Key Genes and Signaling Pathways Contribute to the Pathogenesis of Diabetic Nephropathy

Hailing Yang,1* Dede Lian,2* Xiaofei Zhang,3 Hongjun Li, Guangda Xin5

Introduction. Diabetic nephropathy (DN) is a serious complication of diabetes mellitus involving damage to the capillaries in the glomerulus. This study aimed to explore key genes and signaling pathways participate in the progression of DN.

Methods. Two gene expression profile datasets GSE1009 and GSE30528 downloaded from Gene Expression Omnibus (GEO) were used to analyze the differentially expressed genes (DEGs) between DN samples and controls. Coupled two-way clustering (CTWC) and correspondence analysis were performed to explore the potential functions of DEGs. Then, Gene Ontology (GO) terms and pathways associated with DEGs were identified, followed by constructing of the co-expressed gene network and module. Ultimately, the regulatory network based on the DEGs, miRNAs and transcription factors (TFs) was established.

Results. Total 283 common DEGs were identified from the two datasets, including 219 down-regulated ones (bone morphogenetic protein 7 (BMP7), decay accelerating factor (CD55) and coagulation Factor V (F5) etc.) and 64 up-regulated ones (inhibin beta c subunit (INHBC) and colony stimulating factor 1 receptor (CSF1R) etc.). The miRNA-TF regulatory network was established with three miRNAs, 8 TFs and 58 DEGs. Besides, three significant pathways including cytokine-cytokine receptor interaction, complement and coagulation cascades and TGF-beta signaling pathways were identified.

Conclusion. BMP7, CD55, CSF1R, INHBC and F5 are likely to take crucial roles in the pathogenesis of DN.

INTRODUCTION

Diabetic nephropathy (DN), a progressive kidney disease, is a serious complication of diabetes mellitus involving damage to the capillaries in the glomerulus.1 It is characterized by extracellular matrix (ECM) accumulation, tubulointerstitial degeneration, and fibrosis correlated with a sharp decline in the glomerular filtration rate.2 Currently, over 380 million people are affected by DN worldwide. The international diabetes federation has estimated that this number is expected to increase to 592 million by 2035.3 DN is the primary cause of morbidity and mortality in diabetic patients and leads to end-stage renal disease.4, 5 As such, it is urgent to explore the mechanism of DN for its further prevention and treatment.

Numerous studies showed that pathogenic factors of hypertension, hyperglycemia, hyperlipidemia and inflammatory response are involved in the development of DN.6, 7 Hyperglycemia mainly
initiates the pathological process of DN via pathways such as polyol pathway.\textsuperscript{8} Increasing evidences show that pro-inflammatory cytokines play critical roles in pathogenesis of DN, such as interleukin 1 (IL-1), interleukin 6 (IL-6) and tumor necrosis factor-\(\alpha\) (TNF-\(\alpha\)).\textsuperscript{9-11} Additionally, various kinases and oxidative stress mediators can also activate the process of DN.\textsuperscript{12} However, clinical diagnosis and treatment on the basis of these genes and pathways for the management of DN remain unsatisfactory. Thus, the aim of this study is to identify key genes and pathways related to DN based on two datasets GSE1009 and GSE30528.

In the current study, differentially expressed genes (DEGs) in DN were identified by screening two datasets GSE1009 and GSE30528. Furthermore, coupled two-way clustering analysis (CTWC) was performed to confirm the specificity of DEGs. Later, we selected common DEGs with consistency through correspondence analysis, and then analyzed the potential functions of common DEGs with consistency through gene ontology (GO) functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. Thereafter, co-expressed gene network and module were established. Ultimately, the regulatory network on the basis of DEGs, miRNAs and transcription factors (TFs) was constructed.

**MATERIALS AND METHODS**

**Data preprocessing and DEGs screening**

The gene expression profiles of GSE1009\textsuperscript{13} and GSE30528\textsuperscript{14} were downloaded from National Center of Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) on the basis of platform of GPL8300 and GPL571. The two datasets were tissue samples from glomerulus, including 6 samples (3 samples of controls and 3 samples of DN patients) and 22 samples (13 samples of controls and 9 samples of DN patients) respectively.

