Kidney & Blood Pressure Research

Kidney Blood Press Res 2018;43:568-581
DOI: 10.1159/000488830
Published online: 11 April, 2018
Accepted: 28 March, 2018

Original Paper

Integrative Bioinformatics Analysis Provides Insight into the Molecular Mechanisms of Chronic Kidney Disease

Le-Ting Zhou[a] Shen Qiu[a] Lin-Li Lv[a] Zuo-Lin Li[a] Hong Liu[a] Ri-Ning Tang[a] Kun-Ling Ma[a] Bi-Cheng Liu[a]

[a]Institute of Nephrology, Zhong Da Hospital, Southeast University School of Medicine, Nanjing, Jiangsu, China

Key Words
Chronic kidney disease • Bioinformatics • Molecular mechanisms • Microarray • Protein–protein interaction network

Abstract
Background/Aims: Chronic kidney disease (CKD) is a worldwide public health problem. Regardless of the underlying primary disease, CKD tends to progress to end-stage kidney disease, resulting in unsatisfactory and costly treatment. Its common pathogenesis, however, remains unclear. The aim of this study was to provide an unbiased catalog of common gene-expression changes of CKD and reveal the underlying molecular mechanism using an integrative bioinformatics approach. ***Methods:*** We systematically collected over 250 Affymetrix microarray datasets from the glomerular and tubulointerstitial compartments of healthy renal tissues and those with various types of established CKD (diabetic kidney disease, hypertensive nephropathy, and glomerular nephropathy). Then, using stringent bioinformatics analysis, shared differentially expressed genes (DEGs) of CKD were obtained. These shared DEGs were further analyzed by the gene ontology (GO) and pathway enrichment analysis. Finally, the protein-protein interaction networks(PINs) were constructed to further refine our results. ***Results:*** Our analysis identified 176 and 50 shared DEGs in diseased glomeruli and tubules, respectively, including many transcripts that have not been previously reported to be involved in kidney disease. Enrichment analysis also showed that the glomerular and tubulointerstitial compartments underwent a wide range of unique pathological changes during chronic injury. As revealed by the GO enrichment analysis, shared DEGs in glomeruli were significantly enriched in exosomes. By constructing PINs, we identified several hub genes (e.g. OAS1, JUN, and FOS) and clusters that might play key roles in regulating the development of CKD. ***Conclusion:*** Our

Bi-Cheng Liu

Institute of Nephrology, Zhong Da Hospital, Southeast University School of Medicine, Nanjing, Jiangsu (China)
Tel. +86-025-83262422, E-Mail liubc64@163.com

study not only further reveals the unifying molecular mechanism of CKD pathogenesis but also provides a valuable resource of potential biomarkers and therapeutic targets.

Introduction

Chronic kidney disease (CKD) is a worldwide public health problem, affecting 12% of all adults in the United States and 10.8% in China [1, 2]. Regardless of the underlying primary disease, CKD tends to progress to end-stage kidney disease, resulting in unsatisfactory and costly treatment. Therefore, it is important to identify the common key genes and pathways involved in CKD to improve early recognition and prevent disease progression.

Our knowledge of the underlying molecular mechanisms of CKD primarily originates from studies based on chronic kidney injury models and reviews summarizing common pathophysiological features. Although animal models are essential for understanding the pathogenesis of CKD, they can diverge from real patients and are open to misinterpretation [3]. The interpretation of the literature in reviews is sometimes subjective and lacks quantitative metrics, which may induce bias.

In recent years, bioinformatics has emerged as a powerful tool for providing comprehensive insights into the molecular mechanisms of disease and identifying potential biomarkers and therapeutic targets [4-6]. Using comprehensive pathway analysis and subsequent verification, Butz et al. identified three novel factors, including AHR, GRHL2, and KIAA0101, as potential biomarkers of clear cell renal cell carcinoma [5]. By reanalyzing seven samples from Ove26 (diabetic) and FVB (nondiabetic) mice, Wang et al. found several significant nodes and differentially expressed transcriptional factors that may serve as novel targets for diabetic kidney disease (DKD) [6]. However, owing to the tremendous amount of data generated by high-throughput technologies, bioinformatics analyses based on a few samples may be at high risk of false-positive results.

To this end, we applied robust algorithms and performed an integrative analysis of over 250 microarray datasets of healthy renal tissues and those with various types of established CKD, including DKD, hypertensive nephropathy (HN), and glomerulonephritis (GN), to reveal the underlying molecular mechanisms of CKD. As the glomeruli and tubules may undergo different pathological processes during kidney injury [7], the molecular patterns of the glomerular and tubulointerstitial compartments were analyzed separately. To our knowledge, this is the first integrative bioinformatics analysis of this topic.

Materials and Methods

Datasets collection

Gene expression datasets were compiled from the web resource GEO (https://www.ncbi.nlm.nih.gov/geo/) in May 2017. For datasets from normal human glomeruli and tubules, the primary search criteria were "(glomerulus OR glomeruli OR glomerular)" and "(tubulus OR tubules OR tubular)," respectively. Then, we searched for datasets of diseased glomeruli and tubules. For DKD and HN, the primary search criteria were "(diabetic nephropathy OR diabetic kidney disease OR DKD OR DN)" and "(hypertensive nephropathy OR hypertensive kidney disease OR hypertensive renal disease OR HN)", respectively. We chose IgA nephropathy (IgAN), membranous nephropathy (MN), and focal segmental glomerular sclerosis (FSGS) as representative diseases of GN, with the primary search criteria of "(membranous nephropathy OR membranous glomerulonephritis OR MN)", "(IgA nephropathy OR IgAN)", and "(focal segmental glomerular sclerosis OR FSGS)", respectively. To minimize platform variation, the platform filter criterion was set to "Affymetrix U133". The summary and sample description of each series were examined by two investigators, and those related to our research objectives were further included in the analysis.

Microarray datasets preprocessing and differentially expressed gene (DEG) identification

Data preprocessing and DEG identification were performed using Bioconductor in R according to the following steps [8]:

Before included in the analysis, the quality of each dataset was examined using the general quality control (QC) stats in the simpleaffy package and RNA degradation analysis in the affy package, and unqualified datasets were discarded.

For analysis involving different types of microarrays, the 22, 215 Affymetrix identifiers shared by both the Human Genome U133 Plus 2.0 Array and Human Genome U133A Array were extracted. Datasets of the same category were pooled to expand the sample size. Then, the relative log expression graph was used to evaluate the consistency among datasets, and those with significant bias were discarded.

The robust multi-array average method in the affy package was used to preprocess the original data: firstly, the raw intensity values were background corrected, log2 transformed, and quantile normalized. Secondly, the principal component analysis was used to detect potential batch effects of pooled datasets. Lastly, a linear model was fit to the normalized data to obtain an expression measure for each probe set on qualified dataset [9].

The empirical Bayes method was used to select DEGs for each disease. Statistically significant DEGs were defined as those with p-values < 0.05 after adjustment by the Benjamini-Hochberg method and fold changes > 1.5. Subsequently, DEGs shared among the five diseases were extracted.

