

## Original Article

# The prognostic value of cancer stage-associated genes in clear cell renal cell carcinoma

Xiao-Sheng Ding<sup>1\*</sup>, Yi-Chun Hua<sup>1\*</sup>, Bing-Xuan Han<sup>2\*</sup>, Juan An<sup>1</sup>, Li-Li Zhou<sup>1</sup>, Wei-Ran Xu<sup>1</sup>, Hui Shi<sup>1</sup>, Xi-Xi Zheng<sup>1</sup>, Wei-Wei Shi<sup>3</sup>, Xiao-Yan Li<sup>1</sup>

<sup>1</sup>Department of Oncology, Beijing Tiantan Hospital, Capital Medical University, Beijing 100070, China;

<sup>2</sup>Department of Physical Education, Shanxi Medical University, Taiyuan 030001, Shanxi, China; <sup>3</sup>Department of Oncology, PLA General Hospital, Beijing 100853, China. \*Equal contributors.

Received May 15, 2023; Accepted August 1, 2023; Epub August 15, 2023; Published August 30, 2023

**Abstract:** Objectives: Clear cell renal cell carcinoma (ccRCC) is a highly prevalent subtype of malignant renal tumor, but unfortunately, the survival rate remains unsatisfactory. The aim of the present study is to explore genomic features that are correlated with cancer stage, allowing for the identification of subgroups of ccRCC patients with high risk of unfavorable outcomes and enabling prompt intervention and treatment. Methods: We compared the gene expression levels across ccRCC patients with diverse cancer stages from The Cancer Genome Atlas (TCGA) database, which revealed characteristic genes associated with tumor stage. We then extracted prognostic genes and used least absolute shrinkage selection operator (LASSO) regression to select four genes for feature extraction and the construction of a prognostic risk model. Results: We have identified a total of 171 differentially expressed genes (DEGs) that are closely linked to the tumor stage of ccRCC through difference analysis. A prognostic risk model constructed based on the expression levels of ZIC2, TFAP2A-AS1, ITPKA, and SLC16A12 holds significant prognostic value in ccRCC. The results of the functional enrichment analysis imply that the DEGs are mainly involved in the regulation of immune-related signaling pathways, and therefore may have a significant function in immune system regulation of ccRCC. Conclusions: Our study has successfully identified significant DEGs between high- and low-staging groups of ccRCC using bioinformatics methods. The construction of a prognostic risk model based on the expression levels of ZIC2, TFAP2A-AS1, ITPKA, and SLC16A12 has displayed promising prognostic significance, indicating its valuable potential for clinical application.

**Keywords:** Clear cell renal cell carcinoma, differentially expressed genes, cancer staging, prognosis, bioinformatics

## Introduction

Cancer staging is a critical factor in the diagnosis and treatment of cancer, providing valuable insights into clinical management and prognosis at the individual level. At the population level, cancer staging helps analyze tumor burden and related patterns over time, facilitating the development of cancer control policies and the evaluation of the effectiveness of cancer screening programs [1-3]. Additionally, population-based cancer staging analyses can facilitate the identification of disparities in cancer prognosis [1, 4]. Renal cell carcinoma (RCC) is a complex malignancy that exhibits intricate genetic and molecular alterations, some of which have been strongly linked to patient

prognosis [5]. However, the traditional tumor, node, metastasis (TNM) staging system fails to fully capture the heterogeneity and complexity of RCC, leading to inaccurate prognostic predictions. Therefore, more comprehensive staging systems that incorporate genetic and molecular factors are necessary to accurately predict the outcome of RCC patients. Such systems have the potential to improve clinical decision-making and ultimately enhance patient outcomes.

RCC is a common malignancy in North America, with an incidence of 12 per 100,000 and a peak incidence at ages of 60-70 [6]. Despite medical advancements, the incidence of RCC remains high and continues to rise. In the

## Stage-associated genes predict ccRCC prognosis

United States alone, it is estimated that in 2022, there would be approximately 79,000 new cases and 13,920 deaths from RCC [7]. Clear-cell renal cell carcinoma (ccRCC) is the dominant histologic subtype of RCC, and the 5-year survival rates decline as the disease advances to higher stages. Incidental detection due to the increased use of abdominal imaging has led to a rise in the incidence of localized RCC in developed countries, with nearly 70% of tumors being discovered incidentally [8, 9]. Although clear cell carcinoma is the least malignant among kidney cancers, it is often mixed with granular cell carcinoma and spindle cell carcinoma in clinical practice, and microscopic grading is actually very difficult. Therefore, it is crucial to explore the genomic characteristics of renal clear cell carcinoma through staging. Currently, no validated biomarkers have been included in clinical guidelines for screening and prognosis of patients with RCC. Hence, our study aimed to identify candidate biomarkers by targeting tumor staging-related gene expression data in renal clear cell carcinoma.

### Methods

#### *Sample collection and screening*

Using The Cancer Genome Atlas (TCGA) database available at <http://xena.ucsc.edu/>, we gained access to RNA sequencing (RNA-seq) data and associated clinical information for individuals diagnosed with ccRCC. Patients who lacked survival data were excluded from the clinical information screening, and 529 patients ultimately qualified for further analysis.

#### *Difference analysis*

We defined stage I and II patients as low-staging and stage III and IV patients as high-staging. Using the R package “limma”, we derived a list of genes that were differentially expressed between the high- and low-staging groups. The filtering criteria were set as false discovery rate  $< 0.050$  and  $|\log_2FC| > 1.000$ .

#### *Consensus cluster analysis*

Initially, 171 DEGs were identified from previous studies in the context of ccRCC. Subsequently, unsupervised clustering analysis was conduct-

ed on 529 cases of ccRCC based on the expression levels of these 171 genes. The number of sample clusters was determined using the consensus clustering algorithm. The “ConsensusClusterPlus” R package was utilized to conduct cluster analysis on ccRCC patients and identify three distinct clusters. We performed log-rank test to detect discrepancies in the survival curves between the two clusters. To ensure the stability and reliability of the classification, we ran the analysis for 50 cycles.

#### *Construction of the prognostic model*

We extracted the expression matrix of the 171 DEGs and then identified genes associated with the prognosis of ccRCC by applying univariate Cox regression analysis ( $P < 0.05$ ). Next, multivariate Cox regression analysis and least absolute shrinkage selection operator (LASSO) regression analysis were conducted to identify stage-related genes and construct a prognostic model. The risk score was calculated using the following formula: Risk score = Coef gene  $\times$  Expr gene, where Coef gene represents the correlation between a gene and patient survival in ccRCC, and Expr gene represents the expression level of the gene. Each patient with ccRCC was assigned a risk score based on this formula. Utilizing the cutoff function available in the survival package, patients were categorized into high-risk and low-risk groups. A ROC curve was generated using the “timeROC” package to assess the accuracy of the model.