We downloaded the raw CEL data and used the oligo package (ver.1.40.2) (http://www.bioconductor.org/packages/oligo.html) in R (ver.3.1.3) language to pre-process all the data by performing background correction, conversion of original data and quartile data normalization.\textsuperscript{15,16} Afterwards, DEGs in the two datasets were screened via the limma (ver.3.32.5) (http://bioconductor.org/packages/limma.html) package. Ultimately, the setting of thresholds were \(|\logFC|\) larger than .585 and false discovery rate (FDR) value less than .05.

**CTWC for DEGs**

CTWC can be used to gather the genes which have approximate expression and it is convenient for further research. In this study, the gene expression of the same tissue was significantly different in different states of disease.\textsuperscript{18} We extracted the expression values of DEGs in each samples from standardized transcriptome, then the CTWC\textsuperscript{19,20} of expression values were performed based on the Euclidean distance\textsuperscript{21} via the heatmap package\textsuperscript{22} (ver.1.0.8) (https://cran.r-project.org/package=heatmap) in R (ver.3.1.3) language. The results were presented with a heat map.

**Correspondence analysis of DEGs in GSE1009 and GSE30528 profiles**

All DEGs screened from GSE1009 and GSE30528 profiles were analyzed via correspondence analysis. First, the similarities and differences of DEGs sets were compared, and the results were presented with the Venn diagram. Combining the results of DEGs in two datasets, the common DEGs that were significantly different in the two datasets were selected for the next research. Next, the consistency (up-regulated simultaneously or down-regulated simultaneously) of common DEGs in GSE1009 and GSE30528 were compared. Afterwards, the Pearson correlation coefficient was calculated, and DEGs with consistency in the two datasets were selected from the common DEGs for further research. Finally, in order to show the similarity of gene expression in two datasets, the CTWC based on the expression values of DEGs with consistency were performed.

**Functional and pathway analysis of DEGs with consistency**

Database for Annotation, Visualization and Integrated Discovery (DAVID, ver.6.7) (https://david.ncifcrf.gov/)\textsuperscript{23,24} gene functional classification tool has been developed for relating the functional terms with gene lists by clustering algorithm. In the present study, the DEGs with consistency were divided into significant up-regulated and significant down-regulated gene sets according to the \(\logFC\) value. Then, significant correlations of GO analysis and KEGG pathway analysis were
performed using the DAVID. In the current study, the p-value was measured by hypergeometric distribution. The significant screening threshold was set as $P$-value < .05.

Co-expressed gene network construction and module partition

Based on the expression values of DEGs with consistency in GSE1009 and GSE30528, the Pearson correlation coefficient of expression value between each two DEGs were calculated. Subsequently, only gene pairs with correlation coefficient larger than .8 in the two datasets were retained. At the end, network of co-expressed gene was constructed and displayed with Cytoscape (ver. 3.3) (http://www.cytoscape.org/). Meanwhile, the module partition and function annotation for co-expressed gene network were performed using the Molecular Complex Detection (MCODE, http://apps.cytoscape.org/apps/mcode, parameter: Degree cutoff = 2, Node score cutoff = 0.2, K-core = 2) plugin of Cytoscape (ver. 3.3) and Biological Networks Gene Ontology tool (BINGO) with threshold value of adjusted $P$ < .05.

Retrieval of miRNA associated with DN

We used the miR2 Disease database (http://watson.compbio.iupui.edu:8080/miR2Disease/index.jsp) to search miRNAs associated with DN. “Diabetic nephropathy” was acted as a key word in the database for searching the DN-correlated miRNAs that have been confirmed by report. Each item of miR2 Disease contains detailed information about the connection of miRNA and disease, such as the ID of miRNA, designation of disease, a brief description of relationship between miRNA and disease, detection methods and references for miRNA expression. After obtaining miRNAs related to disease, we retrieved the target genes directly associated with miRNA via miRanda (http://www.microrna.org/microrna/home.do). The target genes regulated by the miRNA associated with disease were mapped to the DEGs with consistency. Ultimately, the regulatory network of gene and miRNA associated with DN was constructed.

