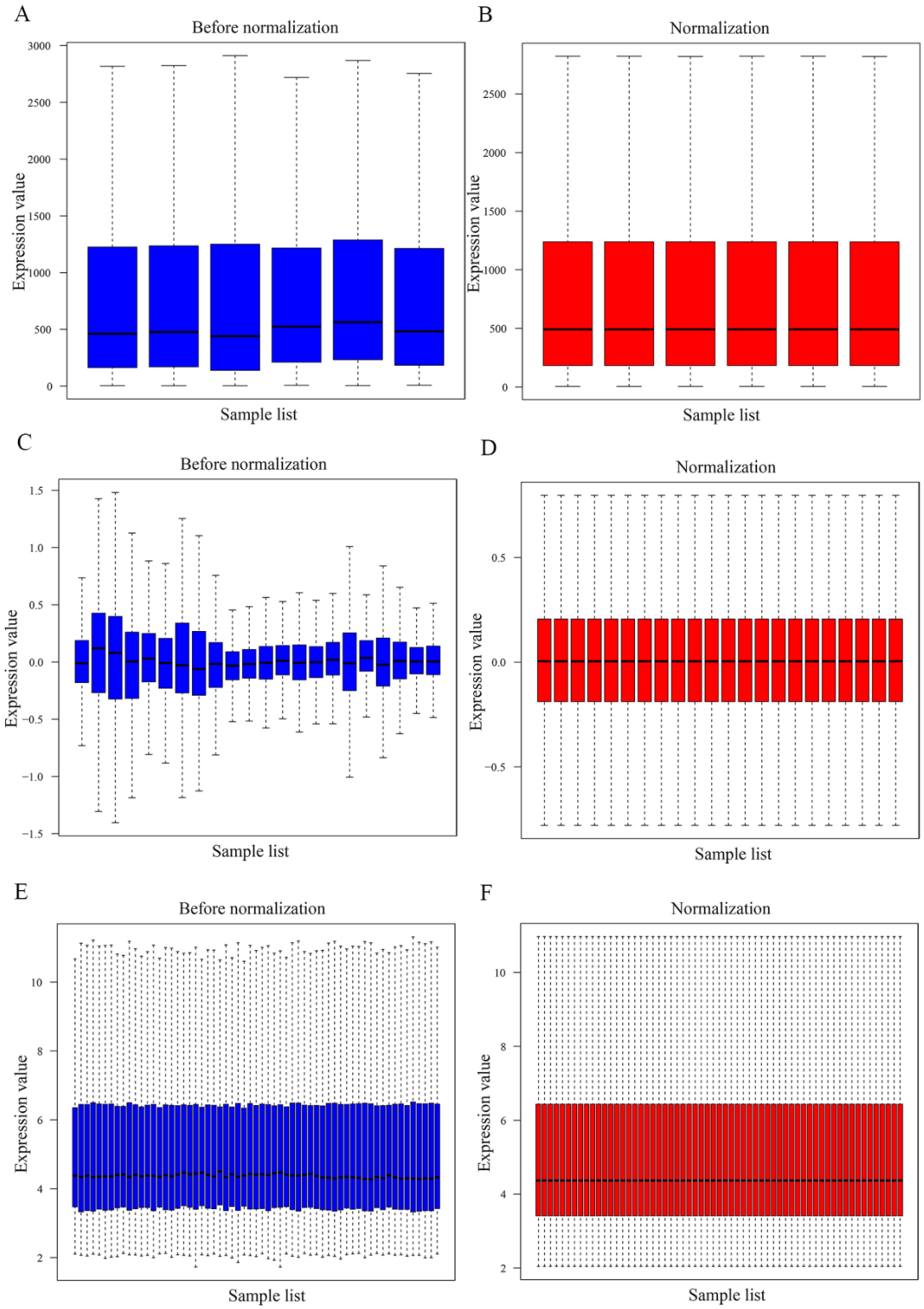

Figure 2: Normalization of gene expression. **(A-B)** Normalization of the GSE1009 data set. **(C-D)** Normalization of the GSE30528 data set. **(E-F)** Normalization of the GSE96804 data set. Blue represents data before normalization, and red represents data afternormalization.

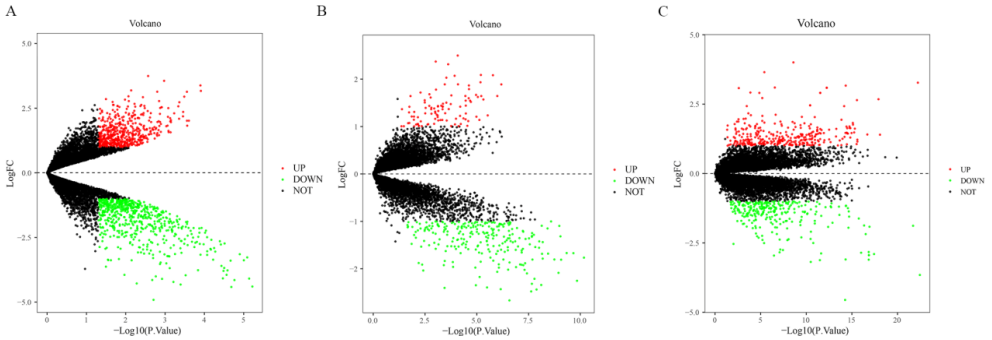

Figure 3: Differentially expressed genes between the two groups of samples in each dataset. **(A)** GSE1009, **(B)** GSE30528, **(C)** GSE96804. The red dots represent the upregulated genes based on an $P < 0.05$ and | log fold change (FC)| > 1; the green dots represent the downregulated genes based on an $P < 0.05$ and | log fold change (FC)| > 1; the black spots represent genes with no significant difference inexpression.

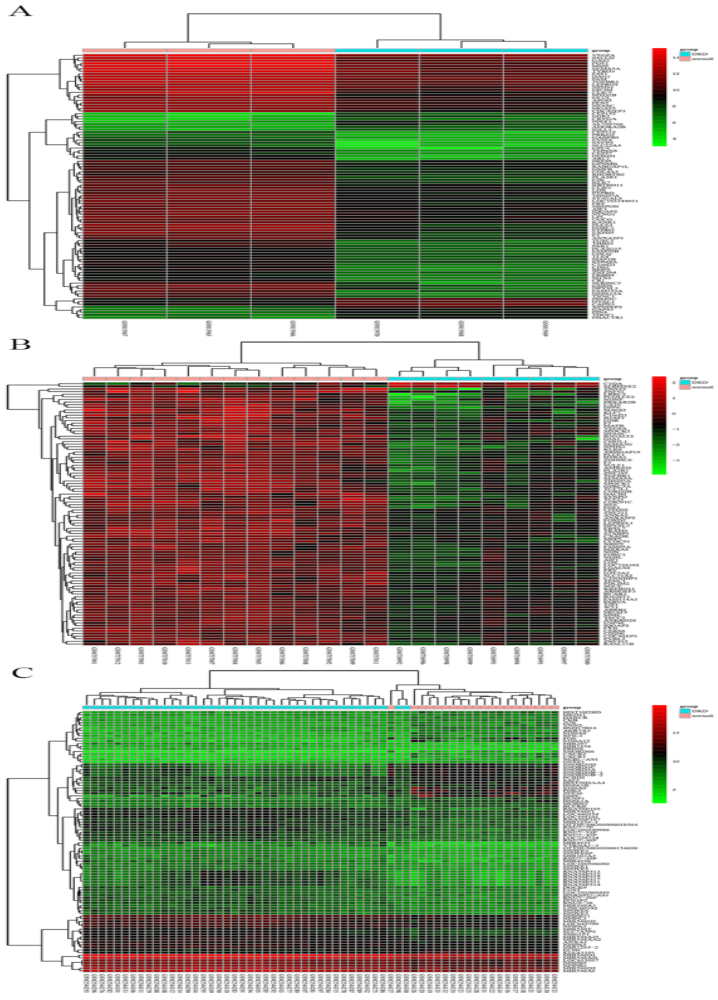

Figure 4: Cluster heat map of the top 100 DEGs. **(A)** GSE1009, **(B)** GSE30528, **(C)** GSE96804. Red indicates relative upregulation of gene expression; green indicates the relative downregulation of gene expression; black indicates no significant change in gene expression; and gray indicates that the signal intensity is not high enough to detect.