#### *Functional analysis*

Gene set enrichment analysis (GSEA) was conducted to explore the potential biological processes and pathways in which the DEGs might participate. The DEGs were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses using the “clusterProfiler” package in R (version 4.2.1) for GSEA. For GSEA, curated reference gene sets from the c5.all.v7.2.symbols.gmt MgDB file were selected. The pathways that were enriched were ranked according to their normalized enrichment score (NES) and corrected  $p$  value. Furthermore, functional or pathway terms with an adjusted  $p$ -value below 0.05 and a false discovery rate (FDR) lower than 0.25 were deemed significantly enriched.

## Stage-associated genes predict ccRCC prognosis

### *Immune analysis*

For the immune cell infiltration analysis, the CIBERSORTx algorithm was utilized to estimate the levels of 22 different immune cell types. The association between the risk score and levels of tumor-infiltrating immune cells, as determined via CIBERSORTx analysis, was subsequently evaluated.

### *Survival analysis*

Kaplan-Meier plots were generated to visualize the survival outcomes of patients, and the log-rank test was conducted using the “survival” package to evaluate the statistical significance of the differences in survival between groups. Additionally, we performed Cox regression analyses, including both univariate and multivariate models, to identify genes that had an independent impact on patient outcomes. We performed a multivariate Cox regression analysis, including prognostic factors that showed a *p* value of less than 0.1 in the previous univariate Cox regression analysis. We created a forest plot to present the results of the Cox regression analysis using the “survminer” R package.

### *Statistical analysis*

All statistical tests were performed using R 4.2.1. The Wilcoxon rank-sum test or Student's *t* test was employed to evaluate discrepancies between groups, as appropriate. Pearson or Spearman correlation tests were utilized for correlation analysis, depending on the nature of the data. Kaplan-Meier plots were generated, and log-rank tests were conducted to determine the significance of survival curve variations. Statistical significance was defined as a two-sided *P* value of 0.05.

## **Results**

### *Differentially expressed genes between stages of renal clear cell carcinoma*

As shown in **Figure 1**, we set stage 1 and stage 2 patients with ccRCC as low grade and stage 3 and stage 4 patients as high grade and conducted difference analysis on the two types of patients. A total of 171 differentially expressed genes were obtained. As shown in **Figure 1A**, the most significantly upregulated genes in high-grade patients included PLEKHS1, LINC01234, SLC38A5, PITX1, MFSD2A,

AC104654.2, PAEP, SAA1, PTPRH, and IL20RB, and the downregulated genes included LIN7A, HMGCLL1, SCN4A, AGTR1, RP11-133F8.2, ENAM, SLC5A8, and CD300LG. According to the obtained list of 171 genes, we also drew an expression heatmap of these genes, which is depicted in **Figure 1B**. The analysis revealed notable differences in the expression levels of these 171 differentially expressed genes across patients at various stages.

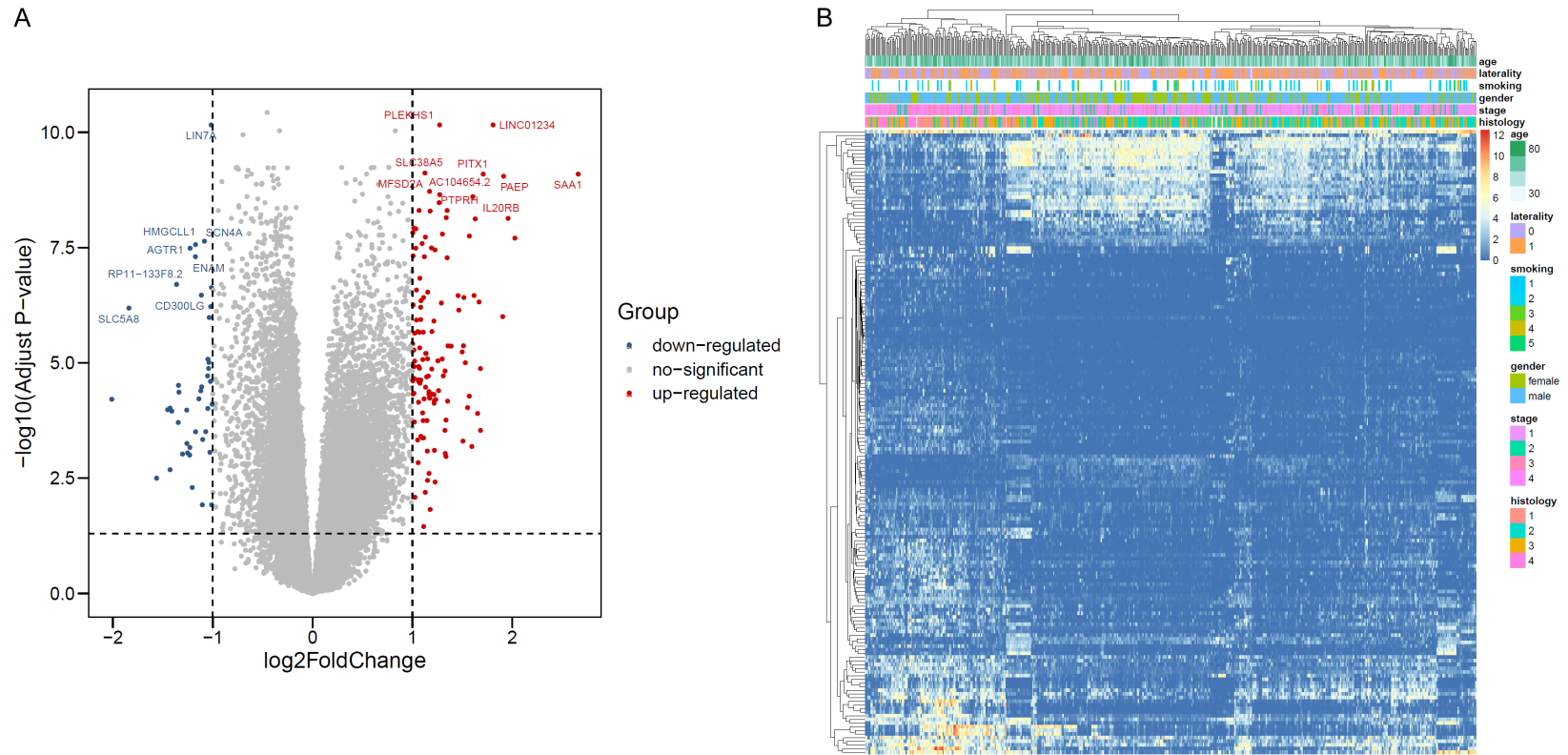
### *Consensus clustering according to differentially expressed genes*

Through differential analysis, 171 significant DEGs were obtained. The expression levels of these genes represented the gene expression characteristics of ccRCC patients in different tumor stages, and the expression levels of these 171 genes varied among tumor stages. We then performed a consensus cluster analysis on the expression data of these 171 genes (**Figure 2**). The analysis revealed that patients with renal clear cell carcinoma can be distinctly divided into two groups based on the expression traits of these 171 genes, and the clustering outcomes were associated with the respective tumor stages. Furthermore, we generated a survival curve between the two groups and compared their survival rates. The findings indicated significant differences in survival between the two groups of patients, implying that the 171 genes could offer useful guidance for assessing prognosis and may have clinical applications.