Construction of miRNA-TF regulatory network

To further comprehend the genes that constituted co-expression network, we searched for the TFs significantly correlated with co-expression genes through Web-based Gene Set Analysis Toolkit (WebGestalt, ver. 2017) (http://www.webgestalt.org/option.php). The $P$ value less than .05 was selected as the significance threshold for screening correlated TFs. The target genes regulated by TFs were mapped to the target genes involved in gene-miRNA regulatory network, then the target gene regulatory network of miRNA and TF were constructed. Finally, the regulated genes were analyzed for functional and pathways analysis.

RESULTS

Data preprocessing and DEGs screening

After datasets from GSE1009 and GSE30528 were normalized, we screened DEGs by limma package. Total 1521 genes screened from GSE1009 expressed significantly different between DN patients and healthy controls, including 1008 up-regulated DEGs and 513 down-regulated DEGs. Meanwhile, a total of 1348 DEGs were obtained from GSE30528 dataset, including 414 up-regulated DEGs and 934 down-regulated DEGs.

CTWC for DEGs

We extracted the expression values of significant DEGs from normalized gene expression profiles of GSE1009 and GSE30528, and then heatmaps of CTWC on the basis of expression values were constructed via the pheatmap. As shown in Figure 1, the different types of samples can be separated distinctly by the selected expression values in the two datasets. It is shown that the DEGs screened from those two datasets were characteristics.

Correspondence analysis of DEGs in GSE1009 and GSE30528 profiles

As shown in Figure 2A, 312 common DEGs in the two datasets were obtained by comparison. Among these common DEGs, 283 common DEGs were consistent in the direction of differential expression in the two datasets (Figure 2B). The CTWC results of 283 common DEGs showed that expression values of these common DEGs can also separate the two datasets completely (Figure 2C and 2D). Therefore, we selected the 283 common DEGs for further research.

Functional and pathways analysis of DEGs with the consistency

The 283 common DEGs were divided into 219
Figure 1. The heatmap of clustering analysis of differentially expressed genes (DEGs) from GSE1009 (A) and GSE30528 (B).

Figure 2. A, Venn diagram of differentially expressed genes (DEGs) from GSE1009 and GSE30528 comparison; B, Scatter diagram of DEGs with consistency from GSE1009 and GSE30528. The red dots stand for the DEGs whose expression were up-regulated simultaneously in two datasets; while the green dots stand for the DEGs whose expression were down-regulated simultaneously in two datasets; the grey dots stand for the DEGs whose differentially expressed directions were inconsistent. The heatmap of clustering analysis of DEGs with consistency from GSE1009 (C) and GSE30528 (D).
down-regulated ones (BMP7, CD55 and F5 etc.) and 64 up-regulated ones (INHBC and CSF1R etc.) according to logFC. Then we performed GO and pathway analyses for up-regulated DEGs and down-regulated DEGs, respectively. The results showed that down-regulated genes were enriched in 23 GO functions and 6 KEGG pathways (Figure 3A and 3C). Meanwhile, up-regulated genes were enriched in 22 GO functions and 6 KEGG pathways (Figure 3B and 3D).

Co-expressed gene network construction and module partition
Total 2184 co-expressed gene pairs were obtained and used to construct the co-expressed gene network. As shown in Figure 4, there were 208 gene nodes (176 down-regulated genes and 32 up-regulated genes) and 2184 connection edges (74 negative correlation connection edges and 2110 positive correlation connection edges) in co-expressed gene network. Besides, total 4 co-expressed gene modules were obtained via MCODE and BINGO plugin (data not shown).

Retrieval of miRNA associated with DN
We obtained three miRNAs (hsa-miR-377, hsa-miR-216a, and hsa-miR-217) associated with DN through the miR2 Disease database. Afterwards, genes with expression correlation in co-expression network were mapped to target genes of these three miRNAs. A total of 142 pairs of connection between miRNA and target genes were obtained and used to construct the miRNA regulatory network. As shown in Figure 5, total 99
Figure 4. Network of co-expressed differentially expressed genes (DEGs) with consistency. The green lines represent the connection of negative correlation gene pairs, while the red lines represent the connection of positive correlation gene pairs. The change of nodes color from green to red presents the change of logFC from negative to positive. The regular triangle and inverted triangle present significantly up-regulated DEGs and down-regulated DEGs respectively.

