Identification of CFTR as a novel key gene in chromophobe renal cell carcinoma through bioinformatics analysis

SHENG WANG1,2, ZHI-HONG YU2 and KE-QUN CHAI2

1The Second Clinical Medical College, Zhejiang Chinese Medicine University; 2Department of Oncology, Tongde Hospital of Zhejiang, Hangzhou, Zhejiang 310053, P.R. China

Received November 6, 2018; Accepted May 30, 2019

DOI: 10.3892/ol.2019.10476

Abstract. Chromophobe renal cell carcinoma (chRCC), the third most common histological subtype of RCC, comprises 5-7% of all RCC cases. The aim of the present study was to identify potential biomarkers for chRCC and to examine the underlying mechanisms. A total of 4 profile datasets were downloaded from the Gene Expression Omnibus database to identify differentially expressed genes (DEGs). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses of DEGs were performed with the Database for Annotation, Visualization and Integrated Discovery. A protein-protein interaction (PPI) network was constructed to predict hub genes. Hub gene expression within chRCC across multiple datasets, as well as overall survival, were investigated by utilizing the Oncomine platform and UALCAN dataset, separately. A total of 266 DEGs (88 upregulated genes and 168 downregulated genes) were identified from 4 profile datasets. Integrating the results from the PPI network, Oncomine platform and survival analysis, CFTR was screened as a key factor in the prognosis of chRCC. GO and KEGG analysis revealed that 266 DEGs were mainly enriched in 17 terms and 9 pathways. The present study identified key genes and potential molecular mechanisms underlying the development of chRCC, and CFTR may be a potential prognostic biomarker and novel therapeutic target for chRCC.

Introduction

Renal cell carcinoma (RCC) is globally the most prevalent cancer affecting the kidney in adults (1). It was reported that ~64,000 new cases were diagnosed in 2017 in the USA (1), and this value has risen by 2-4% each year (2). Chromophobe RCC (chRCC) is the third most common histological subtype of RCC (3), comprising 5-7% of all RCC cases (4). Due to advances in technology for the diagnosis and treatment of chRCC, the 5-year survival rate of chRCC is >75% (5) and the outcome is typically favorable when compared with that of other subtypes (6). However, patients with this disease still have a 5-10% probability of eventually developing progression and metastasis (7). Therefore, it is essential to identify tumor-specific biomarkers and the underlying molecular mechanisms of chRCC, which may be conducive to improved risk assessment of the disease, guiding clinical decision-making, and developing novel diagnostic and therapeutic strategies for chRCC.

The molecular pathogenesis of cancer is complex, involving the inactivation and mutation of tumor suppressor genes and the activation of oncogenes (8). Recently, bioinformatics analysis using high-throughput platforms has emerged as an efficacious approach to identifying new targets and comprehending the underlying molecular mechanisms of carcinoma (9). For instance, Cao et al (10) reported that five genes, COL1A2, COL1A1, COL4A1, THBS2 and ITGA5, which they determined to be significantly overexpressed in gastric cancer (GC), were associated with the prognosis of GC and were potential biomarkers and therapeutic targets for GC. Wang et al (11) identified 227 differentially expressed genes (DEGs) between breast cancer and normal breast tissues, and found that the hub gene NDC80 may be a key prognostic factor and potential target.

In the present study, three raw gene chips [GSE6280 (12), GSE11151 (13) and GSE15641 (14)] were downloaded from the NCBI-Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) in order to detect the DEGs between chRCC tissues and normal renal tissues. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis (15) and Gene Ontology (GO) functional annotation analysis (16) was applied. A protein-protein interaction (PPI) network was subsequently generated to identify hub genes associated with chRCC. To further confirm the association between the hub genes and chRCC, Oncomine dataset (https://www.oncomine.org) and UALCAN (http://ualcan.path.uab.edu) analyses were
performed to examine the expression of the hub genes and associated patient survival rates.

Materials and methods

Microarray data. A total of 3 profile datasets (GSE6280, GSE11151 and GSE15641) were downloaded from the GEO database, a public functional genomics database. The platform for GSE6280 and GSE15641 was GPL96, (HG-U133A) Affymetrix Human Genome U133A Array, and the platform for GSE11151 was GPL570, (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array. The raw data consisted of 11 chRCC tissues (1 in GSE6280, 4 in GSE11151 and 6 in GSE15641) and 32 matched normal tissues (6 in GSE6280, 3 in GSE11151 and 23 in GSE15641).

