Identification of Transcription Factors related to Diabetic Tubulointerstitial Injury

Jialu Liu  
The Second Xiangya Hospital

Guangzhong Duan  
Hunan Communication Engineering Polytechnic

Wenxia Yang  
The Second Xiangya Hospital

Shumin Zhang  
The Second Xiangya Hospital

Fuyou Liu  
The Second Xiangya Hospital

Youming Peng  
The Second Xiangya Hospital

Lin Sun  
The Second Xiangya Hospital

Yu Liu  
The Second Xiangya Hospital

Li Xiao  
Li Xiao (xiaolizndx@csu.edu.cn)  
https://orcid.org/0000-0002-5488-5583

Research article

Keywords: diabetic nephropathies, tubulointerstitial injury, transcription factor, computational biology, regulatory network

Posted Date: August 1st, 2019

DOI: https://doi.org/10.21203/rs.2.12236/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background: Diabetic nephropathy (DN) is a main cause of chronic renal failure. Despite decades of extensive study, the molecular mechanisms underlying diabetic tubulointerstitial injury remain unclear. We aim to identify key transcription factor genes involved in diabetic tubulointerstitial injury. Methods: A microarray dataset (GSE30122) from Gene Expression Omnibus (GEO) was downloaded. A total of 38 transcription factor genes based on 166 differentially expressed genes (DEGs) were identified by UCSC_TFBS. Results: The regulatory network showed connections between top 10 transcription factors and their target DEGs. Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of targeted DEGs indicated that extracellular space, extracellular exosome, cell surface and complement and coagulation cascades were most significantly enriched. Utilizing Nephroseq v5 online platform, the mRNA expression pattern analysis of transcription factor genes demonstrated that mRNA expression of CDC5, CEBPA, FAC1, HFH1, IRF1, NFE2 and TGIF1 increased in renal tubulointerstium of DN patients compared with normal controls while that of CEBPB and FOXO4 decreased in renal tubulointerstium of DN patients compared with normal controls. Correlation analysis between mRNA expression of transcription factor genes in renal tubulointerstium and clinical features showed that AP1, BACH1, CDC5, FAC1, FOXD1, FOXJ2, FOXO1, FOXO4, HFH1, IRF1, POU3F2, SOX5, SOX9, RSRFC4, S8 and TGIF1 may be related to diabetic tubulointerstitial injury. Conclusions: 1) CDC5, FAC1, FOXO4, HFH1, IRF1 and TGIF1 may be key transcription factor genes. 2) Transcription factors involved in diabetic tubulointerstitial injury may become prospective targets for diagnosis and treatment of DN.

Background

As a common complication of diabetes, diabetic nephropathy (DN) has become a primary cause of end stage renal disease (ESRD) worldwide (1, 2). Prognosis of DN patients is closely related to the severity of renal tubulointerstitial fibrosis (3). Besides, proximal tubulopathy has been viewed as a key initial factor in the progression of DN (4). Certain transcription factors (TFs) have been demonstrated to play an important role in these pathologies (5). For instance, myocardin-related transcription factor A (MRTF-A) can promote transcription of type I and II collagen in an epigenetic manner (6). Hypoxia-inducible factor (HIF-1) has been reported to mediates renal tubulointerstitial fibrosis in a murine model of type 1 diabetes (7). Despite decades of extensive study, the molecular mechanisms underlying diabetic tubulointerstitial injury remain unclear. Thus, it is of great significance to identify key TFs associated with diabetic tubulointerstitial injury, as specific therapeutics can then be developed to target activation of selected TFs.

Recently, bioinformatic methods have been broadly employed to screen differentially expressed genes (DEGs) and transcription factor genes. In this study, a mRNA microarray dataset downloaded from Gene Expression Omnibus (GEO) was used for further analysis. DEGs between renal tubulointerstitial tissues of DN patients and normal controls were selected to predict transcription factor genes. The regulatory network between top 10 transcription factors and their target DEGs was constructed by Cytoscape software. Possible mechanisms on how these TFs might exert their influence on diabetic tubulointerstitial
injury via target DEGs were investigated through Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. The mRNA expression pattern analysis of transcription factor genes as well as correlation analysis between mRNA expression of transcription factor genes in renal tubulointerstium and clinical features of DN were performed using Nephroseq v5 online platform. Taken together, a total of 38 transcription factor genes based on 166 DEGs were identified, which may become potential diagnostic biomarkers and therapeutic targets for diabetic tubulointerstitial injury.

