Identification of Key Transcription Factors Associated with Chronic Kidney Disease for Early Clinical Diagnosis

Wei Zhou
Zhejiang University School of Medicine, National Children's Regional Medical Center
Qingqing Jia
Fudan University
Shujun Wu
Clinical Medicine College of Yangzhou University
Xinyu Wang
The First Affiliated Hospital of Nanjing Medical University
Mingzhu Jiang
Clinical Medicine College of Yangzhou University
Xiaowen Yu
Children's Hospital of Nanjing Medical University
Hanyan Meng
Zhejiang University School of Medicine, National Children's Regional Medical Center
Fei Liu
Zhejiang University School of Medicine, National Children's Regional Medical Center
Jianhua Mao
maojh88@zju.edu.cn

Zhejiang University School of Medicine, National Children's Regional Medical Center

Research Article

Keywords: Chronic kidney disease, Transcription factor, Biomarker, Data analysis

Posted Date: May 6th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4287313/v1

Additional Declarations: No competing interests reported.

Abstract

Background: Chronic kidney disease (CKD) is a global health concern with significant implications for public health and mortality rates, projected to become the fifth leading cause of death by 2040. The search for early diagnostic targets for CKD is imperative. In this study, we concentrated on identifying key transcription factors (TFs) for the early diagnosis of CKD and established a regulatory network between these TFs and their corresponding target genes.

Methods: We conducted microarray data analysis and Gene Set Enrichment Analysis (GSEA) to identify differentially expressed genes (DEGs) and the associated pathways in CKD. We further explored the potential regulatory TFs among DEGs using the TRRUST v2 database and validated the TF-target regulatory relationships through correlation analysis and the JASPAR database. The protein expression of the identified TFs in renal tissues was also assessed.

Results: The analysis identified six TFs, namely HNF4A, CEBPA, CREB1, FOS, HIF1A, and SP1, which demonstrated potential as diagnostic biomarkers for CKD. These TFs showed differentially expressed patterns in CKD and were found to have multiple regulatory relationships with DEGs, indicating their crucial role in the disease process. ROC analysis revealed high predictive efficiency for four of these TFs (CREB1, FOS, HIF1A, and SP1), while the combined predictive efficiency of all TFs was exceptionally high.

Conclusion: Our findings highlight the role of transcription factors in the pathophysiological process of CKD and identify several key TFs with potential for clinical translation as early diagnostic biomarkers for the disease. Further validation and exploration are warranted to leverage the potential clinical utility of these TFs in the early diagnosis and prognosis of CKD.

Introduction

Chronic kidney disease (CKD) has emerged as a significant healthcare issue, affecting individuals across all age groups worldwide, which tend to be become the fifth leading cause of death by 2040[1, 2]. Despite the different etiologies of CKD, its prolonged and incurable process ultimately results in end-stage kidney disease (ESKD), which requires renal replacement therapy[3]. CKD is an independent hazard factor for acute kidney injury (AKI), and AKI further aggravates the development and process of CKD[2, 4].

Renal fibrosis is the most important pathological feature of ESKD, which is characterized by glomerulosclerosis, interstitial fibrosis, and immune cell infiltration in the kidney[5-7]. The renal fibrosis is accompanied by pathological accumulation of extracellular matrix (ECM) proteins, including collagens and fibronectin[8]. However, excessive deposition of ECM in the kidney usually means the late irreversible stage of ESKD. Therefore, the search for early diagnostic targets for CKD is urgently needed.

Transcriptional factors (TFs) play a crucial role in recognizing unique DNA sequences to regulate chromatin structure and gene transcription, thereby establishing a complex system essential for controlling the expression of the kidney genome[9]. Recent studies have indicated that dysregulation of

human TFs is closely related to the onset and progression of CKD[10, 11]. It is well known that each TF, functioning as an upstream regulatory signal, can regulate multiple downstream target genes, which affects different biological process and signaling pathways. Therefore, the changes of the expression of TFs deserve attention that can be recognized as an early signal for CKD. Previous studies have identified some TFs mediating the pathogenesis of CKD, including RUNX2, HIF1A, KLF4, P65, IRF1, SMADs and so on[12-18]. Thus, some key TFs closely associated with CKD have the potential to be the biomarkers for early diagnosis.

In this research, we reanalyzed a classic microarray dataset from Nephroseq database (https://www.nephroseq.org/) to achieve TFs for early diagnosis of CKD[19]. In addition, we established the regulatory network between diagnostic TFs with corresponding target genes. Finally, we evaluated the diagnostic efficiency of these TFs.

Materials And Methods

Microarray Data Analysis

The microarray assay consisting of 8 control kidney samples and 53 CKD kidney samples (GSE66494) was reanalyzed using the online tool “GEO2R”. The operation procedure was followed by the official instructions, as options were set to default settings. The differential expression genes (DEGs) were characterized by “adjust P value < 0.05”.

Gene Set Enrichment Analysis (GSEA) Analysis

The gene symbols and corresponding $\log_2$FC values from GSE66494 were used for GSEA analysis based on the online tool (https://www.omicstudio.cn/tool/)[20].

The exploration of differentially expressed transcription factors (TFs)

The top 500 DEGs of CKD patients were entered into TRRUST v2 (https://www.grnpedia.org/trrust/)[21], which were used for getting possible regulatory TFs in the next step. Among these TFs, some were also DEGs, which can be recognized by the venny tool as we described before[22]. In addition, the TF-target regulatory relationships were further validated by correlation analysis and JASPAR database[23]. The correlation analysis was conducted by using two different online tools from omicstudio (https://www.omicstudio.cn/tool/59) and hiplot (https://hiplot.com.cn/cloud-tool/drawing-tool/detail/646). The intersecting results for the above linear correlation analysis were employed for the next linear correlation analysis.

Statistical analysis

The transcriptome expression value of selected genes was used for the linear correlation analysis based on the Graphpad software. The receiver operator characteristic (ROC) was performed by R software.

Results

Analysis Process

The flowchart was shown in Figure 1. Firstly, DEGs were identified from microarray data from the GEO database. Then, GSEA analysis was conducted. Subsequently, the regulatory TFs of the top 500 DEGs was predicted by TRRUST database. Among these TFs, some were differentially expressed TFs, which regulation of corresponding target genes was verified via linear correlation analysis and JASPAR database. Lastly, the diagnostic efficiency of the selected TFs was evaluated ROC curve.