Gene ontology(GO) and pathway enrichment analysis

Three subdatabases (GOTERM_BP_DIRECT, GOTERM_CC_DIRECT, and GOTERM_MF_DIRECT) from DAVID Bioinformatics Resources 6.8 (https://david.ncifcrf.gov/) were used for GO enrichment analysis. KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/) was used for pathway enrichment analysis, integrating the KEGG pathway (http://www.genome.jp/kegg), Reactome (http://www.reactome.org), BioCyc (http://biocyc.org), and Panther (http://www.pantherdb.org) databases. P-values < 0.05 after adjustment by the Benjamini-Hochberg method were regarded as statistically significant.

Construction of the protein–protein interaction (PPI) network (PIN), recognition of hub genes and cluster analysis.

All established PPIs from four databases, BioGRID (http://thebiogrid.org/), DIP (http://dip.doe-mbi.ucla.edu/), HPRD (http://www.hprd.org/), and STRING (http://string-db.org), were used to construct PINs based on selected DEGs. Cytoscape software was utilized to visualize PINs and analyze the characteristics of each node. Those with the highest degrees, defined as the number of edges incident to the nodes, were recognized as hub genes. The MCODE plug-in of Cytoscape was used to perform cluster analysis and identify key clusters. An enrichment analysis based on the Reactome database was performed on the selected clusters to reveal related pathways.

Results

Characteristics of microarray datasets

After screening and quality control tests, a total of 255 qualified microarray datasets were included in the final analysis: 40, 7, 15, 43, 18, and 23 datasets for normal, DKD, HN, IgAN, MN and FSGS glomeruli, respectively, and 22, 11, 21, 25, 18, and 12 datasets for the corresponding tubules, respectively (Table 1). All cases were from subjects with biopsy-proven CKD, and particularly, all subjects with DKD had diabetic nephropathy. Tissue sources of controls were either kidney biopsy from pre-transplant living donor or patients after tumor nephrectomy. These datasets were from GSE37463 [10], GSE47185 [11], GSE35489 [12], GSE21785 [13], GSE20602 [14], GSE50469 [15] and GSE69438 [16]. All datasets were generated by either Human Genome U133 Plus 2.0 Array or Human Genome U133A Array.

Table 1. Datasets included in the study after quality control

Variables	Number of cases	Number of controls	Resources	Platforms
DKD Glomeruli	7	18	GSE37463, GSE47185	Affymetrix U133 Plus 2.0
DKD Tubules	11	22	GSE35489, GSE47185	Affymetrix U133 A
HN Glomeruli	15	22	GSE47185, GSE37463, GSE21785, GSE20602	Affymetrix U133 A
HN Tubules	21	22	GSE47185, GSE37463	Affymetrix U133 A
IgAN Glomeruli	43	22	GSE50469, GSE37463, GSE21785, GSE20602	Affymetrix U133 A
IgAN Tubules	25	22	GSE35489, GSE47185	Affymetrix U133 A
MN Glomeruli	18	22	GSE47185, GSE37463, GSE21785, GSE20602	Affymetrix U133 A
MN Tubules	18	22	GSE35489, GSE47185, GSE69438	Affymetrix U133 A
FSGS Glomeruli	23	40	GSE37463, GSE47185, GSE20602, GSE21785	Affymetrix U133 A Affymetrix U133 Plus 2.0
FSGS Tubules	12	22	GSE35489, GSE47185	Affymetrix U133 A

Identification of shared DEGs of CKD

After comparing gene expression patterns between diseased and normal glomeruli, as shown in Fig. 1, 1753 (upregulated, n = 1430; downregulated, n = 323), 1192 (upregulated, n = 739; downregulated, n = 453), 854 (upregulated, n = 492; downregulated, n = 362), 2863 (upregulated, n = 1627; downregulated, n = 1236), and 1298 (upregulated, n = 1017; downregulated, n = 287) genes were identified as DEGs in DKD, HN, IgAN, MN, and FSGS, respectively. As shown in Table 2, 176 genes were identified as shared DEGs among all diseases, with 104 upregulated (e.g. FCN1, CX3CR1) and 72 downregulated (e.g. APOM, SLC13A1) vs. normal glomeruli.

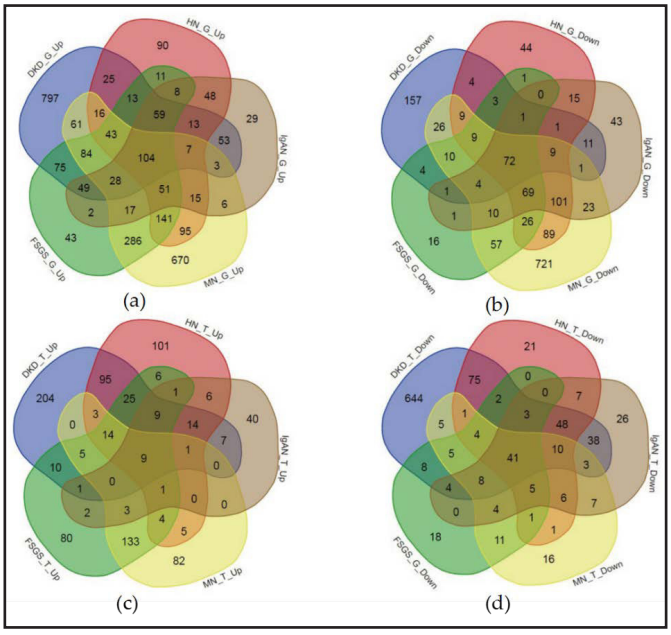

Fig. 1. Identification of shared DEGs in CKD. Different colors represent different primary diseases. Overlapping areas represent shared DEGs. (a) Upregulated genes in glomerular compartments in various types of CKD; (b) downregulated genes in glomerular compartments in various types of CKD; (c) upregulated genes in tubulointerstitial compartments in various types of CKD; (d) downregulated genes in tubulointerstitial compartments in various types of CKD.

In tubules, 1297 (upregulated, n = 397; downregulated, n = 900), 520 (upregulated, n = 294; downregulated, n = 226), 305 (upregulated, n = 94; downregulated, n = 211), 389 (upregulated, n = 260; downregulated, n = 129), and 418 (upregulated, n = 303; downregulated, n = 115) genes were identified as DEGs in DKD, HN, IgAN, MN, and FSGS, respectively. As shown in Table 2, 50 genes were shared DEGs, with nine upregulated (e.g., COL3A1, MARCKS) and 41 downregulated (e.g.

GDF15, RGS3) vs. normal tubules. In addition, nine genes were significantly differentially expressed both in diseased glomeruli and tubules (Table 2). A complete list of the DEGs in each disease is presented in Supplementary Table 1. For all supplemental material see www.karger.com/doi/10.1159/000488830.

