Identification of Hub Genes in Diabetic Nephropathy by an Integrated Bioinformatic Analysis

Chenchun Chen
Guangxi Medical University https://orcid.org/0000-0002-2598-5585
Xiaoqiang Qiu (xqqiu9999@163.com)
Guangxi Medical University

Research Article

Keywords: diabetic nephropathy (DN), differentially expressed genes (DEGs), weighted gene co-expression network analysis (WGCNA), hub gene

Posted Date: July 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-703904/v1

Identification of hub genes in diabetic nephropathy by an integrated bioinformatic analysis

Chenchun Chen [1], Xiaoqiang Qiu [2]

Department of Epidemiology and Health Statistics, School of Public Health,Guangxi Medical University, Nanning 530021, Guangxi, China.

Correspondence address: Xiaoqiang Qiu; xqqiu9999@163.com

Abstract

Background: Diabetic nephropathy (DN) is a progressive kidney disease caused by damage to the capillaries in the kidneys' glomeruli. The dysregulation of genes plays a significant role in the progression of DN.

Methods: In the present study, gene expression profiles were analyzed to identify the key genes and pathways involved in DN. Gene expression profiles of 9 patients with DN and 11 normal subjects were downloaded from the Gene Expression Omnibus (GEO) database. Subsequently, differentially expressed genes (DEGs) were identified and subjected to functional enrichment analysis. In addition, protein-protein interaction (PPI) networks were established and hub genes were identified.

Results: As a result, a total of 424 DEGs were identified (113 were upregulated and 311 were downregulated). In the protein-protein interaction network, four hub modules and 30 hub genes were identified. To explore potential associations between gene and DN clinical features and to identify hub genes, the top 25% of genes with the greatest variance in the gene expression profiles were extracted for weighted correlation network analysis (WGCNA). There were ten genes (RNASE6, CD1C, SASH3,

COL1A2, MS4A6A, CD163, CLEC10A, MOXD1, IQGAP2, GHR) identified as significant DN-associated genes. Furthermore, the expression level of these hub genes was confirmed in the GSE96804 dataset.

Conclusions: These findings provide new insight into DN pathogenesis, which may enhance our fundamental knowledge of the molecular mechanisms underlying this disease.

Key words: diabetic nephropathy (DN), differentially expressed genes (DEGs), weighted gene co-expression network analysis (WGCNA), hub gene

Background

The prevalence of diabetes and its complications poses a major threat to global health, has contributed tremendously to the burden of mortality and disability (1). Acute metabolic complications of diabetes associated with mortality include hyperglycemia and coma due to hypoglycemia(2). While the most devastating consequence of diabetes is its long-term vascular complications(3, 4). These complications are wide-ranging and result, at least in part, from vascular damage caused by chronically elevated blood glucose levels(5). Diabetic microvascular complications (nephropathy, retinopathy, and neuropathy), which are long-term complications that affect small blood vessels, usually affect those with a chronic or uncontrollable disease, but they can also be observed in those who have been diagnosed or have not yet made a diagnosis of diabetes(6).

Diabetic nephropathy (DN) is one of the most fatal long-term complications of diabetes and a leading cause of chronic kidney disease (7). The main signatures of DN usually include glomerular scarring, proteinuria, a progressive decline in renal function, and

even end-stage renal disease (ESRD), which are attributed to tubular interstitial fibrosis, hypertrophy, and dilatation of the glomerular mesentery, thickening of the glomerular basement membrane, loss of foot cell peduncles, and inflammation due to monocyte and macrophage infiltration (7, 8). The pathophysiology of DN is complex, involving interactions between genetic factors, epigenetic factors, and the environment. Diabetogenic stimuli, such as high blood glucose levels; advanced glycation end products (AGEs); growth factors including transforming growth factor β1 (TGF-β1) and platelet-derived growth factor and inflammatory cytokines, which have been implicated in the pathogenesis of DN due to their detrimental effects on multiple renal cell types(9-11). Although the pathophysiology of DN is continually being elucidated, the underlying molecular mechanisms of DN progression are not fully understood.

Advances in histological techniques and the integration of high-dimensional data through systems medicine approaches can provide the molecular mechanism of action for drugs and disease progression pathways (12). Genomic data related to various diseases are stored in public repositories, which can be easily accessed to obtain meaningful information and make novel discoveries (13). Transcriptomic analysis during the development of DN may be of great potential value for timely diagnosis and timely treatment to prevent progression to end-stage renal disease.

In the present study, transcriptomic data of human DN from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) were analyzed with RRA to identify differentially expressed genes (DEGs) between DN tissues and normal subjects (14). These DEGs were adopted into the Gene Ontology (GO) and Kyoto Encyclopedia of

Genes and Genomes (KEGG) analyses. The protein-protein interaction (PPI) network was further constructed to understand cellular mechanisms and interactions between cell's molecular constituents of selected genes. In addition, the top 25% of genes with the greatest variance in the dataset were extracted to perform weighted gene co-expression network analysis (WGCNA). Ten genes were selected as key genes based on the screening conditions of module membership (MM) >0.8 and gene significance (GS)>0.7. Finally, the expression level of these key genes was validated in another expression profiling of DN in GSE96804.

Methods

Datasets selection and data preprocessing

The study design was conducted in the form of a flow diagram (Figure 1). Two appropriate gene expression profiles were downloaded from the Gene Expression Omnibus (GEO) database. The selection datasets were as follows: the GSE30528 (Affymetrix Human Genome U133A 2.0 Array) dataset, with a total of 9 DN samples and 11 normal samples, and the GSE96804 (Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version) dataset, with a total of 41 DN samples and 20 normal samples. In addition, the clinical characteristics of sample information are available online.

Identification of reliable DEGs

The series matrix file of glomeruli samples with normal and DN in the dataset GSE30528 was obtained from the GEO. The R package "limma" was utilized for data analysis, linear models, and screen the DEGs between DN samples and normal samples

(15). Genes with adjusted p-value < 0.05 and |log2fold change (FC)| >1 were considered as significant DEGs. “Pheatmap” (R package) was performed to visualize the expression patterns of the top 200 DEGs (top 100 up-regulated genes and top 100 down-regulated genes according to adjusted P).

Function enrichment and pathway analysis

To study the functional annotation and signaling pathways of the DEGs, Gene Ontology (GO) term enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were practiced with the R package “clusterprofiler”(16). The GO categories of biological process (BP), molecular function (MF), and cellular component (CC) were shown separately. The GO term enrichment and KEGG pathway analysis with adjusted p-value < 0.05 were considered statistically significant.

PPI network of DEGs

To explore the group of proteins encoded by the list of genes, the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (http://www. string-db.org/) was utilized to assess PPI information. Meanwhile, the cytohubba plug-in of Cytoscape software (NRNB, National Resource for Network Biology, US) was applied to PPI network visualization and hub genes selection (17).

