Identification of CD2, CCL5 and CCR5 as potential therapeutic target genes for renal interstitial fibrosis

Chuanjie Zhang¹#, Xin Hu²#, Feng Qi³#, Jun Luo⁴, Xiao Li⁵

¹Department of Urology, Ruijin Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai 200025, China; ²First Clinical Medical College of Nanjing Medical University, Nanjing 210029, China; ³Department of Urology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China; ⁴Department of Urology, Shanghai Fourth People’s Hospital Affiliated to Tongji University School of Medicine, Shanghai 200081, China; ⁵Department of Urology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing 210009, China

Contributions: (I) Conception and design: C Zhang, X Li; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: None; (V) Data analysis and interpretation: C Zhang, X Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Jun Luo. Department of Urology, Shanghai Fourth People’s Hospital affiliated to Tongji University School of Medicine, Shanghai 200081, China. Email: abell_ll_luojun@sina.com; Xiao Li. Department of Urology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing 210009, China. Email: leex91@163.com.

Background: We aimed to explore potential gene biomarkers of renal interstitial fibrosis (RIF) due to a lack of effective and non-invasive methods for diagnosis.

Methods: Three data sets (GSE22459, GSE76882 and GSE57731) including 350 samples were acquired from Gene Expression Omnibus (GEO) database. We used bioconductor limma package to perform background adjustment. Cluster analysis was conducted by ‘edgeR’ package to identify the differentially expressed genes (DEGs). We generated heat maps with using heatmap package in R software. Function annotation of genes was performed by Gene Ontology (GO) enrichment analysis. STRING (Search Tool for the Retrieval of Interacting Genes) database was employed to construct the protein-protein interaction (PPI) network and the results were visualized by Cytoscape 3.6.1. At last, we applied Graphpad Prism 7.0. to explore the correlation between three hub genes and pathological degrees of RIF.

Results: By applying the “edgeR” package in R, we detected 116 DEGs with three data sets. These genes were enriched in 19 GO biological process categories. Three main hub genes (CD2, CCL5 and CCR5) were identified after construction of PPI network. In Pearson correlation coefficient, CD2, CCL5 and CCR5 was found to hold higher expression patterns in RIF samples based on independent data set GSE57731. Besides, their gene expression levels were found significantly positive correlation with the degree of RIF (CD2: P<0.05, r=0.29; CCL5: P<0.05, r=0.31; CCR5: P<0.05, r=0.38).

Conclusions: CD2, CCL5 and CCR5 might serve as potential early biomarkers of RIF. The mechanism between these genes and RIF remains to be further studied.

Keywords: Renal transplantation; renal interstitial fibrosis (RIF); diagnosis; gene; Gene Expression Omnibus (GEO)
Introduction

Most chronic kidney disease (CKD) would finally turn into End-Stage Renal Disease (ESRD), impacting 12% of all adults in the United States and 10.8% in China (1). By 2030, the number of ESRD patients who need maintenance hemodialysis would increase to about 5.4 billion (2). Currently, renal transplantation is recognized as the best treatment for ESRD as life expectancy in renal transplant patients has improved significantly (3). Over the past 20 years, although the success rate and 1-year survival rate of renal transplantation has improved, long-term survival rate remained dissatisfied. By report, long-term survival rate was 68% for living-related renal transplantation and 51% for cadaveric renal transplantation (4).

Renal interstitial fibrosis (RIF) of the transplanted renal, as a pathophysiological process, acts as the main pathological feature of renal transplantation failure, whose clinical manifestation was the decrease of glomerular filtration rate (5). RIF is related to antibody-mediated rejection (AMR) (6), epithelial to mesenchymal transition (EMT) (7) and excessive accumulation of extracellular matrix (ECM) (8). Presently, examination-renal biopsy is still the gold standard for the diagnosis of RIF. However, as an invasive method, biopsy difficulty, biopsy-sampling error and diagnostic deviation of pathologist may affect diagnosis results, which limits its application. Recently, some serum biomarkers including urinary-specific matrix metalloproteinase (MMP)-generated collagen type I and III degradation fragments (9) might help predict the progression of RIF. However, the specificity of serum biomarkers is still unsatisfied due to interference of other renal diseases such as glomerular nephritis and acute kidney injury. Therefore, it is particularly important to find effective and non-invasive methods for early diagnosis of RIF, which could help adjust the therapeutic regimen timely so as to improve the function of transplanted renal.