Methods

Microarray data information. Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo) is a public genomics data repository storing abundant high throughput gene expression data (8). The series of GSE30122(9) was downloaded from GEO database which based on GPL571(Affymetrix Human Genome U133A 2.0 Array) platform. This microarray datas include 24 normal controls and 10 renal tubulointerstitial tissue samples from DN patients.

Data preprocessing and differential expression analysis. The raw data were preprocessed by log2 transformation and Z-score normalization. The expression level of genes with more than one probe was averaged. DEGs (adjusted P-value < 0.05 and | log FC | (fold change) > 1) between renal tubulointerstitial tissues of DN patients and healthy controls were screened by limma package(10) in R software. Afterwards, volcano plot of DEGs was drawn by gplots package in R software.

Identification of transcription factor genes and regulatory network construction of top 10 transcription factors. Transcription factor genes (adjusted P-value < 0.05) involved in diabetic tubulointerstitial injury were selected by UCSC_TFBS on Database for Annotation, Visualization and Integrated Discovery (DAVID). The regulatory network between top 10 transcription factors and their target genes was visualized by Cytoscape (version 3.7.0) software based on the data from UCSC_TFBS.

Gene ontology(GO) and pathway analyses. As an online bioinformatics database, Database for Annotation, Visualization and Integrated Discovery (DAVID 6.8, http://david.ncifcrf.gov) (11, 12) provides comprehensive functional annotation information on multiple genes. GO enrichment analysis covers categories of biological processes (BP), cellular component (CC) and molecular function (MF) (13). KEGG is a widely used database in conducting pathway analysis(11). GO enrichment and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analyses of targeted genes that regulated by identified transcription factors were performed using DAVID online tools. Gene count >2 and P<0.05 were set as the cutoff value.

Statistical analysis. mRNA expression pattern (9, 14, 15) of transcription factor genes in renal tubulointerstium in DN patients compared with normal controls was analyzed by Nephroseq v5 online platform (http://v5.nephroseq.org). Also, Pearson correlation analysis between transcription factor genes and glomerular filtration rate (GFR) (9, 15), serum creatine level (SCR) (14, 15), proteinuria(15), body weight (15) and body mass index (BMI) in renal tubulointerstium in DN patients was performed.
Insignificant results are not shown. Comparisons between 2 groups were performed using unpaired the Student’s t test. A two-tailed value of $P < 0.05$ was considered statistically significant.

**Results**

Screening of DEGs involved in diabetic tubulointerstitial injury. To identify DEGs related to diabetic tubulointerstitial injury, the mRNA expression microarray (GSE30122) was downloaded from GEO. After normalization of the raw microarray data (Figure 1a), 166 DEGs associated with diabetic tubulointerstitial lesions were identified using limma package as shown in the volcano plot (Figure 1b). Among them, 159 genes were upregulated and 7 genes were downregulated.

Identification of transcription factor genes and regulatory network construction of top 10 transcription factors. To determine transcription factor genes related to diabetic tubulointerstitial injury, UCSC_TFBS online tool on DAVID was employed to identify transcription factor genes that regulate DEGs. As shown in Figure 2a, a total of 38 transcription factor genes were indicated to be involved in diabetic tubulointerstitial injury. Top 10 transcription factors and their target DEGs were applied to create the regulatory network via Cytoscape software. The regulatory network consisted of 500 interactions between 10 transcription factors and 116 DEGs (Figure 2b).

GO enrichment analysis of targeted DEGs. To investigate biological roles of DEGs moduled by 38 transcription factors, GO enrichment analysis was conducted via DAVID. Myelination ($P=5.43E-05$), glomerulus development ($P=5.64E-05$), lung development ($P=6.01E-05$), extracellular matrix organization ($P=6.74E-05$) and heart development ($P=2.41E-04$) were the top5 significant enrichment of biological process (Figure 3a). Extracellular space ($P=3.04E-08$), extracellular exosome ($P=6.00E-06$), cell surface ($P=1.71E-05$), slit diaphragm ($P=3.34E-05$) and proteinaceous extracellular matrix ($P=1.07E-04$) were the top5 significant enrichment of cell component (Figure 3a). Glycosaminoglycan binding ($P=5.39E-04$), heparin binding ($P=6.42E-04$), extracellular matrix binding ($P=0.0016$), phospholipase inhibitor activity ($P=0.0043$) and tropomyosin binding ($P=0.0069$) were the top5 significant enrichment of molecular function (Figure 3a).

KEGG pathway analysis of targeted DEGs. To explore signaling pathways of DEGs moduled by identified transcription factors, KEGG pathway analysis was performed via DAVID. Figure 3b showed that these DEGs were primarily enriched in complement and coagulation cascades ($P=5.60E-04$), tight junction ($P=0.0105$) and cell adhesion molecules (CAMs) ($P=0.0129$).