Expression analysis of DEGs. All raw data were processed with the R version 3.5.1 software package (https://www.r-project.org/). The ‘limma’ package (http://www.bioconductor.org/packages/release/bioc/html/limma.html) in R was utilized for data normalization. The Affy package (http://www.bioconductor.org/packages/release/bioc/html/affy.html) was utilized for gene differential expression analysis. Genes with log fold-change (FC) >1 and P<0.05 were considered to be DEGs.

GO enrichment analysis. The Database for Annotation, Visualization and Integrated Discovery (DAVID) (15) (https://david-d.ncifcrf.gov; version 6.8) provides a comprehensive set of functional annotation tools for investigators to better understand the biological significance of certain genes. Based on DAVID, GO analysis, including analysis of cellular component (CC), molecular function (MF) and biological process (BP) terms, was performed. P-values of <0.01 and gene counts of >10 were considered significant thresholds.

KEGG analysis. KOBAS (16) (http://kobas.cbi.pku.edu.cn; ver. 3.0), a web server for gene or protein functional annotation and functional gene set enrichment, was used for pathway enrichment analysis. Pathways with P-values of <0.01 were screened as statistically significant.

PPI network. With the confidence level >0.7 and ‘Homo sapiens’ as a limit, a PPI of DEGs was gathered from the Search Tool for the Retrieval of Interacting Genes/Proteins (17) (https://string-db.org; ver.10.5). The network visualization software CytoScape version 3.6 (https://cytoscape.org) was utilized to generate PPI networks. The top 10 genes were subsequently selected and considered to be hub genes using the plug-in unit CytoHubba.

Expression and survival analysis of hub genes. The Oncomine platform featuring scalability, high quality, consistency and standardized analysis was utilized to investigate hub gene expression within chRCC across multiple datasets. Patients were divided into low- and high-expression groups according to the median gene expression. UALCAN (18), a user-friendly, interactive web resource for analyzing cancer transcriptome data based on The Cancer Genome Atlas dataset, was utilized to construct an overall survival analysis for the hub genes.

Results

Identification of DEGs in chRCC. After normalization, a total of 266 overlapping DEGs (Fig. 1 and Table SI) were identified from 3 profile datasets (GSE6280, GSE11151 and GSE15641), including 88 upregulated genes and 178 downregulated genes (logFC>1 and P<0.05). The heatmaps of the top 20 DEGs and the results of the normalization of each dataset are presented in Fig. 2.

GO enrichment analysis. All DEGs were input into the online tool DAVID to perform GO analysis. The results demonstrated that, for CC, DEGs of chRCC were mainly enriched in 10 terms, including ‘extracellular exosome’, ‘plasma membrane’, ‘extracellular region’ and ‘extracellular matrix’. For MF, DEGs were mainly enriched in 2 terms, namely ‘calcium ion binding’ and ‘heparin binding’, while for BP, DEGs were mainly enriched in 5 terms, namely ‘cell adhesion’, ‘extracellular matrix organization’, ‘skeletal system development’, ‘aging’ and ‘angiogenesis’ (Fig. 3).

KEGG analysis. After gene ID conversion, all DEGs were uploaded to KOBAS to analyze the pathways at the functional level. There were 9 KEGG pathways associated with enriched DEGs, comprising ‘pathways in cancer’ and ‘metabolic pathways’, among others (Fig. 4).

PPI network. In the PPI network (Fig. 5), red, green and violet nodes represent upregulated genes, downregulated genes and other human proteins interacting with DEGs, respectively. Using the plug-in unit cytoHubba, 10 hub genes with the highest degree of interaction were screened (Fig. 6), including 8 upregulated genes and 178 downregulated genes.

Comparison of hub genes across multiple analyses. The results of hub gene expression level analysis in chRCC revealed that the expression of KIT, CFTR and ALDOA had differences among different analysis datasets (Fig. 8; Fig. S1).

Survival analysis. The overall survival analysis of the 10 hub genes demonstrated that only high expression levels of CFTR...
Figure 2. Normalization and cluster heatmaps of the top 20 DEGs in each dataset. (A) Normalization and cluster heatmaps of the top 20 DEGs in GSE6280. (B) Normalization and cluster heatmaps of the top 20 DEGs in GSE11151. (C) Normalization and cluster heatmaps of the top 20 DEGs in GSE15641. DEGs, differentially expressed genes.
were associated with a worse survival rate in patients with chRCC (Fig. 9; Fig. S2).

Discussion

chRCC is the third most common histological subtype of RCC, behind clear cell RCC and papillary RCC (3); it accounts for 5-7% of all RCC cases (4). Although patients with chRCC have a better prognosis compared with other subtypes, the long-term outcomes are highly variable and there is a 5-10% probability of eventually developing metastasis (7). Therefore, it is essential to identify the tumor-specific biomarkers and the underlying molecular mechanisms of chRCC, which may be conducive to developing novel diagnostic and therapeutic strategies for chRCC. Microarray analyses with high-throughput sequencing technologies have been widely used to determine potential diagnostic and therapeutic targets in the progression of cancer (19,20).