Figure 5. Regulatory networks of differentially expressed genes (DEGs) -miRNAs. The regular triangle and inverted triangle present significantly up-regulated DEGs and down-regulated DEGs respectively. The quadrangle presents miRNA. The change of triangle color from green to red presents the change of logFC from negative to positive.
nodes (3 miRNAs and 96 genes (87 down-regulated and 9 up-regulated) ) and 142 connection edges (50 connection edges correlated with has-miR-377, 52 connection edges correlated with has-miR-217 and 40 connection edges correlated with has-miR-216a) were contained in miRNA regulatory network.

Construction of miRNA-TF regulatory network

To further comprehend the target genes regulated by miRNA associated with DN in miRNA regulatory network, we obtained 8 TFs significantly correlated with genes in miRNA regulatory network through WebGestalt (Table 1). Then the regulatory network of miRNA and TF was constructed. As shown in Figure 6, there were 69 nodes (3 miRNAs, 8 TFs and 58 DEGs (3 up-regulated and 55 down-regulated) and 211 connection edges (90 miRNA regulate target gene connections and 121 TF regulate target gene connections) in miRNA-TF regulatory network. Then we performed GO and pathway analysis for genes in miRNA-TF regulatory network. As shown in Table 2, the genes (BMP7, INHBC, CSF1R, CD55 and F5) were prevailingly enriched in 3 different KEGG pathways, such as cytokine-cytokine receptor interaction, complement and coagulation cascades and transforming growth factor (TGF) signaling pathways. Besides, genes were also enriched in 10 GO terms, including negative regulation of cell proliferation, cell adhesion and kidney development etc.

Table 1. List of Significant Transcription Factors (TFs)

<table>
<thead>
<tr>
<th>TF</th>
<th>ID</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>E12</td>
<td>DB_ID: 2409</td>
<td>rawP = 2.92e-19; adjP = 1.81e-17</td>
</tr>
<tr>
<td>FOXO4</td>
<td>DB_ID: 2416</td>
<td>rawP = 8.08e-19; adjP = 2.50e-17</td>
</tr>
<tr>
<td>MAZ</td>
<td>DB_ID: 2430</td>
<td>rawP = 8.15e-14; adjP = 1.26e-12</td>
</tr>
<tr>
<td>FREAC2</td>
<td>DB_ID: 2417</td>
<td>rawP = 6.21e-14; adjP = 1.26e-12</td>
</tr>
<tr>
<td>NFAT</td>
<td>DB_ID: 2437</td>
<td>rawP = 1.37e-13; adjP = 1.70e-12</td>
</tr>
<tr>
<td>AP1</td>
<td>DB_ID: 2402</td>
<td>rawP = 1.95e-12; adjP = 1.51e-11</td>
</tr>
<tr>
<td>PAX4</td>
<td>DB_ID: 2445</td>
<td>rawP = 3.97e-12; adjP = 2.73e-11</td>
</tr>
<tr>
<td>CHX10</td>
<td>DB_ID: 2408</td>
<td>rawP = 1.15e-10; adjP = 7.13e-10</td>
</tr>
</tbody>
</table>

Note. adj stands for adjust.
DISCUSSION
DN is one of a major cause of morbidity and mortality in diabetes mellitus.33 Although numerous studies have been carried out to explore the pathogenesis of DN, it is still not elucidated completely. In this paper, we performed comprehensive bioinformatical analysis to obtain more target genes and pathways involved in the development of DN. The results of this study indicated that three pathways, cytokine-cytokine receptor interaction, complement and coagulation cascades and TGF-beta signaling pathways, might play pivotal roles in DN. Importantly, five genes (BMP7, INHBC, CSF1R, CD55 and F5) enriched in the three pathways and regulated by three miRNAs (hsa-miR-377, hsa-miR-216a, and hsa-miR-217) were likely to participate in the pathogenesis of DN.