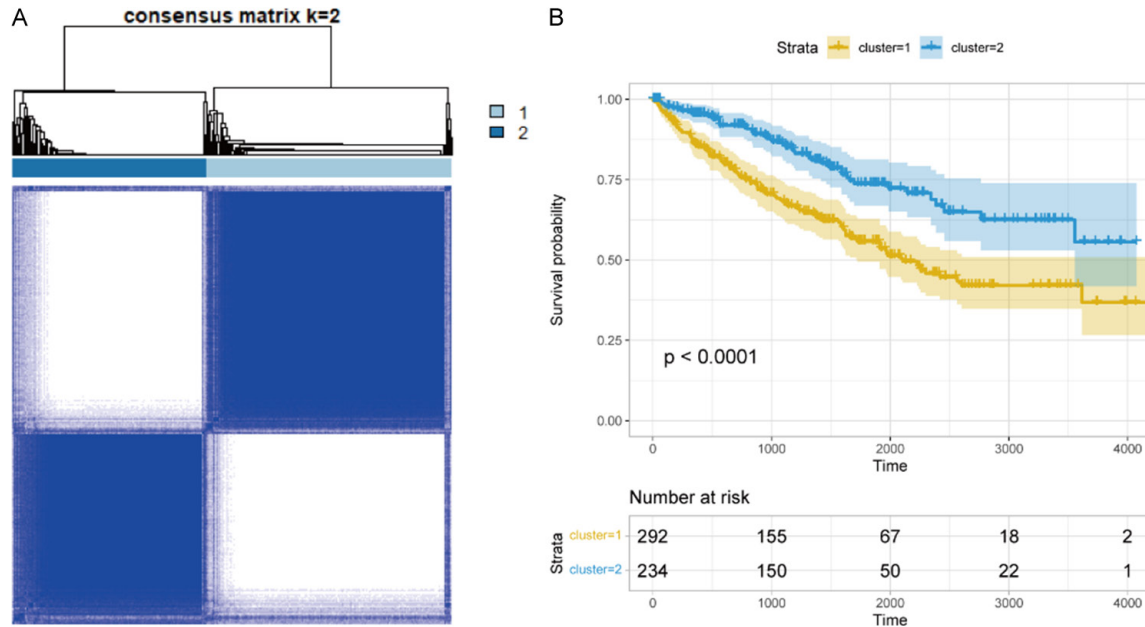
### *Functional analysis of DEGs*

To better understand the biological functions of these 171 DEGs, we analyzed the signaling pathways in which they may be involved by GSEA to speculate on their roles in the occurrence and development of ccRCC (**Figure 3**). We selected the relevant channels in the GO, KEGG and REACTOME databases for analysis. The results showed that the upregulated genes were enriched in the immunoglobulin complex circulating pathway and the immunoglobulin receptor binding pathway. In the KEGG database, upregulated genes were enriched in the primary immunodeficiency and systemic lupus erythematosus signaling pathways. According to the REACTOME database, upregulated genes were enriched in the pathway “CD22-mediated B-cell receptor (BCR) regulation and antigen-

## Stage-associated genes predict ccRCC prognosis



## Stage-associated genes predict ccRCC prognosis



**Figure 2.** Consensus cluster. A: The consensus cluster of the clear cell renal cell carcinoma (ccRCC) patients by 171 differentially expressed genes (DEGs) expression levels; B: The survival difference between the two clusters.

activated BCRs, leading to the generation of second messengers". The identified DEGs are primarily associated with the regulation of immune-related signaling pathways, suggesting that these genes may be involved in the functional regulation of the immune system in ccRCC.

### Construction of the prognostic model

Cox-LASSO regression was used to fit the expression data of 171 genes with the prognostic survival data and delete the redundant genes with high collinearity (**Figure 4**) by LASSO regression. A total of 4 prognostic genes were obtained, namely, ZIC2, TFAP2A-AS1, ITPKA, and SLC16A12, which showed statistically significant prognostic value in ccRCC patients. To evaluate the overall prognostic significance of these 4 genes, risk factors were constructed by summing the value of the products of the expression levels of these prognostic genes and their regression coefficients. The higher the score of the patient, the higher the risk. Then, we plotted the survival curves to compare the low-score and high-score patients, which showed a significant difference.

### Prognostic significance of risk factors

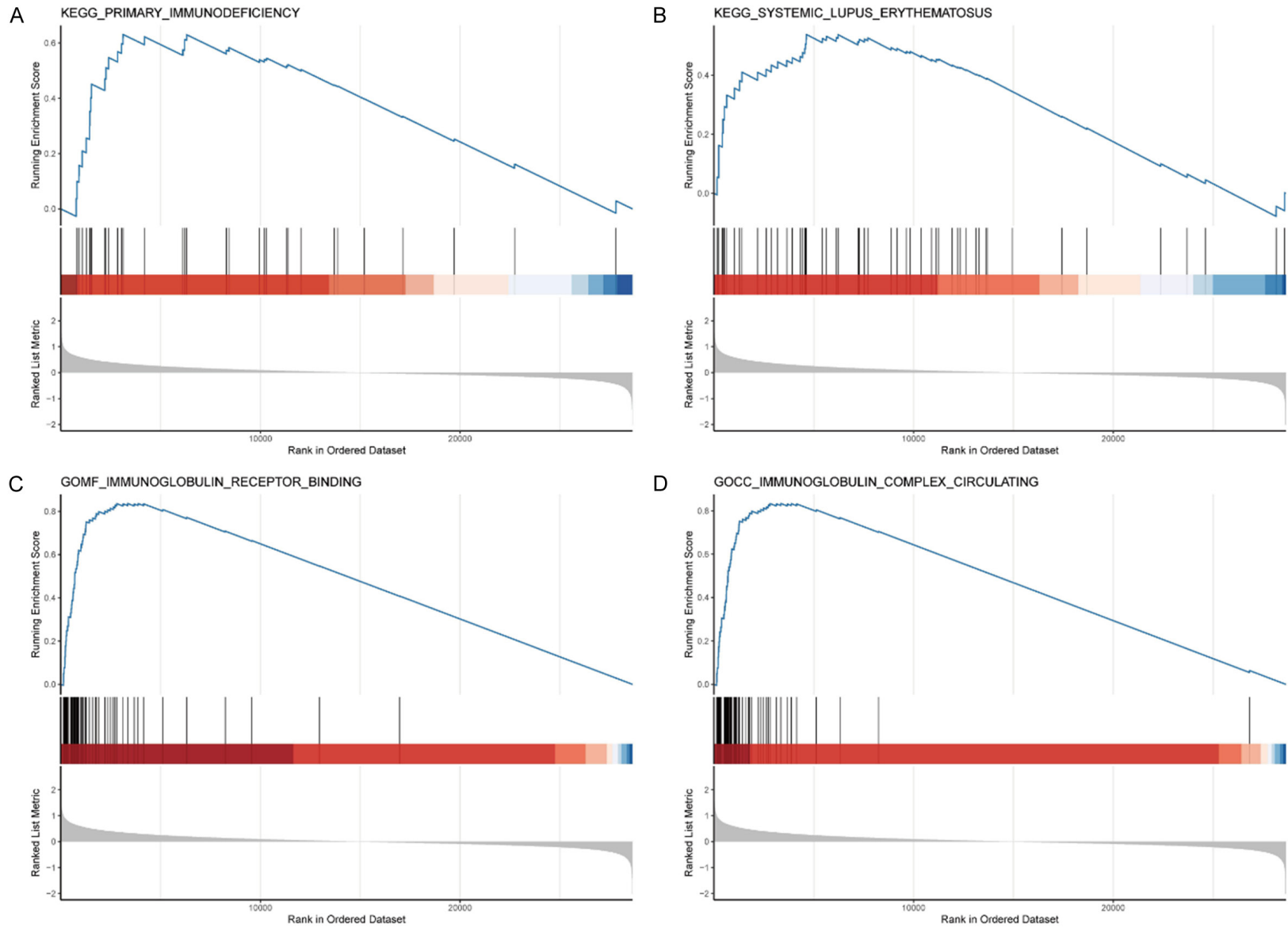
To understand the prognostic significance of the risk factor model, we performed multivari-