Identification of DEGs Between Normal and CKD Samples and KEGG Analysis

We selected GSE66494 dataset to conduct deep data mining, because many DEGs from the dataset had been proved to be important pathogenic genes related to CKD, such as PXR, ROCK2 as our research team and other researchers identified before[24-26]. Gene expression levels were compared between the CKD and control samples, the results from GEO2R were shown in Figure 2A-D. Considering the continuous updates of the KEGG database, we conducted functional analysis based on the newest background genesets using GSEA method. A series of up-regulated and down-regulated pathways were shown in Figure 2E, which can be summarized as four categories, including down-regulated energy metabolism, up-regulated proinflammatory and immune infiltration, profibrotic and proliferative pathways and renal dysfunction. As is widely known, renal interstitial fibrosis is the most significant pathological manifestation in CKD patients at the end-stage[27]. Activated TGFβ/smads signaling and wnt/β-catenin signaling are canonical pathways associated with renal fibrosis[12, 28, 29]. Therefore, the dataset met the standard for further analysis.

Screening and identifying key TFs among DEGs

As the microarray dataset reflects changes in transcription levels, we looked up the upstream regulatory TFs for these mRNA changes. TRRUST2.0 is a manually curated database containing 8,444 TF-target regulatory relationships of 800 human TFs[21]. All of the regulatory network were from recorded literature, which meant more solid than bioinformatics prediction and high-throughput sequencing data. Then, 38 probable regulatory TFs were achieved from TRRUST meeting the criterion “$p < 0.05$” (Supplementary Figure 1). Among these TFs, 20 were differentially expressed in CKD group (Figure 3A). Furthermore, the above TFs with less than 3 target genes were excluded, only 13 TFs were used for the next validation (Figure 3B).

The expression of screening TFs and the linear correlation between TFs and matching targets

Among these 13 TFs, heatmap showed that 7 were significantly up-regulated, while the others were significantly down-regulated (Figure 4A). Each TF had its own target gene, which was achieved from TRRUST database. Furthermore, we adopted two ways to conduct the correlation between TFs and corresponding target genes (Figure 4B & C). Each transcription factor (TF) directly interacts with the

promoter region of one or multiple target genes, thereby influencing the expression of mRNA in those genes. This suggests a potential linear correlation between the TF and its corresponding target genes. We found that many TFs had positively or negatively related target genes (Figure 4D-R) containing HNF4A, CEBPA, CREB1, FOS, HIF1A and SP1. Among these TFs, HIF1A, SP1 and CREB1 were up-regulated in human CKD kidney tissue, while the others were down-regulated.

The potential binding sites of key TFs

JASPAR database provided a potential binding sites between TFs and target genes. According to the previous analysis, we focused on the specific TFs and their corresponding binding sites. From the Figure 5A, SP1 had 6 binding sites with the promoter region of FOS, 5 binding sites with COL11A2, 1 binding site with PRSS50, 2 binding sites with NR4A1, 1 binding site with CHI3L1, 3 binding sites with FBLN1,

2 binding sites with EPOR and 5 binding sites with LCAT. CREB1 had only 1 binding site with FOXA1. HNF4A had 2 binding sites with the promoter region of ABCC6, 1 with SLC22A6, 1 with MMP7. CEBPA had 2 binding sites with PCK2. FOS had 1 binding site with FOXA1. HIF1A had 5 binding sites with ASS1. The high-frequency binding base sequences corresponding to each transcription factor was shown in Figure 5B-K.

The protein expression of hub TFs in renal tissues

We observed hub TFs in the normal kidney tissue based on the Human Protein Atlas Database. All of the selected TFs were expressed at a protein level in the normal kidney. However, they distributed in the different area of the kidney. HNF4A existed mainly in the nuclear area of renal tubules (Figure 6A). CEBPA was expressed in both nucleus and cytoplasm of renal tubules (Figure 6B). CREB1 distributed in the nucleus of both glomerulus and renal tubules (Figure 6C). Interestingly, FOS was almost expressed in the whole kidney (Figure 6D). On the contrary, HIF1A was expressed at a low protein level in the glomerulus (Figure 6E). SP1 mainly distributed in the nucleus of glomerulus (Figure 6F).

Predictive efficiency of core TFs in CKD patients

ROC analysis was employed to evaluate the predictive efficiency of the six genes mentioned. Four of their AUC values were greater 0.9, including CREB1, FOS, HIF1A and SP1 (Figure 7C-F), while the AUC value of HNF4A and CEBPA was just 0.658 and 0.5802 (Figure 7A & B), and the summary result was shown in Figure 7G. The combined predictive efficiency of the above TFs was shown in Figure 7H, which AUC is equal to 1.

Discussion

CKD is a kind of progressive disease whose clinical assessment is usually based on laboratory examination and pathological examination. Over the past years, proteinuria, eGFR, Scr, BUN and some other indexes have reflected the function of kidney. However, when these indexes show significant

pathological changes, the kidney has progressed to the end stage. Therefore, a strategy for early diagnosis of CKD needs to be developed urgently.

With the rapid development of the technology of multiple omics, molecular diagnosis has been applied to clinical practice. Many researchers have found some key biomarkers highly related to CKD. Zhang et al identifies that XDH is positively correlative with the damage of kidney[30]. MMP2 and MMP9, which are closely related to glomerulonephritis as we have identified before, are also important biomarkers for CKD[22, 31].

Different from the former diagnostic models, in this study, we focused on searching some diagnostic biomarkers for CKD from transcription factors because of their upstream regulatory effects. We used TRRUST database to predict the regulatory TFs for DEGs and got a series TFs, among which some were also differentially expressed genes. Further analysis revealed that 6 of them had multiple regulatory relationships with DEGs including HNF4A, CEBPA, CREB1, FOS, HIF1A and SP1, which were supported by linear correlation analysis and evidences from JASPAR database. HNF4A, namely hepatocyte nuclear factor 4-alpha, is associated with the differentiation of proximal tubules, whose deletion or mutation will result in Fanconi renotubular syndrome[32-34]. HNF4A is also down-regulated in the unilateral ureteral obstruction model, which is a classic model using for researching CKD[35]. Additionally, liver-specific HNF4A-deficient mice progress to liver fibrosis, which can be rescued by applicability of HNF4A mRNA therapeutics[36, 37]. CEBPA, namely CCAAT/enhancer-binding protein alpha, was down-regulated in our analysis. Interesting, conversely, it is reported to be up-regulated in the UUO kidney[38]. However, another research supports that the expression of CEBPA is repressed by TGFβ, which contributed to the initiation of endothelial-to-mesenchymal transition in the endothelium[39]. Cyclic AMP-responsive element-binding protein 1 (CREB1) can be induced by TGFβ, which can result in matrix metalloproteinase and fibronectin accumulation[40-42]. FOS, also called protein c-FOS, is increased in the UUO model for 12 days[43]. FOS is a vital component of TGFβ/SMADs signaling, which mediates cell proliferation[44, 45]. However, our analysis shows that FOS is up-regulated in human renal tissues. More samples may be needed to include in order to test whether FOS can be a biomarker for CKD at the early stage. HIF1A, also named hypoxia-inducible factor 1-alpha, can promote renal fibrosis in some kidney diseases[46, 47]. Therapeutic strategy targeting HIF-1α can protect the kidney from AKI to CKD progression[48]. SP1 is increased in glomerular or proximal tubular tissues in glomerulonephritis and obstructive nephropathy, whose expression is positively correlative to p-Smad2/3[49]. In addition, the expression of SP1 is positively relative to collagen I[50].