The TGF-beta superfamily contains a group of secreted peptides, such as the bone morphogenetic protein (BMP) family and the TGF-beta family.34 Among these secreted peptides, TGF-beta signaling is likely to serve a vital part in the proceeding of DN. Previous investigations suggested that TGF-β1, a member of TGF-beta family, is a major cause for ECM accumulation.35 Furthermore, a new study found that dencichine performed a potential therapeutic effect on DN by down-regulating TGF-β/Smad signaling in DN glomeruli.36 Moreover, our results also revealed that genes in miRNA-TF regulation network were significantly enriched in TGF-beta signaling pathway. These results demonstrated that activation of the TGF-beta signaling pathway might implicate in the pathogenesis of DN. Consequently, we speculated that genes (BMP7, INHBC, etc.) involved in the TGF-beta signaling pathway might also associated with DN.

BMP-7, a member of the TGF-beta superfamily, plays a vital role in the progression of kidney and regulation of nephrogenesis.37 Consistent to our results, a previous study found that expression of BMP-7 was decreased at advanced stage of DN.38 In the present study, BMP-7 was significantly enriched in the TGF-beta signaling pathway. These findings also indicated that BMP-7 might be related to progression of DN through its interaction with TGF-beta signaling pathway. There is no evidence to prove the function of INHBC in DN. However, INHBC also belongs to TGF-beta superfamily.39 Besides, INHBC was regulated by miR-377 in this study, which has been confirmed to take a critical role in the pathophysiology of DN.31 Therefore, INHBC might be related to DN mediated by TGF-beta signaling pathway and miR-377.

Kelly et al. demonstrated that the renal injury in DN is mediated by activation of complement system.40 Our results showed that CD55 and F5 were enriched in complement and coagulation cascades pathway. Thus, we speculate that CD55 and F5 might take a vital part in the development of DN.

Table 2. Functional and Pathways Analyses for Differentially Expressed Genes (DEGs) in miRNA-transcription Factors (TFs) Regulatory Network

<table>
<thead>
<tr>
<th>Term</th>
<th>Count</th>
<th>P</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0008285—negative regulation of cell proliferation</td>
<td>9</td>
<td>5.56E-05</td>
<td>COL4A3, CBLB, BMP2, NDN, ABI1, CD24, GPNMB, CXADR, BMP7</td>
</tr>
<tr>
<td>GO:0042127—regulation of cell proliferation</td>
<td>12</td>
<td>1.35E-04</td>
<td>COL4A3, CBLB, BMP2, NDN, CD24, GPNMB, CXADR, BMP7</td>
</tr>
<tr>
<td>GO:0007167—enzyme linked receptor protein signaling pathway</td>
<td>8</td>
<td>2.76E-04</td>
<td>BMP2, NDN, MYO1E, VEGFA, ABI1, BMP7, CD24, FK1, CSF1R</td>
</tr>
<tr>
<td>GO:0050767—regulation of neurogenesis</td>
<td>5</td>
<td>0.003489</td>
<td>BMP2, MAP1B, NPTN, CD24, BMP7</td>
</tr>
<tr>
<td>GO:0051094—positive regulation of developmental process</td>
<td>6</td>
<td>0.003838</td>
<td>LPL, BMP2, MAP1B, NPTN, CD24, BMP7</td>
</tr>
<tr>
<td>GO:0007155—cell adhesion</td>
<td>9</td>
<td>0.004365</td>
<td>COL4A3, F5, GNE, FERM1, NPTN, CD24, GPNMB, CXADR, CD2AP</td>
</tr>
<tr>
<td>GO:0022610—biological adhesion</td>
<td>9</td>
<td>0.004403</td>
<td>COL4A3, F5, GNE, FERM1, NPTN, CD24, GPNMB, CXADR, CD2AP</td>
</tr>
<tr>
<td>GO:0009967—positive regulation of signal transduction</td>
<td>6</td>
<td>0.004933</td>
<td>BMP2, F3, VEGFA, CD24, BMP7, CITED2</td>
</tr>
<tr>
<td>GO:0001822—kidney development</td>
<td>4</td>
<td>0.005367</td>
<td>TCF21, BMP2, MYO1E, BMP7</td>
</tr>
<tr>
<td>GO:0060284—regulation of cell development</td>
<td>5</td>
<td>0.00734</td>
<td>BMP2, MAP1B, NPTN, CD24, BMP7</td>
</tr>
<tr>
<td>hsa04060: Cytokine-cytokine receptor interaction</td>
<td>5</td>
<td>0.032411</td>
<td>BMP2, VEGFA, INHBC, BMP7, CSF1R</td>
</tr>
<tr>
<td>hsa04350: TGF-beta signaling pathway</td>
<td>3</td>
<td>0.042629</td>
<td>BMP2, INHBC, BMP7</td>
</tr>
<tr>
<td>hsa04610: Complement and coagulation cascades</td>
<td>3</td>
<td>0.043133</td>
<td>CD55, F5, F3</td>
</tr>
</tbody>
</table>