ate Cox regression for the four genes included in the model and plotted a graph by integrating the risk factors with other clinical features (**Figure 5**). ITPKA gene expression was significantly associated with prognosis, while SLC16A12 expression tended to be associated with prognosis. High ITPKA expression was associated with worse outcomes for patients with kidney cancer, and SLC16A12 was a risk factor for patients with RCC (although not statistically significant). Subsequently, we investigated the predictive ability of the model for renal cancer prognosis by drawing the ROC curve. The ROC curve showed that this model demonstrated excellent accuracy in predicting the 1-year, 3-year and 5-year survival rates of ccRCC patients. The prediction accuracy for the 1-year survival rate reached 70%. The prediction accuracy for 3-year survival rate reached 66%. The prediction accuracy for the 5-year survival rate reached 69%.

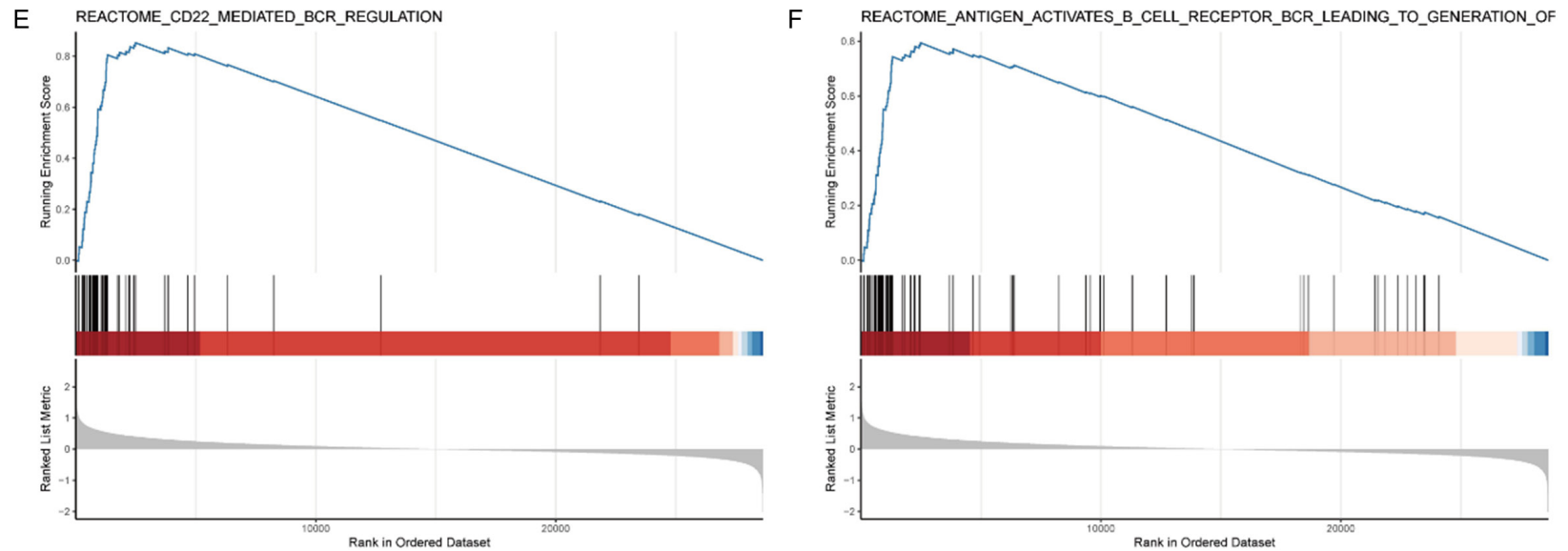
### Association of immune infiltration levels with risk factors

We explored the association between the risk score and the levels of immune components in the renal clear cell carcinoma microenvironment through correlation analysis (**Figure 6**). The findings demonstrated a significant correlation between risk score and the infiltration lev-

