

A novel five-zinc finger-related-gene signature related to tumor immune microenvironment allows for treatment stratification and predicts prognosis in clear cell renal cell carcinoma patients

Feilong Zhang

Beijing Chaoyang Hospital

Jiyue Wu

Beijing Chaoyang Hospital

Jiandong Zhang

Beijing Chaoyang Hospital

Peng Cao

Beijing Chaoyang Hospital

Zeja Sun

Beijing Chaoyang Hospital

Wei Wang (✉ weiwang0920@163.com)

Beijing Chaoyang Hospital <https://orcid.org/0000-0003-2642-3338>

Research

Keywords:

Posted Date: February 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-207677/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Clear cell renal cell carcinoma (ccRCC) is one of the most prevalent renal malignant tumors, which survival rate and quality of life of ccRCC patients are not satisfactory. Therefore, identification of prognostic biomarkers of ccRCC patients will contribute to early and accurate clinical intervention and treatment, and then improve their prognosis.

Methods

We downloaded the original expression data of mRNAs from The Cancer Genome Atlas database and the zinc finger(ZNF)-related genes (ZRGs) from UniProt online database. Differentially expressed ZRGs (DE-ZRGs) was screened from tumor and adjacent nontumor tissues and functional enrichment analysis was conducted out. A five-ZRG signature were constructed by univariate Cox regression, least absolute shrinkage and selection operator and multivariate Cox regression. Furthermore, we screened out independent prognosis-related factors to build a nomogram by univariate and multivariate Cox regression. Potential biological pathways of five ZRGs were analyzed by Gene Set Enrichment Analysis (GSEA). Then, we further quantitatively analyze immune infiltration and evaluate tumor microenvironment by single sample GSEA. Finally, drug sensitivity of ccRCC patients was analyzed by the Genomics of Drug Sensitivity in Cancer database.

Results

TRIM59, VAV3, ZNF189, AGAP9 and PYGO1 were screened to be significantly associated with the prognosis of ccRCC patients. Through incorporated risk score and clinical parameters, we constructed a nomogram, which showed a good prognostic performance for ccRCC patients.

1. Introduction

Kidney cancer is the most common malignant cancer of urogenital system with over 400000 cancer new cases and 175000 cancer deaths worldwide per year [1]. In 2020, there is expected to more than 73000 new cases of kidney cancer and 14000 kidney cancer deaths in the United States [2]. Renal cell carcinoma (RCC), the most common type of kidney cancer, originates from renal tubular epithelial cells, which accounts for more than 90% of all kidney cancer cases [3]. RCC is consist of several subtypes with unique characteristics, of which clear cell RCC (ccRCC) comprise the majority of histological subtypes [4]. Local ccRCC is mainly treated with partial or radical nephrectomy, but approximately 30% of patients with localized ccRCC eventually develop metastasis postoperatively, which seriously affects the overall survival (OS) rate in ccRCC patients [5]. Identification of the high risk of ccRCC patients timely is more conducive to early and accurate clinical intervention and treatment. Therefore, it is imperative to explore

the effective specific biomarkers that can predict prognosis of ccRCC patients accurately and develop novel treatments for this fatal disease.

Zinc finger (ZNF) domains exist in at least 5% of human proteins, and ZNF proteins (ZNFs) are one of the most abundant proteins in human genome, which was first discovered due to the characteristics of DNA-binding proteins [6–8]. Currently, it is increasingly reported that ZNFs can specifically bind to a large variety of substrates such as RNAs, proteins, lipids and post-translational modifications. The diversity of their functions indicates that they may play potential roles in signal transduction, transcriptional regulation, cell migration, and DNA double-strand break repair [9–11]. In addition, the diverse roles of ZNFs in gene expression regulation suggest that they may also play an indispensable role in initiation and progression of cancers [12]. Numerous studies have confirmed that the vital roles of ZNFs in occurrence and prognosis of RCC. Marc Henrion et al. conducted a genome-wide association study and found that the ZEB2 gene is another common susceptibility locus for RCC, and the genetic variation in ZEB2 can affect the risk of RCC [13]. Furthermore, studies have shown that ZEB2, RBCK1 and ZFX, as independent prognostic biomarkers, are associated with lower OS time of RCC patients [14–16]. However, few studies have reported a relation between ZNFs and prognosis of ccRCC patients. Methylation of the ZNF278 gene promoter indicated a poor prognosis of ccRCC [17]. As a prognosis-related factor, low expression of CASZ1 was significantly correlated with the unsatisfactory progression-free survival and OS of patients with ccRCC [18]. However, there are insufficient studies on the potential biological function of ZNFs in the progression of ccRCC and their correlation with the prognosis of ccRCC patients. Further in-depth exploration will help to understand the mechanism involved in ccRCC ontogeny and progression and develop efficient and accurate prognostic biomarkers.