Co-expression network analysis of DEGs

The R package Weighted correlation network analysis “(WGCNA)” is a comprehensive collection of R functions for performing various aspects of weighted correlation network analysis. In this study, the top 25% of genes with the greatest variance in the dataset GSE30528 were extracted as the input data for subsequent WGCNA. The

adjacency matrix was converted into a topological overlap matrix (TOM). Genes were divided into different gene modules based on TOM-based difference dissimilarity measures. Here, the power of $\beta = 19$ (scale-free $R^2 = 0.89$) was set as the soft threshold to ensure a scale-free network. The modules with the highest correlation with clinical features were selected to explore their function. Genes in the module with gene significance (GS) > 0.3 and module membership (MM) > 0.8 were defined as key genes.

Statistical analysis

The results were presented as mean ± standard error mean (S.E.M.). Statistical analysis was performed with IBM SPSS Statistics 25.0 software (SPSS, Inc., Chicago, IL, USA). Significant differences were measured by one-way ANOVA for two groups of data followed by a Tukey's posthoc comparison. P value<0.05 was considered statistically significant.

Results

Identification of DEGs

The workflow for identification, validation, function enrichment, and pathway analysis of DEGs was shown in Figure 1A. The DEGs screening criteria were set in the limma package, and a total of 113 up-regulated and 311 down-regulated significant DEGs were identified from the genes of glomeruli samples with DN patients and normal subjects. The volcano plot displayed the distribution of DEGs between DN and normal glomeruli samples (Figure 1B). Compared to normal subjects, the gene C1QA was the most significant up-regulated gene (P value=2.66E-08, adjusted P value= 7.11E-06), followed by SERPINE2 (P value=3.10E-07, adjusted P value=3.53E-05) in DN samples.

C1QA-associated activation of the complement system was found in patients with diabetic nephropathy (18). Immunostaining for C1QA was also found in the renal cortex and medulla oblongata of sheep with acute kidney injury (19). Meanwhile the genes ZNF415 (P value=7.44E-11, adjusted P value=8.09E-07), LOX (P value=1.81E-10, adjusted P value=8.09E-07) and CA10 (P value=2.31E-10, adjusted P value=8.09E-07) were the most significant down-regulated genes in DN samples. The top 100 up-regulated and down-regulated DEGs were shown in the heatmap (Figure 1C).

Function enrichment of DEGs

To further understand the function of the identified DEGs for diabetic nephropathy, the DEGs were subjected to perform GO analyses in biological process, molecular function, and cellular component. The bar plot was implemented to visualize the analysis of enriched GO terms. In the terms of biological processes, kidney development, renal system development, urogenital system development, and positive regulation of cell adhesion were considered as the significant enrichment (Figure 2A). The actin-binding and cell adhesion molecule binding was the most significantly enriched GO term in molecular function (Figure 2B). What's more, for cellular components, extracellular matrix and adherens junction were the most significantly enriched GO terms (Figure 2C).

Pathway analysis of DEGs

To explore the molecular interaction, reaction, and relation networks in diabetic nephropathy. The DEGs were enriched in the KEGG pathway database. The KEGG pathway enrichment analysis indicated that the DEGs were notably accumulated in the

PI3K-Akt signaling pathway and the Focal adhesion pathway (Figure 3A). Based on accumulated evidence, the PI3K-Akt signaling pathway is necessary for normal metabolism and its imbalance leads to obesity and the development of type 2 diabetes(20).

Construction of PPI network

To gain insight into cellular physiology in normal and diabetic nephropathy states, the PPI networks of DEGs were constructed and presented in Figure 4A. The minimum required interaction score was set at the highest confidence (confidence=0.9) and disconnected nodes in the network were hidden. The Cytoscape plugin cytoHubba provided 11 topological analysis methods to rank nodes in a network by the network features (17). To explore important nodes/hubs and fragile motifs in an interactome network, the top 30 essential genes ranked by Maximal Clique Centrality (MCC) scores were selected. Four hub modules and top 30 essential genes were presented in the PPI network (Figure 4B). The top 10 genes ranked by MCC were C3, GPR183, GPR18, P2RY14, CCL5, CXCL6, CCL19, SST, ADRA2A, and FN1.

Construction of gene co-expression modules

To find the key modules that were most relevant to the clinical features of diabetic nephropathy, the top 25% of genes with the greatest variance in the dataset GSE30528 were extracted to perform WGCNA. Clinical characteristics were retrieved from sample information online. The soft thresholding power β was first calculated, to which the co-expression similarity is raised to calculate adjacency. Here, we set the soft thresholding power β as 19 in the ensuing analysis. The scale independence reached

0.89 (Figure 5A) and with a relatively high-average connectivity (Figure 5B). Then, we constructed the gene network and identified modules using a one-step network construction function. Finally, we eventually identified 11 gene co-expression modules (Figure 6A).

Analysis of gene co-expression modules

The relationship between identified modules was mapped (Figure 7A). Subsequently, we transformed the weighted adjacency matrix into a topological overlap matrix (TOM) and computed the corresponding heterogeneity to minimize noise and spurious association. The heat map depicted the TOM between selected genes included in the analysis. Lighter colors indicated low overlap, while progressively darker reds indicated increasing overlap. Higher TOM values suggest that a pair of genes are more likely to connect and to a set of shared genes. The analysis showed that gene expression was relatively independent between modules.

Identification of key modules

The modules and clinical characteristics were correlated for searching the most significant associations. Our results indicated that the module magenta was most significantly negatively correlated with diabetic nephropathy (correlation coefficient=0.92, P-value =8E-10), while module green was most significantly positively correlated with diabetic nephropathy (correlation coefficient=0.72, P-value =2E-04) (Figure 8A). Based on the correlation coefficient results we identified the green module as the key module for further analysis.

Identification and validation of key genes

There were 120 genes were included in module green. By setting module membership (MM) >0.8 and gene significance (GS) >0.7, 10 genes were selected for key genes: RNASE6, CD1C, SASH3, COL1A2, MS4A6A, CD163, CLEC10A, MOXD1, IQGAP2, GHR (Table 1). Then, gene expression profiles from the GSE96804 dataset were used to validate the expression of these key genes in diabetic nephropathy. As shown in Figure 9A-J, the expression of the up-regulated key genes RNASE6, CD1C, SASH3, COL1A2, MS4A6A, CD163, CLEC10A, AND MOXD1 in diabetic nephropathy were significantly elevated compared with normal subjects. While the expression of IQGAP2 and GHR in diabetic nephropathy was remarkably down-regulated. Furthermore, the regulation of these key genes was all consistent with the relevance between gene expression and clinical characteristics. These results show that the selected key genes are generally differentially expressed in diabetic nephropathy.