Recently, mRNAs have raised great interest to be detected as biomarkers for RIF. Cao et al. pointed that vimentin (VIM) Mrna presented the highest change fold of 9.99 in RIF patients, indicating it might be assumed as a novel non-invasive biomarker to monitor the process of RIF (10). Moreover, studies indicated that RIF was associated with immune inflammatory signaling pathways such as TGF-β/Smad (11) and Wnt/β-catenin (12) pathway. However, limited mechanism of RIF has been discovered. Accordingly, we aimed to identify the differentially expressed genes (DEGs) using the Gene Expression Omnibus (GEO) database. Then we conducted Gene Ontology (GO) function enrichment analysis and protein-protein interaction (PPI) network analysis to excavate the potential biological action mechanisms and hub genes for further exploration.

Methods

Data acquisition

The data of RIF and normal samples were derived from the GEO database. The result was shown in Table 1. There were 24 interstitial fibrosis and tubular atrophy (IFTA) and 16 interstitial fibrosis and tubular atrophy with inflammation (IFTA_i) samples in the GSE22459 data set; 43 IFTA samples, 10 IFTA_i and 54 interstitial fibrosis and tubular atrophy with acute rejection (IFTA_AR) samples in the GSE76882 data set. They were then applied to identify the DEGs. The GSE57731 data set (including 45 normal samples and 14 IFTA samples) was mainly employed to verify the relationship between hub gene and the stage of RIF.

Data pretreatment

Raw data was downloaded from GEO database (CEL files). We used Bioconductor ‘limma’ package to perform background adjustment. Through the platform files and the Chip signal data which comes from gene chip, gene (Gene Symbol) and probe identification (ID) were selected. If a probe set was corresponded to multiple or 0 genes, the data of that probe set was omitted. If multiple probe sets were mapped to the same gene, the gene expression profile data was calculated by the arithmetic mean value of multiple probe sets.
Identification of DEGs

Cluster analysis was conducted to explore the change of gene expression patterns between IFTA samples and normal samples. It was carried out by ‘edgeR’ package in R software (13). |log2 fold-change| >2 and the P value of false discovery rate (FDR) <0.05 was settled as a threshold (14). Through application of Jaccard distance and complete linkage, we analyzed the clustering and formed clusters. Significance analysis of microarrays was applied to identify the DEGs. Besides, we generated heat maps using ‘pheatmap’ package in R software.

Function enrichment

GO enrichment analysis was then conducted by DAVID (database for annotation, visualization, and integrated discovery) (15) with FDR <0.01 to screen significant functional pathway enriched in the DEGs. We applied GO plot package of R software to explore the GO terms.

Establishment of PPI network

For the purpose of verifying the potential relationships of DEGs, STRING (Search Tool for the Retrieval of Interacting Genes) database (16) was employed to construct the PPI network to find hub genes. The modification and visualization of network was performed by Cytoscape 3.6.1 based on the JAVA software (17).

Correlation analysis

We used Graphpad Prism 7 to calculate the Spearman’s rank correlation coefficient to explore the correlation between three hub genes (CD2, CCL5 and CCR5) and pathological degrees of RIF.

Results

DEG patterns

To investigate the differences in gene expression between IFTA samples and normal samples, we used the sample set GSE22459 which contained 25 normal samples, 24 IFTA samples and 16 IFTA_i samples to identify 219 DEGs. Among them, 55 genes were lower expressed, while 164 genes were higher expressed of IFTA samples. At the same time, another sample set GSE76882 which included 99 normal samples, 43 IFTA samples, 10 IFTA_i samples and 54 IFTA_AR samples was used to verify 504 DEGs (184 was downregulated and 320 was upregulated). Next, 116 overlapping genes (84 up-regulated and 32 down-regulated) were obtained by intersecting the two sets of DEGs (Figure 1A). Then we performed cluster analysis for all 116 DEGs. Only the top 60 genes with the most significant differences were selected for visualization in Figure 1B. It suggested that these genes had obvious different gene expression patterns between normal renal samples and IFTA samples.