Herein, we constructed a prognosis-related ZNF-related-gene (ZRG) signature of ccRCC patients by bioinformatics methods. First, we obtained the differentially expressed ZRGs (DE-ZRGs) related to ccRCC using TCGA database and UniProt online database, and constructed a prognosis-related five-ZRG signature of ccRCC. Then, we evaluated and verified the prognostic performance of the five-ZRG model through many methods. By integrating the five-ZRG signature and other prognosis-related clinical parameters of patients, we established a more accurate nomogram for ccRCC patients. In addition, GO and KEGG function enrichment analysis on DE-ZRGs and GSEA analysis on prognosis-related five ZRGs were performed to explore the biological process involved in oncogenesis and development of ccRCC. Finally, based on high- and low-risk patients classified by five-ZRG, we explore whether there are differences in immune cell infiltration, tumor microenvironment (TME) and drug resistance to multiple anticancer drugs.

2. Material And Methods

2.1 Data sources and screening of DE-ZRGs

Firstly, the mRNAs expression data and the corresponding clinical information of ccRCC patients in TCGA database (<http://tcga.cancer.gov>) were downloaded for using. 525 patients were finally included after

excluding patients without complete clinical information and Table S1 showed the characteristics of these ccRCC patients. The expression data of ccRCC mRNAs are derived from 540 tumor samples and 71 adjacent non-tumor normal samples. Next, we downloaded the ZRGs from the UniProt online database (<https://www.uniprot.org>) and intersected them with the ccRCC mRNAs to obtain the ZRGs expressed in tumor samples and adjacent non-tumor normal samples. Finally, using the “edgeR” package in the R software to identify the DE-ZRGs between tumor and adjacent non-tumor normal tissues [19] with $|\log_2FC| > 0.5$ and adjusted $p < 0.05$ as the threshold.

2.2 Identify a potential prognosis-related ZRGs (Pr-ZRGs) signature

We divided all ccRCC patients into two groups equally. The expression data of one group (Training set) was used to construct a signature of ZRGs with prognostic value, while the expression data of another group (Testing set) was used to validate the signature we constructed. Firstly, we performed a univariate Cox regression analysis ($HR \neq 1$, $p < 0.05$) on the expression data of training set ($n = 264$) to identify the Pr-ZRGs. Then, we conducted the LASSO regression analysis to sub-select these Pr-ZRGs using “glmnet” package of the R software, which can scale down the regression coefficient forcedly [20]. 10-times cross-validations were conducted to increase the reliability and objectivity of the result [21]. Then, a multivariate Cox regression analysis was performed on the result of the LASSO regression analysis to get the optimal prognostic ZRGs signature. Finally, we combined the expression levels of optimal prognostic ZRGs and their regression coefficients to establish a risk score, which is as follows:

Exp represents the expression level of ZRGs of patient i , while the coef means its regression coefficient.

2.3 Evaluation and validation of the ZRGs signature

In the training set, we stratified the patients into high-risk and low-risk groups based on the median risk score. The Kaplan–Meier (KM) survival curve was evaluated to compare the survival differences between high-risk and low-risk patients and time-dependent receiver operating characteristic (ROC) curve was employed to evaluate the prognostic accuracy of ZRGs signature [22]. Besides, we also compare the survival differences of ccRCC patients in different clinical subgroups (including ages, genders, histologic grades and pathologic stages) to measure the prognostic performance of the ZRGs signature.

In order to validate the prognostic power and applicability of the ZRGs signature, we performed the above risk score formula on the patients in the testing set ($n = 261$) and entire set ($n = 525$). Then, we stratified the patients into the above two groups similarly. The KM survival analysis and time-dependent ROC curve were also employed to validate the ZRGs signature.