# Stage-associated genes predict ccRCC prognosis

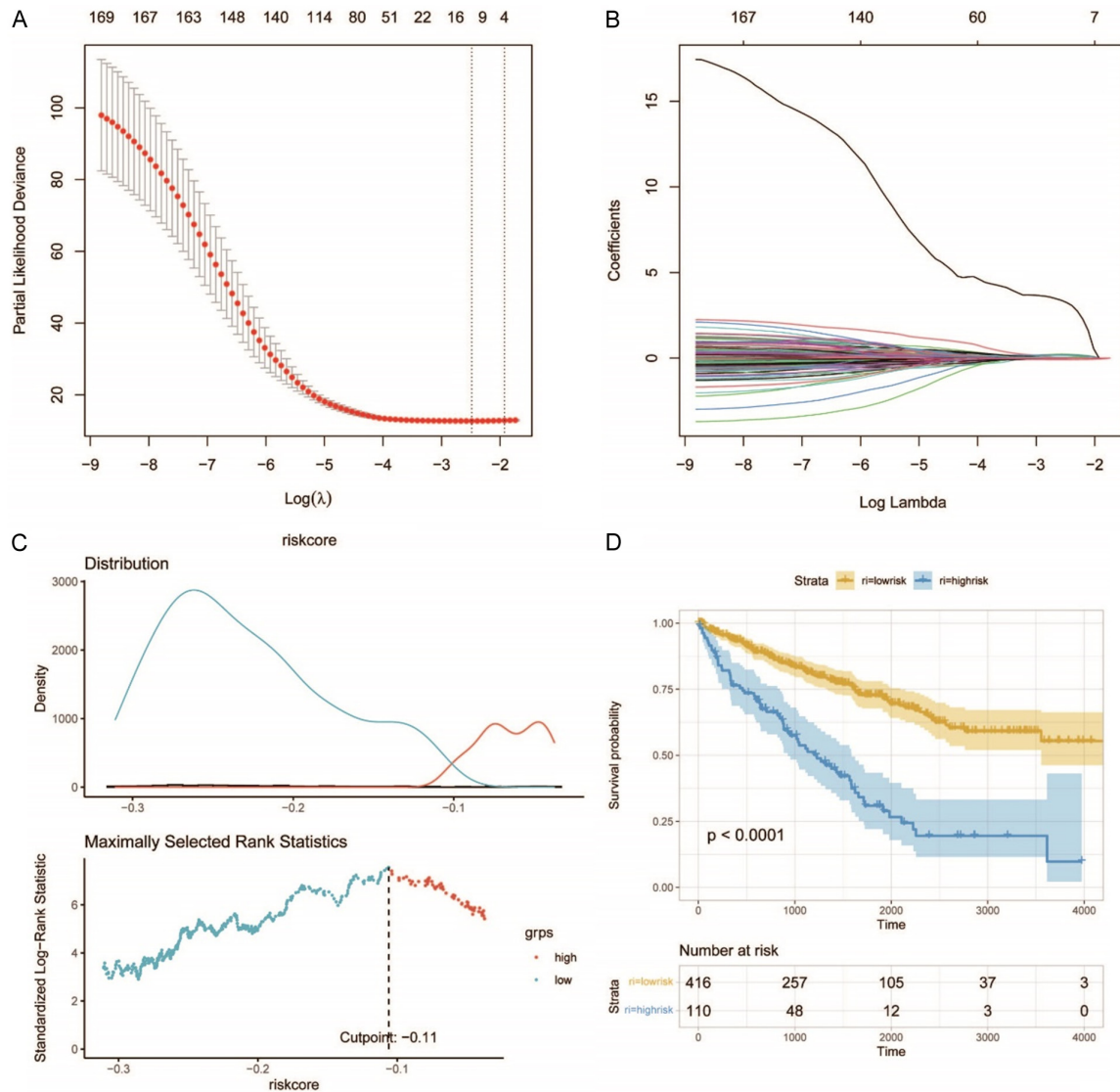


## Stage-associated genes predict ccRCC prognosis



**Figure 3.** Functional analysis. A-F: Gene set enrichment analysis in the primary immunodeficiency, systemic lupus erythematosus, immunoglobulin receptor binding, immunoglobulin complex circulating, CD22 mediated B-cell receptor (BCR) regulation and antigen activates BCR leading to generation of second messenger pathways.

## Stage-associated genes predict ccRCC prognosis



**Figure 4.** Lasso-Cox regression. A: Ten-time cross-validation for tuning parameter selection in the least absolute shrinkage selection operator (LASSO) model; B: LASSO coefficient profiles of the 171 differentially expressed genes (DEGs); C: Distribution plots of the risk score; D: The K-M curves between the high- and low-risk groups.

els of multiple immune cells, such as naive B cells, CD8 T cells, CD4 memory resting T cells, Tregs, and M2 macrophages. The risk score was similar to that of naive B cells, CD4 memory resting T cells, monocytes, M1 macrophages, M2 macrophages, and resting dendritic cells. The risk score had a significant negative correlation with the levels of resting mast cells. The risk score demonstrated a positive correlation with the levels of CD8 T cells, Tfh cells, Tregs, M0 macrophages, and activated dendritic cells.

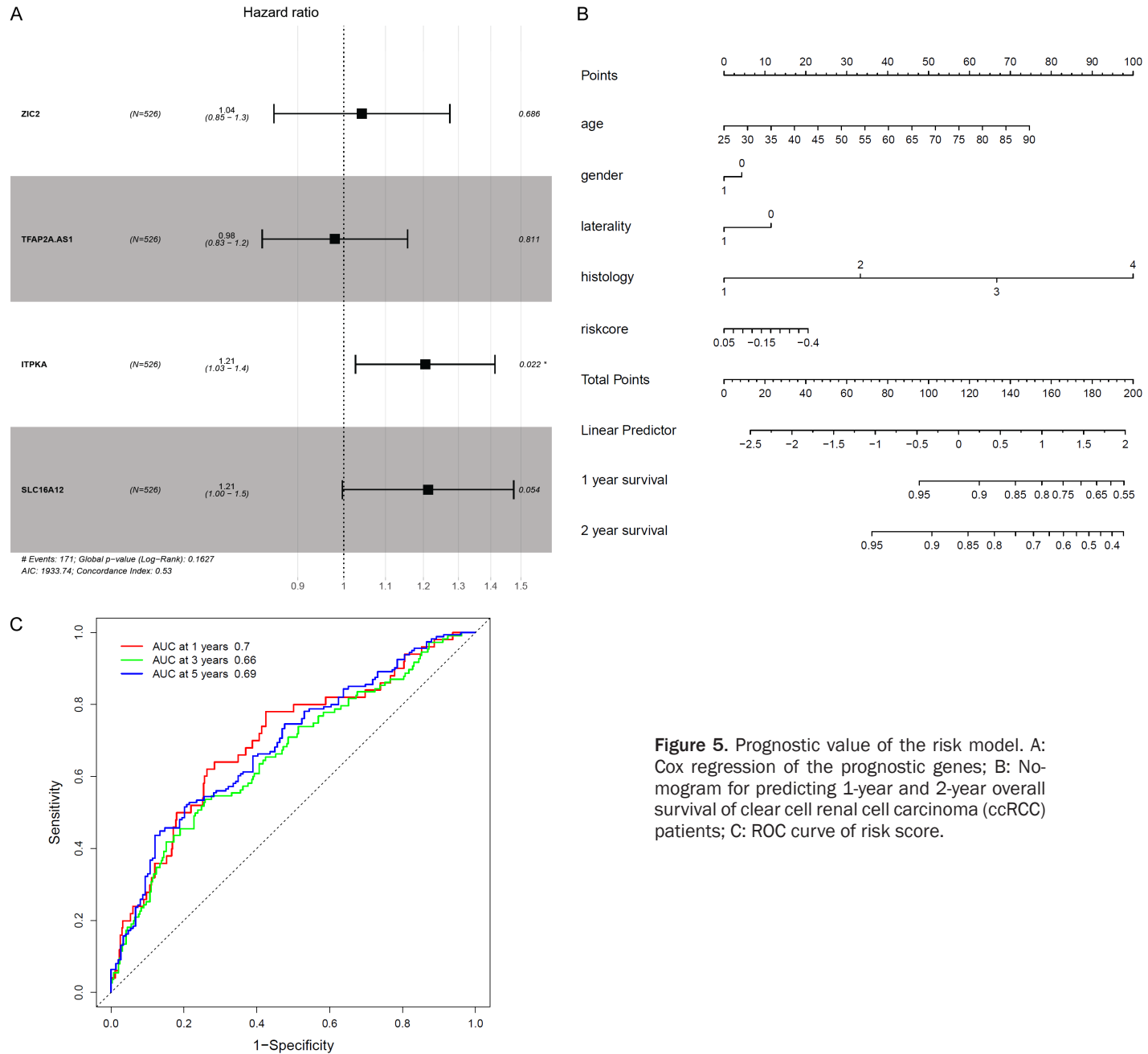
### Discussion

The stage of a tumor is a crucial factor in determining a patient's prognosis. The TNM classifi-

cation system is currently the most commonly employed method of tumor staging, and its significance in predicting prognosis and guiding cancer treatment is indisputable. As precision medicine continues to advance rapidly and new breakthroughs are being made, it is becoming increasingly important to develop more detailed and personalized treatment plans for patients based on their cancer stage. In recent years, several novel prognostic prediction systems have emerged for ccRCC patients. One such system is the Stage, Size, Grade, and Necrosis (SSIGN) Score, which was initially reported in 2002 and has been found to possess superior predictive capability compared to the TNM classification system [10]. Another

