Prognostic Value of Metabolism-Related Genes and Immune Infiltration in Clear Cell Renal Cell Carcinoma

Hanwen Li 1–4
Zengnan Mo 1–4

1Department of Urology, The First Affiliated Hospital of Guangxi Medical University, Nanning, 530021, Guangxi, People’s Republic of China; 2Center for Genomic and Personalized Medicine, Guangxi Medical University, Nanning, 530021, Guangxi, People’s Republic of China; 3Guangxi Collaborative Innovation Center for Genomic and Personalized Medicine, Nanning, 530021, Guangxi, People’s Republic of China; 4Guangxi Key Laboratory for Genomic and Personalized Medicine, Guangxi Key Laboratory of Colleges and Universities, Nanning, 530021, Guangxi, People’s Republic of China

Background: Clear cell renal cell carcinoma (ccRCC) is one of the most prevalent cancers. Thus, it is warranted to detect the status of metabolism-related genes (MRGs) and infiltrating immune cells in ccRCC progression for the prognosis of ccRCC. This research was designed to establish and verify the prognostic signature of ccRCC using MRGs. In addition, we investigated the potential link between the relative proportion of tumor infiltrated immune cells (TIICs) and ccRCC prognosis.

Methods: Sequencing data of metabolism-related gene sets in ccRCC cases were obtained from The Cancer Genome Atlas database (TCGA) and Gene Expression Omnibus Database (GEO). The R Programming Language software packages were applied for differential analysis of MRGs. First, a univariate Cox regression model was applied to determine the MRGs linked with overall survival (OS). Then, the multivariate Cox regression model was applied to establish the prognostic signature. Finally, the CIBERSORT algorithm was used to determine the proportion of TIICs.

Results: Overall, 286 differentially expressed MRGs were identified in the TCGA dataset. Univariate and multivariate Cox regression models were applied to develop a prognostic signature with six MRGs. The predictive capability of the prognostic signature was further verified by TCGA and GEO database. In addition, RS positively correlated with memory B cells, plasma cells, activated memory CD4+ T cells, follicular helper T cells, regulatory T cells, CD8+ T cells, and M0 macrophages, and were negatively associated with resting memory CD4+ T cells, resting dendritic cells, activated dendritic cells, M2 macrophages, monocytes, resting mast cells, and eosinophils.

Conclusion: Herein, a prognostic signature was developed using MRGs for ccRCC prognosis. The proportion of 22 TIICs in ccRCC and the association between TIICs and clinical outcomes were also determined. The identified genes and cells could guide future targeted therapy and immunotherapy.

Keywords: altered metabolism, immunization, clear cell renal cell carcinoma, prognostic markers

Introduction

Renal cell carcinoma (RCC) is the most frequent renal cancer threatening human health. 1 Clear cell renal cell carcinoma (ccRCC), chromophobe renal cell carcinoma (chRCC), and papillary renal cell carcinoma (pRCC), are the most frequently histopathologic subtypes of RCC. 2,3 However, ccRCC is the most prevalent worldwide. 4 ccRCC is invasive cancer originating predominantly from the proximal renal tubular epithelium. 5 About 140,000 ccRCC-related deaths occur annually
worldwide. Although ccRCC can be successfully treated by surgery or radiofrequency ablation if detected at an early stage, the disease is usually asymptomatic and is occasionally diagnosed by imaging. In contrast, up to 1/3 of patients have metastases, and metastatic ccRCC has a high mortality rate due to poor response to radiotherapy and chemotherapy. For example, tyrosine kinase inhibitors currently used in ccRCC therapy act on vascular endothelial growth factor receptor (VEGFR) and target tumor angiogenesis. For example, tyrosine kinase inhibitors currently used in ccRCC therapy act on vascular endothelial growth factor receptor (VEGFR) and target tumor angiogenesis. Although tyrosine kinase inhibitors have significantly improved ccRCC prognosis, only 11.7% of patients with metastatic ccRCC can survive for five years. Therefore, it is necessary to identify a novel promising therapeutic and prognostic biomarker for the clinical treatment of ccRCC.