Furthermore, the above TFs separately have their own target genes. These target genes, including FOS, COL11A2, PRSS50, NR4A1, CHI3L1, FBLN1, FOXA1, EPOR, LCAT, ABCC6, SLC22A6, MMP7, PCK2 and ASS1, are also differentially expressed in human CKD kidney. Among these genes, some are reported to be related to renal fibrosis, such as MMP7, NR4A1, LCAT and so on[51-53]. At the same time, ROC analysis reveals that transcription factor CREB1, FOS, HIF1A and SP1 have higher diagnostic efficacy than HNF4A and CEBPA.

In summary, all of the six TFs calculated by our team have the potential to diagnose and evaluate prognosis for CKD at the early stage. Based on the above analysis, SP1, CREB1 and HIF1A seem to have more strong evidence for further clinical transformation.

However, the study has some limitations. First, as the TF-target regulatory network is based on the published literature, quite a few TFs are not reported in TRRUST, there may be other TFs which is able to be used for early diagnosis of CKD. Second, our findings just come from numerous data analysis and documentary evidence, some experiments ought to be performed for validation. For instance, the binding affinity between TFs and corresponding DNA sequence should depend on electrophoretic mobility shift assay or chromatin immunoprecipitation. And also, dual-luciferase reporter assay is necessary for analyzing the gene expression regulated by TFs in the kidney cell model. Third, CKD patients have different primary diseases, this research do not divide CKD patients into diverse subgroups. We will overcome these shortcomings and make a deep exploration in the next work.

Conclusion

In summary, transcription factors play an important role in the pathophysiological process of chronic kidney disease. We have identified 6 TFs closely associated with CKD based on a comprehensive method containing HNF4A, CEBPA, CREB1, FOS, HIF1A and SP1. According to the literature evidence and deep data analysis, we conclude that SP1, CREB1 and HIF1A may have greater potential for clinical translation.

Declarations

Acknowledgements

Not applicable.

Contributions

Jianhua Mao designed the research. Wei Zhou, Qingqing Jia and Shujun Wu performed bioinformatic analysis and wrote the manuscript. Xinyu Wang, Mingzhu Jiang, Hanyan Meng and Fei Liu helped to solve the clinical problems and contributed to data collection. Xiaowen Yu gave some suggestions on the limitations of the current research. All authors read and approved the final manuscript.

Fundings

This work was supported by grants from the Fundamental Research Funds for the Central Universities (226-2023-00056), the Zhejiang Xinmiao Talents Program (2023R401208), the Scientific Research Fund of Zhejiang University (XY2023011) and Key Research and Development Plan of Zhejiang Province (2021C03079), National key research and development program (2021YFC2702002, 2022YFC2705103).

Data availability

The datasets analysed during the current study are available in the GEO Database (GSE66494, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66494). In addition, the datasets analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

All the authors declared that they have no competing interests.

References

1. Webster AC, Nagler EV, Morton RL, et al. Chronic Kidney Disease. Lancet 2017;389:1238-52.
2. Jiang M, Bai M, Lei J, et al. Mitochondrial dysfunction and the AKI-to-CKD transition. Am J Physiol Renal Physiol 2020;319:F1105-F16.
3. Chen TK, Knicely DH, Grams ME. Chronic Kidney Disease Diagnosis and Management: A Review. JAMA 2019;322:1294-304.
4. Chawla LS, Eggers PW, Star RA, et al. Acute kidney injury and chronic kidney disease as interconnected syndromes. N Engl J Med 2014;371:58-66.
5. Tepus M, Tonoli E, Verderio EAM. Molecular profiling of urinary extracellular vesicles in chronic kidney disease and renal fibrosis. Front Pharmacol 2022;13:1041327.
6. Humphreys BD. Mechanisms of Renal Fibrosis. Annu Rev Physiol 2018;80:309-26.
7. Tang PM, Nikolic-Paterson DJ, Lan HY. Macrophages: versatile players in renal inflammation and fibrosis. Nat Rev Nephrol 2019;15:144-58.
8. Li L, Fu H, Liu Y. The fibrogenic niche in kidney fibrosis: components and mechanisms. Nat Rev Nephrol 2022;18:545-57.
9. Lambert SA, Jolma A, Campitelli LF, et al. The Human Transcription Factors. Cell 2018;172:650-65.
10. Ruiz-Ortega M, Rayego-Mateos S, Lamas S, et al. Targeting the progression of chronic kidney disease. Nat Rev Nephrol 2020;16:269-88.
11. Hishikawa A, Hayashi K, Itoh H. Transcription Factors as Therapeutic Targets in Chronic Kidney Disease. Molecules 2018;23.
12. Ma TT, Meng XM. TGF-beta/Smad and Renal Fibrosis. Adv Exp Med Biol 2019;1165:347-64.
13. Li W, Feng W, Su X, et al. SIRT6 protects vascular smooth muscle cells from osteogenic transdifferentiation via Runx2 in chronic kidney disease. J Clin Invest 2022;132.