In the present study, a total of 266 overlapping DEGs, including 88 upregulated genes and 178 downregulated genes, were identified from 3 profile datasets. GO analysis revealed
that 266 DEGs were mainly enriched in 17 terms, including 'extracellular exosome', 'plasma membrane', 'extracellular region', 'extracellular matrix', 'cell adhesion' and 'extracellular matrix organization'. In addition, 266 DEGs underwent KEGG analysis and were shown to be enriched mainly in 9 pathways. In the PPI network, 10 genes with a high degree of interaction were chosen as hub genes, including 3 upregulated genes (KIT, CFTR and ALDOA) and 7 downregulated genes (DCN, COL3A1, CXCL12, CTGF, LUM, TNC and THBS2).

KIT, a receptor tyrosine kinase, can activate several signaling pathways, including the PI3K-Akt signaling pathway (21). Mutations of KIT are associated with gastrointestinal stromal tumors, lung cancer and other tumor types (22). ALDOA, a member of the class I fructose-bisphosphate aldolase protein family, may contribute to tumorigenesis and the progression of pancreatic and lung cancer (23). DCN plays a vital role in tumor suppression, including a stimulatory effect on autophagy and inflammation, and an inhibitory effect on angiogenesis and tumorigenesis, after binding to multiple cell surface receptors (24). CXCL12 is associated with diverse cellular functions, including immune surveillance, tumor growth and metastasis, and the inflammatory response (25). Tang et al (26) reported that high expression of tenascin C, an extracellular matrix protein, was significantly associated with poor disease-free survival in patients with lung cancer. Thrombospondin 2, as a potent inhibitor of tumor growth and angiogenesis, may be involved in cell adhesion and migration (27).

In order to further verify the association between the 10 hub genes and chRCC, the present study compared the expression of 10 hub genes across multiple datasets using the Oncomine platform; 3 genes (KIT, CFTR and ALDOA) were indicated to have differences among the datasets. Furthermore, the overall survival analysis based on UALCAN revealed that high expression levels of CFTR were associated with a worse survival rate in patients with chRCC. In summary, CFTR may be a potential prognostic biomarker and novel therapeutic target for chRCC.

CFTR, a cAMP-activated chloride channel widely distributed in the epithelial cells of various tissues (28), plays an important role in maintaining cell homeostasis and is associated with metabolism (29). Mutations in CFTR are responsible for regulation of epithelial ion and water transport and fluid homeostasis, which affects the epithelial tissue of various organ systems, including the urogenital, respiratory and gastrointestinal systems (30). In addition, CFTR mutation increases the risk of various types of cancer, including lung, breast and colon cancer (31). Xu et al (32) reported that CFTR could promote the aggression of ovarian cancer and that knockdown of CFTR suppressed the aggressive behavior of ovarian cancer. In addition, Peng et al (33) found that higher expression of
CFTR was associated with aggressive behaviors, progression and a poor prognosis in cervical cancer, suggesting that CFTR may be a novel therapeutic target and prognostic indicator for cervical cancer. A number of studies have reported that the
downregulation of CFTR promotes invasion and proliferation and is associated with poor prognosis in several types of cancer, including lung (34), intestinal (35) and esophageal cancer (36). To date, to the best of our knowledge, there have been no studies on the association between CFTR and RCC. Therefore, further experimental investigation is required to examine the influence of CFTR mutations on RCC, both in vivo and in vitro.

In conclusion, based on integrated bioinformatics analysis, the present study identified 266 DEGs. It was indicated that CFTR may be involved in the progression and poor prognosis of chRCC, and that it may function as a novel therapeutic target and prognostic biomarker for chRCC. These results improve our understanding of chRCC at the molecular level. However, further investigation of CFTR both in vivo and in vitro are required to confirm the findings of this study, in order to verify the functions and elucidate the underlying mechanisms of chRCC.

Acknowledgements
Not applicable.

Funding
This study was supported by Ke-qun Chai Inheritance Studio of National Prominent Chinese Medicine Doctor, State Administration of Traditional Chinese Medicine (Hangzhou, China; grant no. 2A21533).

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
SW and KC conceived and designed the study. ZY and SW performed the experiment. SW wrote the paper. KC and ZY reviewed and edited the manuscript. All authors approved the manuscript and agree to be accountable for all aspects of the research with regard to ensuring that the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics approval and consent to participate
Not applicable.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.
References