The mRNA expression pattern of transcription factor genes in diabetic renal tubulointerstitial. To find out the mRNA expression pattern of selected transcription factor genes, relevant analysis was performed by Nephroseq v5 online platform. The results demonstrated that the mRNA expression of $CDC5$, $CEBPA$, $FAC1$, $HFH1$, $IRF1$, $NFE2$ and $TGIF1$ increased in renal tubulointerstium of DN patients compared with normal controls while that of $CEBPB$ and $FOXO4$ decreased in renal tubulointerstitial of DN patients compared with normal controls (Figure 4).
Association between mRNA expression of transcription factor genes in renal tubulointerstium and clinical features of DN. To explore clinical significance of identified transcription factors in DN, correlation analysis between transcription factor genes and clinical features of DN was conducted by Nephroseq v5 online tool. Firstly, the results showed that mRNA expression of AP1, BACH1, CDC5, FAC1, FOXJ2, IRF1, POU3F2, SOX5, SOX9 and TGIF1 in renal tubulointerstium reversely correlated with GFR in DN patients (Figure 5), suggesting that those transcription factor genes may contribute to the progression of DN. Meanwhile, the mRNA expression of FOXO1 and FOXO4 in renal tubulointerstium positively correlated with GFR in DN patients (Figure 5), indicating that the two transcription factor genes may play a renoprotective role in DN. Secondly, the mRNA expression of AP1, BACH2, FOXD1, FOXJ2 and IRF1 in renal tubulointerstium positively correlated with SCR in DN patients (Figure 6), suggesting that those transcription factor genes may promote the progression of DN. Besides, the mRNA expression of FOXO4, RSRC4 and S8 in renal tubulointerstium negatively correlated with SCR in DN patients (Figure 6), indicating that the three transcription factor genes may have renoprotective roles in DN. Thirdly, the mRNA expression of CDC5 in renal tubulointerstium negatively correlated with proteinuria in DN patients (Figure 7a). Besides, the mRNA expression of CDC5 and FOXO4 in renal tubulointerstium negatively correlated with weight of DN patients (Figures 7b-c). Moreover, the mRNA expression of HFH1 in renal tubulointerstium positively correlated with body mass index in DN patients (Figure 7d).

Discussion

Diabetic nephropathy is a globally leading cause of chronic renal failure. In recent years, diabetic tubulopathy has been recognized to have crucial roles in the development of DN(16, 17). Transcription factors regulate —turn on and off— genes via binding to specific DNA sequences, which are vital for various pathophysiological processes(18). Such TFs also play key roles in diabetic tubulointerstitial injury as TGF-β(19), HIF-1(20) and MRTF-A(6). Although vigorous efforts have been made, the underlying mechanisms of diabetic tubulointerstitial injury still await clarification. The widespread use of microarray technology and bioinformatic methods enables us to identify key transcription factor genes involved in diabetic tubulointerstitial injury, which might yield additional interventional strategies for DN.

A total of 38 transcription factor genes based on 166 DEGs between renal tubulointerstitial tissues of DN patients and normal controls were predicted via UCSC_TFBS. The regulatory network showed connections between top 10 transcription factors and their target DEGs. GO enrichment analysis of targeted DEGs demonstrated that extracellular space, extracellular exosome and cell surface were most significantly enriched. The extracellular space refers to the part of a multicellular organism outside the cells, in which extracellular matrix presents. Diabetic tubulointerstitial fibrosis is characterized by increasing deposition of extracellular matrix in the extracellular space (21). Besides, particular molecules derived from extracellular exosomes have been suggested to serve as potential diagnostic biomarkers in DN including AQP2(22), AQP5(22) and let-7c-5p(23). The loss of molecular binding events between cell surfaces is also involved in diabetic tubulointerstitial fibrosis(24). KEGG pathway analysis of targeted DEGs showed that these DEGs were primarily mapped to complement and coagulation cascades, tight junction and cell adhesion molecules. Existing findings support that activated complement system and
procoagulant events make a contribution to diabetic tubulointerstitial injury\(^{25-28}\). An in vitro study conducted in Madin-Darby canine kidney (MDCK) cell line has demonstrated that exposure to high glucose can result in a significant perturbation of the tight junction associated tubular barrier\(^{29}\). Moreover, cell adhesion molecules such as VCAM-1\(^{30}\) and ICAM-1\(^{31}\) have been reported to play an important role in diabetic tubulointerstitial injury. Together, all these publications are consistent with our results.

