Original Article
Evaluating the microRNA-target gene regulatory network in renal cell carcinomas, identification for potential biomarkers and critical pathways

Jun Li1*, Jian-Hua Huang3*, Qing-Hua Qu1, Qier Xia1, Deng-Shan Wang2, Lei Jin2, Chang Sheng1

1Department of Urology, Pudong New Area People’s Hospital, Shanghai, China; 2Department of Pathology, Pudong New Area People’s Hospital, Shanghai, China; 3Department of Urology, Shanghai Tenth People’s Hospital, Tongji University, Shanghai, 200072, China. *Equal contributors.

Received March 2, 2015; Accepted May 6, 2015; Epub May 15, 2015; Published May 30, 2015

Abstract: Variant microRNA (miRNA) expression is a character of many cancer types. The combined analysis of miRNA and messenger RNA (mRNA) expression profiles is crucial to identifying links between deregulated miRNAs and oncogenic pathways. The aim of this study was to screen several novel genes associated with renal cell carcinoma (RCC), and analyze the gene functions and signal pathways which were critical to RCCs with DNA microarray. The gene expression profile of GSE6344 was downloaded from Gene Expression Omnibus database, including 10 RCC samples and 10 healthy controls. Compared with the control samples, differentially expressed genes (DEGs) of RCC was identified. The selected DEGs were further analyzed using bioinformatics methods. Gene ontology (GO) enrichment analysis was performed using Gene Set Analysis Toolkit and protein-protein interaction (PPI) network was constructed with prePPI. Then, pathway enrichment analysis to PPI network was performed using WebGestalt software. We found that a total of 521 DEGs were down-regulated and 473 DEGs were up-regulated in RCC samples compared to healthy controls. A total of 15 remarkable enhanced functions and 17 suppressed functions were identified. PPI nodes of high degrees, such as RHCG, RALYL, SLC4A1, UMOD and CA9, were obtained. The DEGs were classified and significantly enriched in cytokine and cytokine receptor pathway. The hub genes we find from RCC samples are not only biomarkers, but also may provide the groundwork for a combination therapy approach for RCCs.

Keywords: Microarray profiling, oncogenes, potential biomarkers, critical pathways, renal cell carcinoma

Introduction
Renal cell carcinoma (RCC), is a kidney cancer that originates in the lining of the proximal convoluted tubule, and known to be the most lethal of all the genitourinary tumors [1, 2]. It is also the most common type of kidney cancer in adults, responsible for approximately 80% of cases [3]. RCC is relatively resistant to radiation therapy and chemotherapy, although some cases respond to immunotherapy [4]. So it is clear that early diagnosis and medical intervention seems vital in decreasing mortality and promoting total quality of life, novel molecular markers about kidney cancer that can help individually evaluate risk of patient outcome and predict the prognosis are urgently required, as well as the prediction of therapy effect and advocating personalized treatment [5, 6]. The major difficulty in RCC is the constitution that this disease is not one entity but rather a collection of different types of tumors, each possessing distinct genetic characteristics, histological features, and, to some extent, clinical phenotypes [7].

MicroRNAs (miRNAs) are a class of small non-coding RNAs that can repress gene expression through translational repression or messenger RNA deadenylation and decay by base pairing to partially complementary sites [8, 9]. miRNAs were shown to negatively regulate gene expression at the post-transcriptional level by binding to the 3’-untranslated region (3’-UTR) of target mRNAs. miRNAs have been shown to be involved in tumour progression and metastasis in kidney and other carcinomas [10]. Next-generation small RNA-Sequencing (sRNA-Seq) allows for unbiased quantitative and qualitative scnRNA profiling [11]. When compared to
miRNA array platforms, sRNA-Seq additionally enables the discovery of novel miRNAs as well as the detection of other differentially expressed sncRNAs like small nucleolar RNAs (snoRNAs) and transfer RNA (tRNA)-derived fragments that can mimic miRNA function [12].

Several microarray based studies have demonstrated 21 to 34 differentially expressed miRNAs between ccRCC and normal kidney tissue [13]. SRNA-Seq studies reported more than 100 differentially regulated miRNAs, some of which might serve as diagnostic and prognostic markers [14]. Nevertheless, these studies lack detailed information about miRNA targets and bioinformatical analysis is often only focused on miRNAs currently known to miRbase [15].