14. Yang Y, Yu X, Zhang Y, et al. Hypoxia-inducible factor prolyl hydroxylase inhibitor roxadustat (FG-4592) protects against cisplatin-induced acute kidney injury. Clin Sci (Lond) 2018;132:825-38.
15. Wen Y, Lu X, Ren J, et al. KLF4 in Macrophages Attenuates TNFalpha-Mediated Kidney Injury and Fibrosis. J Am Soc Nephrol 2019;30:1925-38.
16. Tang TT, Wang B, Li ZL, et al. Kim-1 Targeted Extracellular Vesicles: A New Therapeutic Platform for RNAi to Treat AKI. J Am Soc Nephrol 2021;32:2467-83.
17. Huang Y, Wang S, Zhou J, et al. IRF1-mediated downregulation of PGC1alpha contributes to cardiorenal syndrome type 4. Nat Commun 2020;11:4664.
18. Wang J, Ge S, Wang Y, et al. Puerarin Alleviates UUO-Induced Inflammation and Fibrosis by Regulating the NF-kappaB P65/STAT3 and TGFbeta1/Smads Signaling Pathways. Drug Des Devel Ther 2021;15:3697-708.
19. Nakagawa S, Nishihara K, Miyata H, et al. Molecular Markers of Tubulointerstitial Fibrosis and Tubular Cell Damage in Patients with Chronic Kidney Disease. PLoS One 2015;10:e0136994.
20. Lyu F, Han F, Ge C, et al. OmicStudio: A composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. iMeta 2023;2:e85.
21. Han H, Cho JW, Lee S, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res 2018;46:D380-D86.
22. Zhou W, Sha Y, Zeng J, et al. Computational Systems Pharmacology, Molecular Docking and Experiments Reveal the Protective Mechanism of Li-Da-Qian Mixture in the Treatment of Glomerulonephritis. J Inflamm Res 2021;14:6939-58.
23. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res 2022;50:D165-D73.
24. Yu X, Xu M, Meng X, et al. Nuclear receptor PXR targets AKR1B7 to protect mitochondrial metabolism and renal function in AKI. Sci Transl Med 2020;12.
25. Ming WH, Luan ZL, Yao Y, et al. Pregnane X receptor activation alleviates renal fibrosis in mice via interacting with p53 and inhibiting the Wnt7a/beta-catenin signaling. Acta Pharmacol Sin 2023.
26. You R, Zhou W, Li Y, et al. Inhibition of ROCK2 alleviates renal fibrosis and the metabolic disorders in the proximal tubular epithelial cells. Clin Sci (Lond) 2020;134:1357-76.
27. Romagnani P, Remuzzi G, Glassock R, et al. Chronic kidney disease. Nat Rev Dis Primers 2017;3:17088.
28. Edeling M, Ragi G, Huang S, et al. Developmental signalling pathways in renal fibrosis: the roles of Notch, Wnt and Hedgehog. Nat Rev Nephrol 2016;12:426-39.
29. Schunk SJ, Floege J, Fliser D, et al. WNT-beta-catenin signalling - a versatile player in kidney injury and repair. Nat Rev Nephrol 2021;17:172-84.
30. Zhang Y, Ding X, Guo L, et al. Comprehensive Analysis of the Relationship between Xanthine Oxidoreductase Activity and Chronic Kidney Disease. iScience.

31. Saucedo AL, Perales-Quintana MM, Paniagua-Vega D, et al. Chronic Kidney Disease and the Search for New Biomarkers for Early Diagnosis. Curr Med Chem 2018;25:3719-47.
32. Muto Y, Wilson PC, Ledru N, et al. Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. Nat Commun 2021;12:2190.
33. Marable SS, Chung E, Adam M, et al. Hnf4a deletion in the mouse kidney phenocopies Fanconi renotubular syndrome. JCI Insight 2018;3.
34. Hamilton AJ, Bingham C, McDonald TJ, et al. The HNF4A R76W mutation causes atypical dominant Fanconi syndrome in addition to a beta cell phenotype. J Med Genet 2014;51:165-9.
35. Wu B, Brooks JD. Gene expression changes induced by unilateral ureteral obstruction in mice. J Urol 2012;188:1033-41.
36. Kasano-Camones CI, Takizawa M, Ohshima N, et al. PPARalpha activation partially drives NAFLD development in liver-specific Hnf4a-null mice. J Biochem 2023;173:393-411.
37. Yang T, Poenisch M, Khanal R, et al. Therapeutic HNF4A mRNA attenuates liver fibrosis in a preclinical model. J Hepatol 2021;75:1420-33.
38. Fujiu K, Manabe I, Nagai R. Renal collecting duct epithelial cells regulate inflammation in tubulointerstitial damage in mice. J Clin Invest 2011;121:3425-41.
39. Gao R, Wu Y, Yang Q, et al. The Interaction of Apelin and FGFR1 Ameliorated the Kidney Fibrosis through Suppression of TGFbeta-Induced Endothelial-to-Mesenchymal Transition. Oxid Med Cell Longev 2023;2023:5012474.
40. Liu G, Ding W, Neiman J, et al. Requirement of Smad3 and CREB-1 in mediating transforming growth factor-beta (TGF beta) induction of TGF beta 3 secretion. J Biol Chem 2006;281:29479-90.
41. Yano N, Suzuki D, Endoh M, et al. In vitro silencing of the insulin receptor attenuates cellular accumulation of fibronectin in renal mesangial cells. Cell Commun Signal 2012;10:29.
42. Shan Q, Zheng G, Zhu A, et al. Epigenetic modification of miR-10a regulates renal damage by targeting CREB1 in type 2 diabetes mellitus. Toxicol Appl Pharmacol 2016;306:134-43.
43. Silverstein DM, Travis BR, Thornhill BA, et al. Altered expression of immune modulator and structural genes in neonatal unilateral ureteral obstruction. Kidney Int 2003;64:25-35.
44. Zhang Y, Feng XH, Derynck R. Smad3 and Smad4 cooperate with c-Jun/c-Fos to mediate TGF-beta-induced transcription. Nature 1998;394:909-13.
45. Singhal PC, Sharma P, Sanwal V, et al. Morphine modulates proliferation of kidney fibroblasts. Kidney Int 1998;53:350-7.
46. Zhao H, Han Y, Jiang N, et al. Effects of HIF-1alpha on renal fibrosis in cisplatin-induced chronic kidney disease. Clin Sci (Lond) 2021;135:1273-88.
47. Li X, Yang S, Yan M, et al. Interstitial HIF1A induces an estimated glomerular filtration rate decline through potentiating renal fibrosis in diabetic nephropathy. Life Sci 2020;241:117109.
48. Li ZL, Wang B, Lv LL, et al. FIH-1-modulated HIF-1alpha C-TAD promotes acute kidney injury to chronic kidney disease progression via regulating KLF5 signaling. Acta Pharmacol Sin

2021;42:2106-19.

49. Kassimatis TI, Nomikos A, Giannopoulou I, et al. Transcription factor Sp1 expression is upregulated in human glomerulonephritis: correlation with pSmad2/3 and p300 expression and renal injury. Ren Fail 2010;32:243-53.
50. Jiang L, Zhou Y, Xiong M, et al. Sp1 mediates microRNA-29c-regulated type I collagen production in renal tubular epithelial cells. Exp Cell Res 2013;319:2254-65.
51. Hirohama D, Abedini A, Moon S, et al. Unbiased Human Kidney Tissue Proteomics Identifies Matrix Metalloproteinase 7 as a Kidney Disease Biomarker. J Am Soc Nephrol 2023;34:1279-91.
52. Tveitaras MK, Skogstrand T, Leh S, et al. Matrix Metalloproteinase-2 Knockout and Heterozygote Mice Are Protected from Hydronephrosis and Kidney Fibrosis after Unilateral Ureteral Obstruction. PLoS One 2015;10:e0143390.
53. Tao Y, Tang C, Wei J, et al. Nr4a1 promotes renal interstitial fibrosis by regulating the p38 MAPK phosphorylation. Mol Med 2023;29:63.