Metabolic changes are ubiquitous in several human cancer tissues. For instance, the Warburg effect converts glucose into lactic acid despite the presence of oxygen, causing proliferation and increasing biosynthetic demands of cancer cells. Similar changes occur in the tumor microenvironment (TME) of ccRCC. Multi-omics studies have discovered some metabolic alterations in ccRCC that included the tricarboxylic acid (TCA), pentose phosphate, and phosphoinositide 3-kinase (PI3K) pathways, and it has been shown that changes in specific intermediates of these pathways are associated with poor prognosis. These findings indicate that metabolic alteration is associated with ccRCC development and prognosis. ccRCC is characterized by dysregulation of cellular metabolism changes, including increasing glycolysis and fat deposition. Therefore, metabolic abnormalities provide new insights into ccRCC development, providing novel targets for therapeutic intervention and diagnosis. Interestingly, immunity plays an essential role in the progression of various cancers. The immune escape ability or immune surveillance are associated with effective amplification and metastasis of cancer. However, the effects of different types of immune cells in the initiation and progression of ccRCC are unclear.

Herein, the differentially expressed MRGs were obtained from TCGA, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Ontology (GO) databases. First, the univariate Cox model was applied to pick MRGs associated with the OS of ccRCC. Lasso and multivariate Cox regression models were then applied to construct a prognostic signature with six key MRGs. Finally, we used the CIBERSORT algorithm to estimate the proportion of TIICs in cancer and normal tissues of the TCGA database. The above results were further validated via the GEO database. The objective of this research is to determine therapeutic targets and prognostic biomarkers for ccRCC patients.

Materials and Methods

Raw Data

The transcriptome data and clinical characteristics of 539 ccRCC and 72 normal samples were obtained from the TCGA database (https://portal.gdc.cancer.gov/). GSE53757 and GSE29609 datasets were obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Pathways with gene symbols were downloaded from KEGG (https://www.kegg.jp/kegg/) and GO (http://geneontology.org/). MRGs were extracted for further analysis.

Screening of Differentially Expressed MRGs

Limma package of R software was utilized to screen the differentially expressed MRGs. ggplot2 package was used for the heatmap and volcano plots. The thresholds were set as |logFC| > 1 and false discovery rate as (FDR) < 0.05.

Construction of Metabolism-Related Prognostic Signature

Univariate Cox regression model was used to obtain prognostic MRGs, and Lasso regression and multivariate Cox regression model (stepwise models) were used for further screening. A prognostic signature was then built. We then computed the risk score for each patients as follows: Risk score(RS) = (expression of mRNA1 *Coefficient of mRNA1) + (expression of mRNA2 *Coefficient of mRNA2) + (expression of mRNA3 *Coefficient of mRNA3). ccRCC cases were grouped using the median value of RS, namely high-RS group, and low-RS group. Finally, Kaplan-Meier survival analysis was conducted using the R package survival. The R packages survivalROC, survminer, and rms were used to plot ROC curves, nomogram, and a calibration curve of prognostic signature, respectively. Statistical significance was set at P < 0.05.

TIICs Analysis

The CIBERSORT algorithm was used to estimate the proportion of 22 immune cells by identifying the
cell composition of cancer tissues from their
genetic expression.\textsuperscript{22} CIBERSORT algorithm was also
used to calculate the proportion of immune cells in all
samples in the TCGA-KIRC cohort. Samples
with P<0.05 were screened for further analysis.
Spearman rank correlation was used to analyze the
relationship between risk score (RS) and the proportion
of immune cells. Kaplan-Meier was applied to analyze
the correlation between the proportion of immune
cells and overall survival (OS) and the relationship
between immune cells and the prognostic signature in
cCRCC patients. P <0.05 was considered statistically
significant.