Among 38 transcription factor genes, \textit{CDC5}, \textit{FAC1}, \textit{FOXO4}, \textit{HFH1}, \textit{IRF1} and \textit{TGIF1} were not only differentially expressed between renal tubulointerstitial tissues of DN patients and normal controls, but also closely related to clinical features of DN. Thus, these 6 candidates may be key transcription factor genes involved in diabetic tubulointerstitial injury. Forkhead box O4 (FOXO4) is a transcription factor involved in the modulation of hypoxia inducible factor 1 subunit alpha (HIF1A)\(^{32}\), cell cycle\(^{33}\) and insulin signaling pathway\(^{34}\). It has already been recognized as a key transcriptional regulator in DN\(^{35}\). Intriguingly, a previous study demonstrated that the induction of FOXO4 was responsible for podocyte apoptosis mediated by advanced glycation end products\(^{36}\). However, the results of our study suggested that FOXO4 may have a renoprotective role in diabetic tubulointerstitial injury, raising the possibility that one transcription factor may exert a distinctive effect on different parts of the kidney.

Yet, there is still no report on the association between other 5 transcription factor genes (\textit{CDC5}, \textit{FAC1}, \textit{HFH1}, \textit{IRF1} and \textit{TGIF1}) and diabetic nephropyathy. Cell division cycle 5 like (CDC5) is a DNA-binding protein that regulates cell cycle\(^{37}\). \textit{FAC1}, also named as bromodomain PHD finger transcription factor (\textit{BPTF}), is a transcription factor gene related to chromatin remodeling\(^{38}\). HFH1 (Forkhead box Q1,FOXQ1) has been reported to mediate epithelial-mesenchymal transition in various human cancers\(^{39}\). Interferon regulatory factor 1 (IRF1) is a transcription factor regulating multiple cellular processes, especially for the modulation of interferon (IFN) and IFN-inducible genes\(^{40}\). TGF\(_B\) induced factor homeobox 1 (TGIF1) can act as a transcriptional corepressor of SMAD2\(^{41}\) and suppress the function of retinoid X (RXR) receptor\(^{42}\).

\textbf{Conclusions}

In conclusion, this study was intended to search for key transcription factor genes related to diabetic tubulointerstitial injury. A total of 38 transcription factor genes based on 166 DEGs were screened by UCSC\_TFBS, which may provide new insights into pathogenesis and potential drugable targets for DN. Of them, \textit{CDC5}, \textit{FAC1}, \textit{FOXO4}, \textit{HFH1}, \textit{IRF1} and \textit{TGIF1} may be key transcription factor genes. Further experimental studies are needed to confirm our results and delineate biofunctions of those TFs related to diabetic tubulointerstitial injury.

\textbf{Abbreviations}

DN patients: patients with diabetic nephropathy; ESRD: end stage renal disease; DN: Diabetic nephropathy; TFs: transcription factors; MRTF-A: myocardin-related transcription factor A; HIF-1: hypoxia-
inducible factor; GEO: Gene Expression Omnibus; DEGs: differentially expressed genes; KEGG: Kyoto Encyclopedia of Genes and Genomes; UCSC_TFBS: a category provided by DAVID; DAVID: The Database for Annotation, Visualization and Integrated Discovery; GFR: glomerular filtration rate; SCR: serum creatine level.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All data supporting the results of this article are included within the article.

Competing interests

The authors declare no competing interests.

Funding

This work was supported by grants (No.81570658 and No.81570622) from the National Natural Science Foundation of China.

Authors' contributions

JL and GD designed the research and drafted the manuscript. WY and SZ acquired data, analyze data and Statistical analysis. FL, YP, SL, YL and LX revised manuscript for important intellectual content. All authors read and approved the final manuscript.

Acknowledgement

We thank all the patients who participated in the study.

References


Figures
Figure 1

Box plot of normalized data and volcano plot analysis. a Box plot of normalized data from 34 samples. b Volcano plot analysis of DEGs. Red dots represent upregulated genes and green dots represent downregulated genes.
Figure 2

Transcription factor genes identified by UCSC_TFBS and network construction of top 10 transcription factors. a Bar plot of 38 transcription factor genes. Numbers in the bar represent the amount of DEGs modulated by corresponding transcription factors. b Regulatory network of top 10 transcription factors. Yellow circles represent transcription factors. Red and green circles represent upregulated and downregulated DEGs respectively.
Figure 3

GO enrichment and KEGG pathway analyses of DEGs modulated by identified transcription factors. a GO enrichment analysis for targeted DEGs. b KEGG pathway analysis for targeted DEGs. GO, Gene Ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes.
Figure 4