Discussion

Diabetic nephropathy is a highly prevalent complication of diabetes mellitus (DM), which is influenced by both environmental and genetic factors (8). Clinically, patients with DN generally have a poorer prognosis compared with patients without DN. Recent studies on DN have had limited success, in part because not all patients diagnosed with DN have renal dysfunction as a result of their diabetes mellitus. In the development and progression of diabetic nephropathy, genetic factors and signal transduction pathways influence the expression of genes and phenotypes associated with DN. Therefore, the escalating rates of DN indicate the need for a more in-depth understanding of the underlying molecular mechanisms to explore susceptibility modules and genes to

identify better therapies for this disease (8, 21).

In the present study, 424 genes were differentially expressed in glomeruli samples with DN patients and normal subjects. Among these 424 DEGs, 113 genes were upregulated and 311 genes were downregulated. The complement system was identified as one of the significantly regulated pathways in diabetic kidneys according to the analysis of glomerular and tubular (22). Similar to the previous results, the expression of C1QA was markedly upregulated in DN samples. In addition, analysis of the types of immune cells in the glomerulus of diabetic mice reveals that the cells in the immune clusters are predominantly macrophages, showing high expression of typical macrophage markers such as C1QA, Cd74, and Adgre1 (23).

GO term enrichment analysis and KEGG pathway analysis for functional annotation were performed with both upregulated and downregulated genes. GO term analysis revealed that the DEGs were mainly associated with cell adhesion molecule binding, glycosaminoglycan binding, actin binding, and enzyme inhibitor activity. To date, four families of cell adhesion molecules have been identified: globulins, selectins, integrins, and immune Globulin CAM superfamily (24). Leukocytes bound to the activated endothelium via cell adhesion molecule and its receptors and then migrated into the tissue, where they then began the inflammatory process (25). With the further induction of these pathways, massive microvasculature is injured and eventually leads to the complications observed in T2DM patients (26). The KEGG pathway enrichment analysis indicated that the DEGs were accumulated in the phosphatidylinositol-3-kinase/protein kinase B (PI3K-Akt) signaling pathway, Focal adhesion, Regulation of

actin cytoskeleton, Rap1 signaling pathway, and MAPK signaling pathway. Recent experimental evidence suggests that PI3K-Akt is involved in ROS effects on the activation of Nrf2 by oxidative stress (27). Diabetic nephropathy may be improved by antioxidant action and thus decreased ROS production (28). These studies further confirm that the PI3K-Akt signaling pathway plays a critical role in the progression of DN (29, 30).

Moreover, the PPI networks were constructed, and 30 outstanding genes were identified by Maximal Clique Centrality scores. The top 10 genes (C3, GPR183, GPR18, P2RY14, CCL5, CXCL6, CCL19, SST, ADRA2A, and FN1) were considered weighty. Some of these genes have been generally demonstrated in previous studies for their overexpression and relevance in DN. In several animal models of diabetes, Ig and C3 deposition, which was associated with macrophage infiltration, is thought to be in glomeruli and glomerular capillaries (31, 32). In addition, recent studies have shown that the association between the expression of inflammatory genes such as CCL2 and CCL19 and the development of DN can also be replicated in mice, and genetic or pharmacological intervention studies have shown that increased levels of cytokines contribute to the development of DN (33, 34). However, some of these selected genes, such as GPR18, SST, and ADRA2A were still lacking exploration for their functions in DN.

WGCNA is a systems biology method for describing pairwise relationships between gene transcripts and a comprehensive collection of R functions for performing aspects of weighted correlation network analysis (35). Through intensive and systematic

reanalysis of the GSE30528 dataset, we determined that the green module is significantly associated with clinical traits in DN patients. To find more significative genes in DN, we screened 68 genes out from the green module with a cut-off of module membership (MM) >0.8 and gene significance (GS) >0.7. After filtering for GS and MM values, we eventually obtained 10 hub genes (RNASE6, CD1C, SASH3, COL1A2, CD163, CLEC10A, MOXD1, IQGAP2, GHR). Among these ten genes, RNASE6, CD1C, and COL1A2 have been demonstrated to exert essential roles in the pathogenesis of DN (36-38). And CD163 was considered as an early biomarker of nephropathy in Swedish patients with diabetes of 15–34 years of age (39). Finally, the expression level of these hub genes was validated in the GSE96804 dataset. The results of expression level were similar by clinical features.

Conclusion

In summary, a series of bioinformatics analyses of DN samples were performed in the present study. We selected DEGs obtained from the GSE63514 dataset and explored their functions and pathways that may be involved in the initiation and progression of DN. Furthermore, hub genes were identified according to the analysis of WGCNA, revealing an important role of the glomeruli in the pathological mechanisms of diabetic nephropathy. The potential of these key genes for diagnostic, prognostic, and therapeutic targeting deserves further exploration and demonstration.

Availability of data and materials

All data can be accessed in the GEO database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE30528,https://www.ncbi.nl

m.nih.gov/geo/query/acc.cgi?acc=GSE96804).

References

1. Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. Nat Rev Endocrinol. 2018;14(2):88-98.

2. Campbell JE, Drucker DJ. Islet alpha cells and glucagon--critical regulators of energy homeostasis. Nat Rev Endocrinol. 2015;11(6):329-38.

3. Paneni F, Beckman JA, Creager MA, Cosentino F. Diabetes and vascular disease: pathophysiology, clinical consequences, and medical therapy: part I. Eur Heart J. 2013;34(31):2436-43.

4. Li JP, Wei W, Li XX, Xu M. Regulation of NLRP3 inflammasome by CD38 through cADPR-mediated Ca(2+) release in vascular smooth muscle cells in diabetic mice. Life Sci. 2020;255:117758.

5. Forbes JM, Cooper ME. Mechanisms of diabetic complications. Physiol Rev. 2013;93(1):137-88.

6. Valencia WM, Florez H. How to prevent the microvascular complications of type 2 diabetes beyond glucose control. BMJ. 2017;356:i6505.

7. Falkevall A, Mehlem A, Palombo I, Heller Sahlgren B, Ebarasi L, He L, et al. Reducing VEGF-B Signaling Ameliorates Renal Lipotoxicity and Protects against Diabetic Kidney Disease. Cell Metab. 2017;25(3):713-26.

8. Kato M, Natarajan R. Epigenetics and epigenomics in diabetic kidney disease and metabolic memory. Nat Rev Nephrol. 2019;15(6):327-45.

9. Sharma K, Ziyadeh FN. Hyperglycemia and diabetic kidney disease. The case for transforming growth factor-beta as a key mediator. Diabetes. 1995;44(10):1139-46.

10. Kato M, Yuan H, Xu ZG, Lanting L, Li SL, Wang M, et al. Role of the Akt/FoxO3a pathway in TGF-beta1-mediated mesangial cell dysfunction: a novel mechanism related to diabetic kidney disease.