Identification of Integrated Degs

DEGs among three different datasets were integrated by using the RRA method ($P < 0.05$ and $|logFC| > 1$). The RRA method is based on the assumption that each gene in each dataset is randomly arranged. If the gene ranks high in all datasets, the associated P value is lower, and the possibility of differential gene expression is greater. Through rank analysis, 249 integrated DEGs, consisting of 58 upregulated genes and 191 downregulated genes, were identified by the RRA method. The top 20 upregulated genes and the top 20 downregulated genes were mapped to a heatmap, as shown in Figure 5.

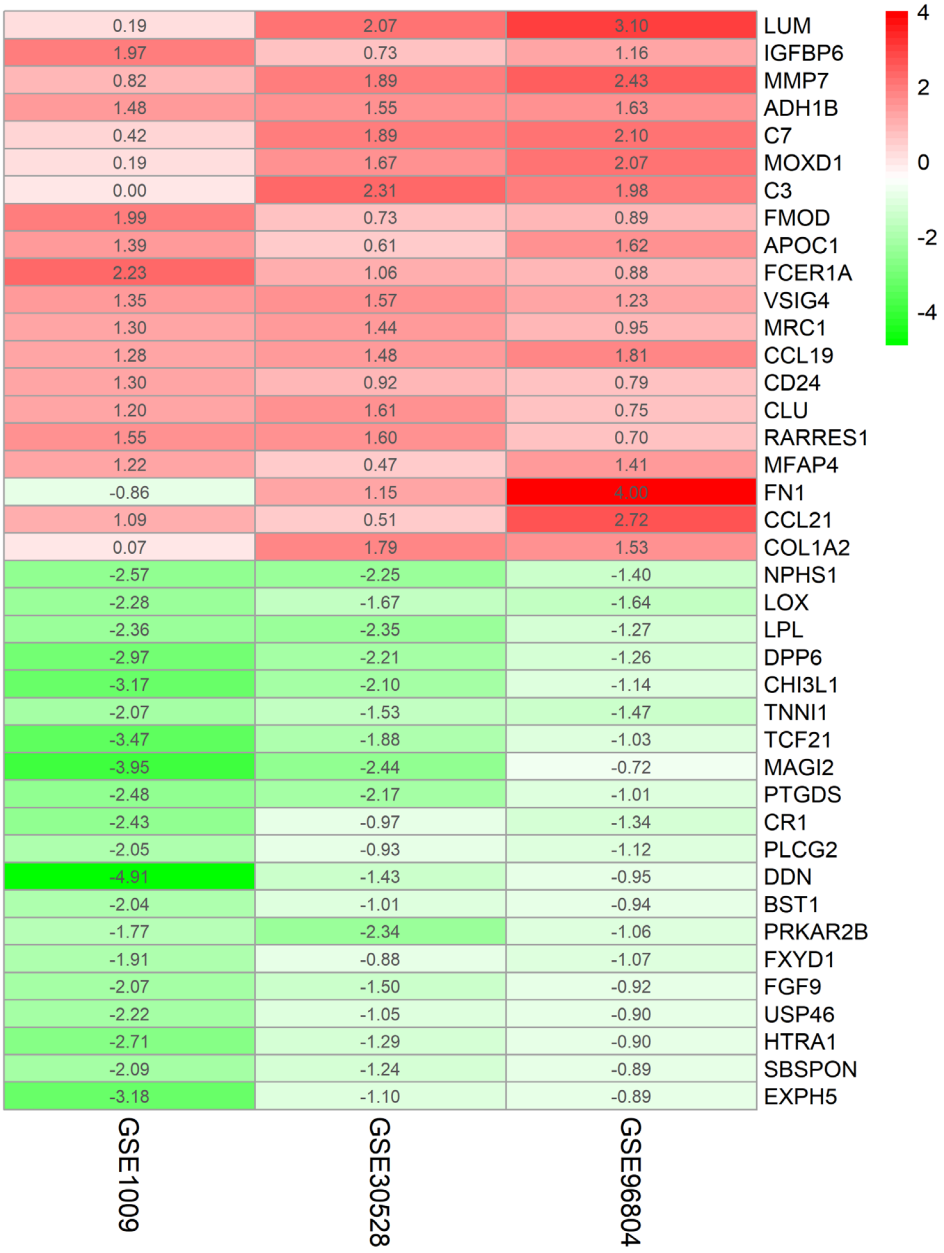

Figure 5: Log FC heatmap of each expression microarray. The abscissa represent the GEO IDs, the ordinate represents the gene name, the red represents log FC > 0, the green represents log FC < 0 and the value in the box represents the log FC value.

GO Functional Enrichment Analysis of Degs

GO functional enrichment analysis of upregulated genes and downregulated genes was divided into Molecular Function (MF), Cellular Component (CC), and Biological Process (BP) to explore the molecular mechanisms of DEGs involved in the development of DN. The three GO results are shown in Figure 7 and Figure 8, and the top 5 most significantly enriched items for each part are summarized according to their P values in Table 1 and Table 2. The GO BP analysis revealed that these upregulated genes were mainly enriched in the immune response, inflammatory response, response to wounding, defense response, and cell surface receptor-linked signal transduction. The GO CC analysis for the upregulated DEGs showed that most were mainly concentrated in the extracellular region, extracellular region part, extracellular space, proteinaceous extracellular matrix, and extracellular matrix. The most upregulated DEGS revealed by the GO MF analysis were mainly involved in enzyme inhibitor activity, endopeptidase inhibitor activity, peptidase inhibitor activity, carbohydrate binding and cytokine activity (Figure 6A and Figure 7A).

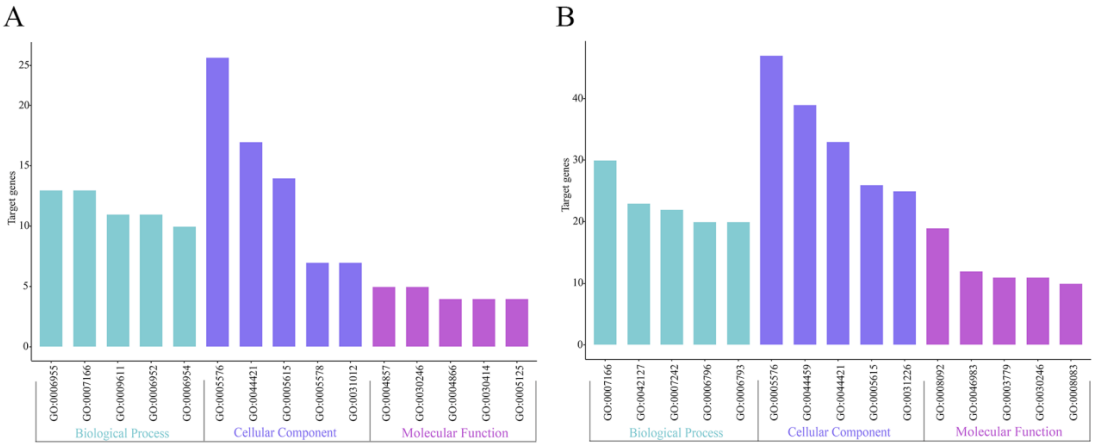

Figure 6: Top 15 enriched GO terms. **(A)** Upregulated DEGs with the top 15 enriched GO terms. **(B)** Downregulated DEGs with the top 15 enriched GO terms.

