Analysis of Ferroptosis-Related Gene Expression and Prognostic Factors of Renal Clear Cell Carcinoma Based on TCGA Database

Sijia Ma¹
Mingming Zhao¹
Jiao Fan¹
Meiying Chang¹
Zhiyu Pan¹
Ziyan Zhang¹
Shunxuan Xue¹
Qi Li²
Yu Zhang¹

¹Department of Nephrology, Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing, People’s Republic of China; ²Department of Clinical Laboratory, Xiyuan Hospital, China Academy of Chinese Medical Sciences, Beijing, People’s Republic of China

Introduction: Renal clear cell carcinoma (ccRCC) is a common tumor of the urinary system, most of which are primary malignant tumors with high metastatic rate and remaining incurable. Ferroptosis is a newly discovered form of iron-dependent programmed cell necrosis in recent years, which is inextricably linked to the occurrence and development of tumors progression. Due to the complexity of the interaction between genes in ccRCC, the research on the pathogenesis of ccRCC is still not remarkably accurate. Therefore, whether ferroptosis-related genes (FRGs) can play a role in predicting prognosis in ccRCC needs to be discussed.

Methods: We entered the Cancer Genome Mapping Project (TCGA) database and downloaded the relevant genes and clinical research data of ccRCC patients. Lasso Cox regression was used to construct a multi-gene prognostic model in the TCGA cohort. R language software was used for drawing pictures related to our study.

Results: Most of the genes involved in ferroptosis (86.2%) existing differences between the tumor and normal tissues in the TCGA public gene database. In terms of univariate Cox regression analysis, 20 differentially expressed genes (DEGs) were associated with prognosis and survival (P<0.05). A prognostic model of 12 FRGs was constructed, and patients were segmented into two different groups depending on how risky they are. Considering overall survival, the high-risk group is dramatically lower than the low-risk group (P<0.001). In multivariate Cox regression analysis, risk scores and stage turned out be an independent prognostic factor (P<0.001). GO and KEGG analysis and ssGSEA analysis of DEGs revealed that these genes were related to immune-related pathways (P<0.05).

Conclusion: This study established and identified the changes in FRGs expression and prognostic factors of ccRCC, which can be helpful for prognosis evaluation and clinical treatment of this disease.

Keywords: renal clear cell carcinoma, ferroptosis, TCGA database, the prognosis

Introduction
The incidence and mortality of renal cell carcinoma (RCC) are rising year after year recently, which seriously affects human health. RCC as a common and refractory malignant tumor of urinary, is among the 10 most common cancers in human,¹ which comprises a diverse group of malignancies arising from the nephron, accounting for about 3% of all malignancies.² The incidence of RCC varies regionally, Northern Europe and North America encountered with the highest rates whereas Asia and Africa with the lowest. Approximately 3–5% of the incidence of RCC is believed to be genetic.³ The main types of RCC are divided
into three broad categories, including clear cell RCC (ccRCC), papillary RCC, and chromophobe RCC. The most familiar type of RCC is ccRCC, accounting for about 75% of the pathological types of renal carcinoma, and has the worst prognosis in the common epithelial neoplasms of the kidney. The incidence of RCC is rising year by year globally. Although early diagnosis and successful urological surgery with partial or total nephrectomy can save lives, only about 10% of patients with RCC show characteristic clinical symptoms. Unfortunately, more than 60% of patients are incidentally found on routine ultrasound. There are many causes for the high mortality rate in RCC, the main of them is that traditional chemotherapy and radiotherapy are unavailability to a large extent in treating all RCC subtypes. Therefore, tumor screening and prevention measures are largely dependent on the causes of tumor occurrence. Between 25% and 30% of ccRCC patients are found to have metastases at the time of their first visit. Although the mechanism of cancer occurrence and development has been deeply researched, the pathogenesis and carcinogenesis of ccRCC remain unclear. Therefore, given the high morbidity and mortality of ccRCC, exploration of molecular markers with prognostic value for immune response in patients with ccRCC, as well as finding the development of novel prognostic models remains critical.