GO enrichment analysis of shared DEGs

A total of 23 GO terms were significantly enriched among DEGs in diseased glomeruli [Fig. 2; molecular function (MF), n = 2; cellular component (CC), n = 14; biological process (BP), n = 7] as determined by adjusted p-values (Benjamini-Hochberg). According to the CC terms, DEGs were significantly enriched for genes encoding proteins involved in extracellular exosomes, with 80 DEGs in this category (Supplemental Table 2). Other enriched CCs included the integral component of the plasma membrane, extracellular space, and MHC class II protein complex. We further extracted DEGs involved in exosomes and performed a pathway enrichment analysis, which showed that these DEGs were mainly associated with metabolic and PPAR signaling pathways (Supplemental Table 3). DEGs in diseased glomeruli were also significantly enriched for several BPs and MFs (Fig. 2), including the interferon (IFN)-γ signaling pathway, antigen processing and presentation, MHC class II receptor activity, and platelet degranulation.

For diseased tubules, 37 GO terms were significantly enriched (Fig. 3 and Supplemental Table 4; CC, n = 2; BP, n = 8; MF, n = 27). DEG-encoded proteins mainly localize to the nucleus and transcription factor complex. The enriched BPs included mostly responses to stimuli (e.g., cAMP, cytokines, and calcium ions) and regulation of cell death, cell cycle, and proliferation. Enriched MF terms were mainly related to molecule binding and transcription factor activity.

Pathway enrichment analysis

Using four databases (KEGG pathway, Reactome, BioCyc and Panther), a total of 265 and 243 pathway terms were found to be significantly enriched in diseased glomeruli and tubules, respectively (Supplementary Tables 5 and 6). The 30 most significantly enriched pathway terms are shown in Fig. 4 and 5. For glomeruli, several metabolic, immune response, and

Table 2. Shared DEGs identified in CKD

Shared DEGs	Gene Names
Upregulated in diseased glomeruli (n=104)	FAM129A, FKBP11, LYN, HCK, FABP4, NMI, MCAM, AGTR1, IFI16, IER5, ISG20, SOSTDC1, ACTA2, FABP5, TAGLN, ICAM2, ARPC1B, S100A4, COL1A2, CRIP1, COL3A1, CD53, COL4A1, FRZB, CASP1, APOBEC3G, FGL2, HLA-F, GBP1, SRGN, C8orf4, DKK3, TRPV2, EED, SYT11, DNMT1, PCSK5, EPB41L2, POSTN, HLA-DPB1, SOX18, MYH11, CFD, CX3CR1, PXDN, LPAR6, PYCARD, ITGB2, FCN1, TM4SF1, CTSS, SLCO3A1, OAS1, TCF4, CDR2L, WFDC1, ARHGEF6, CORO1A, MICB, TYROBP, HLA-DQB1, GLIPR1, HLA-DRA, GPM6B, FEZ1, PHLDA2, HLA-DMA, TSPAN13, IFI44, PCDH12, CD52, Sept4, RFTN1, ECM1, CASP4, SOX17, HLA-DPA1, FGR, ADGRE5, GBP2, NETO2, HLX, HBB, FN1, SERPINH1, TPM1, PSMB9, FSCN1, GATA3, HCLS1, MEF2C, CALHM2, NRP1, ALOX5, TGFBI, PLK2, THBD, PSMB10, HTR2B, PTPRC, HLA-DMB, CD14, KIAA1462, CAMK1
Downregulated in diseased glomeruli (n=72)	XPNPEP2, SLC17A1, RIDA, NR4A2, GK, SLC13A1, CALB1, BHMT, MAOA, ETNK2, KL, SLC22A6, CDHR5, CYP4F3, ESRRG, PBLD, SLC34A1, MT1G, SLC22A8, SLC22A11, ALDH6A1, BHMT2, ZFP36, EGF, DUSP1, PCK1, DPYS, SLC6A13, CTH, AFM, CYP4A11, NR4A3, APOM, SLC7A8, MGAM, ALDH1L1, HPD, SLC12A3, SLC5A12, CYP27B1, FABP1, AK4, APOH, UPB1, SLC7A9, AZGP1, CPNE6, CA2, FBP1, GLYAT, DDC, PLG, ARG2, HAO2, RNF186, LPL, DIO1, CYP2B6, FOSB, SLC27A2, G6PC, ADH6, ALB, ALDOB, SLC13A3, KCNK5, AGMAT, GATM, SLC4A4, PIPOX, PRODH2, UGT2B28
Upregulated in diseased tubules(n=9)	PTPRC, HBB, TNFAIP8, IFI16, COL3A1, PROM1, MARCKS, FN1, COL6A3
Downregulated in diseased tubules (n=41)	GDF15, RGS3, GADD45A, ATF3, NR4A1, HBEGF, FOS, JUNB, RHOB, APOLD1, NFIL3, BHLHE40, EGR1, KLF6, SERPINE1, SOCS2, ZFAND5, CEBPD, JUND, ZFP36, LDLR, SGK1, ADM, DUSP1, FOSB, PTP4A1, NFKBIA, TRIB1, KLF10, TFRC, GEM, NR4A3, JUN, EGR3, MAFF, PDK4, MCL1, CTGF, RETREG1, GADD45B, ZNF165
Dysregulated both in diseased glomeruli and tubules (n=9)	IFI16, COL3A1, ZFP36, NR4A3, DUSP1, FOSB, HBB, FN1, PTPRC

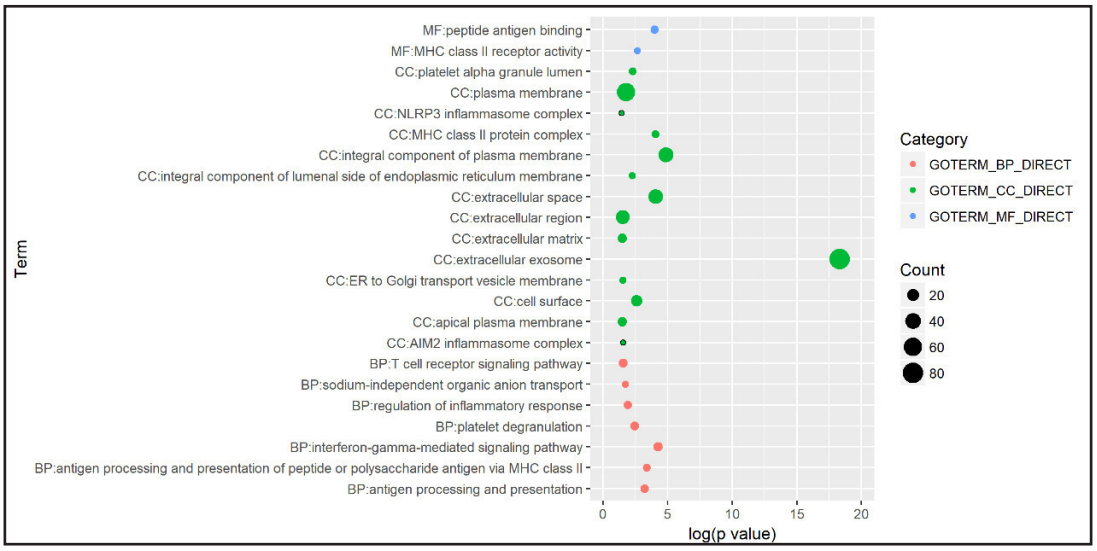

Fig. 2. Significantly enriched GO terms of DEGs in glomerular compartments in CKD. Log refers to logarithm to base 10.