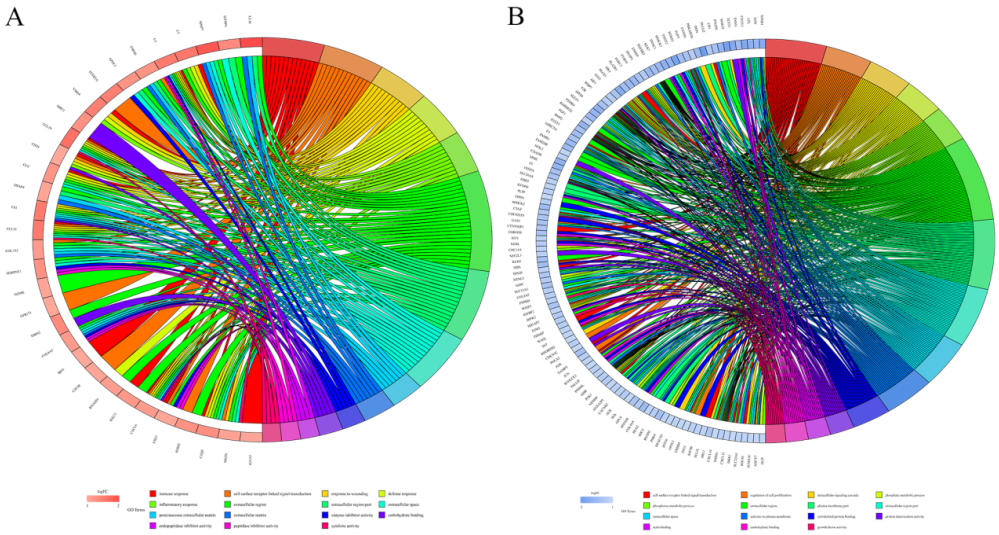

Figure 7: Distribution of integrated DEGs in DN for different GO-enriched functions. **(A)** Upregulated DEGs. **(B)** Downregulated DEGs.

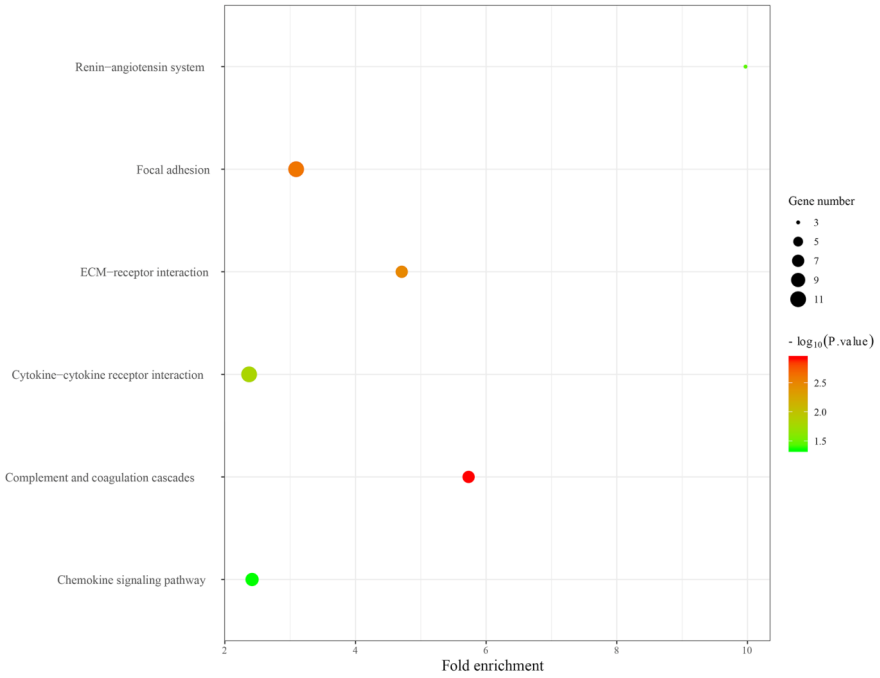

Figure 8: KEGG pathway enrichment analysis of the integrated DEGs.

Category	Term	Count	P-value
BP	immune response	13	3.60E-06
BP	inflammatory response	10	1.75E-06
BP	response to wounding	11	1.34E-05
BP	defense response	11	4.79E-05
BP	cell surface receptor linked signal transduction	13	0.02586
CC	extracellular region	24	2.04E-08
CC	extracellular region part	17	5.07E-08
CC	extracellular space	14	2.98E-07
CC	proteinaceous extracellular matrix	7	7.44E-04
CC	extracellular matrix	7	0.001099
MF	enzyme inhibitor activity	5	0.007045
MF	endopeptidase inhibitor activity	4	0.008032
MF	peptidase inhibitor activity	4	0.009301
MF	carbohydrate binding	5	0.017667
MF	cytokine activity	4	0.017824

Notes:BP, biological process.CC, cellular component.MF, molecular function

Table 1: Upregulated genes top 15 enriched GO terms.

Category	Term	Count	P-value
BP	regulation of cell proliferation	23	8.21E-05
BP	phosphate metabolic process	20	0.014594
BP	phosphorus metabolic process	20	0.014594
BP	cell surface receptor linked signal transduction	30	0.045569
BP	intracellular signaling cascade	22	0.047253
CC	extracellular region part	33	3.23E-08
CC	extracellular space	26	2.49E-07
CC	extracellular region	47	1.63E-06
CC	plasma membrane part	39	0.005308
CC	intrinsic to plasma membrane	25	0.00594
MF	cytoskeletal protein binding	19	1.86E-05
MF	growth factor activity	10	9.84E-05
MF	actin binding	11	0.004366
MF	carbohydrate binding	11	0.007678
MF	protein dimerization activity	12	0.047603

Notes:BP, biological process.CC, cellular component. MF, molecular function

Table 2: Downregulated genes top 15 enriched GO terms.

The downregulated DEGs were mainly involved in theregulation of cell proliferation, phosphate metabolic processes, phosphorus metabolic processes, cell surface receptor-linked signal transduction, and intracellular signaling cascades. GO CC analysis revealed that these genes were mainly concentrated in the extracellular region part, extracellular space, extracellular space, plasma membrane part, and intrinsic to plasma membrane. GO MF analysis for the downregulated DEGsshowed that these genes were mainly enriched in cytoskeletal protein binding, endopeptidase inhibitor activity, actin binding, carbohydrate binding and protein dimerization activity (Figure 6B and Figure 7B).

KEGG Pathway Analysis of Degs

KEGG pathway analysis of the integrated DEGs showed that these genes wereenriched in a total of 6 pathways ($P<0.05$), including the complement and coagulation cascades, focal adhesion, ECM-receptor interaction, cytokine-cytokine receptor interaction, renin-angiotensin system, and chemokine signaling pathways (Figure 8). The DEG-KEGG pathway network was visualized with Cytoscape, as shown in Figure 9. The results indicate that DEGs may participate in the progression and development of DN by regulating these pathways.

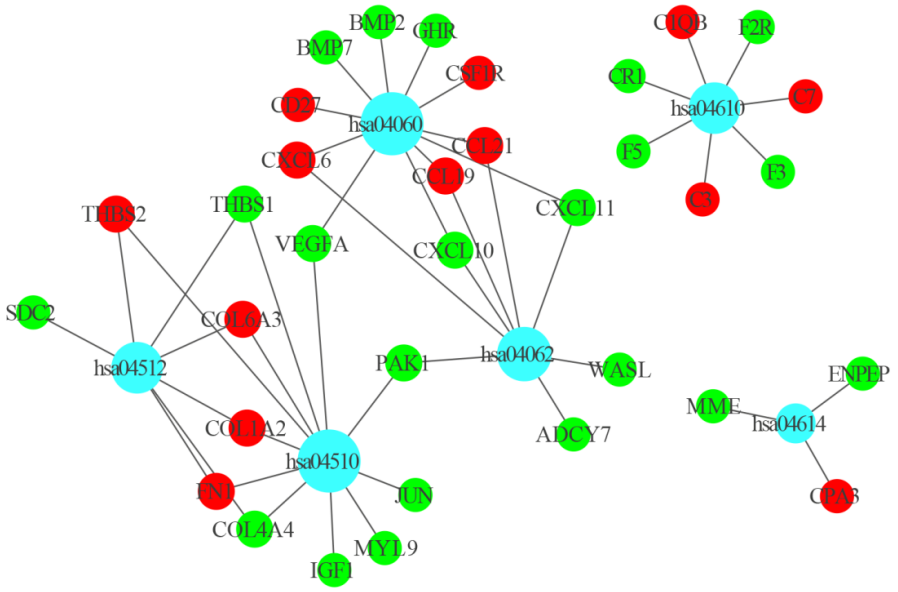

Figure 9: Network map of enriched pathways. Blue represents the pathways, red represents the upregulated genes and green represents the downregulated genes.