Function annotation of DEGs

GO analysis was conducted to obtain the biological functions of 116 DEGs to understand which signaling pathways might serve as an important role in RIF. As shown in Figure 2, results showed that these DEGs were enriched in 19 GO biological process categories. Some pathways were related to inflammation and immunity such as T cell activation, antigen processing and presentation, natural killer cell mediated immunity, myeloid leukocyte activation, adaptive immune response and leukocyte chemotaxis, implying that there might be abnormal inflammatory and immune responses in IFTA samples.

Identification of hub genes via PPI network

We constructed a PPI network to investigate the underlying relationships among these genes, thus detecting 97 nodes and 143 edges. What’s more, Cytoscape 3.6.1 was applied to screen hub genes which were pivotal nodes in this network. As shown in Figure 3, CD2, CCL5 and CCR5 acted as the main hub genes.

Three hub genes (CD2, CCL5 and CCR5) correlated with pathological degrees of RIF

We selected the independent sample set GSE57731 from the GEO database to further explore the relationship between hub genes with the pathological degree of RIF. Since most RIF events occurred in the first 3 months after renal transplantation, we used the GSE57731 sample set which came from the blood samples collected in 3 months. CD2, CCL5 and CCR5 were found to kept higher expression patterns in high-fibrosis samples. Moreover, their gene expression levels were found significantly positive correlation with the degree of RIF (Figure 4, CD2: P<0.05, r=0.29; CCL5: P<0.05, r=0.31; CCR5: P<0.05, r=0.38).
Discussion

Renal transplantation is the most mature and reliable method for the treatment of ESRD, which could effectively improve the quality of life and survival time of patients (18). However, the long-term survival of allogeneic renal transplantation remained unsatisfactory (19). A cohort study illustrated that RIF was an important time-dependent histologic lesion, occurring in 70% of biopsy samples in 3 months on post-transplantation (20). Immune cells (macrophages, T lymphocytes and mast cells) (21) and profibrogenic cytokines (TGF-1 and CTGF) (22) were involved in the occurrence and development of RIF. Screening the early potential targets for diagnosis was necessary as it was reported that early graft function held a significant influence on long-term renal function (23).

Nowadays, both diagnosis and intervention of RIF were challenges for clinicians. Glomerular filtration rate and
serum creatinine clearance had limited clinical utility in assessing renal histopathological changes and fibrosis after renal transplantation. In addition, the non-invasive diagnosis of RIF was tough with the interference of complex immune regulatory network and non-immune factors (24). As serum markers were easily influenced by other renal diseases, most recent discoveries of diagnostic biomarkers have arisen in the field of genetic molecules such as microRNA (miRNA). For example, MiR-433 was an important component of TGF-β/Smad3 pathways and could enhance TGF-β1-induced fibrosis by enhancing the antizyme inhibitor—Azin1 in vitro and in vivo (25). Jenkins et al. found that miR-

**Figure 3** PPI network analysis of 116 genes with RIF. The red diamond represents the three hub genes. PPI, protein-protein interaction; RIF, renal interstitial fibrosis.

**Figure 4** CD2, CCL5 and CCR5 expression was positively correlated with the degree of RIF. RIF, renal interstitial fibrosis.
192 mediated the development of tubulointerstitial fibrosis via repression of E-box repressors such as ZEB1, causing an activation of akt kinase signaling to increase the expression of collagen (26).

In this study, we identified 116 DEGs between IFTA samples and normal samples based on data sets downloaded from GEO database. All samples were obtained directly from fibrous renal tissue to improve their accuracy. Among these DEGs, GO analysis was conducted to verify some pathways of inflammation and immunity such as T cell activation and natural killer cell mediated immunity. Finally, we detected three DEGs (CD2, CCL5 and CCR5) which may be the potential molecular mechanisms and biomarkers by performing PPI network.

DEGs were found enriched in some pathways related to the inflammation and immunity. “T cell activation” was the most enriched GO terms in BP. A study suggested that the recruitment and activation of CD4+ T lymphocyte could be the critical step of the onset of RIF so that the suppression of CD4+ T lymphocyte differentiation to Th2 could be a potential therapeutic intervention (27). Besides, ‘Natural killer cell mediated immunity’ term was also enriched. NK cells were subsets of lymphocytes and played an important role in regulating the immune response. Law et al. found that the number of total NK cells (CD3− CD56+) were detected higher in RIF tissues compared with tissues without fibrosis, indicating NK cells was a vital factor to fibrotic process (28). Other certified immune-related terms included antigen processing and presentation, myeloid leukocyte activation, adaptive immune response and leukocyte chemotaxis.