Figures

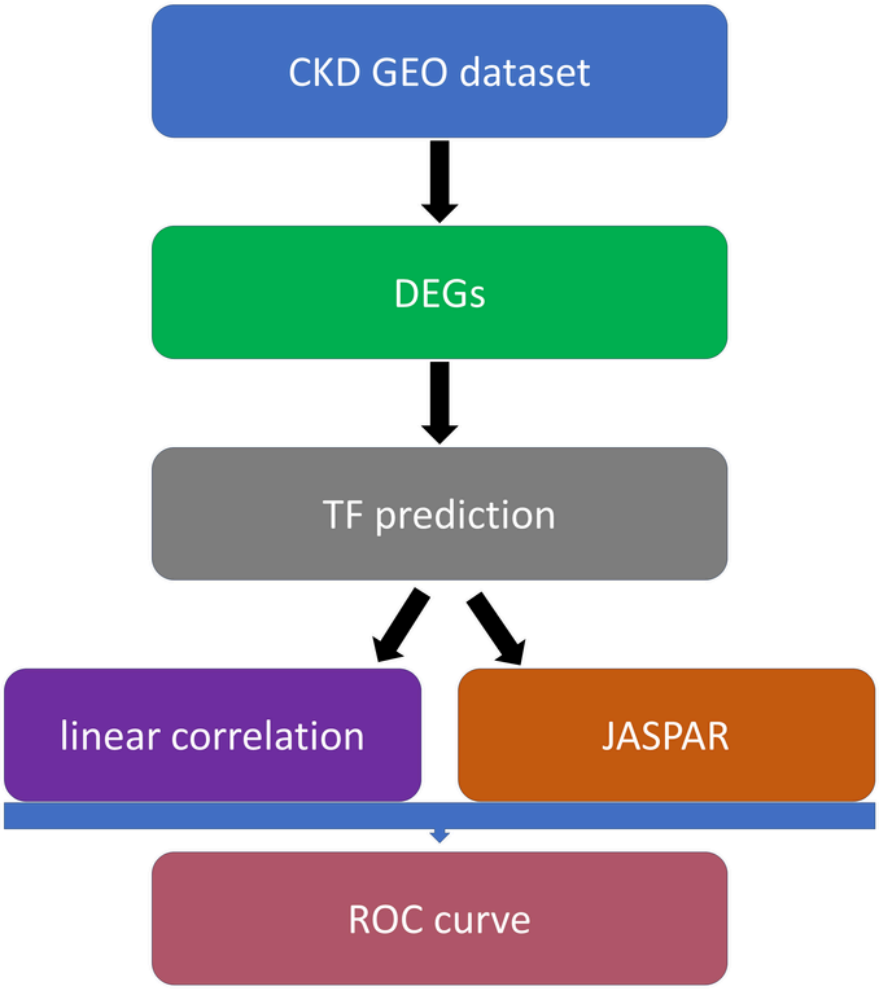

Figure 1

The flowchart of the whole strategy.

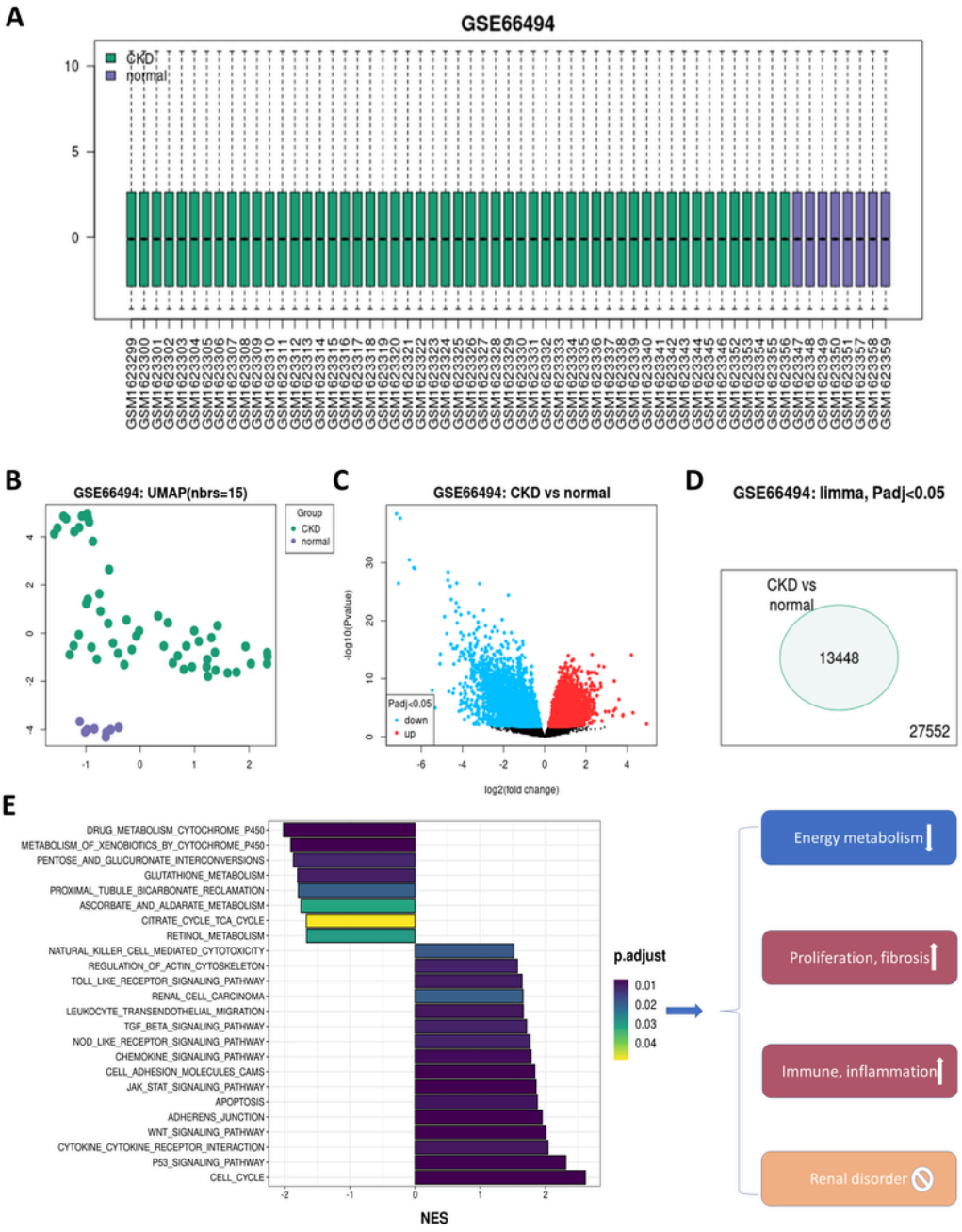

Figure 2

Differential gene expression in the GSE66494 dataset and the KEGG pathways involved. A. Boxplot of dataset. B. UMAP plot of dataset. C. Volcano plot of dataset. D. The amounts of differential genes of dataset. E. The up-regulated and down-regulated pathways of dataset based on GSEA analysis.