Construction of Protein-Protein Interaction (PPI) Network and Module Analysis

The 249 integrated DEGs were analyzed, and a PPI network was constructed. The network was displayed and analyzed using Cytoscape. The top 10 hub genes were identified based on degree values (Table 3). Moreover, 9 modules were obtained in the PPI network by using MCODE, and the 2 most important modules were selected according to the score (Figure 10). The 2 modules were subjected to KEGG pathway enrichment analysis in the DAVID database (Table 4). Module 1 was mainly enriched in focal adhesion and ECM-receptor interactions, and module 2 was mainly enriched in cytokine-cytokine receptor interactions, the chemokine signaling pathway, the Toll-like receptor signaling pathway and the TGF-beta signaling pathway.

Gene symbol	Gene description	LogFC	Degree
ALB	Albumin	-1.30	53
FN1	Fibronectin 1	1.43	50
VEGFA	Vascular Endothelial Growth Factor A	-1.52	49
IGF1	Insulin Like Growth Factor 1	-1.14	32
JUN	Jun Proto-Oncogene, AP-1 Transcription Factor Subunit	-1.09	27
FOS	FosProto-Oncogene, AP-1 Transcription Factor Subunit	-1.54	25
CTGF	Connective Tissue Growth Factor	-1.40	25
C3	Complement C3	1.43	21
COL1A2	Collagen Type I Alpha 2 Chain	1.12	20
CLU	Clusterin	1.18	20

Note: FC, fold change

Table 3: The degree values of the top 10 hub genes.

Module	Pathway	Count	P-value	Genes
Module1	Focal adhesion	5	4.1E-3	COL6A3, VEGFA, IGF1, THBS2, FN1
	ECM-receptor interaction	4	3.2E-3	COL6A3, THBS2, SDC2, FN1
Module2	Cytokine-cytokine receptor interaction	7	6.00E-05	BMP2, CCL21, CCL19, CXCL6, CXCL11, BMP7, CXCL10
	Chemokine signaling pathway	6	1.40E-04	ADCY7, CCL21, CCL19, CXCL6, CXCL11, CXCL10
	Toll-like receptor signaling pathway	4	2.90E-03	FOS, JUN, CXCL11, CXCL10
	TGF-beta signaling pathway	3	2.60E-02	BMP2, BMP7, THBS1

Table 4: KEGG enrichment of genes in the top 2 modules.

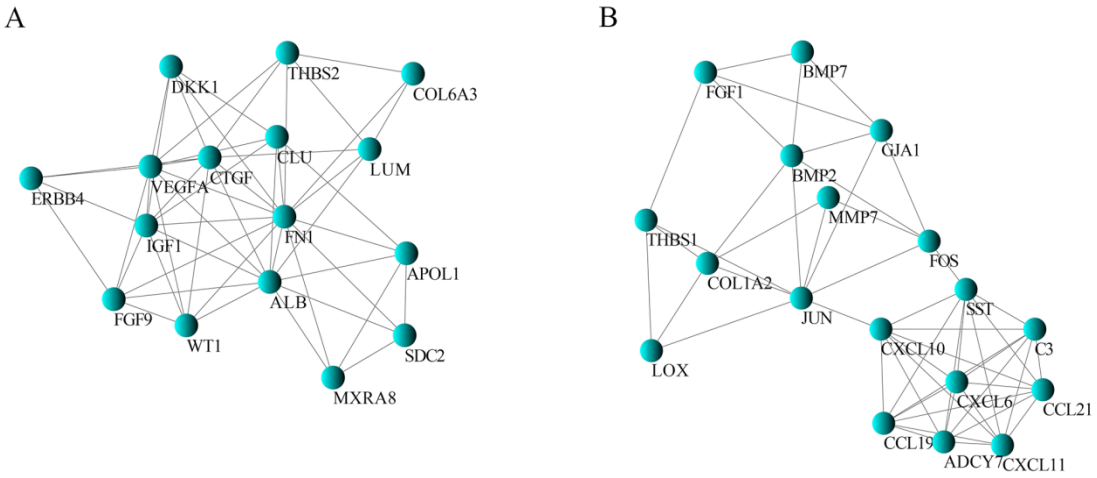

Figure 10: PPI network of module 1 and module 2. The cycles represent genes, lines represent interactions between gene-encoded proteins.

Association Between The Hub Genes And Clinical Features Of DN

We examined the differential expression of hub genes between DN patients and healthy living donors using Nephroseq v5 (Figure 11). The expression levels of C3, CLU, COL1A2, and FN1were elevated (Figure 12A - D), while the expression levels of FOS, IGF1, ALB, VEGFA, JUN and CTGF were decreased in DN renal tissues compared with healthy kidney tissues (Figure 11E - J). The correlation between the hub genes and the GFR of DN patients was analyzed (Figure 12). The expression levels of C3, COL1A2, FN1, CLU, and JUN were negatively correlated with GFR in DN renal tissues (Figure 12A - E), suggesting that these genes may contribute to the occurrence and progression of DN. Nonetheless, the expression levels of VEGFA, ALB, CTGF, FOS, and IGF1 were positively correlated with GFR (Figure 12F - J), suggesting that they may contribute to the maintenance and improvement of renal function. In addition, the correlation between hub genes and Scr in DN patients was determined (Figure 11). The expression levels of C3, CLU, COL1A2, IGF1, and JUN were positively correlated with Scr in DN renal tissues (Figure 13A-E), whereas that of VEGFA was negatively associated with Scr (Figure 13F). Therefore, the expression changes of these six genes may result in the occurrence and development of DN.