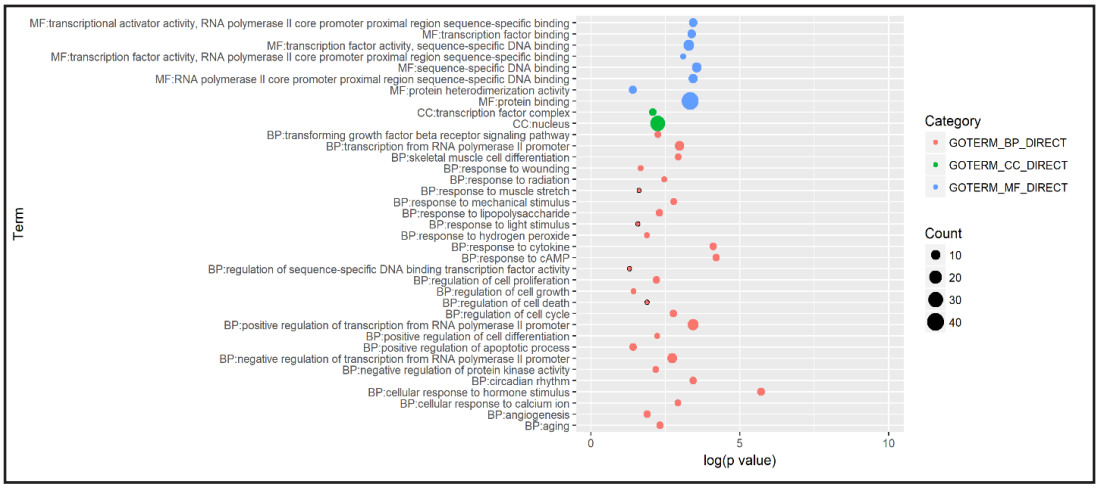

Fig. 3. Significantly enriched GO terms of DEGs in tubulointerstitial compartments in CKD. Log refers to logarithm to base 10.

signaling pathways (e.g., metabolism of lipids/lipoproteins and amino acids, SLC-mediated transmembrane transport, and IFN-γ pathway) were found to be dysregulated. Notably, other diseases such as Staphylococcus aureus infection and autoimmune thyroid disease were also listed as enriched terms, indicating that these disorders may share similar pathophysiological pathways with CKD. In tubules, our analysis highlighted the dysregulation of a wide range of signaling pathways, including the cholecystokinin receptor(CCKR), apoptosis, PI3K-Akt, MAPK, and TGF-β signaling pathways, some of which have been proposed to play critical roles in CKD development.

Key candidate gene and cluster identification based on PINs

Next, we reconstructed PINs for the glomerular and tubulointerstitial compartments based on the DEGs (Fig. 6 and 7), which showed interactions among gene products. As shown in Table 3, using a criterion of >5 degrees of interaction, a total of 10 genes (5 for diseased tubules, 5 for diseased glomeruli) were identified as hub genes, including JUN, EGR1, OAS1, and FN1. According to MCODE scores, two clusters (Fig. 8; cluster 1 for diseased glomeruli,

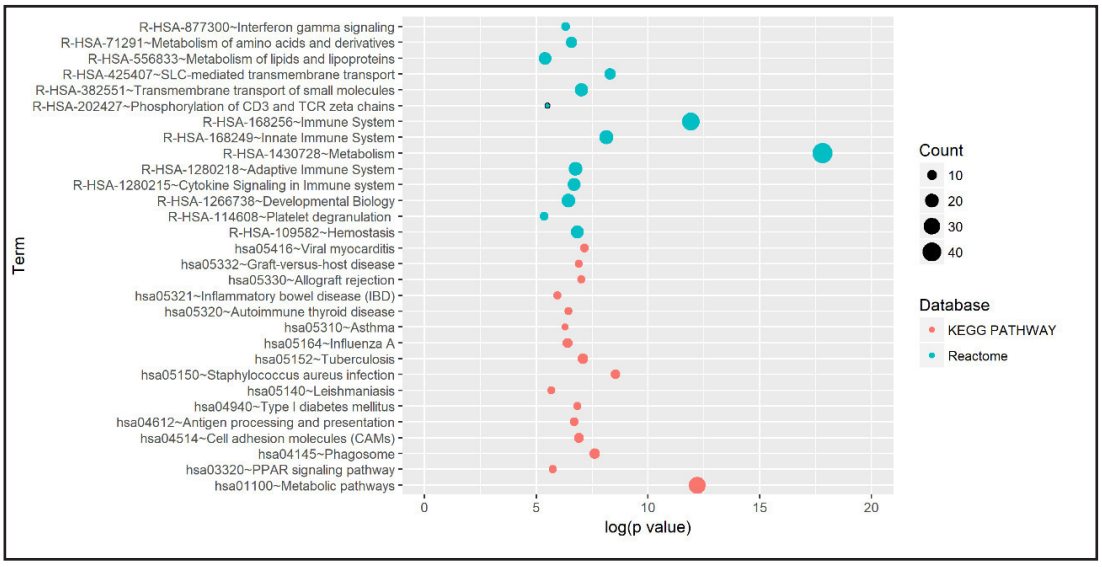

Fig. 4. Significantly enriched pathways of DEGs in glomerular compartments in CKD. Log refers to logarithm to base 10.

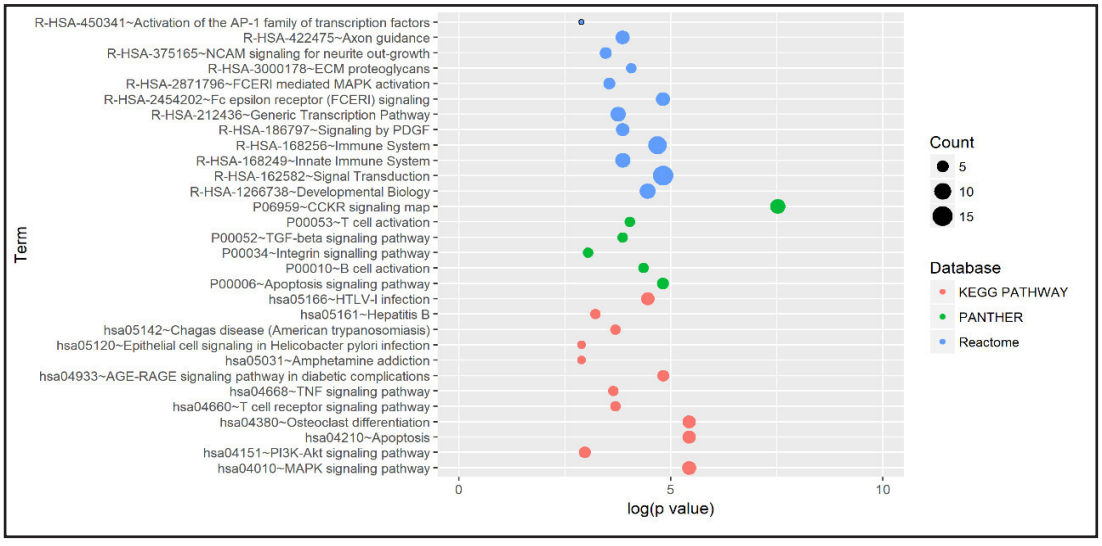

Fig. 5. Significantly enriched pathways of DEGs in tubulointerstitial compartments in CKD. Log refers to logarithm to base 10.