2.4 Construction of nomogram

To determine whether the ZRGs signature is an independent prognosis-related factor of ccRCC. We set risk score, ages, genders, histologic grades and pathologic stages as independent variables, while outcome of prognosis as dependent variable to perform univariate Cox regression analysis. We further

performed multivariate Cox regression analysis on these variables with statistical differences ($p < 0.05$) to determine if they are independent prognostic factors of ccRCC. In order to build an accurate and reliable clinical predictive model, we constructed a nomogram by integrating all independent prognostic factors to predict the survival of ccRCC patients in different years (in 1-, 3-, and 5-year) with the “rms” package in the R software [23]. The nomogram can integrate various prognostic factors to predict the occurrence of individual clinical events, it was often used to measure the prognosis of patients.

2.5 Evaluate and validate the nomogram

Calibration plots were used to evaluate the performance of our nomogram. In the calibration plots, X-axis represents the predicted survival rate by the nomogram while Y-axis means the actual survival rate respectively, and the 45° dotted line represented the best outcome. Then, KM survival analysis and time-related ROC analysis were also conducted to measure the predictive power of the nomogram.

2.6 Functional enrichment analysis

To find out the potential biological mechanism correlated with the ontogeny and progression of ccRCC, functional enrichment analysis was performed on DE-ZRGs and optimal prognostic ZRGs. As for the DE-ZRGs, the “clusterProfiler” package in the R software (P-value of < 0.05 was considered statistically significant) was used to perform the gene ontology (GO) annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis [24].

Gene Set Enrichment Analysis (GSEA) was conducted to identify the potential biological pathways of the optimal prognostic ZRGs that were finally screened out. Patients with high- or low- risks were set as the phenotypes to be compared and the annotated gene set file (c2.cp.kegg.v7.2.symbols.gmt) was used for reference. The number of replacements was set to 1000, FDR < 0.05 and $p < 0.05$ was considered statistically significant [25].

2.7 Analysis of immune infiltration and evaluation of tumor microenvironment

The expression levels of immune cell-specific marker genes in each tissues were quantified based on the single sample gene set enrichment analysis (ssGSEA) algorithm. The marker genes of the main immune cells were obtained from the previous studies [26]. Using the “GSVA” package in the R software to assess these genes to identify whether there were differences in immune cell infiltration and immune cell function between the high- and low-risk patients. Besides, the “ESTIMATE” algorithm was conducted to calculate the immune score and the stroma score through the “estimate” package in the R software, and the formula described by Yoshihara et al was used to infer the tumor purity [27].

2.8 Drug sensitivity analysis of the ZRGs signature

The mRNA expression profile of the ccRCC cell lines and the half maximal inhibitory concentration (IC50) of the anticancer drugs were obtained from Genomics of Drug Sensitivity in Cancer (GDSC) database and

employed to evaluate drug sensitivity analysis between high- and low-risk groups [28]. $P < 0.05$ was set as the threshold.

2.9 Statistical analysis

All statistical analyses were performed by the R software (version 3.6.3) in our study. Statistical tests were all two sides and $p < 0.05$ were statistically significant.

3. Results

3.1 DE-ZRGs in ccRCC

Figure 1 showed the flowchart of this study. Firstly, a total of 55266 mRNAs were obtained from the raw mRNA sequencing data of the TCGA dataset, while a total of 1818 ZRGs were obtained from the UniProt online database. By intersecting the ccRCC mRNAs and ZRGs, a total of 1753 ccRCC ZRGs were finally obtained (Fig. 2A). Principal component analysis (PCA) showed that the expression profile of ZRGs between tumor and adjacent non-tumor tissues were significantly different (Fig. 2B). Subsequently, we screened out a total of 271 DE-ZRGs, including 175 up-regulated DE-ZRGs and 96 down-regulated DE-ZRGs (Fig. 2C and Table S2). Figure 2D showed the heatmap of 271 DE-ZRGs in tumor and adjacent nontumor tissues.

3.2 Identify a Pr-ZRGs signature in the training set

According to our screening criteria, a total of 99 Pr-ZRGs were screened in the training set (Table S3). In order to screen out the ZRGs with better prognostic value, we performed LASSO regression with 10-times cross-validations and obtained 11 candidate ZRGs (Fig. 3A, B). Then, these candidate ZRGs were further performed on multivariate Cox regression analysis. Finally, five ZRGs with the optimal prognostic value were identified (Fig. 3C). Figure 3D showed the expression levels of the optimal Pr-ZRGs in tumor and adjacent nontumor tissues. The prognostic power of each optimal Pr-ZRGs was shown in Fig. 4. Furthermore, we established a risk score formula through combining the expression levels of the five optimal Pr-ZRGs and their corresponding regression coefficients, and calculated the risk score of all 525 patients according to the formula. The formula is as follows: $(\text{Risk score} = \text{Exp}(\text{TRIM59}) * 0.327921772 - \text{Exp}(\text{VAV3}) * 0.077385777 - \text{Exp}(\text{ZNF189}) * 0.050279296 + \text{Exp}(\text{AGAP9}) * 0.21099254 + \text{Exp}(\text{PYGO1}) * 0.321962062)$