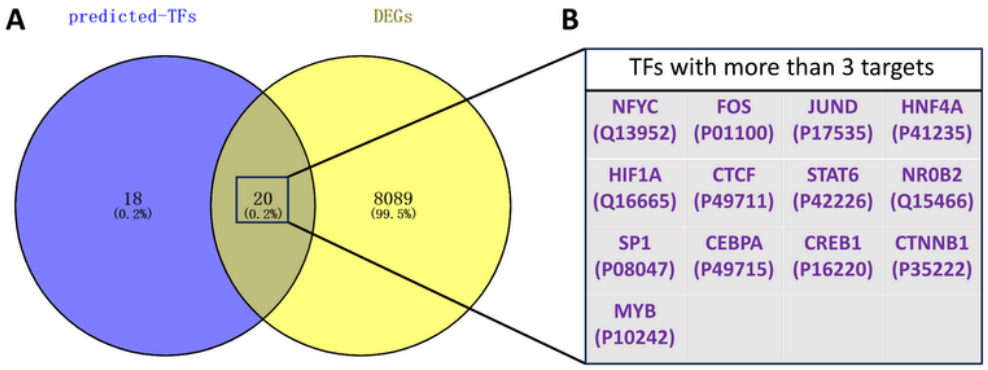

Figure 3

The differentially expressed TFs in CKD kidney. A. The venny diagram between predicted TFs and DEGs in the CKD kidney. B. Differentially expressed TFs with more than 3 target genes and their corresponding protein coding in the UniProt database.

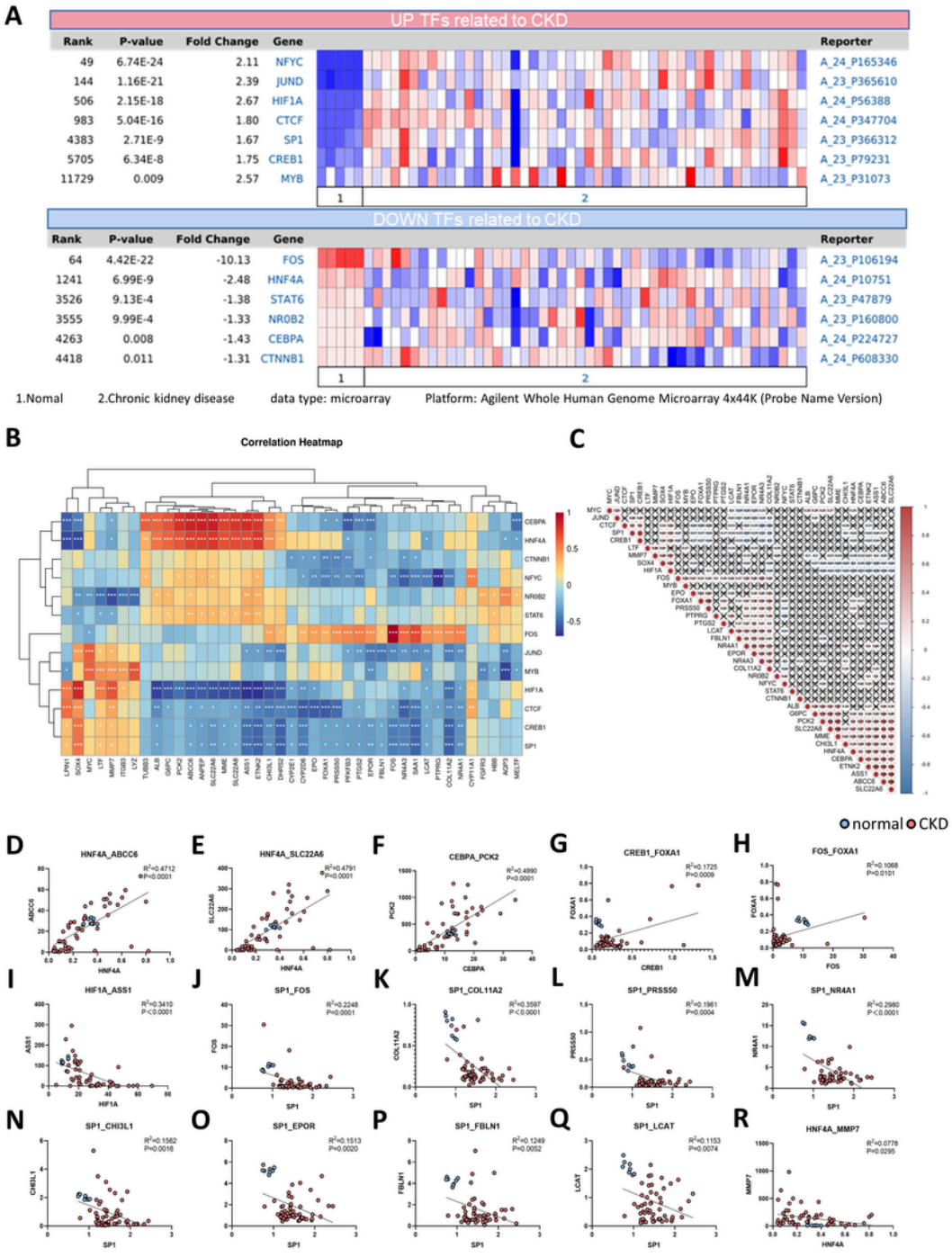

Figure 4

The up-regulated and down-regulated TFs in CKD and their relative linear analysis with corresponding target genes. A. Heatmap based on Nephroseq database reveals the up-regulated and down-regulated TFs in CKD. B-C. The linear analysis between TFs and their corresponding targets via two different methods. D-R. The linear analysis between each TF and each target gene.