cluster 2 for diseased tubules) were identified from the PINs. Cluster 1 consisted of 16 nodes and 31 edges and was mainly associated with platelet dysfunction, abnormal interleukin signaling, and extracellular matrix organization, as revealed by pathway enrichment analysis. Cluster 2 had 17 nodes and 38 edges and was significantly associated with SMAD2/SMAD3/SMAD4 heterotrimer-regulated transcription, TGF-β receptor complex signaling, and interleukin signaling. A complete list of the enriched pathways of each cluster is presented in Supplementary Table 7.

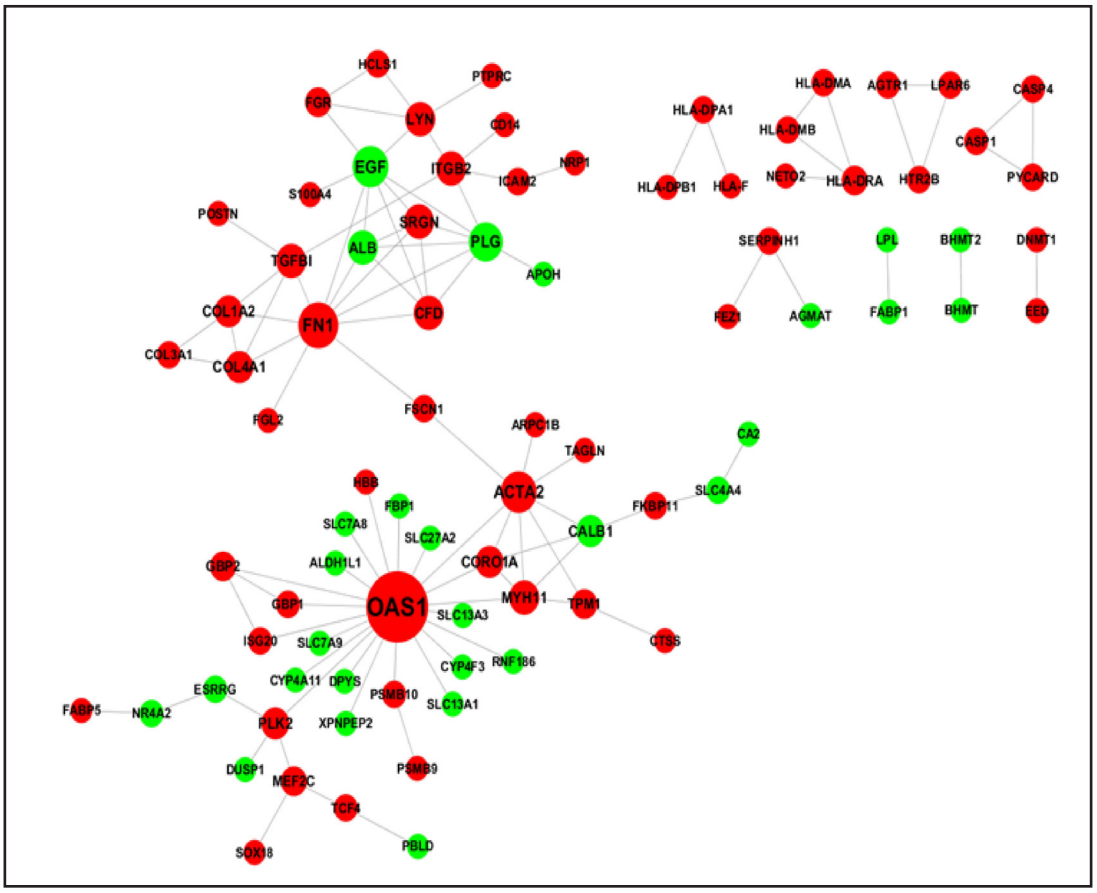

Fig. 6. PIN reconstruction of glomerular compartments in CKD. Red, upregulated nodes; green, downregulated nodes. Node size indicates degrees, which is defined as the number of edges incident to each node.

Fig. 7. PIN reconstruction of tubulointerstitial compartments in CKD. Red, upregulated nodes; green, downregulated nodes. Node size indicates degrees, which is defined as the number of edges incident to each node.

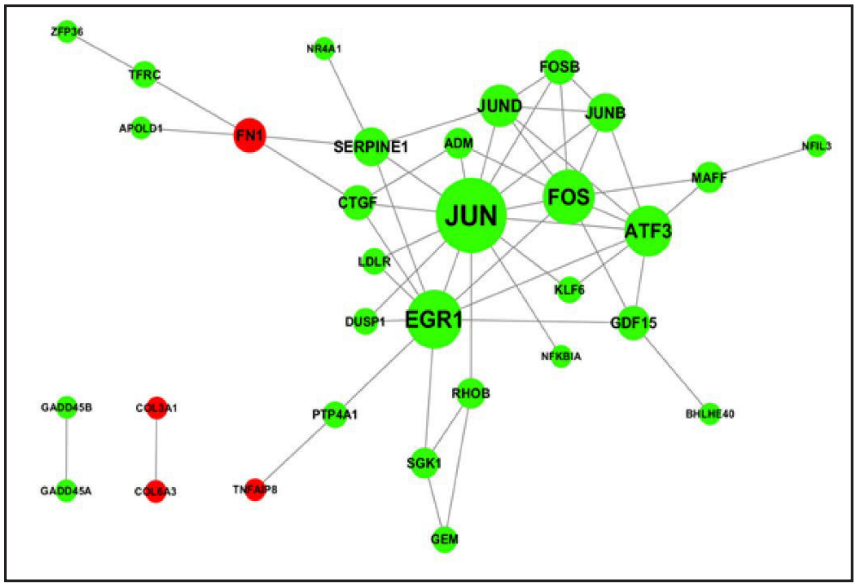

Table 3. Hub genes identified in PINs as possessing >5 degrees

Gene Names	Full names	Descriptions	Degrees
JUN	Jun Proto-Oncogene	Downregulated in diseased tubules	14
EGR1	Early Growth Response 1	Downregulated in diseased tubules	10
FOS	Fos Proto-Oncogene	Downregulated in diseased tubules	9
ATF3	Activating Transcription Factor 3	Downregulated in diseased tubules	8
JUND	JunD Proto-Oncogene	Downregulated in diseased tubules	6
OAS1	2'-5'-Oligoadenylate Synthetase 1	Upregulated in diseased glomeruli	21
FN1	Fibronectin 1	Upregulated in diseased glomeruli	10
EGF	Epidermal Growth Factor	Downregulated in diseased glomeruli	8
ACTA2	Actin, Alpha 2, Smooth Muscle, Aorta	Upregulated in diseased glomeruli	8
PLG	Plasminogen	Downregulated in diseased glomeruli	7