Ferroptosis is a specific way of regulated cell death (RCD) that relies on oxidized iron, reactive oxygen species and lipid peroxidation products that can be accumulated to lethal levels in the process. As a newly discovered form of RCD, ferroptosis is involved in human development, immunity, aging, and various pathological processes, has a pivotal position in maintaining normal cell and tissue survival. Emerging evidence suggests that several carcinogenic pathways are involved in ferroptosis, which plays an important role in inhibiting tumorigenesis and being explored continuously as a new therapeutic approach to eliminate obstinate cancer cells. Ferroptosis also plays a key role in inhibiting tumorigenesis by removing cells from the environment that are deficient in nutrients or have been impaired by infections or suffered external pressures. In recent years, ferroptosis is inseparable from a variety of kidney diseases, such as acute kidney injury, polycystic kidney disease, and RCC. Research showed that RCC is particularly impressionable to ferroptosis, ccRCC can become sensitive to ferroptosis cell death by inhibiting glutathione synthesis and ultimately preventing tumor growth, such as silence of GPX4.

There has new evidence that immunotherapy is a premier treatment of choice for metastatic ccRCC. Studies have shown that increased ferroptosis and immune activation can be synergically enhanced in killing cancer cells. For example, depletion of NCOA4, an autophagy-some gene involved in ferritin phagocytosis, can weaken ferroptosis by eliminating iron accumulation. However, reports about the role of FGRs in the prognosis of patients with ccRCC are still in the minority.

In this study, the mRNA gene expression profile of ccRCC patients in TCGA database was analyzed, and the influence of age, gender and other significantly differentially expressed genes related to ferroptosis on the prognosis of patients was explored. The purpose is to provide a new reference index for prognostic risk stratification, prognosis assessment and treatment strategy selection in patients with ccRCC.

Materials and Methods

Database

After login TCGA website, we organized a choice for kidney renal clear cell carcinoma (KIRC), which is equivalent to ccRCC, download as of February 1, 2021 (https://portal.gdc.cancer.gov), 530 cases of transcriptome patients with KIRC data. Perl software was used to convert the genetic data of related samples into matrix files. According to the Ensembl database (http://www.ensembl.org) convert “Ensembl_Stable_id” into “Gene Symbol” in the mRNA file, and then organized it into a unique one. There involved 72 clinically relevant normal samples and 539 tumor samples. The samples were mainly from people with KIRC in America. Use normalized reads to count values. Since the TCGA database is publicly accessible to obtain open access data, no ethical review or approval is required for this study. TCGA’s data is publicly queryable. The study followed TCGA’s data access policy and publication guidelines.

Construction and Verification of a Novel Prognostic FRG Survival Model

A total of 60 genes related to ferroptosis were obtained by reading the literature. The “Limma” package of R language was applied to process the relevant data. We performed Wilcox differential analysis on FRGs from normal and tumor samples, and set false detection rates (FDR) as 0.05, Benjamini and Hochberg (BH) method...
was performed to estimate the FDR value. After the correction and adjustment of P value by BH method, the corrected threshold \( p_{adj}<0.05 \), and the information of differentially expressed genes (DEGs) was preserved after filtering. The overall survival (OS) rates for different genes were calculated based on univariate Cox analysis, and the “Survival” of R language software was performed to establish a prognosis model. Used the “Venn” package in R language to draw the Venn map of ferroptosis-related DEGs and prognostic genes, as the same time preserved the related information of the intersection genes. Prognostic risk characteristics were assessed by using the “GLMnet” package and the “Survival” package of R language, based on the minimum absolute contraction and selection operator (Lasso) method.\(^7\) The formula of specific risk score is as follows: \( \text{esum} = (\text{expression of the selected FRG} \times \text{corresponding coefficient}) \), the gene expression coefficients were output during cross validation after adjusting for minimum error points.\(^6\) Then, the risk score of each ccRCC patient was calculated on the basis of the formula, and the median score was taken as the cut-off value to divide them into low-risk and high-risk groups. “Survminer” and “Survival” of R language were used to analyze the data of patients in the high and low-risk groups, and survival curves were drawn to analyze whether there was a difference in survival rate between the two groups. The 3-year ROC curve was plotted with the “timeRoc” of R package to evaluate the accuracy of related genes in estimating patients’ survival time. Based on the different gene expression of patients in the two groups, the R package “Rtsne” and “ggplot2” were used to reduce the dimension of the data, PCA and t-SNE analysis were performed to investigate whether there were distribution differences between the two risk groups of patients.