A

Transcription factor	Target gene	Score	Relative score	Sequence ID	Start	End	Strand	Predicted sequence
Total 18 putative site(s) on the positive strand were predicted with relative profile score threshold 95%								
MA0079.2.SP1	FOS	14.453586	0.995298923	NC_000014.9:75276828-75278828	1582	1591	+	CCCCTCCCCC
MA0079.3.SP1	FOS	15.43146	0.975286119	NC_000014.9:75276828-75278828	1581	1591	+	TCCCCTCCCCC
MA0079.2.SP1	FOS	13.213753	0.964074552	NC_000014.9:75276828-75278828	159	168	+	CCCCGCCTCC
MA0079.2.SP1	FOS	13.213753	0.964074552	NC_000014.9:75276828-75278828	1601	1610	+	CCCCGCCTCC
MA0079.3.SP1	FOS	14.486494	0.963397361	NC_000014.9:75276828-75278828	158	168	+	TCCCCGCCTCC
MA0079.3.SP1	FOS	14.459636	0.963059453	NC_000014.9:75276828-75278828	1600	1610	+	CCCCCGCCTCC
MA0079.1.SP1	COL11A2	11.045893	0.972161372	NC_000006.12:c33162694-33160694	1783	1792	+	GAGGCGGGGT
MA0079.1.SP1	COL11A2	11.045893	0.972161372	NC_000006.12:c33162694-33160694	1801	1810	+	GGGGCTGGGT
MA0079.3.SP1	COL11A2	15.093589	0.971035308	NC_000006.12:c33162694-33160694	869	879	+	CCCCCGCCCCT
MA0079.3.SP1	COL11A2	14.962341	0.969384064	NC_000006.12:c33162694-33160694	874	884	+	GCCCCTCCTCC
MA0079.2.SP1	COL11A2	13.027086	0.959373479	NC_000006.12:c33162694-33160694	875	884	+	CCCCTCCTCC
MA0079.3.SP1	PRSS50	13.716568	0.95371081	NC_000003.12:c46712117-46710117	1392	1402	+	TTCCCTCCCCC
MA0079.2.SP1	NR4A1	14.453586	0.995298923	NC_000012.12:52020832-52022832	1093	1102	+	CCCCTCCCCC
MA0079.3.SP1	NR4A1	15.43146	0.975286119	NC_000012.12:52020832-52022832	1092	1102	+	TCCCCTCCCCC
MA0079.5.SP1	CHI3L1	14.750445	0.975201446	NC_000001.11:c203178931-203176931	1059	1067	+	GGGGAGGGG
MA0079.2.SP1	FBLN1	13.213753	0.964074552	NC_000022.11:45500883-45502883	1939	1948	+	CCCCGCCTCC
MA0079.3.SP1	FBLN1	15.706597	0.978747657	NC_000022.11:45500883-45502883	1938	1948	+	GCCCCGCCTCC
MA0079.5.SP1	FBLN1	13.402603	0.951780456	NC_000022.11:45500883-45502883	1732	1740	+	GGGGAGGAG
Total 8 putative site(s) on the positive strand were predicted with relative profile score threshold 90%								
MA0018.2.CREB1	FOXA1	9.00289	0.906041171	NC_000014.9:c37589552-37587552	349	356	+	TGACATCA
MA0079.1.SP1	EPOR	9.171796	0.906789836	NC_000019.10:c11377207-11375207	1670	1679	+	GAGGGTGGGT
MA0079.2.SP1	EPOR	10.835842	0.904188444	NC_000019.10:c11377207-11375207	1760	1769	+	CCCTGCCTCC
MA0079.5.SP1	LCAT	12.212311	0.931097155	NC_000016.10:c67939750-67937750	129	137	+	GGGGCGGTG
MA0079.5.SP1	LCAT	12.212311	0.931097155	NC_000016.10:c67939750-67937750	140	148	+	GGGGCGGTG
MA0079.3.SP1	LCAT	11.298021	0.923282724	NC_000016.10:c67939750-67937750	99	109	+	GCTCCACCCAC
MA0079.1.SP1	LCAT	9.226104	0.908684185	NC_000016.10:c67939750-67937750	140	149	+	GGGGCGGTGA
MA0079.1.SP1	LCAT	9.026768	0.90173102	NC_000016.10:c67939750-67937750	138	147	+	GGGGGGCGGT
Total 7 putative site(s) on the positive strand were predicted with relative profile score threshold 85%								
MA0114.2.HNF4A	ABCC6	8.896688	0.853294207	NC_000016.10:c16149565-16147565	753	767	+	GAGGACTTGGTTCTT
MA0114.4.HNF4A	ABCC6	11.323141	0.872754521	NC_000016.10:c16149565-16147565	1104	1116	+	ACCAAATTCCAGC
MA0114.2.HNF4A	SLC22A6	10.119499	0.869570966	NC_000011.10:c62976597-62974597	140	154	+	CTTGCCTTTGCCCCT
MA0114.4.HNF4A	MMP7	11.524467	0.876765433	NC_000011.10:c102520508-102518508	228	240	+	TGGAAAGTCCAAT
MA0102.3.CEBPA	PCK2	7.1550646	0.88370774	NC_000014.9:24092171-24094171	1191	1201	+	GTTTCCTCATT
MA0102.3.CEBPA	PCK2	6.1691127	0.871895949	NC_000014.9:24092171-24094171	1865	1875	+	GTTACATCATG
MA0476.1.FOS	FOXA1	5.975023	0.868628017	NC_000014.9:c37589552-37587552	3	13	+	GGAGACTCATG
Total 5 putative site(s) on the positive strand were predicted with relative profile score threshold 80%								
MA1106.1.HIF1A	ASS1	6.545462	0.846239267	NC_000009.12:130442707-130444707	1003	1012	+	CCACGTGGCT
MA1106.1.HIF1A	ASS1	5.385243	0.818920892	NC_000009.12:130442707-130444707	1367	1376	+	GGACTTGCTC
MA1106.1.HIF1A	ASS1	4.8551283	0.80643888	NC_000009.12:130442707-130444707	1604	1613	+	GCAGGTGCCA
MA1106.1.HIF1A	ASS1	4.7103024	0.803028827	NC_000009.12:130442707-130444707	32	41	+	CTAGGTGCTG
MA1106.1.HIF1A	ASS1	4.5892386	0.800178276	NC_000009.12:130442707-130444707	18	27	+	CCACCTGCTC

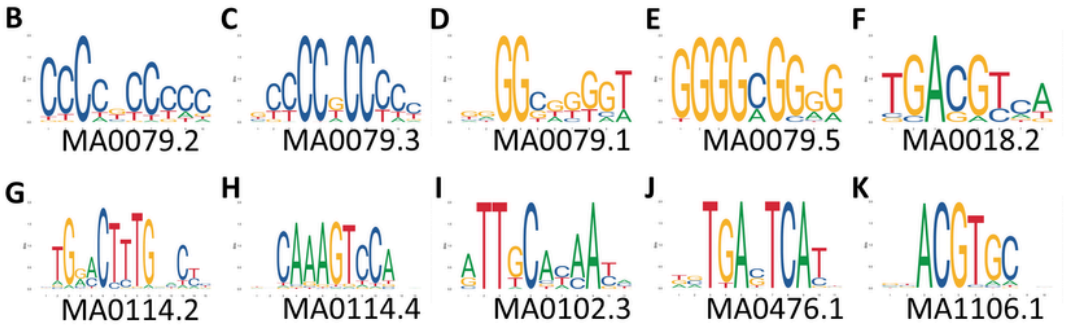

Figure 5

The binding sites between selected TFs and target genes. A. The combing prediction between TFs and target sequences by JASPAR. B-E. The binding sites information for SP1. F. The binding sites information for CREB1. G-H. The binding sites information for HNF4A. I. The binding sites information for CEBPA. J. The binding sites information for FOS. K. The binding sites information for HIF1A. The larger the letter, the greater the frequency of the corresponding base