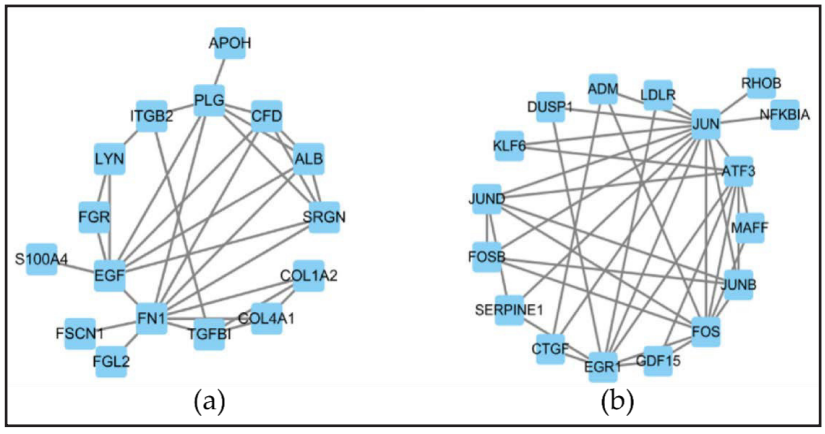

Fig. 8. Cluster analysis of PINs. The MCODE plug-in of Cytoscape was used to identify (a) cluster 1 from glomerular compartments and (b) cluster 2 from tubulointerstitial compartments as key regulatory networks in CKD.

Discussion

CKD arises from several heterogeneous disease pathways that chronically affect the function and structure of the kidney [17]. Many studies have demonstrated that these primary diseases share a wide range of common pathological pathways that play essential roles in disease development and progression. Previous gene expression studies on CKD have been limited by either imperfect disease models or small sample sizes. Here, we conducted an integrative bioinformatics analysis to identify shared key genes and pathways. The use of large sample size and robust algorithms ensure the reliability of the identified DEGs, as well as the subsequent enrichment analysis and PIN reconstructions. Our study not only provides insights into the molecular mechanisms of CKD but also acts as a resource for promising candidate biomarkers and therapeutic targets to facilitate the diagnosis and treatment of CKD. To our knowledge, this is the largest integrative bioinformatics analysis of this topic.

By comparing the transcriptomic profiles of the diseased tissues from different types of CKD, we identified hundreds of compartment-specific DEGs associated with CKD. In addition to several previously known key CKD genes (e.g., TGFBI, COL4A1, and FCN1), a lot of additional transcripts were also identified.

According to the GO enrichment analysis, glomerular DEGs are significantly enriched for genes involved in membrane structures, especially extracellular exosomes. Exosomes are a type of extracellular vesicles 30–100 nm in diameter that play a pivotal role in cell–cell communication [18]. Emerging studies have shown that exosomes act as key regulators of kidney injury [19]. Our previous studies have also found that transcriptomic changes in urinary exosomes can be utilized as discriminative biomarkers of mild renal fibrosis [20,

21], indicating their potential roles as the regulators and biomarkers of kidney injury. These enriched DEGs, as shown by pathway enrichment analysis, are mainly associated with metabolic and PPAR signaling pathways. PPAR signaling plays an important role in renal diseases by regulating a wide range of processes including glucose and lipid metabolism, inflammation, and immunity [22]. Whether DEG products exacerbate kidney injury via an exosome-mediated pathway remains to be explored. In this analysis, three key components of the NLRP3 inflammasome, namely PYD and CARD domain containing protein (PYCARD), caspase 1 (CASP1), and caspase 4 (CASP4), are significantly upregulated in diseased glomeruli. Previous studies have found that NLRP3 inflammasome activation, both in immune cells and kidney intrinsic cells, may induce inflammation in glomerular compartments and contribute to the onset of glomerular injury [23]. Taken together, these data show that NLRP3 inflammasome activation might be an essential event during CKD onset.

Pathway enrichment analysis reveals that CKD development is mainly associated with metabolic disorders, aberrant immune responses, and abnormally activated signaling pathways in glomerular compartments. Metabolites are considered to be the link between genotypes and phenotypes and act as the final proxies for physiological homeostasis [24]. Our analysis suggests that extensive metabolic disorders involving the amino acids, lipids, nucleotides, and carbohydrates of the glomeruli are common features of CKD. In fact, over 100 amino acids, purine and pyrimidine metabolites can be detected in at least 80% of participants with type 1 diabetes with impaired renal function [25]. The PIN cluster analysis further identified cluster 1, composed of 16 nodes, as the key regulatory network in the glomerular compartment. Enrichment analysis reveals that cluster 1 is strongly associated with platelet dysfunction, abnormal interleukin signaling, and extracellular matrix organization. CKD is well documented to be associated with a disturbed balance of the coagulation system [26]. Our data suggest that abnormal platelet function may be a major contributor to the imbalance. Similarly, a recent study observed markedly enhanced platelet activity in CKD mice, which could be attenuated by inhibiting MAPK signaling via klotho administration [27]. Regarding hub genes, FN1 and ACTA2 are well documented key regulators of renal fibrosis. EGF is also identified as a hub gene, which has attracted increasing interest recently. Both intrarenal and urinary expression of EGF was found to be positively correlated with eGFR[16]. EGF pathway has also been recognized as a novel therapeutic target of CKD progression [28]. Besides, OAS1 was found to have the most links with other nodes in the glomerular PIN. Early studies have shown that OAS1 is associated with hepatitis C virus infection, multiple sclerosis, and type 1 diabetes [29-31]. However, the relationship between OAS1 and CKD remains to be established.

In recent years, tubular epithelial cells (TECs) have attracted increasing attention as the primary sensor and effector during the development of CKD. TECs, especially the proximal TECs, are particularly vulnerable to injury, ultimately resulting in cell death and tubule atrophy [32]. Consistently, our analysis shows that the DEGs in tubulointerstitial compartments are significantly enriched in GO terms such as responses to stimuli and regulation of cell death, cell cycle, and proliferation. Among the enriched pathways in diseased tubules, CCKR signaling is at the top of the list. CCKR signaling is a well-known regulatory pathway in the digestive tract that exerts anti-inflammatory effects. Through microarray analysis and subsequent validation, a recent study showed that cholecystokinin is also expressed in the kidney and has renoprotective effects through its anti-inflammatory activity in diabetic nephropathy [33]. Our data further suggest that CCKR signaling may also play a regulatory role in other CKDs. Other enriched signaling pathways in diseased tubules include the apoptosis, TGF-β, MAPK, TNF, integrin, and PI3K-Akt signaling pathways, which have been well studied in the context of CKD [34, 35]. Renal interstitial fibrosis is regarded as a common pathological pathway in CKD. Our PPI analysis further illustrates that the TGF-β/Smad signaling pathway might act as a master regulator. Recent studies have shown considerable cross-talks between TGF-β/Smad signaling and some other pathways, such as the MAPK and Wnt/β-catenin pathways, in the regulation of renal interstitial fibrosis progression [36]. Notably, the five hub genes

of the tubulointerstitial PIN all encode transcription factors, four of which are subunits of AP-1 transcription factors (JUN, FOS, ATF3, and JUND) with a complex role in regulating cell life and death [37]. Interestingly, a recent study demonstrated that many fibrotic diseases share the activation of JUN (also known as c-Jun) in fibroblasts, leading to enhanced cell proliferation and migration [38]. Considering that tubular injury leads to epithelial cell cycle arrest [39], it is not surprising that the level of JUN mRNA was significantly downregulated in diseased tubules. Based on our data, further examination of the role of AP-1 transcription factors in CKD progression is warranted.