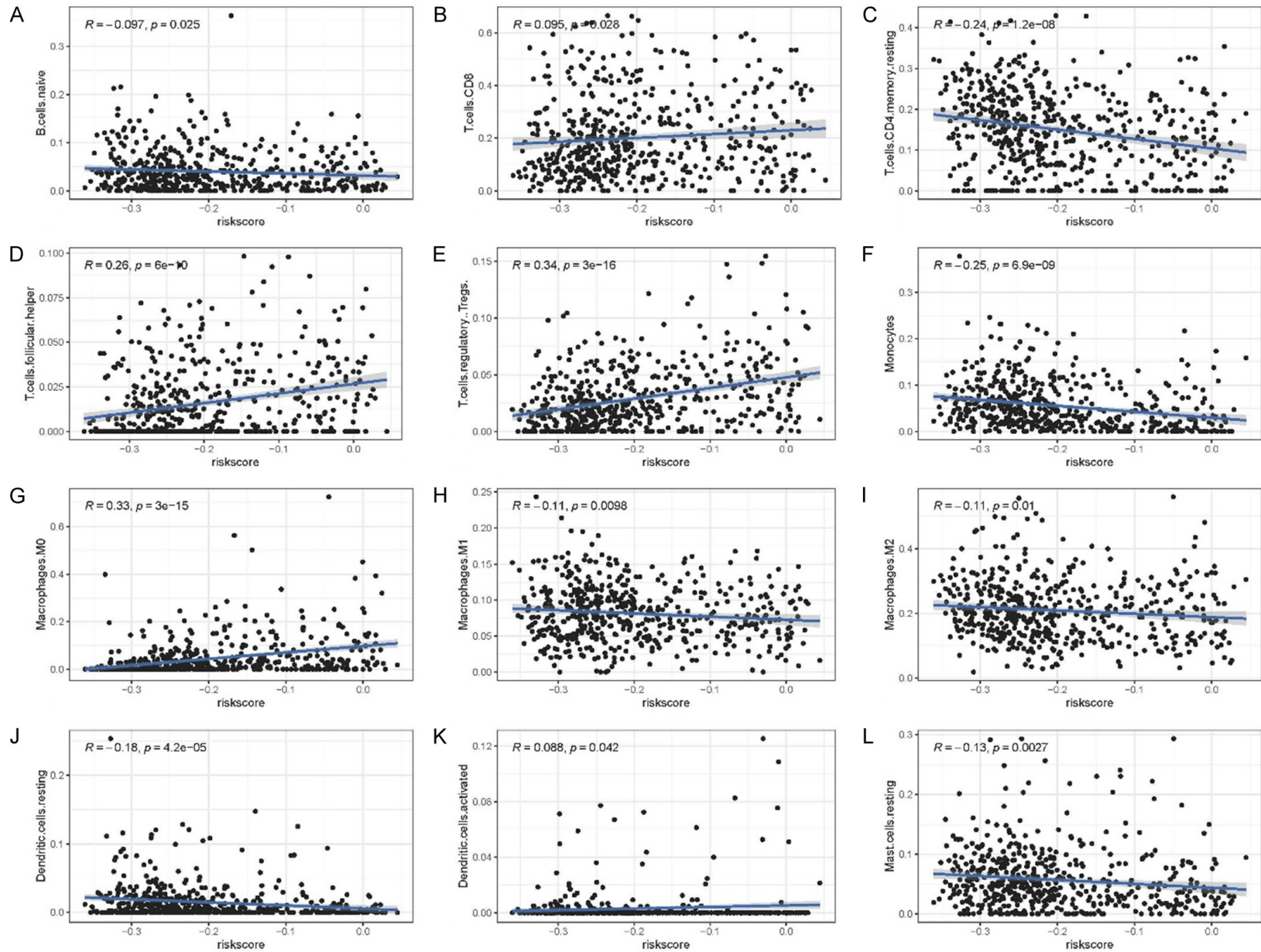


## Stage-associated genes predict ccRCC prognosis



**Figure 5.** Prognostic value of the risk model. A: Cox regression of the prognostic genes; B: Nomogram for predicting 1-year and 2-year overall survival of clear cell renal cell carcinoma (ccRCC) patients; C: ROC curve of risk score.

## Stage-associated genes predict ccRCC prognosis



## Stage-associated genes predict ccRCC prognosis

**Figure 6.** The correlation between risk score and TILs. A-L: The correlation between risk score and naïve B cells, CD8 T cells, resting memory CD4 T cells, Tfh cells, Tregs, Monocytes, MO macrophages, M1 macrophages, M2 macrophages, resting DC, activated DC and resting mast cells.

staging system, which takes tumor grade, N stage, and patient performance status into consideration, has been shown in studies to be effective in differentiating between RCC patients with varying survival probabilities [11]. However, these schemes rely on parameters that are subjectively evaluated by expert pathologists, and predictions may be influenced by interobserver variability. Therefore, more concise and practical tools are urgently needed to improve prognostic predictions for ccRCC patients. To address this issue, we explored the gene expression profiles of different stages and established a gene expression-based prognostic prediction model for renal clear cell carcinoma. We combined stage I and stage II renal clear cell carcinoma patients into one group and compared them with stage III and IV patients, which led to the identification of 171 DEGs associated with tumor stage. We then extracted prognostic genes from these characteristic genes and selected four genes to develop a prognostic risk model. Our study has shown that our model has significant prognostic ability in ccRCC, enabling the identification of high-risk subgroups of ccRCC patients to guide timely interventions and treatments.