**GO and KEGG Enrichment Analysis of Related Differential Genes**

The software package “Clusterprofiler” of R was applied to perform gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) results for visualization of the different genes between the two different risk patients, and relevant graphs were drawn subsequently. The “GSVA” and “GSEABASE” of R package were used to score the relevant differential genes, explore the different immune cells and immune functions of the two groups, and draw the relevant graphs.\(^9\)

**Statistical Analysis**

To determine whether there are differences in gene expression between tumor and normal tissues, we used Student’s \( t \)-test to analyze the corresponding tissue. We aimed to compare the difference of gene proportion among patients with different tissues by Chi-square test. SsGSEA scores of immune cells or pathways were compared between the two risk level groups by using Mann–Whitney test with adjusted \( p \) values after using BH method. The OS rate between the patients with different risk levels was identified by Kaplan–Meier method and bilateral Log rank test. In order to confirm the independent predictors of OS, univariate and multivariate Cox regression analyses were utilized. R software (Version 4.0) was used to complete the statistical analysis, the result was considered to statistically significant while \( P<0.05 \), and all \( P \) values were double-tailed distribution.

**Results**

To systematically describe our study, we developed a flow chart (Figure 1). As a result, we enrolled a total of 537 KIRC patients from the TCGA-LIHC cohort finally. The detailed clinical features of these patients, such as age, sex, and tumor stage are introduced in Table 1.

**Identification of Differential Prognostic Genes Associated with Ferroptosis**

According to various relevant indicators, the FRGs in the TCGA cohort were screened out and identified in Figure 2. Notably, the gene expression of FRGs was discrepantly expressed between tumor tissues and neighbouring normal tissues, which encountered statistical significance. Most of the genes associated with ferroptosis were differentially expressed (50/58, 86.2%) (Figure 2A), considering the criteria for FDR \(<0.05 \) and \([\log_2 \text{(fold change)}]>1 \), we finally obtained 26 regulated FRGs, heatmaps of these FRGs between normal tissue and ccRCC were illustrated (Figure 2B), with blue squares representing the low expressed genes and red for the high expressed genes. After univariate Cox regression analysis, 26 genes of them were confirmed to be related to OS (Figure 2C). Therefore, these 26 genes were included in the subsequent prognostic analysis. The interaction network diagram among these genes hinted out \( ACACA, FADS2, HMGCR \) and \( ACSL3 \) as the central genes (obtained by MCC algorithm in Cytoscape software) (Figure 2D). Furthermore, Figure 2E shows the correlations between these genes.
Construction of the TCGA Prognostic Model

According to the expression arrays of 26 genes obtained from the previous analysis, the prognosis model was established by Lasso Cox regression analysis. After regression analysis, the optimal 12 genes and their corresponding correlation coefficients were obtained (Figure 3). The risk score was expressed as the corresponding gene multiplied by its coefficient, such as \((\text{AKR1C1} \times -0.0527) + (\text{CARS1} \times 0.7691) + \ldots + (\text{NOX1} \times 0.2608)\). After removing samples with incomplete clinical data, patients were grouped into two different risk groups, according to the median risk score in the available data, both high- and low-risk groups involved 263 cases (Figure 3A). As illustrated in Figure 3B, patients with a high-risk score had shorter survival times but higher mortality rates than those with low-risk scores. When another survival analysis was performed, for instance, the likelihood of survival shown by the Kaplan–Meier curve, survival was significantly higher in the low-risk group than in the high-risk group (P<0.001, Figure 3C). The ROC curve was applied to evaluate the predictive property of the OS risk score. The area under the curve (AUC) was 0.782 within 1 year, 0.740 within 2 years, and 0.743 within 3 years (Figure 3D). After PCA and T-SNE analysis, FRGs were reduced their dimension, then we found that patients in different risk groups were distributed in two different orientations (Figure 3E and F).