3.3 Evaluation and validation of the five-ZRG signature

We stratified patients into high-risk ($n = 132$) and low-risk group ($n = 132$) based on the median risk score (1.063) of the training set (Fig. 5A). The survival status of the patients was shown in Fig. 5B, while the heatmap of the expression levels of the five optimal Pr-ZRGs was shown in Fig. 5C. The KM survival curve shows that the high-risk patients had significantly worse overall survival (OS) than low-risk patients (Fig. 5D). The ROC curve shows that the five-ZRG signature had a reliable prognostic accuracy with the

area under the curve (AUC) was 0.812 (Fig. 5E). Besides, we measured the prognostic performance of the signature between different clinical subgroups and the results also showed that the high-risk patients had significantly worse OS than low-risk patients (Fig. 6A-D).

We further applied the risk score formula to the testing set and the entire set to validate the constructed five-ZRG signature. Consistent with the training set, both sets presented that the OS of high-risk groups was significantly lower than that of low-risk groups (Fig. 7A, B). The AUC of the testing set was 0.652 and that of the entire set was 0.745 (Fig. 7C, D). In summary, the signature constructed by us had good prognostic power for ccRCC patients.

3.4 Construction, calibration and validation of the nomogram

After univariate and multivariate Cox regression analysis, the risk score based on the five-ZRG signature and other clinical parameters (ages of patients, histologic grades and pathologic stages) were identified as the independent prognostic factors for ccRCC (Fig. 8A, B). Then, a nomogram consists of the risk score, ages of patients, histologic grades and pathologic stages was constructed to predict the survival of ccRCC patients at 1-, 3-, and 5-year (Fig. 8C).

The calibration plots revealed that the predicted outcome of the nomogram are in good agreement with the actual outcome at 1-, 3-, and 5-year (Fig. 8D). Besides, the KM survival analysis and the time-dependent ROC curve indicated that the nomogram has a better prognostic performance than the five-ZRG signature (Fig. 8E, F) and consistent results were got in the testing and entire set (Fig. 9A-D).

3.5 Functional enrichment analysis

The GO and KEGG enrichment analysis were performed on the 271 DE-ZRGs to identify the potential biological functions. The results of GO showed that DE-ZRGs were enriched in some biological processes including regulation of GTPase activity, transcription initiation from RNA polymerase α promoter, and intracellular receptor signaling pathway, etc. (Table S4 and Fig. 10A). KEGG pathway enrichment showed that DE-ZRGs were mainly enriched in five signaling pathways, including Herpes simplex virus 1 infection, Endocytosis, Phosphatidylinositol signaling system, Choline metabolism in cancer, and Glycerolipid metabolism (Table S5 and Fig. 10B).

As shown in Fig. 10C, the result of GSEA indicated that primary immunodeficiency, intestinal immune network for IgA production, cytokine-cytokine receptor interaction, base excision repair and p53 signaling pathway were closely associated with the worse OS of the high-risk patients (Fig. 10C).

3.6 Immune infiltration and tumor microenvironment related to five-ZRG signature

Immune infiltration analysis showed that the percentages of aDCs, T helper cells, pDCs, CD8⁺ T cells, Macrophages, Tfh, Th1 cells, Th2 cells, and TIL cells of the high-risk patients were significantly higher than those in the low-risk patients, while the percentages of iDCs and Mast cells were significantly lower (Fig. 11A). The immune function analysis showed that the scores of HLA, CCR, check-point, cytolytic activity, inflammation-promoting, APC co-stimulation and parainflammation were significantly higher in the high-risk patients, while Type χ IFN Response was significantly lower (Fig. 11B).

The tumor microenvironment analysis revealed that the high-risk patients had significantly higher immune scores and stromal scores than the low-risk patients, while the tumor purity scores of the high-risk patients were lower (Fig. 11C). The above results all indicated that there was a significant immune heterogeneity between the two groups stratified based on the five-ZRG signature, which provided a reference for exploring tumor development and immunotherapy.