Through LASSO regression analysis, we identified several genes linked to the prognosis of renal clear cell carcinoma, with four key genes being particularly relevant. These genes are ZIC2, ITPKA, SLC16A12, and TFAP2A-AS1. ZIC2 functions as a transcription factor and belongs to the ZIC family, regulating cellular processes such as cell proliferation, differentiation, and migration [12]. Although the role of ZIC2 in ccRCC is not yet fully elucidated, studies have suggested that it may contribute to the progression of various cancer types, such as nasopharyngeal carcinoma, bladder cancer, and liver cancer [13-15]. Specifically, Wu et al. demonstrated that ZIC2 can promote the overexpression of RUNX2 through upregulation of expression levels as part of the Zic2/Runx2/NOLC1 signaling axis, leading to the malignant proliferation and migration of tumor cells in renal clear cell carcinoma [16]. Liu et al. identified ITPKA as a protein that promotes cell motility and increases the metastatic capacity of

tumor cells [17]. Expression of the ITPKA gene was significantly upregulated in stage IV compared to stage I renal clear cell carcinoma patients [17]. Similarly, Zhu et al. found that ITPKA promotes the growth, migration, and invasion of renal cell carcinoma via the activation of the mTOR signaling pathway and is overexpressed in RCC patients [18]. Guan et al. reported that elevated expression of ITPKA is linked to greater vascular infiltration and shorter survival time in liver cancer patients [19]. SLC16A12 is a member of the solute carrier group of membrane transport proteins, and its precise function is not yet clearly understood [20]. Mei et al. have demonstrated that decreased levels of SLC16A12 expression may be linked to an adverse prognosis in individuals with renal clear cell carcinoma, suggesting that SLC16A12 might play a critical role as a tumor suppressor [21]. The SLC16A family of transport proteins has been associated with the development and advancement of various forms of cancer, primarily by regulating the transmembrane transportation of monocarboxylic acids, including but not limited to lactic acid, pyruvate, and butyric acid [22-25]. In the development of a prognostic model for renal clear cell carcinoma, a long non-coding RNA, TFAP2A-AS1, was included alongside protein-coding RNA as a risk factor. Studies have suggested that TFAP2A-AS1 serves as a tumor suppressor and is linked to a favorable prognosis in breast cancer. Furthermore, Zhou et al. demonstrated that in breast cancer, TFAP2A-AS1 functions as a miRNA sponge for miR-933, which results in the degradation of smad2 mRNA and impedes the growth and migration of breast cancer cells [26]. Interestingly, in a different study conducted by Jiang et al., high levels of TFAP2A-AS1 expression were found to be associated with a poor prognosis in ccRCC patients [27]. The observed contrasting association between TFAP2A-AS1 expression levels and prognosis in patients with breast cancer and ccRCC may be explained by the differences in the biological behavior of the tumor and may highlight the complex and varied roles played by lncRNAs in different cancers [28-30]. Based on various studies that have identified the roles of the genes comprising the predictive model in

cancer progression, we can conclude that there is a significant theoretical basis for the application value of its constituent genes in patients with kidney cancer. The crucial role of ITPKA should also not be overlooked, as indicated by multivariate Cox regression analysis, which demonstrated its significant prognostic value. This conclusion is also supported by several additional publications. Therefore, the prognostic model proposed in this study holds promising guiding significance for predicting the prognosis of RCC patients, and the ITPKA gene, as one of its components, may warrant further investigation as a potential target for the future treatment of kidney cancer.

To explore the biological functions and underlying mechanisms associated with the prognosis of renal clear cell carcinoma, we performed functional enrichment analysis of DEGs using the GO, KEGG, and REACTOM databases. Our findings suggest that these genes may play a role in regulating immune function in individuals with renal clear cell carcinoma. Specifically, we observed that upregulated genes were enriched in the immunoglobulin complex circulating pathway and the immunoglobulin receptor binding pathway. Moreover, examination of the KEGG database revealed that the upregulated genes were enriched in signaling pathways pertaining to primary immunodeficiency and systemic lupus erythematosus. Furthermore, analysis of the REACTOME database revealed that the upregulated genes were enriched in the “CD22-mediated BCR regulation and antigen activation of BCRs, leading to the generation of second messengers”. Our study also demonstrated a significant relationship between risk factors and immune cell invasion levels, including CD4 memory resting T cells, CD8 T cells, Tregs, and M2 macrophages. Interestingly, our analysis revealed that the risk factor score was closely correlated with the levels of naive B cells, CD4 memory resting T cells, monocytes, M1 macrophages, M2 macrophages, and resting dendritic cells, whereas the score was significantly negatively correlated with the infiltration levels of resting mast cells. Moreover, the levels of resting mast cells were positively correlated with those of CD8 T cells, Tfh cells, Tregs, M0 macrophages, and activated dendritic cells. These findings are of interest, as previous studies have proposed that the prognostic genes

included in our risk model may be linked to immune infiltration levels. For instance, Lv et al. discovered that ZIC2 could exert a substantial influence on tumorigenesis by modulating TMB and MSI, manipulating cancer immune infiltrating cells, and regulating immune checkpoints, MMRs, DNA methyltransferase genes, and N6-methyladenosine (M6A) RNA methylation, all of which could ultimately impact patient prognosis [31]. Similarly, Sun et al. reported a significant correlation between ZIC2 mRNA expression levels and the infiltration of immune cells, including B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells, in the immune microenvironment of liver cancer [32]. Overall, we suggest that the DEGs identified in this study might have a crucial function in controlling the immune system and microenvironment composition in ccRCC. As shown by the strong interpretability of our proposed model, the calculated risk score reflects the microenvironment characteristics of the tumor and exhibits regulatory effects on several key pathways of biological function regulation, thus highlighting specific functional changes in tumors.

Nevertheless, there are some limitations to our research. First, the data we used were predominantly from the TCGA database, which mostly includes data on Caucasian patients. As such, the generalizability of our findings to Asian populations requires further investigation. Second, as a retrospective study, it may be difficult to adjust the research design in real time. Finally, we would like to emphasize that although our study utilized bioinformatics algorithms to identify genes associated with prognosis in renal clear cell carcinoma, the results reported here are still speculative and should be confirmed through laboratory experiments. Despite these limitations, our study presents credible and noteworthy findings regarding the genetic differences in ccRCC patients with varying disease stages. As such, we believe our results warrant further attention and continued investigation.

### Conclusion

By employing bioinformatics tools, we have identified ZIC2, TFAP2A-AS1, ITPKA, and SLC-16A12 as genes that are closely linked to the tumor stage of renal clear cell carcinoma.

Moreover, the construction of a prognostic risk model based on the expression levels of these four genes has displayed promising prognostic significance, indicating its valuable potential in clinical settings.

### Acknowledgements

The current study was supported by the grants from National Natural Science Foundation of China (No. 81974361) and Beijing Municipal Education Commission-Beijing Natural Science Foundation Union Program (KZ2020100250-46).

### Disclosure of conflict of interest

None.