Results
Identification of Differentially Expressed MRGs
Through the metabolic-related pathway in the KEGG and
GO databases, we obtained a total of 940 MRGs. Their
expression matrices were obtained from the TCGA-KIRC
cohort for subsequent differential analysis. A total of 286
differentially expressed MRGs (109 upregulated and 177
downregulated) were obtained from tumor and normal
samples in the TCGA database, using |logFC| > 1 and
FDR< 0.05 as the threshold. These were expressed in
a heatmap (Figures 1A and B). Next, the 286 differentially
expressed MRGs were subjected to pathway enrichment

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Identification of a metabolism-related signature and gene functional enrichment analysis. (A) Volcano plot of DEMRGs: red indicates upregulated MRGs, green indicates downregulated MRGs, and black indicates MRGs that were not significantly differentially expressed. (B) Heatmap of MRGs in KIRC and normal kidney tissues in the TCGA database; Green to red represents the progression from low expression level to high expression level. (C) GO-term function enrichment analysis of MRGs. BP: biological process, CC: cellular component, MF: molecular function. (D) KEGG analysis of MRGs.}
\end{figure}
analysis in GO and KEGG database, which pointed out that the MRGs were highly enriched in small molecule catabolic process, mitochondrial matrix, and coenzyme binding (Figure 1C). KEGG pathway analysis of carbon, purine, arachidonic acid, and retinal metabolism is shown in Figure 1D. Taken together, functional enrichment analyses demonstrated that the differentially expressed MRGs were invariably enriched in the catabolic process of nutrients and mitochondria-related physiological functions.

Construction of Metabolism-Related Prognostic Signature

Univariate Cox regression models were applied to obtain differentially expressed MRGs affecting patient’s OS. A total of 96 MEGs were identified. Lasso regression (Figure 2A and B) and multivariate Cox regression models (Figure 2C) were then applied to confirm the key MRGs. Six MRGs, including PLA2G6, PAFAH2, PYCR1, CYP3A7, RRM2, and ALDH6A1, were identified. Afterward, a metabolism-related prognostic signature was constructed using the six MRGs.

Verification of the Prognostic Signature Using TCGA and GEO Databases

A nomogram was constructed to predict the OS of ccRCC (Figure 3A). The calibration curve of the nomogram showed the predicted survival rates for each OS are highly consistent with observed data. (Figure 3B–D). The RS of ccRCC cases in the TCGA was calculated, dividing the cases into either high- or low-RS groups according to the median value. Heatmap showing the expression level of genes in the metabolism-related prognostic signature, RS distribution, and survival status of ccRCC cases are demonstrated in Figure 3E–G. Kaplan–Meier survival analysis showed that the clinical outcome of the patients was worse in the high-RS group than in the low-RS group in the TCGA database (P < 0.001; Figure 3H). The area under the curve (AUC) for OS of the prognostic signature was 0.740, 0.729, and 0.756, indicating high sensitivity and specificity (Figure 3I). Similar results were further validated using the KIRC cases in GEO databases (P = 0.008; Figure 4). Then we performed univariate and multivariate Cox regression analysis models to evaluate the independent prognostic ability of the signature. The results point out that age, grade, stage, and riskScore were markedly associated with OS (Figure 5A; p < 0.001), and these factors could present as independent prognostic predictor (Figure 5B; p < 0.01).

Analysis of Critical MRGs Expression Levels in the Prognostic Signature

mRNA expression levels were analyzed to determine the transcription level of the six MRGs using the TCGA database. PLA2G6, PYCR1, and RRM2 expression levels were significantly upregulated in tumor tissues compared to normal tissues. However, ALDH6A1, CYP3A7, and PAFAH2 expression levels were significantly downregulated in normal tissues (Figure 6). The high expression level of PYCR1, PLA2G6, and RRM2 was associated with markedly poorer OS, while the high expression level of ALDH6A1, CYP3A7, and PAFAH2 was correlated with better survival outcomes (Figure 7). Overall, these results reveal that these MRGs play distinct roles in ccRCC progression.