309 J Am Soc Nephrol. 2006;17(12):3325-35.

310 11. Ruggenenti P, Cravedi P, Remuzzi G. The RAAS in the pathogenesis and treatment of diabetic

311 nephropathy. Nat Rev Nephrol. 2010;6(6):319-30.

312 12. Lambers Heerspink HJ, Oberbauer R, Perco P, Heinzel A, Heinze G, Mayer G, et al. Drugs meeting

313 the molecular basis of diabetic kidney disease: bridging from molecular mechanism to personalized

314 medicine. Nephrol Dial Transplant. 2015;30 Suppl 4:iv105-12.

315 13. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive

316 for functional genomics data sets--update. Nucleic Acids Res. 2013;41(Database issue):D991-5.

317 14. Kolde R, Laur S, Adler P, Vilo J. Robust rank aggregation for gene list integration and meta-analysis.

318 Bioinformatics. 2012;28(4):573-80.

319 15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression

320 analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

321 16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes

322 among gene clusters. OMICS. 2012;16(5):284-7.

323 17. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-

324 networks from complex interactome. BMC Syst Biol. 2014;8 Suppl 4:S11.

325 18. Bus P, Chua JS, Klessens CQF, Zandbergen M, Wolterbeek R, van Kooten C, et al. Complement

326 Activation in Patients With Diabetic Nephropathy. Kidney Int Rep. 2018;3(2):302-13.

327 19. Palviainen MJ, Junnikkala S, Raekallio M, Meri S, Vainio O. Activation of complement system in

328 kidney after ketoprofen-induced kidney injury in sheep. Acta Vet Scand. 2015;57:15.

329 20. Huang X, Liu G, Guo J, Su Z. The PI3K/AKT pathway in obesity and type 2 diabetes. Int J Biol

330 Sci. 2018;14(11):1483-96.

21. Reidy K, Kang HM, Hostetter T, Susztak K. Molecular mechanisms of diabetic kidney disease. J Clin Invest. 2014;124(6):2333-40.

22. Woroniecka KI, Park AS, Mohtat D, Thomas DB, Pullman JM, Susztak K. Transcriptome analysis of human diabetic kidney disease. Diabetes. 2011;60(9):2354-69.

23. Fu J, Akat KM, Sun ZG, Zhang WJ, Schlondorff D, Liu ZH, et al. Single-Cell RNA Profiling of Glomerular Cells Shows Dynamic Changes in Experimental Diabetic Kidney Disease. Journal of the American Society of Nephrology. 2019;30(4):533-45.

24. Aricescu AR, Jones EY. Immunoglobulin superfamily cell adhesion molecules: zippers and signals. Curr Opin Cell Biol. 2007;19(5):543-50.

25. Schnoor M, Alcaide P, Voisin MB, van Buul JD. Crossing the Vascular Wall: Common and Unique Mechanisms Exploited by Different Leukocyte Subsets during Extravasation. Mediators Inflamm. 2015;2015:946509.

26. Hadi HA, Suwaidi JA. Endothelial dysfunction in diabetes mellitus. Vasc Health Risk Manag. 2007;3(6):853-76.

27. Deng X, Rui W, Zhang F, Ding W. PM2.5 induces Nrf2-mediated defense mechanisms against oxidative stress by activating PIK3/AKT signaling pathway in human lung alveolar epithelial A549 cells. Cell Biol Toxicol. 2013;29(3):143-57.

28. Yay A, Akkus D, Yapislar H, Balcioglu E, Sonmez MF, Ozdamar S. Antioxidant effect of carnosine treatment on renal oxidative stress in streptozotocin-induced diabetic rats. Biotech Histochem. 2014;89(8):552-7.

29. Piao Y, Yin D. Mechanism underlying treatment of diabetic kidney disease using Traditional Chinese Medicine based on theory of Yin and Yang balance. J Tradit Chin Med. 2018;38(5):797-802.

30. Lee WC, Li LC, Ng HY, Lin PT, Chiou TT, Kuo WH, et al. Urinary Exosomal MicroRNA Signatures in Nephrotic, Biopsy-Proven Diabetic Nephropathy. J Clin Med. 2020;9(4).

31. Wehner H, Hohn D, Faix-Schade U, Huber H, Walzer P. Glomerular changes in mice with spontaneous hereditary diabetes. Lab Invest. 1972;27(3):331-40.

32. Chow F, Ozols E, Nikolic-Paterson DJ, Atkins RC, Tesch GH. Macrophages in mouse type 2 diabetic nephropathy: correlation with diabetic state and progressive renal injury. Kidney Int. 2004;65(1):116-28.

33. Breyer MD, Susztak K. The next generation of therapeutics for chronic kidney disease. Nat Rev Drug Discov. 2016;15(8):568-88.

34. Anderberg RJ, Meek RL, Hudkins KL, Cooney SK, Alpers CE, Leboeuf RC, et al. Serum amyloid A and inflammation in diabetic kidney disease and podocytes. Lab Invest. 2015;95(6):697.

35. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

36. Cui C, Cui Y, Fu Y, Ma S, Zhang S. Microarray analysis reveals gene and microRNA signatures in diabetic kidney disease. Mol Med Rep. 2018;17(2):2161-8.

37. Tang SCW, Yiu WH. Innate immunity in diabetic kidney disease. Nat Rev Nephrol. 2020;16(4):206-

38. Kelly KJ, Liu Y, Zhang J, Dominguez JH. Renal C3 complement component: feed forward to diabetic kidney disease. Am J Nephrol. 2015;41(1):48-56.

39. Samuelsson M, Dereke J, Svensson MK, Landin-Olsson M, Hillman M, on the behalf of the DSg. Soluble plasma proteins ST2 and CD163 as early biomarkers of nephropathy in Swedish patients with diabetes, 15-34 years of age: a prospective cohort study. Diabetol Metab Syndr. 2017;9:41.

Acknowledgements

Not applicable.

Funding

This work was supported by the National Key Research and Development Program (2017YFC0907103) and Key research and Nanning Key Research and Development Program (20183039-2).

Author information

Affiliations

Department of Epidemiology and Health Statistics, School of Public Health,Guangxi Medical University, Nanning 530021, Guangxi, China.

Contributions

Contribution to the concept or design of the work: XQQ acquisition, analysis or interpretation of data: CCC. All authors read and approved the final manuscript.

Corresponding author

Correspondence to Xiaoqiang Qiu.

Ethics declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

All the authors declare they have no competing interests.

Tables:

Gene	Gene significance	Module membership
RNASE6	0.82	0.93
CD1C	0.80	0.92
SASH3	0.77	0.91
COL1A2	0.75	0.90
MS4A6A	0.77	0.88
CD163	0.81	0.84
CLEC10A	0.78	0.83
MOXD1	0.76	0.83
IQGAP2	-0.80	-0.80
GHR	-0.79	-0.89

Table 1 Hub genes of the module green. Hub genes were defined as having a gene significance over 0.7 and a module membership over 0.8.