Independent Prognostic Analysis of 12 Genes

To determine whether risk score could be used as an overall prognostic predictor independent of survival, we utilized univariate and multivariate Cox regression analyses to evaluate existing clinical indicators. In univariate Cox regression analysis, age group (>65 vs ≤65 years) showed HR = 1.638, 95% CI = 1.210–2.218, P = 0.001; gender (male/female) group, HR = 0.945, 95% CI = 0.691–1.291, P = 0.721; grade group (G3/G4 vs G1/G2), HR = 2.647, 95% CI = 1.872–3.742; stage group (III/IV vs I/II), HR = 3.938, 95% CI = 2.854–5.432, P < 0.001; risk group (high vs low), HR=3.469, 95% CI = 2.764–4.356.
In multivariate Cox regression analysis, age group (>65 vs ≤65 years) showed HR = 1.455, 95% CI = 1.073–1.971, P = 0.016; grade group (G3/G4 vs G1/G2), HR = 1.394, 95% CI = 0.959–2.026, P = 0.081; stage group (III/IV vs I/II), HR = 2.584, 95% CI = 1.835–3.640, P < 0.001; risk group (high vs low), HR = 2.624, 95% CI = 2.039–3.377 (P < 0.001, Figure 4B). In conclusion, after adjusting for other confounding factors, age, stage and risk scores can be confirmed as the independent prognostic factor.

Functional Analysis
In order to understand the biological functions and mechanisms between DEGs and ccRCC, we used GO enrichment analysis and KEGG pathway analysis on the DEGs underlying the different risk groups in this study. GO enrichment analysis divided into three parts, including cellular component (CC), molecule function (MF) and biological processes (BP), and the main biological processes involved them include a series of immune responses as shown in Figure 5A. (P. adjust < 0.05). Results of the KEGG pathway analysis revealed that DEGs were most significantly enriched in the following several pathways: cytokine–cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, IL-17 signaling pathway (Figure 5B, P. adjust <0.05). Among them, the IL-17 signaling pathway, complement and coagulation cascades, NF-kappa B signaling pathway are obviously related to the mechanism of inflammation and immune response.

To explore whether there were differences in immune-related states between the two groups of patients at different risk, ssGSEA analysis was used to score different immune cells and related immune pathways similarly (Figure 6). In terms of antigen presentation process, there were significant differences in ADCs, Macrophages, Tfh, APC_CO_stimulation, Cytolytic_activity and inflammation-promoting between the two different risk groups, and the high-score group has higher immunization grades than the low-score group. Furthermore, on ssGSEA analysis, the high-risk group also has significantly higher scores of inflammation-promoting and para-inflammation than those in the low-risk group, which correspond to the enrichment of KEGG in the three inflammation-related pathways. In conclusion, among the results with significant differences, except the scores of iDCs, Mast_cells and Type_II_IFN_Response in the high-risk group were less than those in the low-risk group, all other scores were higher in the former than those in the latter.

Discussion
In the existing study, we integrated 60 ferroptosis-related genes (FRGs) through database mining and literature search, and systematically analysed the 60 FRGs expression in ccRCC tissue and its relationship with prognosis of survival through Cox and Lasso regression model, ultimately selected 12 pivotal FRGs related to prognosis, successfully build a ccRCC prognosis prediction model. The result demonstrates that the high-risk group encountered obviously lower OS time than the low-risk group, predicts ccRCC patients within 3 years of survival rate of the area under the ROC curve AUC were greater than 0.7, suggests the model has a valid sensitivity and specificity of ideal model. Prognostic models of 12 FRGs were constructed and validated in an external cohort.