Here we used omiRas to analyze a publicly available dataset (GEO: GSE6344) published by Gumz M et al., comprising twenty sRNA-Seq libraries of ten ccRCCs and ten adjacent control tissues from the same patient in order to identify sncRNAs with deregulated expression across all cases [16]. After outlier detection with principle component analysis samples of nine patients were used for downstream analysis.

We detected 104 sncRNAs as differentially expressed between the groups. Among these were several miRNAs without previous implication in kidney cancer development, like miR-147a. Additionally, we detected seven snoRNAs and two tRNA derived fragments as differentially expressed between ccRCC and control tissues. We connected the deregulated miRNAs to biological pathways composed of differentially expressed genes under potential post-transcriptional control of these miRNAs. To do so, we utilized another publicly available mRNA- Sequencing (RNA-Seq) dataset. The “interaction network tool” of omiRas allows for the construction of interaction networks of miRNAs and mRNAs, interrogating the information from several miRNA-mRNA interaction databases. Therefore, we in silico assigned functions to significantly deregulated miRNAs and defined miRNAs implicated in the carcinogenesis of ccRCC [17, 18].

However, the paramount regulator and distinct molecular mechanism of RCC has yet to be evaluated as the complexity of its pathogenesis. In this present paper, we aimed to screen several novel tumor suppressor genes and explore the molecular mechanism of RCC with a computational analysis. The most significant DEGs which served as potential biomarkers may provide a new sight on RCC clinical therapy.

Materials and methods

Dataset collection and Affymetrix microarray

The gene profile of GSE6344 was downloaded from Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo), which was based on the platform of Affymetrix Human Genome U133B Array condensed [19]. This expression dataset was deposited by Gumz ML. A total of 20 samples were derived from 10 ccRCC patients and 10 age- and gender- matched healthy control subjects.

Data preprocessing and differential expression analysis

Then we converted the probe-level data in CEL files into expression measures. For each sample, the expression values of all probes for a given gene were reduced to a single value by taking the average expression value [20]. After that, we imputed missing data and performed quartile data normalization. The multtest package in R was used to identify differentially expressed genes (DEGs) in patients with RCC compared to healthy controls. To circumvent the multi-test problem which might induce too much false positive results, the Benjamini & Hochberg (BH) method was used to adjust the raw P-values into false discovery rate (FDR) [21]. The FDR less than 0.05 and the absolute logFC larger than 1 were chosen as cut-off criteria.

Gene ontology enrichment analysis

To produce a dynamic, controlled vocabulary that can be applied to all eukaryotes, Gene Ontology (GO) analysis has been used frequently in functional study large scale genomic and transcriptional data [22]. DEGs were separated into two sets based on different expression behavior, and then the GO analysis was performed using Gene Set Analysis Toolkit suit [23]. A P-value less than 0.05 was considered statistically significant.
Regulatory network in clear cell renal cell carcinomas

A

GSE6344/GPL96 selected samples

B

GSE6344/21367_atRHC9

C

GSE6344/21367_atRALY1

expression value

expression value

normal ccRCC

normal ccRCC

Figure 1. Normalization of gene expressions in the ccRCC sample and control samples. The midcourt line in the box was the median value. After normalization, the median values were approximated with an average value of 7, which represents that the gene expression values in different samples are at the same levels and comparable.

Figure 2. Interaction networks of DEGs constructed with prePPI. The line represents that the two genes may have an interaction based on their expression, evolution or structures. A. Hierarchical clustering of miRNA in kidney tissue samples. Kidney tissue samples were clustered according to the expression miRNAs between tumor tissue and normal tissue. Data from each miRNA were median centered. Samples are in columns and miRNAs in rows. N normal, T tumor. The P values for these miRNAs were less than 0.05 in tumor tissue compared with normal tissue. B. PPI network based on SLC family.
Construction and analysis of the protein-protein interaction (PPI) network

Protein-protein interaction network is a database of predicted and experimentally determined protein-protein interactions (PPI) using a Bayesian framework that combines structural, functional, evolutionary and expression information for yeast and human [24]. In this analysis, the most significant up- or down-regulated DEGs were screened, and a PPI network was constructed by collecting interactions from prePPI database.