The mRNA expression pattern of transcription factor genes in renal tubulointerstium of DN patients compared with normal controls. a The mRNA expression of CDC5 increased in DN patients compared with normal controls. b The mRNA expression of CEBPA increased in DN patients compared with normal controls. c The mRNA expression of CEBPB decreased in DN patients compared with normal controls. d The mRNA expression of FAC1 increased in DN patients compared with normal controls. e The mRNA expression of FOXO4 decreased in DN patients compared with normal controls. f The mRNA expression of HFH1 increased in DN patients compared with normal controls. g The mRNA expression of IRF1 increased in DN patients compared with normal controls. h The mRNA expression of IRF1 increased in DN patients compared with normal controls. i The mRNA expression of NFE2 increased in DN patients compared with normal controls. j The mRNA expression of TGIF1 increased in DN patients compared with normal controls.
with normal controls. k The mRNA expression of TGIF1 increased in DN patients compared with normal controls. NC, normal control. DN, diabetic nephropathy. P<0.05 was considered statistically significant.* P<0.05,** P<0.01.

Figure 5

Correlation between mRNA expression of transcription factor genes in renal tubulointerstium and GFR in DN patients. a The mRNA expression of AP1 negatively correlated with GFR (P-value=0.009, r value=...
b The mRNA expression of AP1 negatively correlated with GFR (P-value=-0.010, r value=-0.735). c The mRNA expression of CDC5 negatively correlated with GFR (P-value=0.006, r value=-0.862). d The mRNA expression of CDC5 negatively correlated with GFR (P-value=0.031, r value=-0.678). e The mRNA expression of IRF1 negatively correlated with GFR (P-value=0.002, r value=-0.816). f The mRNA expression of IRF1 negatively correlated with GFR (P-value=0.046, r value=-0.611). g The mRNA expression of POU3F2 negatively correlated with GFR (P-value=0.046, r value=-0.640). h The mRNA expression of SOX9 negatively correlated with GFR (P-value=0.009, r value=-0.743). i The mRNA expression of BACH1 negatively correlated with GFR (P-value=0.013, r value=-0.819). j The mRNA expression of SOX5 negatively correlated with GFR (P-value=0.001, r value=-0.920). k The mRNA expression of SOX5 negatively correlated with GFR (P-value=0.044, r value=-0.823). m The mRNA expression of SOX5 negatively correlated with GFR (P-value=0.044, r value=-0.823). n The mRNA expression of TGIF1 negatively correlated with GFR (P-value=0.010, r value=-0.836). o The mRNA expression of FOXJ2 negatively correlated with GFR (P-value=0.047, r value=-0.609). p The mRNA expression of FOXO1 positively correlated with GFR (P-value=0.027, r value=0.691). q The mRNA expression of FOXO4 positively correlated with GFR (P-value=0.007, r value=0.561). P<0.05 was considered statistically significant. GFR, glomerular filtration rate. MDRD, modification of diet in renal disease. CG, Cockcroft Gault.
Correlation between mRNA expression of transcription factor genes in renal tubulointerstium and SCR in DN patients. 

a) The mRNA expression of AP1 positively correlated with SCR (P-value=0.022, r value=0.679). 

b) The mRNA expression of BACH2 positively correlated with SCR (P-value=0.007, r value=0.626). 

c) The mRNA expression of FOXD1 positively correlated with SCR (P-value=0.021, r value=0.555). 

d) The mRNA expression of FOXJ2 positively correlated with SCR (P-value=0.005, r value=0.653). 

e) The
mRNA expression of FOXO4 negatively correlated with SCR (P-value=0.022, r value= -0.551). f The mRNA expression of RSRFC4 negatively correlated with SCR (P-value=0.016, r value= -0.572). g The mRNA expression of IRF1 positively correlated with SCR (P-value=0.007, r value= 0.630). h The mRNA expression of IRF1 positively correlated with SCR (P-value=0.049, r value= 0.604). i The mRNA expression of S8 negatively correlated with SCR (P-value=0.012, r value= -0.596). P<0.05 was considered statistically significant. SCR, serum creatine.

Figure 7

Correlation between mRNA expression of transcription factor genes in renal tubulointerstium and proteinuria, weight and body mass index in DN patients. a The mRNA expression of CDC5 negatively correlated with proteinuria (P-value=0.007, r value= -0.754). b The mRNA expression of CDC5 negatively correlated with weight (P-value=0.027, r value= -0.920). c The mRNA expression of FOXO4 negatively correlated with weight (P-value=0.045, r value= -0.887). d The mRNA expression of HFH1 positively correlated with body mass index (P-value=0.006, r value= 0.935). P<0.05 was considered statistically significant.