**Address correspondence to:** Xiao-Yan Li, Department of Oncology, Beijing Tiantan Hospital, Capital Medical University, No. 119 Nansihuan West Road, Fengtai District, Beijing 100070, China. Tel: +86-10-59976369; E-mail: lixiaoyan@bjtth.org; Wei-Wei Shi, Department of Oncology, PLA General Hospital, No. 28 Fuxing Road, Haidian District, Beijing 100853, China. Tel: +86-10-66875308; E-mail: shiweiwei301@sina.com

### References

- [1] Cancer Australia. A national cancer data strategy for Australia. 2012.
- [2] National cancer stage at diagnosis data | National Cancer Control Indicators.
- [3] Parkin DM. The role of cancer registries in cancer control. *Int J Clin Oncol* 2008; 13: 102-111.
- [4] Coleman MP, Forman D, Bryant H, Butler J, Rachet B, Maringe C, Nur U, Tracey E, Coory M, Hatcher J, McGahan CE, Turner D, Marrett L, Gjerstorff ML, Johannesen TB, Adolphsson J, Lambe M, Lawrence G, Meechan D, Morris EJ, Middleton R, Steward J and Richards MA; ICBP Module 1 Working Group. Cancer survival in Australia, Canada, Denmark, Norway, Sweden, and the UK, 1995-2007 (the International Cancer Benchmarking Partnership): an analysis of population-based cancer registry data. *Lancet* 2011; 377: 127-138.
- [5] Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* 2013; 499: 43-49.
- [6] Volpe A and Patard JJ. Prognostic factors in renal cell carcinoma. *World J Urol* 2010; 28: 319-327.
- [7] Kidney Cancer: Statistics | Cancer.Net.
- [8] Partin AW, Peters CA, Kavoussi LR, Dmochowski RR and Wein AJ. *Campbell-Walsh-Wein Urology Twelfth Edition Review E-Book*. 2020.
- [9] Capitanio U, Bensalah K, Bex A, Boorjian SA, Bray F, Coleman J, Gore JL, Sun M, Wood C and Russo P. Epidemiology of renal cell carcinoma. *Eur Urol* 2019; 75: 74-84.
- [10] Moch H, Artibani W, Delahunt B, Ficarra V, Knuechel R, Montorsi F, Patard JJ, Stief CG, Sulser T and Wild PJ. Reassessing the current UICC/AJCC TNM staging for renal cell carcinoma. *Eur Urol* 2009; 56: 636-643.
- [11] Zisman A, Pantuck AJ, Dorey F, Said JW, Shvarts O, Quintana D, Gitlitz BJ, deKernion JB, Figlin RA and Belldegrun AS. Improved prognostication of renal cell carcinoma using an integrated staging system. *J Clin Oncol* 2001; 19: 1649-1657.
- [12] Inaguma S, Ito H, Riku M, Ikeda H and Kasai K. Addiction of pancreatic cancer cells to zinc-finger transcription factor ZIC2. *Oncotarget* 2015; 6: 28257-28268.
- [13] Shen ZH, Zhao KM and Du T. HOXA10 promotes nasopharyngeal carcinoma cell proliferation and invasion via inducing the expression of ZIC2. *Eur Rev Med Pharmacol Sci* 2017; 21: 945-952.
- [14] Wang J, Ma W and Liu Y. Long non-coding RNA HULC promotes bladder cancer cells proliferation but inhibits apoptosis via regulation of ZIC2 and PI3K/AKT signaling pathway. *Cancer Biomark* 2017; 20: 425-434.
- [15] Zhu P, Wang Y, He L, Huang G, Du Y, Zhang G, Yan X, Xia P, Ye B, Wang S, Hao L, Wu J and Fan Z. ZIC2-dependent OCT4 activation drives self-renewal of human liver cancer stem cells. *J Clin Invest* 2015; 125: 3795-3808.
- [16] Wu CY, Li L, Chen SL, Yang X, Zhang CZ and Cao Y. A Zic2/Runx2/NOLC1 signaling axis mediates tumor growth and metastasis in clear cell renal cell carcinoma. *Cell Death Dis* 2021; 12: 319.
- [17] Liu Q, Zhao S, Su PF and Yu S. Gene and isoform expression signatures associated with tumor stage in kidney renal clear cell carcinoma. *BMC Syst Biol* 2013; 7 Suppl 5: S7.
- [18] Zhu X, Xu A, Zhang Y, Huo N, Cong R, Ma L, Chu Z, Tang Z, Kang X, Xian S and Xu X. ITPKA1 promotes growth, migration and invasion of renal cell carcinoma via activation of mTOR signaling pathway. *Onco Targets Ther* 2020; 13: 10515-10523.
- [19] Li J, Zhu YH, Huang P, Zhang B, Sun J and Guan XY. ITPKA expression is a novel prognostic factor in hepatocellular carcinoma. *Diagn Pathol* 2015; 10: 136.
- [20] Lin L, Yee SW, Kim RB and Giacomini KM. SLC transporters as therapeutic targets: emerging

## Stage-associated genes predict ccRCC prognosis

- opportunities. *Nat Rev Drug Discov* 2015; 14: 543-560.
- [21] Mei J, Hu K, Peng X, Wang H and Liu C. Decreased expression of SLC16A12 mRNA predicts poor prognosis of patients with clear cell renal cell carcinoma. *Medicine (Baltimore)* 2019; 98: e16624.
- [22] Park SJ, Smith CP, Wilbur RR, Cain CP, Kallu SR, Valasapalli S, Sahoo A, Guda MR, Tsung AJ and Velpula KK. An overview of MCT1 and MCT4 in GBM: small molecule transporters with large implications. *Am J Cancer Res* 2018; 8: 1967-1976.
- [23] Baltazar F, Pinheiro C, Morais-Santos F, Azevedo-Silva J, Queirós O, Preto A and Casal M. Monocarboxylate transporters as targets and mediators in cancer therapy response. *Histol Histopathol* 2014; 29: 1511-1524.
- [24] Doyen J, Trastour C, Ettore F, Peyrottes I, Tousseant N, Gal J, Ilc K, Roux D, Parks SK, Ferrero JM and Pouysségur J. Expression of the hypoxia-inducible monocarboxylate transporter MCT4 is increased in triple negative breast cancer and correlates independently with clinical outcome. *Biochem Biophys Res Commun* 2014; 451: 54-61.
- [25] Pinheiro C, Longatto-Filho A, Azevedo-Silva J, Casal M, Schmitt FC and Baltazar F. Role of monocarboxylate transporters in human cancers: state of the art. *J Bioenerg Biomembr* 2012; 44: 127-139.
- [26] Zhou B, Guo H and Tang J. Long non-coding RNA TFAP2A-AS1 inhibits cell proliferation and invasion in breast cancer via miR-933/SMAD2. *Med Sci Monit* 2019; 25: 1242-1253.
- [27] Jiang W, Guo Q, Wang C and Zhu Y. A nomogram based on 9-lncRNAs signature for improving prognostic prediction of clear cell renal cell carcinoma. *Cancer Cell Int* 2019; 19: 208.
- [28] Schmitt AM and Chang HY. Long noncoding RNAs in cancer pathways. *Cancer Cell* 2016; 29: 452-463.
- [29] Wu H, Yang L and Chen LL. The diversity of long noncoding RNAs and their generation. *Trends Genet* 2017; 33: 540-552.
- [30] Wang KC and Chang HY. Molecular mechanisms of long noncoding RNAs. *Mol Cell* 2011; 43: 904-914.
- [31] Lv Z, Qi L, Hu X, Mo M, Jiang H, Fan B and Li Y. Zic family member 2 (ZIC2): a potential diagnostic and prognostic biomarker for pan-cancer. *Front Mol Biosci* 2021; 8: 631067.
- [32] Sun L, Lin Y, Wang G, Zhang L, Hu L and Lu Z. Correlation of zinc finger protein 2, a prognostic biomarker, with immune infiltrates in liver cancer. *Biosci Rep* 2021; 41: BSR20203115.