Figures:

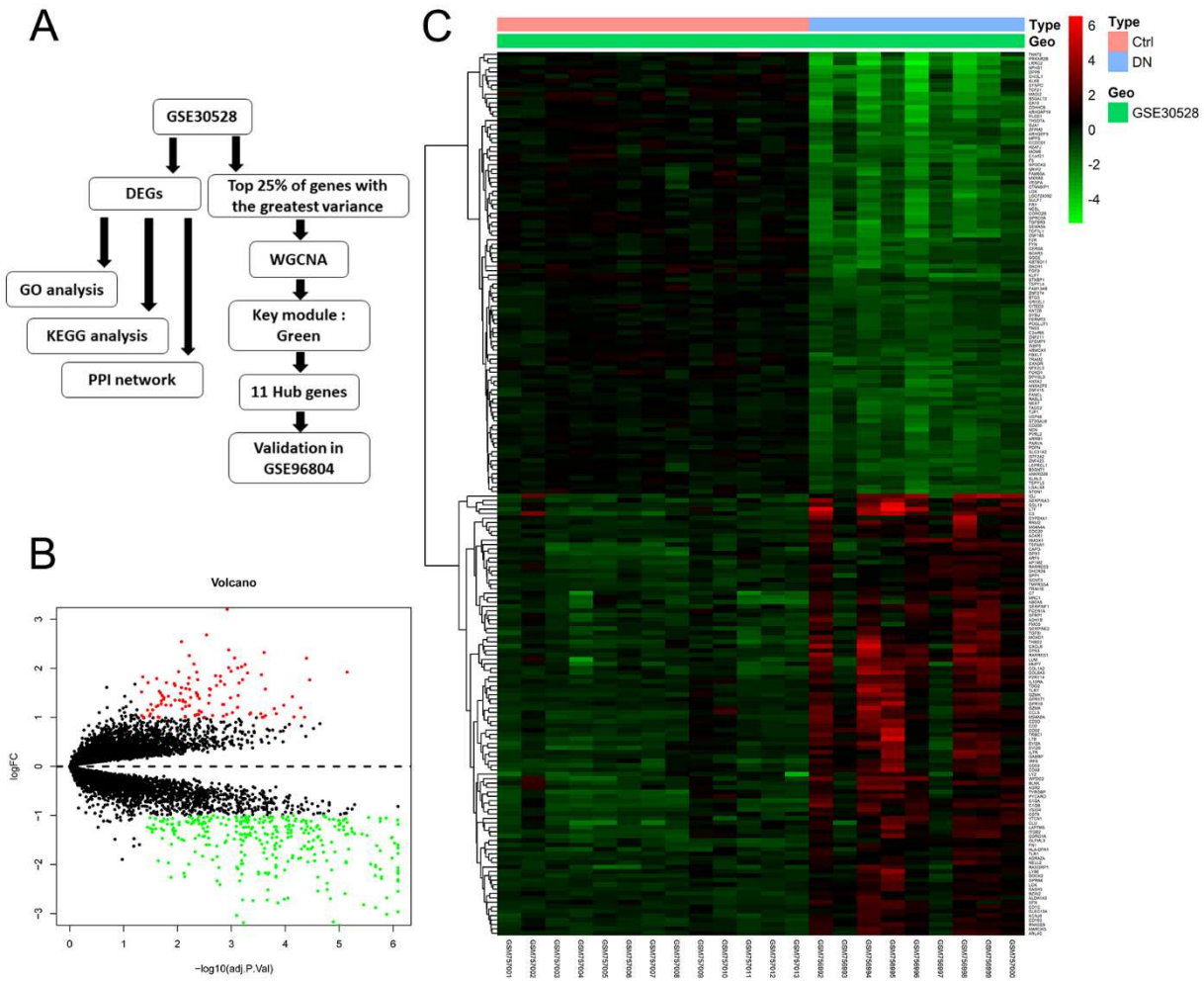
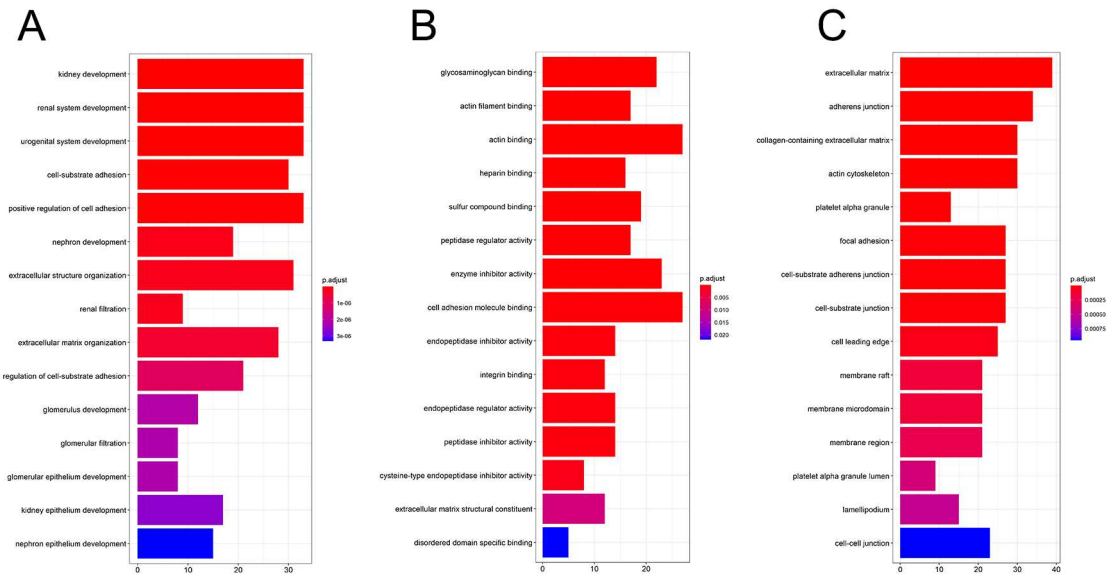


Figure 1 The flow diagram of data preparing, processing, and analysis in our study (A). The volcano map of DEGs, identified in GSE30528, between DN and normal glomeruli samples (B). The red points in the volcano plots represent upregulation and the green plots represent downregulation. Heatmap of the top 200 DEGs (top 100 up-regulated DEGs and top 100 down-regulated DEGs) according to the value of |logFC| (C). The color from green to red in the heatmap shows the process from low expression to high expression.