There are some limitations of our study. First, gene expression changes represent only one type of molecular alteration that occurs in CKD. Many other factors, including gene mutations, non-coding RNAs, proteomic and metabolomic changes, should also be considered to provide a more comprehensive understanding of disease development. Recently, metabolomics has emerged as a promising approach to identify new pathological pathways, and in-depth analysis of both genetic and metabolomic changes is expected to reveal more valuable information of the pathogenesis of CKD [40]. Second, gene expression data, rather than protein expression data, were used to reconstruct PINs, which might introduce bias as the changes observed on gene expression level are not always the same as those on the protein level. Third, despite we have applied several quality control procedures to ensure the consistency of datasets from different studies, heterogeneity among donors might still bring in potential bias. Fourth, our results are based on in silico analysis and therefore require further verification.

Conclusion

Our study not only further reveals the unifying molecular mechanism in CKD pathogenesis but also represents a valuable resource for potential biomarkers and therapeutic targets. We identified hundreds of compartment-specific genes that are differentially expressed in the common primary diseases of CKD. Most of the enrichment analysis results are consistent with our current understanding of the molecular mechanism of CKD. Others, such as the involvement of the dysregulated genes in exosomes, provide promising directions for future research. To further refine our results, we also identified key CKD regulatory networks and hub genes by constructing PINs. Notably, the shared DEGs not enriched in GO terms or PINs may also have critical pathological functions during disease formation and progression. Researchers are strongly encouraged to explore the supplemental materials, which provide further details for reference.

Abbreviations

BP (biological process); CC (cellular component); CKD (chronic kidney disease); DEG (differentially expressed gene); DKD (diabetic kidney disease); FSGS (focal segmental glomerular sclerosis); GN (glomerulonephritis); GO (gene ontology); HN (hypertensive nephropathy); IFN (interferon); IgAN (IgA nephropathy); MF (molecular function); MN (membranous nephropathy); PIN (protein–protein interaction network); PPAR (peroxisome proliferator-activated receptor); PPI (protein–protein interaction); QC (quality control).

Acknowledgements

Author Contributions: L.T.Z. and B.C.L. conceived and designed the study; S.Q., Z.L.L and H.L. collected the datasets; L.T.Z., R.N.T and K.L.M analyzed the data; L.T.Z. and B.C.L. wrote the paper. S.Q. and L.L.L revised the paper.

Funding: This study was supported by National Natural Scientific Foundation Grants 81670696, Clinic Research Center of Jiangsu Province Grant BL2014080, and Postgraduate Research & Practice Innovation Program of Jiangsu Province(KYCX17_0173).

Disclosure Statement

The authors declared they have no conflicts of interest.

References

1 Zhang L, Wang F, Wang L, Wang W, Liu B, Liu J, Chen M, He Q, Liao Y, Yu X, Chen N, Zhang JE, Hu Z, Liu F, Hong D, Ma L, Liu H, Zhou X, Chen J, Pan L, et al.: Prevalence of chronic kidney disease in China: a cross-sectional survey. Lancet 2012;379:815-822.

2 Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, Saran R, Wang AY-M, Yang C-W: Chronic kidney disease: global dimension and perspectives. Lancet 2013;382:260-272.

3 Becker GJ, Hewitson TD: Animal models of chronic kidney disease: useful but not perfect. Nephrology Dialysis Transplantation 2013;28:2432-2438.

4 Zhou LT, Cao YH, Lv LL, Ma KL, Chen PS, Ni HF, Lei XD, Liu BC: Feature selection and classification of urinary mRNA microarray data by iterative random forest to diagnose renal fibrosis: a two-stage study. Sci Rep 2017;7:39832.

5 Butz H, Szabo PM, Nofech-Mozes R, Rotondo F, Kovacs K, Mirham L, Girgis H, Boles D, Patocs A, Yousef GM: Integrative Bioinformatics Analysis Reveals New Prognostic Biomarkers of Clear Cell Renal Cell Carcinoma. Clin Chem 2014;60:1314-1326.

6 Wang WN, Zhang WL, Zhou GY, Ma FZ, Sun T, Su SS, Xu ZG: Prediction of the molecular mechanisms and potential therapeutic targets for diabetic nephropathy by bioinformatics methods. Int J Mol Med 2016;37:1181-1188.

7 Seccia TM, Maniero C, Belloni A, Pessina AC, Rossi G: Glomerular, tubulo-interstitial, and perivascular fibrosis in the kidney are differentially modulated by angiotensin II, endothelin-1, and L-type calcium channel blockade. Hypertension 2008;52:E92-E92.

8 Gao S, Xu S, Fang YP, Fang JW: Prediction of core cancer genes using multi-task classification framework. J Theor Biol 2013;317:62-70.

9 Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-264.

10 Berthier CC, Bethunaickan R, Gonzalez-Rivera T, Nair V, Ramanujam M, Zhang WJ, Bottinger EP, Segerer S, Lindenmeyer M, Cohen CD, Davidson A, Kretzler M: Cross-Species Transcriptional Network Analysis Defines Shared Inflammatory Responses in Murine and Human Lupus Nephritis. J Immunol 2012;189:988-1001.

11 Martini S, Nair V, Keller BJ, Eichinger F, Hawkins JJ, Randolph A, Boger CA, Gadegbeku CA, Fox CS, Cohen CD, Kretzler M, European Renal c DNAB, Cohort CP, Consortium CK: Integrative Biology Identifies Shared Transcriptional Networks in CKD. J Am Soc Nephrol 2014;25:2559-2572.

12 Reich HN, Tritchler D, Cattran DC, Herzenberg AM, Eichinger F, Boucherot A, Henger A, Berthier CC, Nair V, Cohen CD, Scholey JW, Kretzler M: A Molecular Signature of Proteinuria in Glomerulonephritis. PLoS One 2010;5:e13451.

13 Lindenmeyer MT, Eichinger F, Sen K, Anders HJ, Edenhofer I, Mattinzoli D, Kretzler M, Rastaldi MP, Cohen CD: Systematic Analysis of a Novel Human Renal Glomerulus-Enriched Gene Expression Dataset. PLoS One 2010;5:e11545.