The Landscape of Immune Infiltration in ccRCC

The landscape of TIICs in all ccRCC cases was described by applying the CIBERSORT algorithm. A heatmap plot illustrating the relative abundance of TIIC subpopulations is shown in Figure 8A. Differential analysis indicated that tumor tissues harbored a higher level of naive B cells, plasma cells, CD8+ T cells, naive CD4+ T cells, resting memory CD4+ T cells, follicular helper T cells, regulatory T cells (Tregs), M0 macrophages, M1 macrophages, and resting dendritic cells than the normal tissues. However, naive B cells, naive CD4+ T cells, resting memory CD4+ T cells, monocytes, resting dendritic cells, and resting mast cells fractions were relatively lower in the tumor samples than in the normal tissues (Figure 8B). Notably, the relative proportions of TIICs were correlated in tumor tissues (Figure 8C). CD8+ T cells and follicular helper T cells had the most significant positive correlation (Pearson correlation = 0.58), while CD8+ T and resting memory CD4+ T cells had the most significant negative correlation (Pearson correlation = 0.62). Another point worth mentioning is that CD8+ T cells also had a moderate negative correlation with M2 macrophages (Pearson correlation = 0.51). The above results demonstrate that various immune cells interact or constrain each other in the tumor microenvironment of ccRCC, changing dynamically with tumor progression.

Clinical Evaluation by Infiltrating Immune Cell Subsets

Each case was then classified into low- and high-RS groups according to the median level of TIICs. Resting
dendritic and mast cells were significantly associated with a favorable outcome (Figure 9). In contrast, follicular helper T cells and Tregs were associated with a poorer prognosis, suggesting that these TIICs may be vital in inhibiting ccRCC progression and anti-tumor response in the tumor microenvironment.

Correlation Analysis Between the Prognostic Signature and TIICs

Pearson correlation was applied to analyze the correlation between RS and TIICs. On the one hand, RS was positively correlated with the relative level of memory B cells, M0 macrophages, plasma cells, activated
Figure 3 Development of metabolism-related prognostic signature of ccRCC patients in the TCGA database. (A) A nomogram based on the signature in ccRCC patients at 1, 3, and 5 years. (B–D) Calibration curves of the nomogram for the signature at 1, 3, and 5 years. (E) Heatmap of the metabolism-related gene expression in the TCGA database. (F) Risk plot showing each point based on the risk score. (G) Survival status of ccRCC patients. (H) Kaplan-Meier plot for overall survival (OS) based on risk score of the six MRGs-based signature of ccRCC patients in the TCGA database. (I) Time-independent ROC analysis of risk scores for OS prediction in the TCGA database.
memory CD4+ T cells, CD8+ T cells, follicular helper T cells, and Tregs (Figure 10A–G). However, on the other hand, RS was negatively correlated with the relative level of activated dendritic cells, resting dendritic cells, eosinophils, M2 macrophages, resting mast cells, monocytes, and resting memory CD4+ T cells (Figure 10H–N). Indeed, the prognostic signature was significantly associated with TIICs (either positive or negative correlation).

**Discussion**

cRCC is one of the most common neoplasms of the kidney endangering human health worldwide. Mounting evidence has demonstrated that metabolic changes occur
Figure 6: MRGs expression in the TCGA database ($P < 0.05$). (A) Expression level of PLA2G6, (B) PYCR1, (C) RRM2, (D) ALDH6A1, (E) CYP3A7, and (F) PAFAH2.

Figure 7: The relationship between MRGs expression and OS of ccRCC patients in the TCGA database ($P < 0.05$). (A) The relationship between ALDH6A1 and OS, (B) CYP3A7 and OS, (C) PAFAH2 and OS, (D) PYCR1 and OS, (E) PLA2G6 and OS, and (F) RRM2 and OS.
Figure 8 The immune infiltration in ccRCC. (A) Heatmap of the 22 immune cell proportions. (B) Violin plot visualizing the difference in immune infiltration between normal and tumor tissues. (C) Heatmap depicting correlations between infiltrated immune cells in tumor tissues.
Undoubtedly, metabolic-related expression profiles play a critical role in tumor occurrence and development. Previous studies have revealed that metabolic alteration is correlated with ccRCC development. Therefore, new predictive signatures based on MRGs are needed to assess ccRCC prognosis. Herein, a reliable metabolic predictive signature was developed using six MRGs (PLA2G6, PYCR1, RRM2, PAFAH2, CYP3A7, and ALDH6A1) via the TCGA-KIRC cohort. The signature was validated using a dataset in the GEO database.