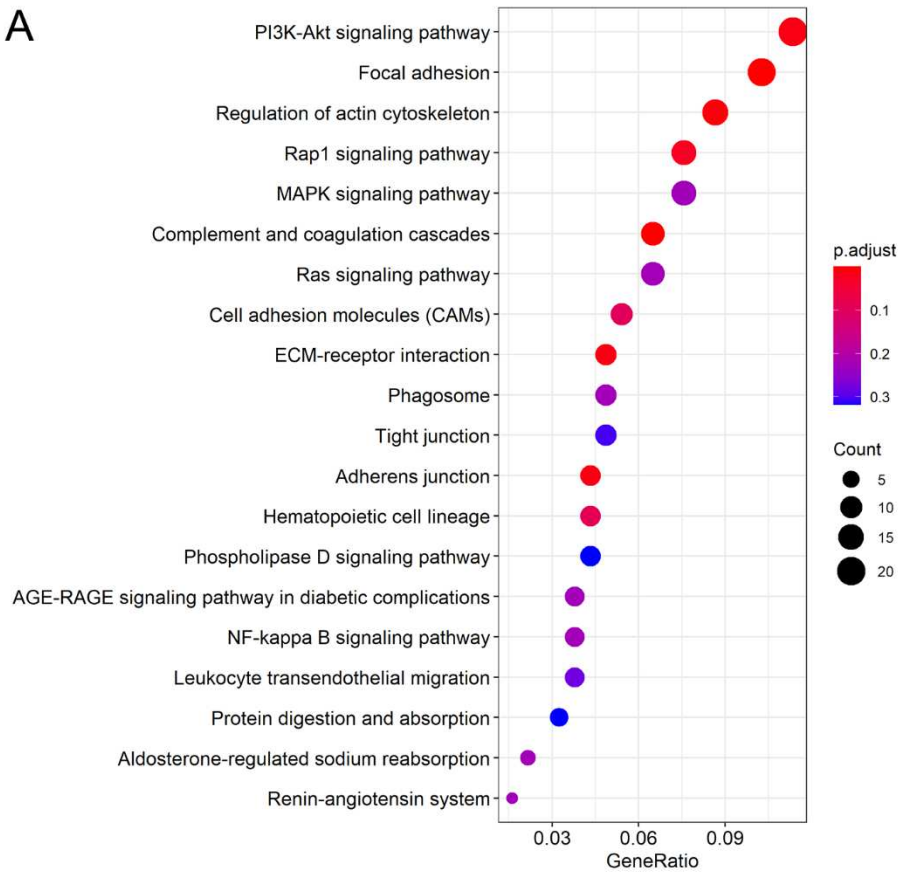
457

458 **Figure 2** Gene Ontology analysis of the DEGs regarding the biological process(A),

459 cellular component(B), and molecular function (C). The top 15 significant terms of

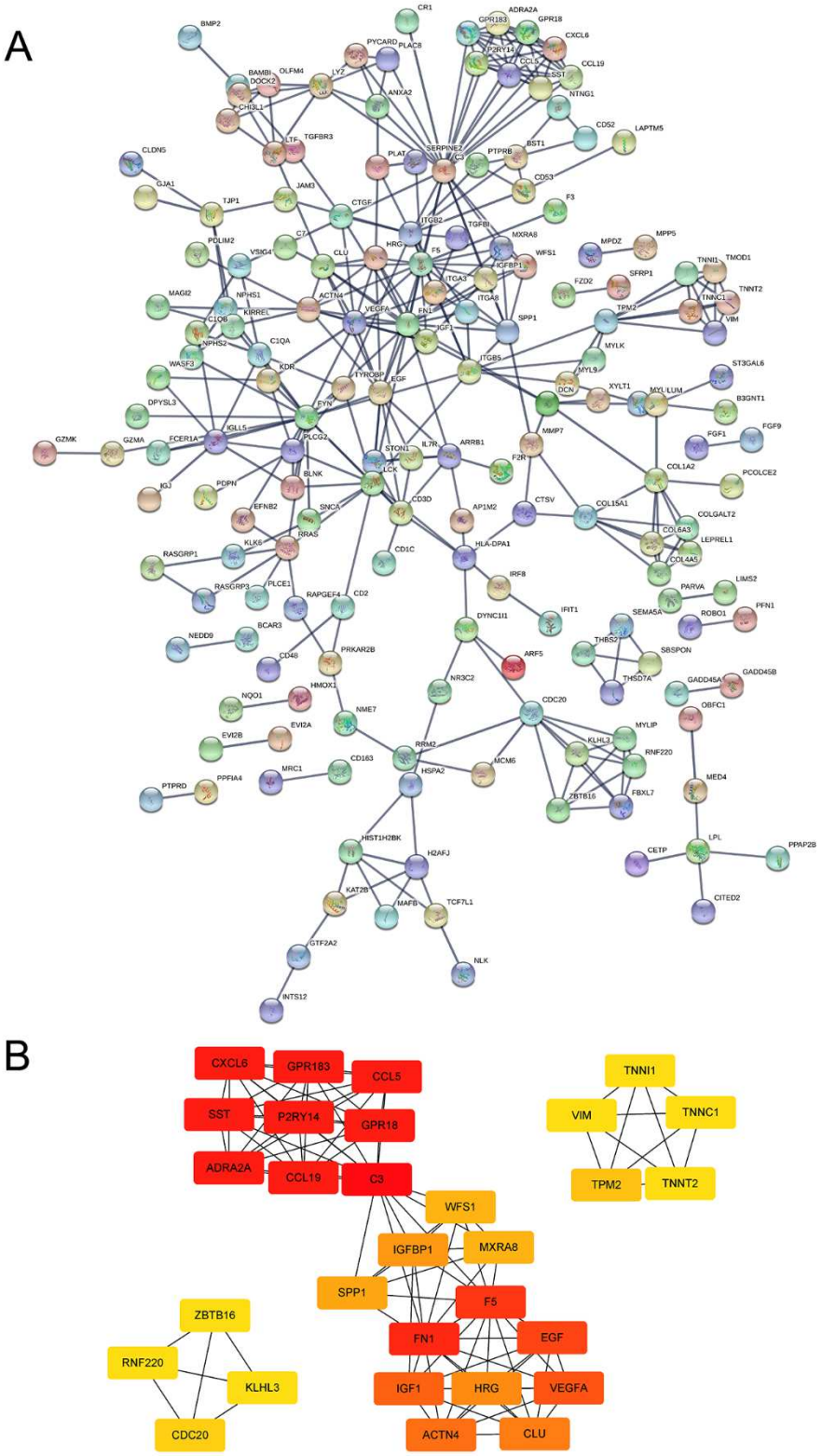
460 enrichment analysis are shown in the bar chart.

461



462

463 **Figure 3** The top 20 terms of KEGG pathway enrichment analysis in the dot plot (A).



464

465 **Figure 4** Protein-protein interaction network of the DEGs (A). Top 30 essential genes

466 in the PPI network (B).