In addition, immune related pathways were found to be enriched in functional analysis. Some genes may induce or inhibit ferroptosis in tumors by regulating certain inhibitors, but their relevance to patient survival outcomes remains unclear. Notably, more than three-quarters (86.2%) of
Figure 2 Identification of candidate FGRs in TCGA cohort. (A) The Venn diagram demonstrated the relationship between prognostic genes and DEGs. (B) The heat map shows the expression difference of 26 intersection genes in different tissues. (C) Univariate regression forest plot of 26 intersection gene expression. (D) The network of 26 candidate genes downloaded from STRING database, stray genes has been removed. (E) Demonstration of candidate gene-related networks, different colors represent different correlation coefficients.
FRGs had differential expression in ccRCC tissues and adjacent non-tumor tissues, and 52% of the DEGs were significantly differentially expressed in univariate regression Cox analysis related to survival prognosis. This suggests that ferroptosis has a potential role in ccRCC. Therefore, it is meaningful to model the prognosis of ccRCC by FRGs.

Figure 3 Construction of survival and prognosis of the 12FGRs in the TCGA cohort. (A) Risk score distribution of ccRCC patients. (B) The distributions of OS-related indicators. (C) Kaplan–Meier curves for different risk groups. (D) The time-dependent ROC curves of patients with different risk levels, AUC of the curves indicated that the model has high accuracy. (E) PCA plot and (F) t-SNE analysis of the prognosis model confirmed that the distribution of patients in the two risk groups was discrete.
The prognosis model constructed in this study was composed of 12 FGRs, including AKR1C1, CARS1, CD44, DPP4, FANCD2, HMGCR, MT1G, NCOA4, SLC7A11, ACACA, NOX1, GOT1. There are relevant studies to confirm that AKR1C1 expression level is strongly correlated between AKR1C1 expression level, cell malignant transformation and resistance to cancer therapy, and AKR1C family serves as a new therapeutic target holding great promise in cancer treatment. CD44 expression has the effect of suppressed ferroptosis through relying upon OTUB1 manner in cancer cells. It is located in the basal membrane and lateral membrane of the kidney, and is a marker for cancer stem cell-like cells in RCC. In healthy individuals, CD44 is almost not expressed in the kidney, but is highly expressed in ccRCC, which may be related to the elevated expression levels of MMP2 and MMP9. Owing to DPP4 activity in human urine is positively correlated with a number of indicators for evaluating the renal function, such as serum creatinine levels, proteinuria, therefore DPP4 activity in urinary could be viewed as a biomarker to predict and judge the occurrence and progression of renal disease. Ferroptosis was inhibited when DPP4 activity was blocked by the tumor suppressor p53. As a nucleoprotein involved in DNA damage repair, FANCD2 can protect bone marrow stromal cells (BMSCs) from ferroptosis-mediated injury. Mechanically, FANCD2 mainly inhibits iron accumulation and lipid peroxidation during fertilization through transcription-dependent and non-dependent mechanisms. FANCD2-deficient BMSCs were more sensitive to ferroptosis induced by erastin and had increased DNA damage.

The expression of MT1G in renal clear cell carcinoma cells was higher than that in other histologic tumor subtypes (P<0.05) and may be an important prognostic factor for ccRCC. It has been reported that when ferritin phagocytosis of NCO4-related genes in ccRCC is in a low expression state, it can reduce immune cell infiltration and lead to impaired IFN-γ receptor signaling pathway. Therefore, NCOA4 is expected to serve as a new biomarker to identify those patients who are potentially suitable for the ferroptosis-induction therapy or combined immunotherapy. Cysteine, as a precursor of glutathione, can be converted from cysteine transported into the cell by the Xc– system. This system is constituted of the light-chain subunit SLC7A11 and the heavy-chain subunit SLC3A2. Extracellular cystine is imported with intracellular glutamate release, on account of SLC7A11 is specific for the Xc– system, SLC7A11 act as an important transporter to maintain the homeostasis of cystine in cells. SLC7A11 also has been found to be overexpressed in renal collecting duct carcinoma compared with matched non-tumor tissues. Some studies have shown that high expression of FANCD2, CARS and SLC7A11, and low expression of HMGCR and NCOA4 are associated with high-risk ccRCC patients similarly, which is the same as the results of our study. In conclusion, these genes may promote or inhibit the progression of ccRCC to some extent, but whether all of these genes can influence the prognosis of ccRCC patients through ferroptosis mechanism remains to be further investigated.