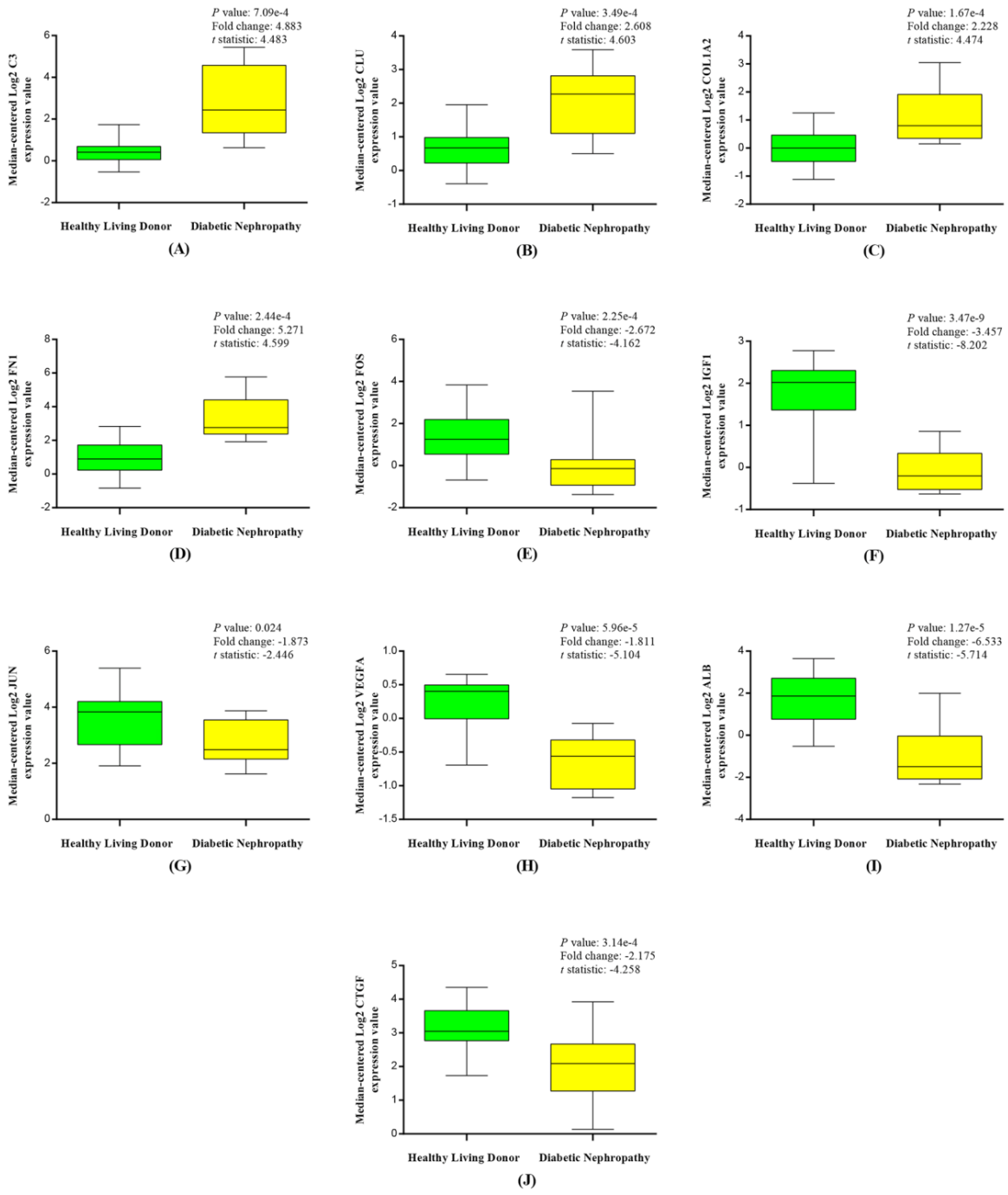

Figure 11: The different expression of hub genes in DN renal tissues and healthy kidney tissues. **(A)** The expression of C3 increased in DN renal tissues. **(B)** The expression of CLU increased in DN renal tissues. **(C)** The expression of COL1A2 upregulated in DN renal tissues. **(D)** The expression of FN1 upregulated in DN renal tissues. **(E)** The expression of FOS decreased in DN renal tissues. **(F)** The expression of IGF1 decreased in DN renal tissues. **(G)** The expression of JUN decreased in DN renal tissues. **(H)** The expression of VEGFA decreased in DN renal tissues. **(I)** The expression of ALB downregulated in DN renal tissues. **(J)** The expression of CTGF downregulated in DN renal tissues.

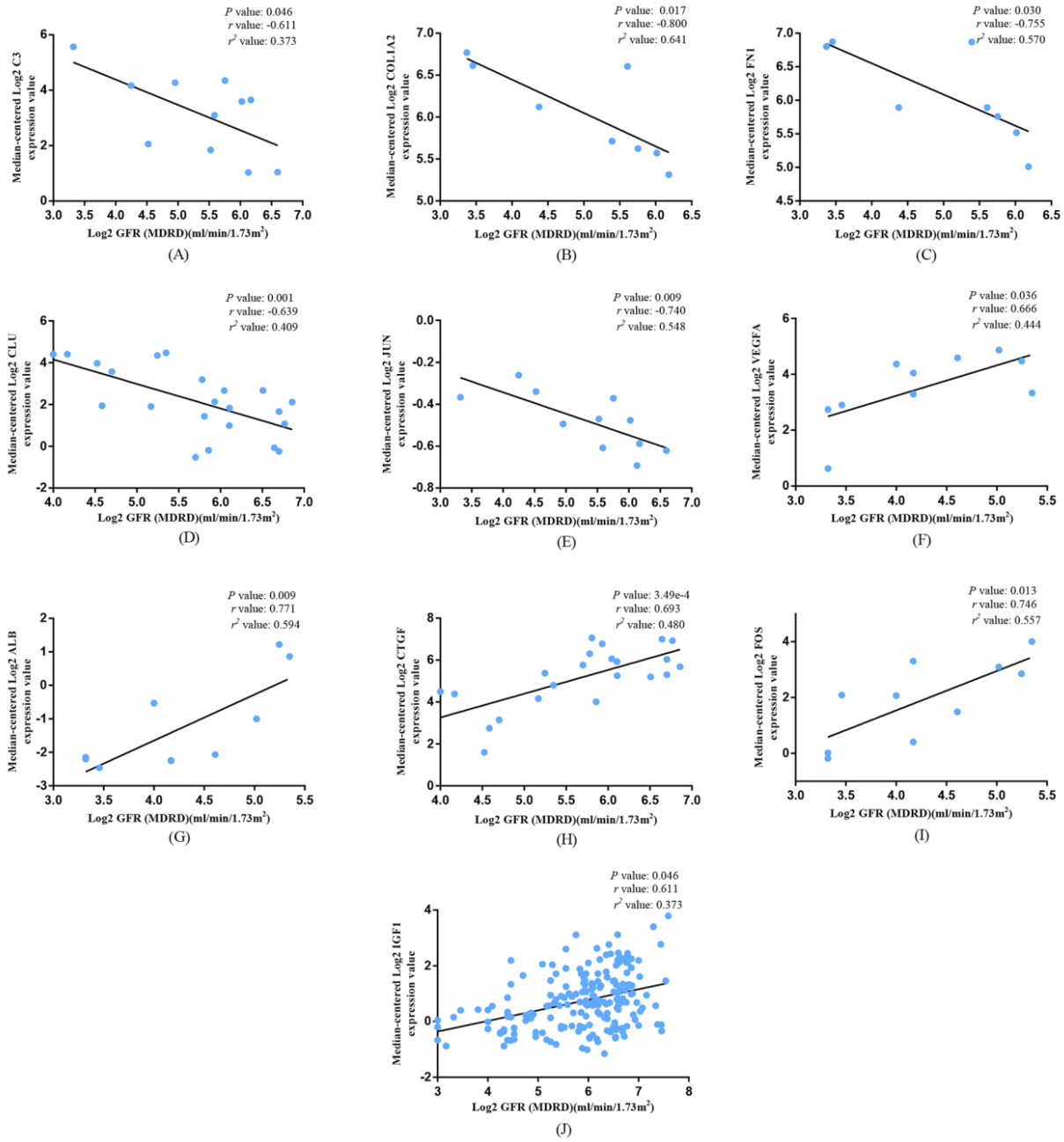

Figure 12: Association between the expression of hub genes and GFR in DN patients. **(A)** The expression of C3 was negatively correlated with GFR (P < 0:046, r = -0:611). **(B)** The expression of COL1A2 was negatively correlated with GFR (P = 0:017, r = -0:800). **(C)** The expression of FN1 was negatively correlated with GFR (P = 0:030, r = -0:755). **(D)** The expression of CLU was negatively correlated with GFR (P = 0:001, r = -0:639). **(E)** The expression of JUN was negatively correlated with GFR (P = 0:009, r = -0:740). **(F)** The expression of VEGFA was positively correlated with GFR (P < 0:036, r = 0:666). **(G)** The expression of ALB was positively correlated with GFR (P = 0:009, r = 0:771). **(H)** The expression of CTGF was positively correlated with GFR (P = 3.49e-4, r = 0:693). **(I)** The expression of FOS was positively correlated with GFR (P < 0:013, r = 0:746). **(J)** The expression of IGF1 was positively correlated with GFR (P = 0:046, r = 0:611).