14 Neusser MA, Lindenmeyer MT, Moll AG, Segerer S, Edenhofer I, Sen K, Stiehl DP, Kretzler M, Grone HJ, Schlondorff D, Cohen CD: Human Nephrosclerosis Triggers a Hypoxia-Related Glomerulopathy. Am J Pathol 2010;176:594-607.

15 Hodgin JB, Berthier CC, John R, Grone E, Porubsky S, Grone HJ, Herzenberg AM, Scholey JW, Hladunewich M, Cattran DC, Kretzler M, Reich HN: The Molecular Phenotype of Endocapillary Proliferation: Novel Therapeutic Targets for IgA Nephropathy. PLoS One 2014;9:e103413.
16 Ju WJ, Nair V, Smith S, Zhu L, Shedden K, Song PXK, Mariani LH, Eichinger FH, Berthier CC, Randolph A, Lai JYC, Zhou Y, Hawkins JJ, Bitzer M, Sampson MG, Thier M, Solier C, Duran-Pacheco GC, Duchateau-Nguyen G, Essioux L, et al.: Tissue transcriptome-driven identification of epidermal growth factor as a chronic kidney disease biomarker. Sci Transl Med 2015;7:316ra193.
17 Webster AC, Nagler EV, Morton RL, Masson P: Chronic Kidney Disease. The Lancet;389:1238-1252.
18 Camussi G, Deregibus MC, Bruno S, Cantaluppi V, Biancone L: Exosomes/microvesicles as a mechanism of cell-to-cell communication. Kidney Int 2010;78:838-848.
19 Erdbrugger U, Le TH: Extracellular Vesicles in Renal Diseases: More than Novel Biomarkers? J Am Soc Nephrol 2016;27:12-26.
20 Lv LL, Cao YH, Ni HF, Xu M, Liu D, Liu H, Chen PS, Liu BC: MicroRNA-29c in urinary exosome/microvesicle as a biomarker of renal fibrosis. American Journal of Physiology-Renal Physiology 2013;305:F1220-F1227.
21 Lv LL, Cao Y, Liu D, Xu M, Liu H, Tang RN, Ma KL, Liu BC: Isolation and Quantification of MicroRNAs from Urinary Exosomes/Microvesicles for Biomarker Discovery. Int J Biol Sci 2013;9:1021-1031.
22 Yang JC, Zhou YF, Guan YF: PPAR gamma as a therapeutic target in diabetic nephropathy and other renal diseases. Curr Opin Nephrol Hypertens 2012;21:97-105.
23 Conley SM, Abais JM, Boini KM, Li PL: Inflammasome Activation in Chronic Glomerular Diseases. Curr Drug Targets 2017;18:1019-1029.
24 Hocher B, Adamski J: Metabolomics for clinical use and research in chronic kidney disease. Nature Reviews Nephrology 2017;13:269-284.
25 Niewczas MA, Mathew AV, Croall S, Byun J, Major M, Sabisetti VS, Smiles A, Bonventre JV, Pennathur S, Krolewski AS: Circulating Modified Metabolites and a Risk of ESRD in Patients With Type 1 Diabetes and Chronic Kidney Disease. Diabetes Care 2017;40:383-390.
26 Lutz J, Menke J, Sollinger D, Schinzel H, Thurmel K: Haemostasis in chronic kidney disease. Nephrology Dialysis Transplantation 2014;29:29-40.
27 Yang K, Du CH, Wang XM, Li FJ, Xu Y, Wang S, Chen SL, Chen F, Shen MQ, Chen M, Hu MJ, He T, Su YP, Wang JP, Zhao JH: Indoxyl sulfate induces platelet hyperactivity and contributes to chronic kidney disease-associated thrombosis in mice. Blood 2017;129:2667-2679.
28 Kok HM, Falke LL, Goldschmeding R, Nguyen TQ: Targeting CTGF, EGF and PDGF pathways to prevent progression of kidney disease. Nature Reviews Nephrology 2014;10:700-711.
29 Qu HQ, Polychronakos C, Type IDGC: Reassessment of the type I diabetes association of the OAS1 locus. Genes Immun 2009;10:S69-S73.
30 O'Brien M, Lonergan R, Costelloe L, O'Rourke K, Fletcher JM, Kinsella K, Sweeney C, Antonelli G, Mills KH, O'Farrelly C, Hutchinson M, Tubridy N: OAS1 A multiple sclerosis susceptibility gene that influences disease severity. Neurology 2010;75:411-418.
31 Garcia-Alvarez M, Berenguer J, Jimenez-Sousa MA, Pineda-Tenor D, Aldamiz-Echevarria T, Tejerina F, Diez C, Vazquez-Moron S, Resino S: Mx1, OAS1 and OAS2 polymorphisms are associated with the severity of liver disease in HIV/HCV-coinfected patients: A cross-sectional study. Sci Rep 2017;7:41516.
32 Chevalier RL: The proximal tubule is the primary target of injury and progression of kidney disease: role of the glomerulotubular junction. American Journal of Physiology-Renal Physiology 2016;311:F145-F161.
33 Miyamoto S, Shikata K, Miyasaka K, Okada S, Sasaki M, Kodera R, Hirota D, Kajitani N, Takatsuka T, Kataoka HU, Nishishita S, Sato C, Funakoshi A, Nishimori H, Uchida HA, Ogawa D, Makino H: Cholecystokinin Plays a Novel Protective Role in Diabetic Kidney Through Anti-inflammatory Actions on Macrophage Anti-inflammatory Effect of Cholecystokinin. Diabetes 2012;61:897-907.
34 Gewin L, Zent R, Pozzi A: Progression of chronic kidney disease: too much cellular talk causes damage. Kidney Int 2017;91:552-560.
35 Liu MN, Ning XX, Li R, Yang Z, Yang XX, Sun SR, Qian Q: Signalling pathways involved in hypoxia-induced renal fibrosis. J Cell Mol Med 2017;21:1248-1259.
36 Meng XM, Nikolic-Paterson DJ, Lan HY: TGF-beta: the master regulator of fibrosis. Nature Reviews Nephrology 2016;12:325-338.

Kidney Blood Press Res 2018;43:568-581
DOI: 10.1159/000488830
Published online: 11 April, 2018

37 Shaulian E, Karin M: AP-1 as a regulator of cell life and death. Nat Cell Biol 2002;4:E131-E136.
38 Wernig G, Chen SY, Cui L, Van Neste C, Tsai JM, Kambham N, Vogel H, Natkunam Y, Gilliland DG, Nolan G, Weissman IL: Unifying mechanism for different fibrotic diseases. Proc Natl Acad Sci U S A 2017;114:4757-4762.
39 Bonventre JV: Primary proximal tubule injury leads to epithelial cell cycle arrest, fibrosis, vascular rarefaction, and glomerulosclerosis. Kidney International Supplements 2014;4:39-44.
40 Menni C, Zierer J, Valdes AM, Spector TD: Mixing omics: combining genetics and metabolomics to study rheumatic diseases. Nature Reviews Rheumatology 2017;13:174-181.