The mechanism of tumor susceptibility to ferroptosis has been a hot research area in the past literature, but the potential regulatory mechanism between tumor immunity and ferroptosis still remains unclear. When GO and KEGG enrichment analysis on the differential risk DEG genes was performed, we found that a number of biological pathways and functions involved in immunity were associated with. Therefore, we can conclude that ferroptosis is tightly

![Figure 4](https://doi.org/10.2147/IJGM.S323511)
linked to tumor immunity, or ferroptosis may affect the progress of ccRCC through immune mechanism. Notably, in the ssGSEA analysis, we found significant differences in the antigen presentation process between the high- and low-risk groups, which may be caused by the ferroptosis cells releasing some signals to attract the antigen presenting cells; thus, the antigen presentation process occurs, leading to the activation of the immune stress response. Recent research also suggests that the immune system may prevent tumorigenesis in part through ferroptosis mechanisms. In addition, CD8$^+$ T cell, macrophages were remarkably higher in the high-risk group, which is possibly because CD8$^+$ T cell accumulation in RCC is associated with disease progression. CD8$^+$ T cells play a central role in immunotherapy, and higher CD8$^+$ T cell levels are associated with poorer prognosis in patients with ccRCC; macrophage migration inhibitory factor (MIF), as a cytokine mediating the interaction between malignant cells and innate immune system, has attracted much attention in recent years for its role in tumorigenesis. Negative expression of MIF has been proved to be an independent prognostic marker in ccRCC patients. In addition, higher risk scores were strongly associated with inflammation. In KEGG functional analysis, FRGs were also enriched in some inflammation-related pathways, such as IL-17 and NF-kB signal pathway. Studies have shown that IL-17 gene and protein expression

Figure 5 GO and KEGG analysis results in the TCGA cohort. (A) GO enrichment analysis revealed the biological processes and molecular functions involved in 26 prognostic-related FRGs; (B) KEGG analysis shown the signaling pathways involved in 26 prognostic-related FRGs.
Figure 6 The boxplot of ssGSEA scores. (A) Score of 16 immune cells in different risk patients with ccRCC. (B) Different expression in different risk level groups of 13 immune-related functions; blue represents low risk, red represents high risk; *P< 0.05; **P< 0.01; ***P< 0.001.
are significantly elevated in the blood and tumors of patients with ccRCC, and some genes promote the progression of ccRCC by regulating the NF-κB signaling pathway. Rhein has been confirmed that it has the ability to inhibit the proliferation of RCC cells through the NF-κB pathway and reduce the migration and invasion of RCC cells. In conclusion, CD8+ T cell, macrophages, and inflammatory responses may be associated with the poor prognosis of ccRCC.

There are still some deficiencies in the process and method of this study. First, the establishment of this prognostic model established in our study is limited by race, region, sample size and other factors. Therefore, the robustness of the model needs to be further improved. Second, due to the limitation of clinical data, this study did not cover the subtypes of kidney renal clear cell carcinoma, so the significance of guiding the treatment strategies for specific subtypes remains to be further studied. In a word, 12 prognostic models of FRGs were constructed, and independent prognostic factors associated with ccRCC were discussed in this study. The potential mechanism of FRGs in ccRCC needs further experimental study and continuous improvement.

Ethical Approval
Our study did not involve any experiment on humans or animals; therefore, no ethical approval was obtained.

Funding
This work was supported by the National Natural Science Foundation of China (grant number 81873300); the Capital Health Research and Development of Special (grant number 2018-2-4173); and the Fundamental Research Funds for the Central public welfare research institutes (grant number ZZ11-023).

Disclosure
The authors declare no conflicts of interest in this work.

References


The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/international-journal-of-general-medicine-journal