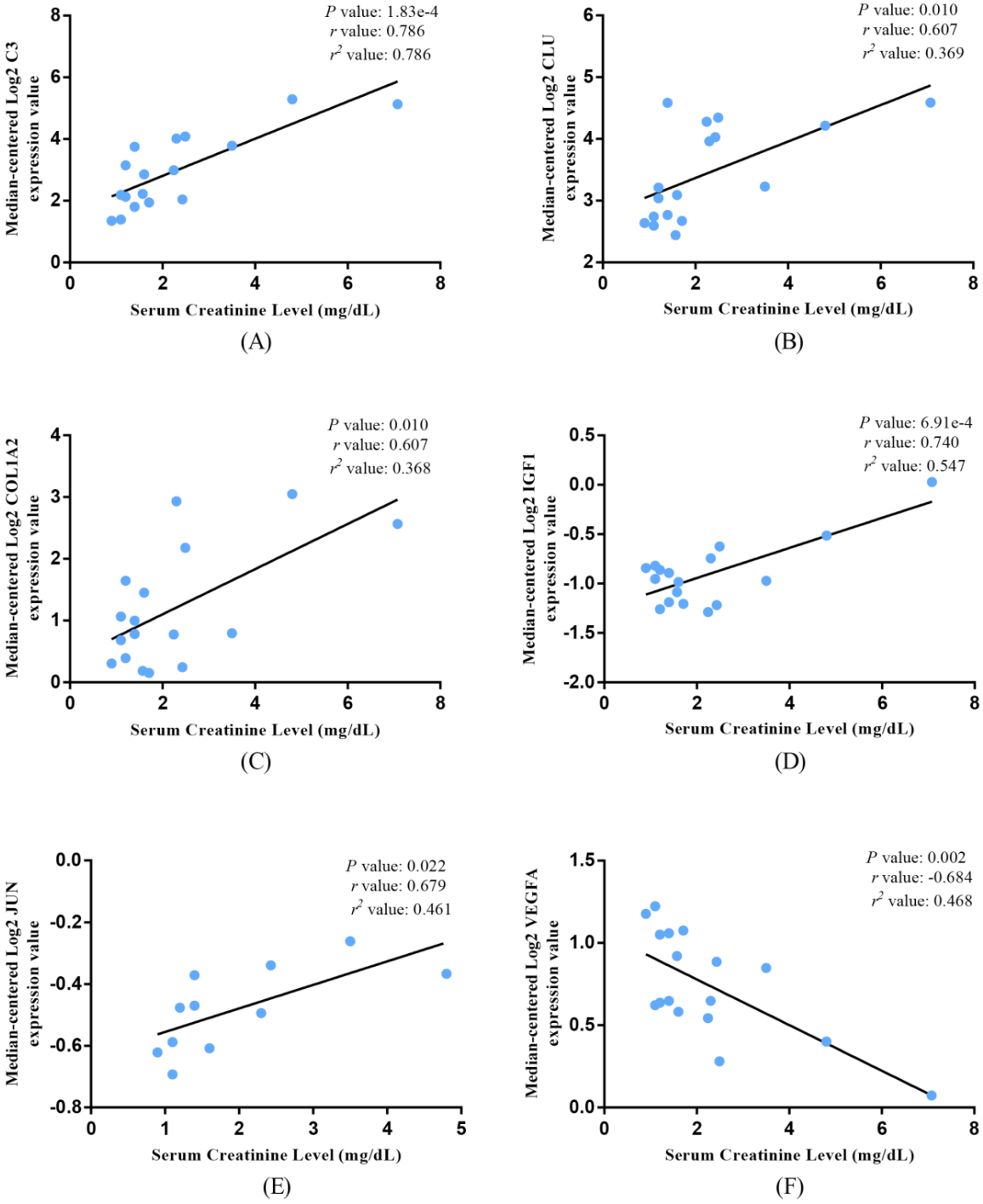

Figure 13: Association between the expression of hub genes and Scr in DN patients. **(A)** The expression of C3 positively correlated with Scr (P = 1:83e-4, r = 0:786). **(B)** The expression of CLU was positively correlated with Scr (P = 0:010, r = 0:607). **(C)** The expression of COL1A2 was positively correlated with Scr (P = 0:010, r = 0:607). **(D)** The expression of IGF1 was positively correlated with Scr (P = 6:91e-4, r = 0:369). **(E)** The expression of JUN was positively correlated with Scr (P = 0:022, r = 0:679). **(F)** The expression of VEGFA was negatively correlated with Scr (P < 0:002, r = -0:684).

Discussion

Diabetic Nephropathy (DN) is one of the major microvasculature components of diabetes mellitus; it is the result of multiple gene interactions and involves a multitude of different pathways. In this study, GO enrichment analysis revealedthat the integrated DEGs were involved in the immune response, inflammatory response, extracellular region, cytokine activity, and growth factor activity, which are associated with fibrosis and inflammation in DN [34,35]. The KEGG pathway analysis revealed that integrated DEGs were mainly enriched in the following top five pathways: complement and coagulation cascades, focal adhesion, ECM-receptor interaction, cytokine-cytokine receptor interactions and the Renin-Angiotensin System (RAS). The complement system is an important differentially regulated pathway in DN glomeruli. Previous studies have shown that the complement system participates in the pathogenesis of DN and may be a therapeutic target [36,37]. Additionally, several studies have demonstrated that activation of the complement system is a vital cause of renal injury in DN [38,39]. ECM-receptor interaction has been proven to have a considerable contribution to the development of DN [40,41]. Activation of RAS is one major cause of renal injury in DN, and RAS dysregulation has been verified to be involved in the inflammatory process of DN [42]. The cytokine-cytokine receptor interaction pathway isrelated to the inflammatory reaction process, and the focal adhesion pathway is closely related to immunologicalstress. Inflammatory reactions and immunization arerelated to the occurrence and development of DN [43,44].