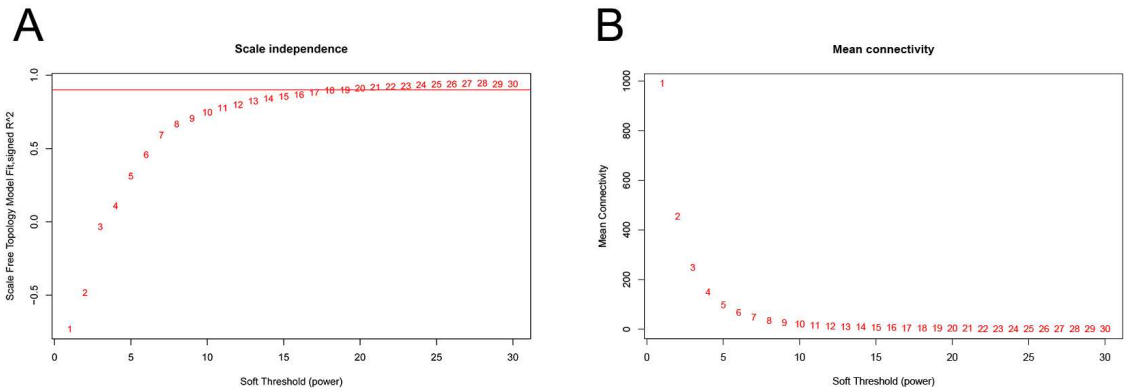
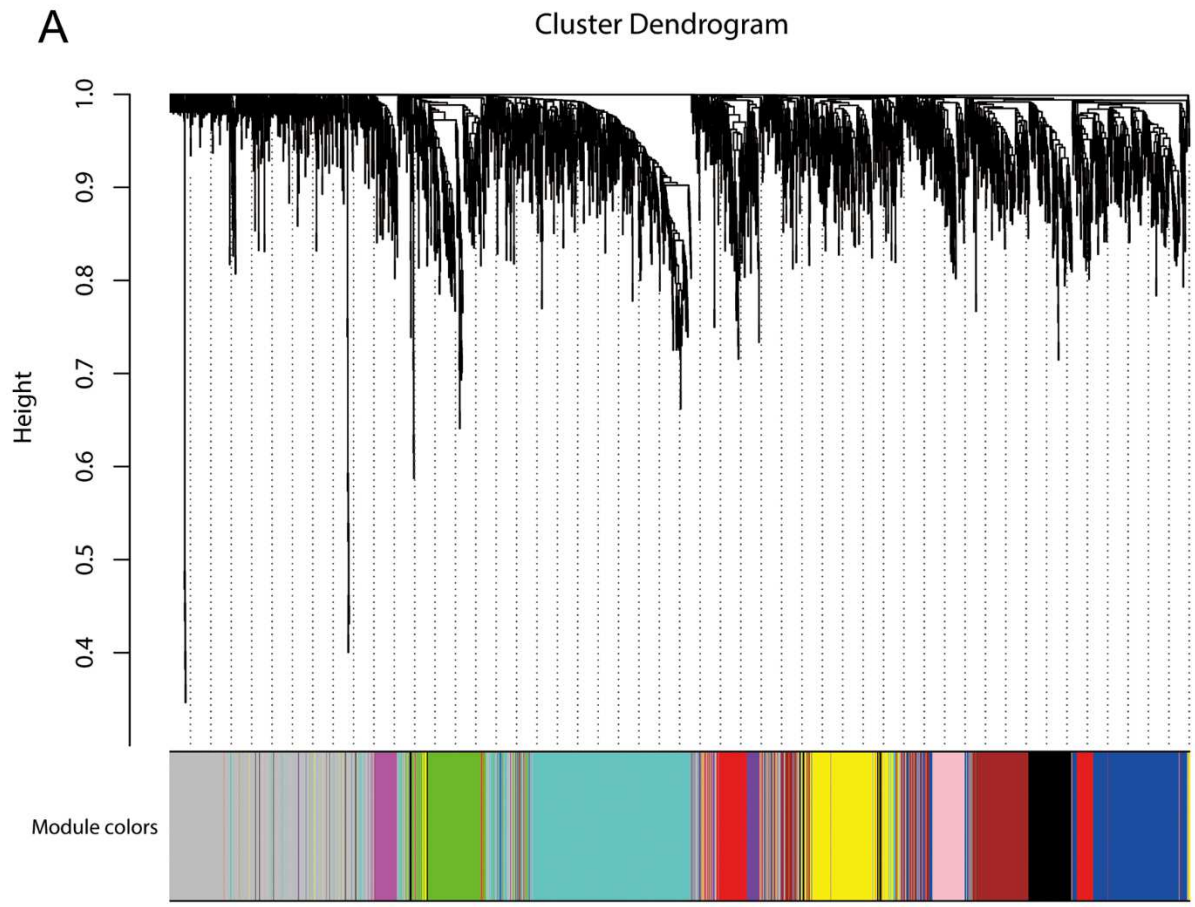


Figure 5 Analysis of network topology for various soft-thresholding powers. The x-axis reflects the soft-thresholding power. The y-axis reflects the scale-free topology model fit index (A). The x-axis reflects the soft-thresholding power. The y-axis reflects the mean connectivity (degree) (B).



473

474 **Figure 6** Clustering dendrogram of genes, with dissimilarity based on the topological

475 overlap, together with assigned module colors (A).