Note. GO stands for Gene Ontology.
of DN. For example, the activation of C3 is strongly associated with DN in rats.\textsuperscript{40} To our knowledge, CD55 can inhibit the activation of C3.\textsuperscript{41} Consequently, down-regulation of CD55 is likely to contribute to the progression of DN through diminishing the inhibition of C3 activation. F5, also known as Factor V Leiden (FVL), serves as a central regulatory role in hemostasis.\textsuperscript{42} Wang et al. found that the mutation of FVL reduced albuminuria in murine diabetic nephropathy and in human type 1 and type 2 diabetic patients.\textsuperscript{43} Furthermore, Peter et al. demonstrated that FVL mutation is relevant for early stages of DN by modifying the glomerular dysfunction.\textsuperscript{44} These results indicated that FVL mutation might have a protective effect in DN. Thus, the abnormal expression of FVL is also likely to take a key part in the pathogenesis of DN.

Wu and co-workers demonstrated that cytokines and their receptors might be applied to predict the progression of DN.\textsuperscript{45} Additionally, Kato et al. observed that miR-217-mediated phosphatase and tensin homologue (PTEN) downregulation might contribute to the activation of protein kinase B (PKB/AKT).\textsuperscript{46} CSF1R, which regulated by miR-217, was enriched in pathways associated with cytokine-cytokine receptor interaction in the current study. Besides, Cannarile et al. revealed that CSF1R plays a crucial role in the proliferation, survival, and motility of macrophages.\textsuperscript{47} A study uncovered that CSF1R in macrophages could activate the protein kinase B (PKB/AKT) through multiple signal transduction pathways.\textsuperscript{48} Moreover, Kattla et al. confirmed that PKB/AKT might serve as a pivotal role in the pathogenesis of DN.\textsuperscript{46} Therefore, we speculate that CSF1R is likely to associate with DN mediated by PKB/AKT and miR-217.

However, the predicted results cannot be verified by laboratory data due to the limitation of sample extraction. In further studies, we will confirm the expression of the above discussed DEGs through establishing the animal model. Also, the interaction of DEGs and regulatory relationship between TFs and DEGs will be verified.

**CONCLUSIONS**

In summary, our results indicated that the complement and coagulation cascades, TGF-beta signaling pathway and cytokine-cytokine receptor interaction pathway were likely to correlate with the progression of DN. Five genes (BMP7, INHBC, CSF1R, CD55 and F5) regulated by the three pathways might serve a crucial role in the pathogenesis of DN. Theses findings might provide a further understanding for the pathogenesis of DN and help for the development of novel therapeutic targets in DN treatment.

**FUNDING**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**DECLARATIONS OF INTEREST**

None.

**REFERENCES**

11. Mccarthy ET, Sharma R, Sharma M, et al. TNF-alpha increases albumin permeability of isolated rat glomeruli through the generation of superoxide. Journal of the...
Key genes and signaling pathways for DN—Yang et al


Correspondance to:
Guangda Xin, MM
Department of Nephrology, China-Japan Union Hospital of Jilin University, No.126 Xiantai Street, Changchun 130033, China
Tel: +86 139 4494 6581
Fax: +86 0431 8464 1026
E-mail: xin51097461@163.com

Received July 2018
Revised September 2018
Accepted November 2018