We also constructed PPI networks with 249 integrated DEGs and identified the following 10 hub genes: ALB, FN1, VEGFA, IGF1, JUN, FOS, CTGF, C3, COL1A2, and CLU.We performed correlation and subgroup analyses among the hub genes and clinical features of DN by using Nephroseq v5.Among these genes, the Albumin (ALB) gene encodes Human Serum Albumin (HSA). HSA is a multifunctional protein that has antioxidant and anti-inflammatory properties and can combine with endocrine compounds and drugs [45]. HSA levels were negatively correlated with urinary protein content [46]. A study conducted by Taguchi et al. showed that the HSA dimer can function as a plasma-retaining agent of fatty acid-conjugated antidiabetic drugs in diabetic nephropathy [47], but the correlation between HSA and DN has not yet been illustrated. We found that ALB expression was decreased in DN renal samples compared with healthy renal samples and was positively correlated with GFR (P = 0:009, r = 0:771). FN1 and CTGF have emerged as potent proinflammatory and profibrotic regulators. It has been reported that FN1 plays crucial roles in the progression of renal fibrosis to DN [48,49], and is significantly upregulated in the glomeruli of both DN patients and mice [50]. We found that FN1 expression was higher in DN renal samples and was negatively correlated with GFR (P = 0:001, r = -0:639).CTGF has emerged as a potential biomarker and therapeutic target of DN [51,52]. The progression of DN can be delayed by inhibiting CTGF expression [53]. We found that CTGF expression was downregulated in DN renal tissues and was positively correlated with GFR (P = 3.49e-4, r = 0:693).VEGFA is the most important vascular endothelial growth factor. VEGFA is mainly expressed in glomerular podocytes and tubular cells, and the mRNA and protein expression levels of VEGFA were significantly downregulated in human renal biopsy specimens with DN [54]. We found that VEGFA expression was decreased in DN renal tissues, was positively correlated with GFR (P < 0:036, r = 0:666), and was negatively correlated with Scr (P < 0:002, r = -0:684). In addition, studies have shown thatVEGFA could serve as a biomarker to identify the progression of DN [55-57]. COL1A2, a fibril-forming collagen that is upregulated in kidney tissues, reflects a higher degree of renal fibrosis in DN [58]. We found that COL1A2 expression was upregulatedin DN renal tissues, was negatively correlated with GFR (P = 0:017, r = -0:800), and was positively correlated with Scr (P = 0:010, r = 0:607).CLUis a glycoprotein that is expressed in many tissues, including the kidney [59]. He et al. found that glomerular CLU is upregulated in both patients with DN and streptozotocin-induced diabetic mice, and overexpressing CLU may protect against oxidative stress-induced apoptosis in podocytes [60]. Moreover, CLU may delay or even halt the progression of DN by modulating Akt-related pathways [61,62]. We found that cl u expression was increased in DN renal tissues, was negatively correlated with GFR (P = 0:001, r = -0:639), and was positively correlated with Scr (P = 0:010, r = 0:607).C3 plays an important rolein the complementary systemand is elevated in DN glomerular injury in DN [63]. Moreover, it has been reported to mediate renalinjury [38]. We found that C3 expression was increased in DN renal tissues, was negatively correlated with GFR (P < 0:046, r = -0:611), and was positively correlated with Scr (P = 1:83e-4, r = 0:786). IGF1 is a peptide growth factor. Both IGF1 and IGF1 receptors may play key roles in the development of DN by mediating cell growth and apoptosis [64,65]. Brittain et al.found thatIGF1 is downregulated in the kidneys of humans with CKD and rodents with DN [66]. We found that IGF1 expression was decreased in DN renal tissues, was positively correlated with GFR (P = 0:046, r = 0:611),and was positively correlated with Scr (P = 6:91e-4, r = 0:369). JUN(c-jun) and FOS(c-fos) are major components of the AP-1 transcription factor. c-jun was progressively increased during renal fibrosis in DN [67]. The overexpression of c-fos and c-jun in Glomerular Mesangial Cells (GMCs) is an important pathologic feature of DN [68]. We found that FOS and JUN expression was decreased in DN renal tissues. The expression of JUN was negatively correlated with GFR (P = 0:009, r = -0:740) and was positively correlated with Scr (P = 0:022, r = 0:679). The expression of FOS was positively correlated with GFR (P < 0:013, r = 0:746). Furthermore, KEGG pathway analysis of modules 1 and 2 suggested that the occurrence and development of DN might be related to these pathways and that regulating the complement cascade, inflammatory reactions, and

immunological stress and inhibiting ECM-receptor interactions might be an effective treatment for DN

Conclusions

In summary, this study revealed that10 hub genes and their related pathways are closely related to the occurrence and development of DN, indicating that these10 hub genes may be regarded as key biomarkers in the diagnosis and treatment of DN. However, our results need to be verified by experiments because our study was conducted on the basis of data analysis. Simultaneously, the correlation between gene expression and clinical manifestations was analyzed, which may provide a novel perspective for the early diagnosis and targeted therapy of DN.

References

1. Rabkin R (2003) Diabetic nephropathy. Clinical cornerstone 5: 1-11.
2. Zhang L, Long J, Jiang W (2016) Trends in Chronic Kidney Disease in China. The New England journal of medicine 375: 905-906.
3. Fineberg D, Jandeleit-Dahm KA, Cooper ME (2013) Diabetic nephropathy: diagnosis and treatment. Nature reviews Endocrinology 9: 713-723.
4. Schrijvers BF, De Vriese AS, Flyvbjerg A (2004) From hyperglycemia to diabetic kidney disease: the role of metabolic, hemodynamic, intracellular factors and growth factors/cytokines. Endocrine reviews 25: 971-1010.
5. Montero RM, Covic A, Gnudi L (2016) Diabetic nephropathy: What does the future hold? International urology and nephrology 48: 99-113.
6. Marshall CB (2016) Rethinking glomerular basement membrane thickening in diabetic nephropathy: adaptive or pathogenic? American journal of physiology Renal physiology 311: F831-f843.
7. Dai H, Liu Q (2017) Research Progress on Mechanism of Podocyte Depletion in Diabetic Nephropathy 2017: 2615286.
8. Federation (2021) ID IDF Diabetes Atlas, 10th ed. Brussels: International Diabetes Federation 2021
9. Lim A (2014) Diabetic nephropathy - complications and treatment. International journal of nephrology and renovascular disease 7: 361-381.
10. Sun YM, Su Y, Li J (2013) Recent advances in understanding the biochemical and molecular mechanism of diabetic nephropathy. Biochemical and biophysical research communications 433: 359-361.
11. Garud MS, Kulkarni YA (2014) Hyperglycemia to nephropathy via transforming growth factor beta. Current diabetes reviews 10: 182-189.
12. Yamagishi S, Matsui T (2010) Advanced glycation end products, oxidative stress and diabetic nephropathy. Oxidative medicine and cellular longevity 3: 101-108.
13. Noh H, King GL (2007) The role of protein kinase C activation in diabetic nephropathy. Kidney international Supplement 2007: S49-53.
14. Wada J, Makino H (2013) Inflammation and the pathogenesis of diabetic nephropathy. Clinical science 124: 139-152.
15. Fan Y, Lee K, Wang N (2017) The Role of Endoplasmic Reticulum Stress in Diabetic Nephropathy. Current diabetes reports 17: 17.
16. Yang D, Livingston MJ, Liu Z (2018) Autophagy in diabetic kidney disease: regulation, pathological role and therapeutic potential. Cellular and molecular life sciences : CMLS 75: 669-688.
17. Panchapakesan U, Pollock C (2016) Long non-coding RNAs-towards precision medicine in diabetic kidney disease? Clinical science 130: 1599-602.
18. Li JY, Yong TY, Michael MZ (2010) Review: The role of microRNAs in kidney disease. Nephrology 15: 599-608.
19. Ruggenenti P, Cravedi P, Remuzzi G (2010) The RAAS in the pathogenesis and treatment of diabetic nephropathy. Nature reviews Nephrology 6: 319-330.
20. Umanath K, Lewis JB (2018) Update on Diabetic Nephropathy: Core Curriculum 2018. American journal of kidney diseases : the official journal of the National Kidney Foundation 71: 884-895.
21. Qi C, Mao X (2017) Classification and Differential Diagnosis of Diabetic Nephropathy 2017: 8637138.
22. Thomas B (2019) The Global Burden of Diabetic Kidney Disease: Time Trends and Gender Gaps. Current diabetes reports 19: 18.
23. Metzker ML (2010) Sequencing technologies - the next generation. Nature reviews Genetics 11: 31-46.
24. S UK, Rajan B, D TK (2020) Involvement of Essential Signaling Cascades and Analysis of Gene Networks in Diabesity 11.
25. Clough E, Barrett T (2016) The Gene Expression Omnibus Database. Methods in molecular biology 1418: 93-110.
26. Geng XD, Wang WW, Feng Z (2019) Identification of key genes and pathways in diabetic nephropathy by bioinformatics analysis. Journal of diabetes investigation 10: 972-984.
27. Li Z, Liu J, Wang W (2020) Investigation of hub genes involved in diabetic nephropathy using biological informatics methods. Annals of translational medicine 8: 1087.
28. Liu X, Li X (2019) Key Genes Involved in Diabetic Nephropathy Investigated by Microarray Analysis. Journal of computational biology : a journal of computational molecular cell biology 26: 1438-1447.
29. Wang Z, Wang Z, Zhou Z (2016) Crucial genes associated with diabetic nephropathy explored by microarray analysis. BMC nephrology 17: 128.
30. Mou X, Zhou DY, Liu YH (2019) Identification of potential therapeutic target genes in mouse mesangial cells associated with diabetic nephropathy using bioinformatics analysis. Experimental and therapeutic medicine 17: 4617-4627.
31. Wang WN, Zhang WL, Zhou GY (2016) Prediction of the molecular mechanisms and potential therapeutic targets for diabetic nephropathy by bioinformatics methods. International journal of molecular medicine 37: 1181-8.
32. Kolde R, Laur S, Adler P (2012) Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics (Oxford, England) 28: 573-580.
33. Song D, Tian J, Hu Y (2020) Identification of biomarkers associated with diagnosis and prognosis of gastroesophageal junction adenocarcinoma-a study based on integrated bioinformatics analysis in GEO and TCGA database. Medicine 99: e23605.