Pathway enrichment analysis

For functional annotation the DEGs, the over-represented KEGG categories in pathways was identified using KAAS [25]. The P-value less than 0.05 was chosen as cut-off criterion. Up and downregulated genes in ccRCC were mapped to functional Gene Ontology (GO) categories using DAVID Bioinformatics Resources 6.7. Genes within enriched (FDR < 0.05) categories were committed to the STRING database to determine protein-protein interactions of their gene products. Additionally, miRNAs that might be causative for the deregulation of genes within the category were detected as described above. PCA and hierarchical clustering of the differentially expressed miRNAs were performed and visualized with R 3.0.2. Networks of genes from the same GO category were visualized with Cytoscape. Visualizations of annotation statistics for each library were taken from omiRas.

Results

Differential gene expression in RCC samples

After normalization, the data was performed differential expression analysis. The genes with P-value less than 0.05 and logFC absolute value over than 1, were considered as DEGs. A total of 521 DEGs were down-regulated and 473 DEGs were up-regulated in RCC samples compared to healthy controls. Among these 984 DEGs, nuclear factor kappa (NF-κB) was the most significantly up-regulated gene and Mitogen-activated protein kinase (MAPK) was down-regulated the most (P=3.1E-04, logFC=7.56; P=3.2E-06, logFC=-9.47, respectively). In addition, cytokine receptors, such as interleukin 17 receptor alpha (IL17RA) and colony stimulating factor 1 receptor (CSF1R), were also up-regulated (P=3.2E-06, logFC=3.17; P=4.6E-04, logFC=2.18, respectively) (Figure 1).

mRNA-protein interaction network and potential miRNA upstream regulators by IPA

Interaction networks of DEGs were constructed as shown in Figure 2. According to the results, interaction relationship of 250 DEGs was obtained. A high level of connectivity was shown in the network constructed based on Mitogen-activated protein kinase (Figure 2). PPI nodes extracellular signal-regulated kinase (ERK), interleukin 6 (IL6), bone morphogenetic protein receptor, type II (BMPR2), P38 MAPK and natural cytotoxicity triggering receptor 1 (NCR1) were of high degrees in the network, all of which were obtained from RCC samples. In addition, the PPI network based on MAPK was also constructed, with imperfect connections due to the absence of several key hubs, as is shown in Figure 2.

Potential biomarkers and critical pathways in RCCs

We performed a functional enrichment analysis using Gene Set Analysis Toolkit and identified 11 remarkable enhanced functions and 13 suppressed functions (Figure 3, FDR < 0.05). The DEGs was classified with GO analysis according to their functions. Among these enhanced functions, the most remarkable up-regulation was immune response based on cytokine and cytokine receptor enrichment (GO: 0006657, FDR = 5.6E-27). The other significant functions included defense response (GO: 0006948, FDR = 6.1E-12), response to wounding (Go: 0009623, FDR = 1.7E-13), leukocyte activation (GO: 0057324, FDR = 1.4E-06), positive regulation of immune system process (GO: 0003214, FDR = 2.6E-12), inflammatory response (GO: 0005658, FDR = 1.4E-07) and so on. Genes, such as toll-like receptor (TLR) family, IL super-family, chemokine (C-X-C motif) ligand 12 (CXCL12), TAP binding protein (TAPBP), C-type lectin domain family 4 member A (CLEC4A), major histocompatibility complex class I (HLA), complement component 3a receptor 1 (C3AR1) and integrin alpha L (ITGAL), were involved in GO analysis.
Regulatory network in clear cell renal cell carcinomas

Figure 3. HGF network in renal cancer samples. Functional enrichment of DEGs between RCC sample and healthy samples (P < 0.05). Among these enhanced functions, the most remarkable up-regulation was immune response based on cytokine and cytokine receptor enrichment (GO: 0006955, FDR = 5.8E-29) and the most evidently downward modulation function in RCC was oxidation reduction which regulated by hormones (GO: 0055114, FDR = 3.9E-36), respectively.
The most evidently downward modulation function in RCC was oxidation reduction which regulated by hormones (GO: 00057216, FDR = 3.7E-24). The other significantly suppressed function included ion transport (GO: 0005421, FDR = 9.7E-05), response to organic substance (GO: 0013021, FDR = 3.0E-05) and generation of precursor metabolites and energy (GO: 0005986, FDR = 7.7E-11). The response to hormone stimulus was also suppressed (GO: 0008627, FDR = 5.4E-01), which in turn weakened the chemical homeostasis.