PLA2G6 is mainly located in mitochondria and is involved in cell differentiation, proliferation, apoptosis, signal transduction, and membrane phospholipid remodeling. Previous researches have shown that PLA2G6 alterations are associated with various autosomal recessive neurodegenerative disorders. PLA2G6 alterations are rare events in kidney cancers and only occurs in small-cell variant renal oncocytoma. In this study, the expression level of PLA2G6 was upregulated in ccRCC tissues and was negatively correlated with OS. This observation indicates PLA2G6 could be a prognostic risk factor for ccRCC and a potential therapeutic target for ccRCC patients. PYCR1 promotes proline synthesis, increases mitochondrial reactive oxygen species (ROS) production, and promotes epithelial-mesenchymal transition (EMT) through induction of mitochondrial ion protease, thus activating angiogenesis and M2 macrophage polarization.

**Figure 9** Survival plots of the median of immune cell subsets with P-value. (A–D) Key immune cells influencing patients’ survival based on the median level of immune cells.
PYCR1 upregulation is correlated with shorter OS time in RCC patients, consistent with this study. RRM2 is a small subunit of Ribonucleotide Reductase and is essential for DNA synthesis and cell proliferation. Research has shown that RRM2 expression has been identified in numerous cancers, such as kidney cancer, colorectal cancer, breast cancer, and prostate cancer, and RRM2 overexpression is associated with a poorer outcome in these cancers. The carcinogenic effect of RRM2 is related to the promotion of EMT and angiogenesis. PAFAH2 encodes an oxidized-phospholipid-selective phospholipase A2, which participates in the metabolism of esterified 8-isoprostaglandin F2α and protects tissues against damage from oxidative stress injury. PAFAH2 expression is downregulated in ccRCC tissues. However, no literature has described the mechanisms of PAFAH2 in improving ccRCC outcome to play a protective role against cancer. It has been hypothesized that the

Figure 10 Scatter plot showing the relationship between infiltrated immune cells and prognostic model (P < 0.05). (A–G) Positive correlations between infiltrated immune cells and prognostic model. (H–N) Negative correlations between infiltrated immune cells and prognostic model.
A large-scale analysis targeting ccRCC tumors confirmed T cells are the main TIICs within the TME. Overall, higher CD8+ T cell infiltration in ccRCC is associated with worse outcomes, implying that the infiltrating CD8+ T cells might play a critical role in TME. Multiple ccRCC studies have confirmed that these infiltrating CD8+ T cells can be suppressed to become dysfunctional. ccRCC typically progresses with T cell infiltration. T cell activation influences the prognosis of ccRCC patients, explaining the positive correlation between activated memory CD4+ T cells and RS and the negative correlation between resting memory CD4+ T cells and RS. Previous research on CD4+ tumor-infiltrating T cells in ccRCC provided evidence of CD4+ T cell-mediated anti-tumor immunity. However, CD4+ T cells are also associated with an immunosuppressed state of the ccRCC. Therefore, the underlying mechanisms of memory CD4+ T cells in ccRCC progression call for further investigation.

Mast cells and eosinophils are primarily pertinent to allergic diseases. Herein, they were associated with a lower risk. Indeed, previous studies have shown that mast cells have pro-tumor activity in some cancers, such as melanoma and cervical cancer. Besides, they have anti-tumor activity in breast cancer and prostate cancer. Nonetheless, the underlying molecular mechanisms of these phenomena warrant further investigations.

**Conclusions**

In conclusion, a prognostic model that independently predicts the prognosis of ccRCC patients was developed using six MRGs. The dedifferentiation algorithm showed significant differences in the cell composition of the TIICs of ccRCC and the association between TIICs and clinical outcomes. These genes and cells may guide future targeted therapies and immunotherapies. Nevertheless, this study did not fully assess the six genes used for formulating the model and the 14 associated infiltrating immune cells. Therefore, further validation is required since our results are based on public databases and computational algorithms.

**Funding**

This work was supported by grants from the National Key Research and Development Program of China (2017YFC0908000), the National Key Research and Development Project (2020YFA0113200), the Major Project of Guangxi Innovation Driven (AA18118016), the Guangxi key Laboratory for Genomic and Personalized Medicine (20-065-33), Natural Science Foundation of China (81770759, 82060145, 31970814).

**Disclosure**

All authors declare no conflict of interests.

**References**