34. Karpman D, Ståhl AL, Arvidsson I (2017) Extracellular vesicles in renal disease. Nature reviews Nephrology 13: 545-562.

35. Tervaert TW, Mooyaart AL, Amann K (2010) Pathologic classification of diabetic nephropathy. Journal of the American Society of Nephrology 21: 556-563.

36. Flyvbjerg A (2017) The role of the complement system in diabetic nephropathy. Nature reviews Nephrology 13: 311-318.

37. Fortpied J, Vertommen D, Van Schaftingen E (2010) Binding of mannose-binding lectin to fructosamines: a potential link between hyperglycaemia and complement activation in diabetes. Diabetes/metabolism research and reviews 26: 254-260.

38. Kelly KJ, Liu Y, Zhang J (2015) Renal C3 complement component: feed forward to diabetic kidney disease. American journal of nephrology 41: 48-56.

39. Acosta J, Hettinga J, Flückiger R (2000) Molecular basis for a link between complement and the vascular complications of diabetes. Proceedings of the National Academy of Sciences of the United States of America 97: 5450-5655.

40. Wang J, Zhang Q, Li S (2020) Low molecular weight fucoidan alleviates diabetic nephropathy by binding fibronectin and inhibiting ECM-receptor interaction in human renal mesangial cells. International journal of biological macromolecules 150: 304-314.

41. Jun JI, Lau LF (2011) Taking aim at the extracellular matrix: CCN proteins as emerging therapeutic targets. Nat Rev Drug Discov 10: 945-963.

42. Feng Q, Liu D, Lu Y (2020) The Interplay of Renin-Angiotensin System and Toll-Like Receptor 4 in the Inflammation of Diabetic Nephropathy 2020: 6193407.

43. Zhang C, Zhu X, Li L (2019) A small molecule inhibitor MCC950 ameliorates kidney injury in diabetic nephropathy by inhibiting NLRP3 inflammasome activation. Diabetes, metabolic syndrome and obesity : targets and therapy 12: 1297-1309.

44. Wada J, Makino H (2016) Innate immunity in diabetes and diabetic nephropathy. Nature reviews Nephrology 12: 13-26.

45. Dockal M, Carter DC, Rüker F (2000) Conformational transitions of the three recombinant domains of human serum albumin depending on pH. The Journal of biological chemistry 275: 3042-3050.

46. Zhao XM, Zhang Y, He XH (2018) Chinese herbal medicine Shenzhuo Formula treatment in patients with macroalbuminuria secondary to diabetic kidney disease: study protocol for a randomized controlled trial. Trials 19: 200.

47. Taguchi K, Chuang VT, Yamasaki K (2015) Cross-linked human serum albumin dimer has the potential for use as a plasma-retaining agent for the fatty acid-conjugated antidiabetic drugs. The Journal of pharmacy and pharmacology 67: 255-263.

48. Schena FP, Gesualdo L (2005) Pathogenetic mechanisms of diabetic nephropathy. Journal of the American Society of Nephrology 16: S30-33.

49. Yasuda F, Mii A, Morita M (2018) Importance of frequency and morphological characteristics of nodular diabetic glomerulosclerosis in diabetic nephropathy. Human pathology 75: 95-103.

50. Kliewe F, Kaling S, Lötzsch H (2019) Fibronectin is up-regulated in podocytes by mechanical stress. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 33: 14450-14460.

51. Nguyen TQ, Roestenberg P, van Nieuwenhoven FA (2008) CTGF inhibits BMP-7 signaling in diabetic nephropathy. Journal of the American Society of Nephrology 19: 2098-2107.

52. Mason RM (2009) Connective tissue growth factor(CCN2), a pathogenic factor in diabetic nephropathy. What does it do? How does it do it? Journal of cell communication and signaling 3: 95-104.

53. Wang S, Li B, Li C (2015) Potential Renoprotective Agents through Inhibiting CTGF/CCN2 in Diabetic Nephropathy. Journal of diabetes research 2015: 962383.

54. Bailey E, Bottomley MJ, Westwell S (1999) Vascular endothelial growth factor mRNA expression in minimal change, membranous, and diabetic nephropathy demonstrated by non-isotopic in situ hybridisation. Journal of clinical pathology 52: 735-738.

55. Mima A (2013) Diabetic nephropathy: protective factors and a new therapeutic paradigm. Journal of diabetes and its complications 27: 526-530.

56. Majumder S, Advani A (2017) VEGF and the diabetic kidney: More than too much of a good thing. Journal of diabetes and its complications 31: 273-279.

57. Dei Cas A, Gnudi L (2012) VEGF and angiopoietins in diabetic glomerulopathy: how far for a new treatment? Metabolism: clinical and experimental 61: 1666-1673.

58. Riser BL, Najmabadi F, Garchow K (2014) Treatment with the matricellular protein CCN3 blocks and/or reverses fibrosis development in obesity with diabetic nephropathy. The American journal of pathology 184: 2908-2921.

59. Nakatani S, Wei M, Ishimura E (2012) Proteome analysis of laser microdissected glomeruli from formalin-fixed paraffin-embedded kidneys of autopsies of diabetic patients: nephronectin is associated with the development of diabetic glomerulosclerosis. Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association 27: 1889-1897.

60. He J, Dijkstra KL, Bakker K (2020) Glomerular clusterin expression is increased in diabetic nephropathy and protects against oxidative stress-induced apoptosis in podocytes. Scientific reports 10: 14888.

61. Kim JH, Kim JH, Jun HO (2010) Protective effect of clusterin from oxidative stress-induced apoptosis in human retinal pigment epithelial cells. Investigative ophthalmology & visual science 51: 561-566.

62. Jun HO, Kim DH, Lee SW (2011) Clusterin protects H9c2 cardiomyocytes from oxidative stress-induced apoptosis via Akt/GSK-3β signaling pathway. Experimental & molecular medicine 43: 53-61.

63. Zhang J, Wang Y, Zhang R (2019) Serum levels of immunoglobulin G and complement 3 differentiate non-diabetic renal disease from diabetic nephropathy in patients with type 2 diabetes mellitus. Acta diabetologica 56: 873-881.

64. Kang BP, Frencher S, Reddy V (2003) High glucose promotes mesangial cell apoptosis by oxidant-dependent mechanism. American journal of physiology Renal physiology 284: F455-466.

65. Yang S, Chintapalli J, Sodagum L (2005) Activated IGF-1R inhibits hyperglycemia-induced DNA damage and promotes DNA repair by homologous recombination. American journal of physiology Renal physiology 289: F1144-1152.

66. Brittain AL, Kopchick JJ (2019) A review of renal GH/IGF1 family gene expression in chronic kidney diseases. Growth hormone & IGF research : official journal of the Growth Hormone Research Society and the International IGF Research Society 48-49: 1-4.

67. Ma FY, Flanc RS, Tesch GH (2007) A pathogenic role for c-Jun amino-terminal kinase signaling in renal fibrosis and tubular cell apoptosis. Journal of the American Society of Nephrology 18: 472-484.

68. Mao CP, Gu ZL (2005) Puerarin reduces increased c-fos, c-jun, and type IV collagen expression caused by high glucose in glomerular mesangial cells. Acta pharmacologica Sinica 26:982-986.