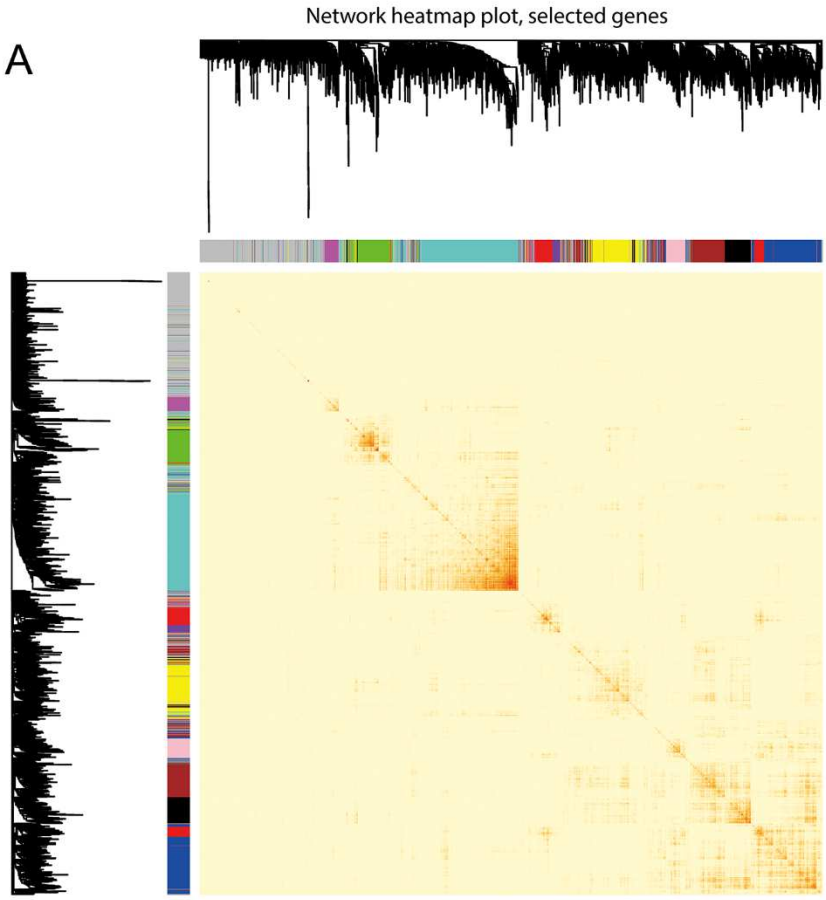
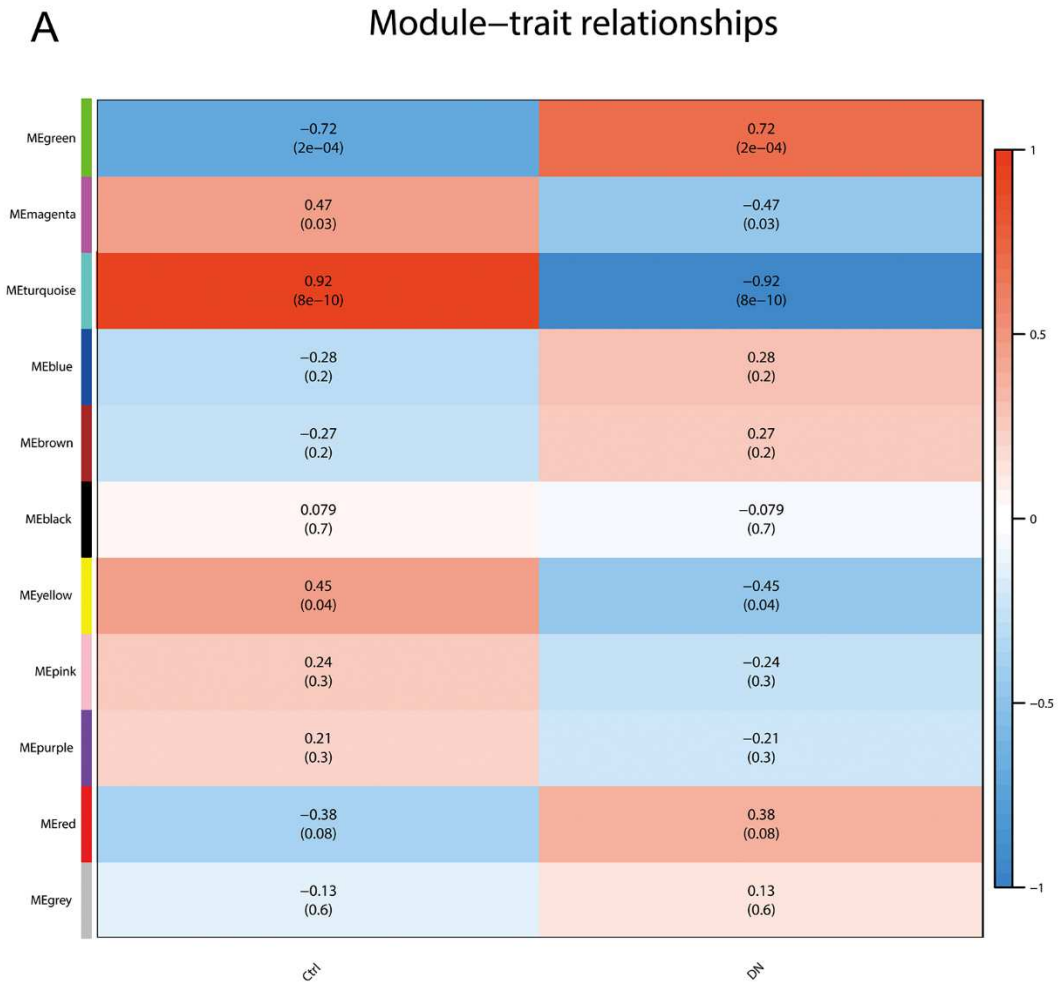


Figure 7 Visualization of the WGCNA network using a heatmap plot. The heatmap depicts the topological overlap matrix (TOM) among all modules included in the analysis. The light color represents a low overlap, and the progressively darker red color represents an increasing overlap (A).

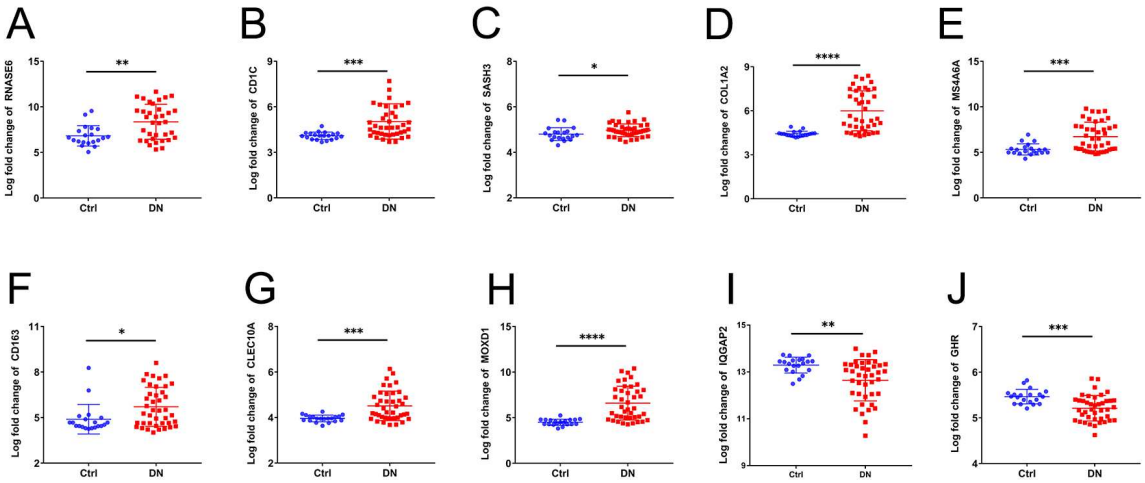


481

482 **Figure 8** Module–trait associations. Each row corresponds to a module, and each

483 column corresponds to a trait. Each cell contains the corresponding correlation and P-

484 value. The table is color-coded by correlation according to the color legend (A).



485

486 **Figure 9** Scatter plot of the expression level of key genes (A-J). The blue scatter reflects

487 the control group, and the red scatter reflects the T2DM group. * P < 0.05,

488 **P<0.01,***P<0.001,****P<0.0001.