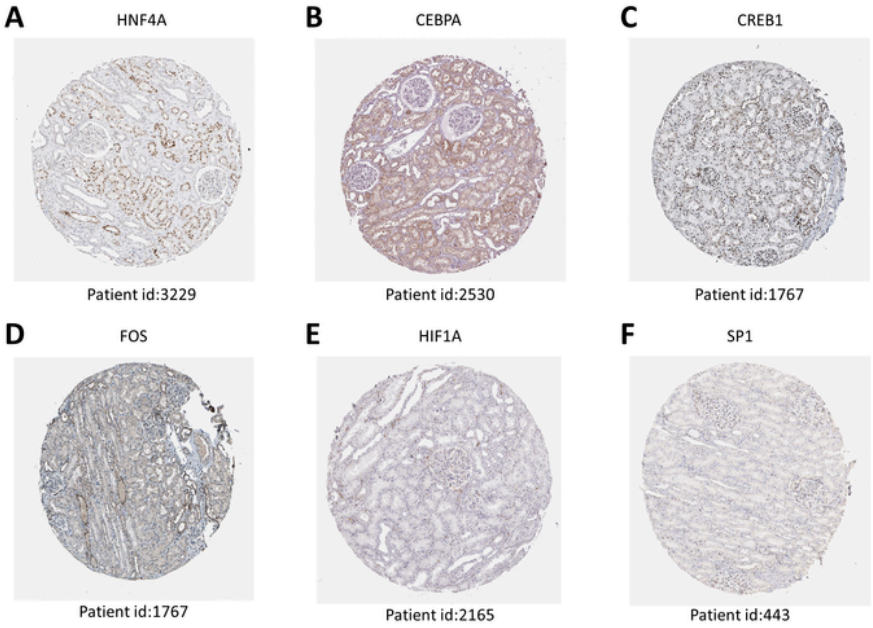

Figure 6

Immunohistochemistry of the hub TFs based on the Human Protein Atlas database. A. The protein expression of HNF4A in the normal kidney tissue. B. The protein expression of CEBPA in the normal kidney tissue. C. The protein expression of CREB1 in the normal kidney tissue. D. The protein expression of FOS in the normal kidney tissue. E. The protein expression of HIF1A in the normal kidney tissue. F. The protein expression of SP1 in the normal kidney tissue.

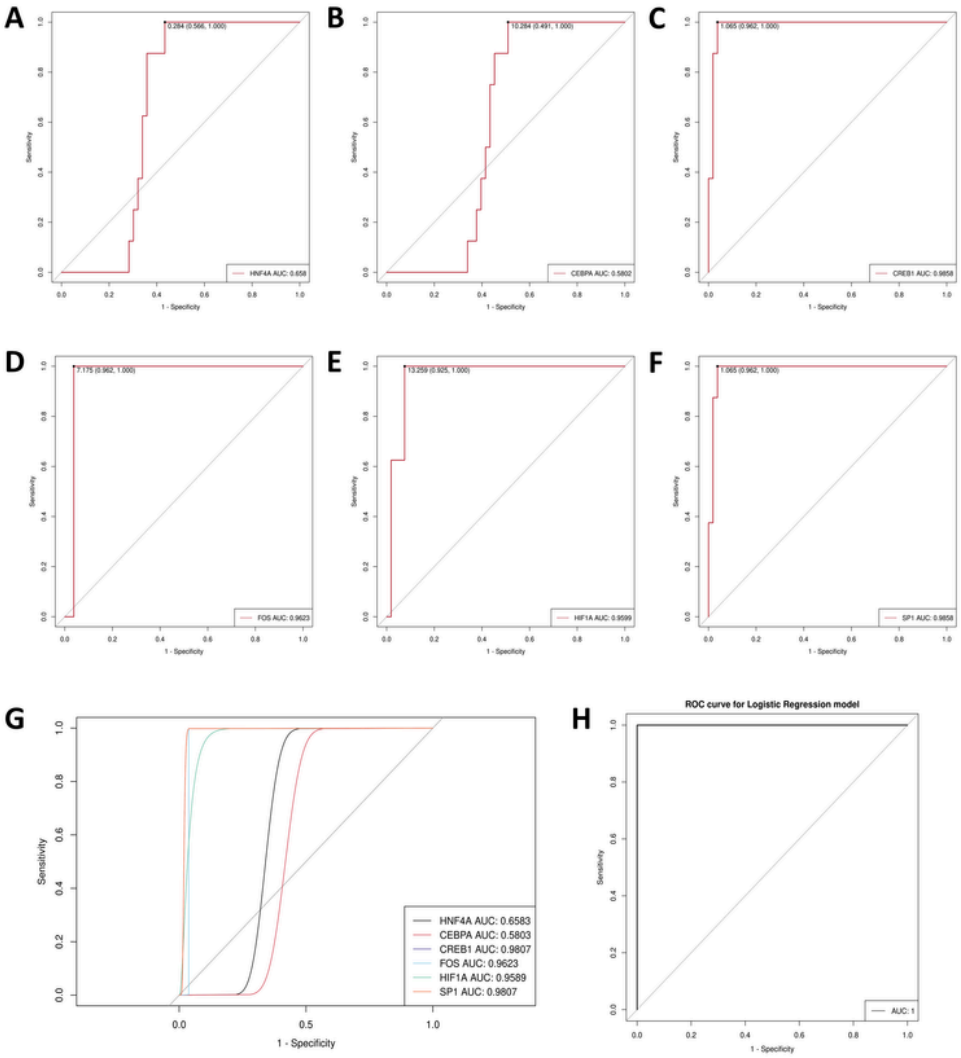

Figure 7

The diagnostic efficacy of 6 key TFs and their combined predictive efficiency in CKD. A. ROC analysis of HNF4A in CKD. B. ROC analysis of CEBPA in CKD. C. ROC analysis of CREB1 in CKD. D. ROC analysis of FOS in CKD. E. ROC analysis of HIF1A in CKD. F. ROC analysis of SP1 in CKD. G. The collected results of the above single result. H. The combined predictive efficiency for the above six TFs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigure1andFigurelegend.docx