CD2 which located in 1p13.1 is highly expressed in RIF. Its encoding protein is a surface antigen found on all peripheral blood T-cells. Moreover, CD2-positive T cells were the predominant infiltrating cells in various glomerulonephritis (29). Relevant studies on CD2AP encoded by CD2 have been reported that CD2AP was also involved in a variety of signaling pathways. By selectively enhancing the expression of TGF-β, CD2AP could activate the TGF-β/smad3 signaling pathway which promoted the occurrence of RIF (30). What’s more, CD2AP was important for maintaining the integrity of slit diaphragm and cytoskeleton (31). Elevated CD2 expression might be a compensatory repair mechanism after renal tissue injury. Besides, the reduction of urinary exosome mRNA of CD2AP in RIF made it possible to be a noninvasive detection biomarker of renal fibrosis (32).

CCL5 was one of several chemokine genes clustered on the q-arm of chromosome 17. In renal transplantation rejection, molecular biology studies have confirmed that the activation of NF-κB signaling enhanced the secretion of TGF-β, upregulating CCL5 in mesangial cells to promote RIF (33). Furthermore, CCL5 used chemokine receptor chemokine (C-C motif) receptor 5 (CCR5) that was found to be highly expressed in our study as a co-receptor (34). CCR5 encoded a member of the beta chemokine receptor family, which was predicted to be a seven transmembrane protein similar to G protein-coupled receptors. A study pointed that inhibition of a potent mitogen for fibroblasts PDGF-C produced a decrease in the expression of the renal chemokines CCL5, accompanied by a significant reduction in expression of CCR mRNA. We hypothesized that CCL5 could be combined with CCR5+ T cells to promote the aggregation of monocytes or macrophages, thus leading to the occurrence of immune inflammatory response. However, further exploration is needed (35). Moreover, research has shown that CCR5 may be a potential drug target for anti-fibrosis therapy (36).

Admittedly, we found that there were significantly different gene patterns between high- and low-fibrosis samples. GSE57731 which was a serological sample was employed to verify it. It was cheaper and safer to get the patients’ serum than to obtain their kidney tissue with less damage to the patients. Gene expression levels of them were remarkably positively correlated with the degree of RIF, indicating that there were changes in expression of these genes in the early stages of disease. Research has reported that the expression of CCL5 in RIF increased with the increase of pathological grading (37). As early graft function held a significant effect on long-term renal function (23), CD2, CCL5 and CCR5 have potential to be the early biomarkers of diagnosis. If the biomarkers acquired from fibrous renal tissue could be verified in serological samples, it would be of great value of medical application and research.

The use of clinical biomarkers is easier and less expensive than direct measurement of the final clinical endpoint. To be honest, there existed some studies which focused on the specific role of these genes in the mechanism of renal fibrosis related to these three genes. It was necessary, but was just the first step to know the pathophysiological relationship between the biomarker and the relevant clinical endpoint. In this research, we emphasized on transforming these basic studies into clinical applications. We hoped that they can be used in screening, diagnosis, characterization and monitoring of RIF.
Our study still has some limitations. First, we were currently unable to conduct the multi-center clinical verification. Second, due to the limited research conditions, we could not perform in vitro or in vivo experiments presently. So, we have limited evidence to explain the mechanism of RIF. Nevertheless, we found three DEGs (CD2, CCL5 and CCR5) which may be the potential molecular mechanisms and markers for further study.

**Acknowledgments**

**Funding:** This work was supported by grants from Excellent young talents in clinical medicine of Hongkou District (HKYQ2018-09), the National Natural Science Foundation of China (No. 81702520), Medical Research Project of Jiangsu Provincial Health and Family Planning Commission (No. H2018052), Research Project of Jiangsu Cancer Hospital (No. ZN201602), and the young talents program of Jiangsu Cancer Hospital (No. 2017YQL-04).

**Footnote**

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**References**


Cite this article as: Zhang C, Hu X, Qi F, Luo J, Li X. Identification of CD2, CCL5 and CCR5 as potential therapeutic target genes for renal interstitial fibrosis. Ann Transl Med 2019;7(18):454. doi: 10.21037/atm.2019.08.62