Cytokine and cytokine receptor pathway enrichment in RCCs

Pathway enrichment analysis of all genes involved in PPI was performed with NF-κB. Only cytokine and cytokine receptor pathway was enriched with a remarkable FDR less than 0.001 in PPI network of MAPK. Genes associated with cytokine and cytokine receptor pathway were chemokine, hematopoietin, WNT, TNF family and TGF-β family proteins genes. Five chemokines genes participated in this pathway, including IL6, NOB1, ERK1/2, JNK, P38 MAPK and Sclcl4. The primary hematopoietins were cytokine receptors, such as SOS, CRK and SLC4A1. TGF-β family proteins, such as PROC, ALB, VEGFRA, E-cadherin, β-catenin and their receptor TGF-β, were also involved in these pathways (Figure 4).

Discussion

Our study links coding- and non-coding transcriptome data of normal and ccRCC tissue from two distinct studies. By the use of several miRNA-mRNA interaction databases available in omiRas we are able to provide new insights into the influence of aberrant miRNA expression on hundreds of deregulated genes [26]. Small non-coding RNA has drawn more attention in the recent years due to their role in the gene transcriptional and posttranscriptional regulation [27]. Nearly, 30% of gene expression in the human body is regulated by miRNAs. Recently, targeted therapies were developed to interfere with the transduction of key signaling pathways or to inhibit the function of tumour-specific molecules in malignant ccRCC [28].

In the present paper, we identified 984 DEGs in RCC patient samples compared to the control samples. DEGs in RCCs were highly associated with immune response, hormone response and defense response, which may play important roles in tumor initialization and migration. Pathway enrichment demonstrates that only cytokine and cytokine receptor pathway was enhanced in RCC patients compared to healthy controls. Genes obtained from the results of pathway enrichment analysis are the subject of our investigation. Among these DEGs, MAPK and NOB1 were the most significantly down-regulated and up-regulated genes, respectively. In addition, cytokine receptors, such as ERK1/2 and P38 MAPK, were also up-regulated.

MAPK signaling pathways can induce either cell proliferation or cell survival depending on the cell type and stimulus, the activation of the MAPK pathway has been associated with renal cancer proliferation [29]. The three main members that integrate the MAPK family in mammalian cells are ERK1/2, JNK and p38 MAPK, which are important in the control of cell differentiation, proliferation and apoptosis [30]. NOB1 protein is a key factor linking the proteasome and cellular growth, and therefore investigation of the NOB1 function will shed some light on the mechanism of growth control by the ubiquitin-proteasome pathway [31]. Perhaps the inhibition of the proteasome leads to stabilization of proteins that increase phosphorylation of the three key components in the MAPK pathway [27].

Others reported that TGF-β1 can also influence signal transduction pathways, such as nuclear factor kappa-B (NF-κB) and PI3K/Akt signal pathway [32, 33]. Although the precise function of TGF-β1 in carcinogenesis is unknown, previous studies showed that TGF-β1 may regulate a set of genes in cell growth and proliferation, which may be important in cancer development and cancer cell proliferation. Due to these functions of E-cadherin, it plays a key tumor suppressor role in suppressing the invasiveness of cancer cells [34, 35]. Cadherin switch is a key change during EMT, during which the normal expression of E-cadherin is replaced by the abnormal expression of N- or P-cadherin [36]. This downregulation of E-cadherin is associated with the release of β-catenin, which then migrates to the nucleus and activates WNT signaling resulting in the EMT and metastasis [37, 38]. Based on the analysis above, E-cadherin
Regulatory network in clear cell renal cell carcinomas
and β-catenin have often been used to monitor the progress of EMT during embryonic development and cancer progression [39, 40].

In conclusion, though the DEGs and relevant genes may provide a new way into the therapy approach for ccRCCs, however, the expression of these genes was not confirmed by real-time PCR and the function in RCCs was not evaluated. The potential biomarkers and hub genes, such as ERK1/2, P38 MAPK, VEGFRA, SCLC4A1 and TGF-β, only provided potential targets for RCC therapy, and further effort to confirm the hypothesis was still in great needed.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Chang Sheng, Department of Urology, Pudong New Area People’s Hospital, 490 South Chuan Huan Road, Shanghai 201200, China. E-mail: shengchang1014@163.com

References