3.7 Drug sensitivity analysis of the two subgroups

As the result, the IC₅₀ of 28 anticancer drugs differed between the two subgroups, and the IC₅₀ of these 28 drugs in the high-risk patients were all higher than low-risk patients (Fig. 12). This result also provided a reference for the treatment of ccRCC in the future.

4. Discussion

ccRCC is considered to be an immunogenic and heterogeneous tumor caused by various carcinogenic factors [29]. Although, ccRCC can be successfully treated with surgical strategies, some of patients will develop metastasis, which is the main cause of death in ccRCC patients [30]. To reduce the mortality of ccRCC patients and improve their prognosis, there is urgently required to screen significantly prognosis-related biomarkers for patients with ccRCC. For the past few years, ZNFs have been found to be closely correlated with various biological processes, and they are significantly involve in the ontogeny, progression and prognosis of many tumors, indicating that ZNFs have the potential to be used as prognosis-related biomarkers. The objective of this study is to construct a group of significantly Pr-ZRGs of ccRCC and to stratify risk of ccRCC patients by ZRGs signature, so as to achieve individual prognosis analysis and then guide early intervention and treatment of ccRCC patients.

We analysed the DE-ZRGs between tumor tissues and adjacent normal renal tissues through TCGA database and 271 DE-ZRGs were finally screened out. Then, through a series of regression analysis in the training set, we constructed a five-ZRG signature, including TRIM59, VAV3, ZNF189, AGAP9 and PYGO1, which is validated an independent prognostic factor of patients with ccRCC. According to the risk score formula based on five-ZRG signature, patients were stratified into two groups (high-risk and low-risk). KM survival analysis showed that the high-risk patients had a significantly lower OS time than that in the low-risk patients. ROC curve showed that the combination of five ZRGs signature had good prognostic accuracy. We further validated these findings in the testing set and entire set. Furthermore, to establish a more reliable and individualized clinical prediction method, we incorporated the five-ZRG signature with other prognosis-related clinical factors, including age, stage and grade, to construct a nomogram. The

predicted outcome of the nomogram in 1-, 3-, and 5-year OS showed a good agreement in actual outcome in the calibration plot, and KM survival analysis and ROC curve indicated that the nomogram also had a better prognostic performance.

Herein, we identified five closely Pr-ZRGs of ccRCC patients. Interestingly, previous studies have reported that they were involved in the initiation, progression and prognosis of numerous tumors. Rongkun Li et al. found that TRIM59 could enhance glycolysis by activating the PI3K/AKT/mTOR pathway, which eventually led to the progression of pancreatic cancer [31]. Studies showed that up-expression of TRIM59 was indicated the poor prognosis of NSCLC and breast cancer [32, 33]. As a new carcinogenic EGFR/STAT3 signal regulator, TRIM59 promoted the occurrence of glioblastoma by inhibiting STAT3 dephosphorylation, which related to poor prognosis of glioblastoma [34]. In addition, TRIM59 might promote prostate cancer cell proliferation and tumor growth through its effect on cell cycle progression [35]. VAV3, a member of the VAV family of oncoproteins, is a carcinogenic gene [36]. Studies showed that overexpression of VAV3 led to abnormal expression of cell cycle control and metastasis-related molecules in colorectal cancer cells by activating PI3K/AKT signal pathway, which might be used as an independent prognosis-related biomarker of colorectal cancer patients [37]. VAV3 is also overexpressed in gastric cancer, and its overexpression is related to poor prognosis and short survival time of patients [38]. Additionally, VAV3 is correlated with the oncogenesis and development of prostate carcinoma by regulating the activity of androgen receptor [39, 40]. PYGO1 is a key regulatory gene of Wnt signaling pathway, which is correlated with the development of kidney and the transcription of colorectal carcinoma cells [41, 42]. Herein, we conducted GO and KEGG analysis of 271 DE-ZRGs. The result of GO analysis showed that DE-ZRGs were mainly associated with the regulation of GTPase activity and the ubiquitin-related proteases activity, suggesting that these ZRGs were involved in the ontogeny and progression of tumors such as migration, invasion and tumor angiogenesis [43, 44]. KEGG enrichment analysis showed ccRCC ontogeny and progression may related to the signal pathways of endocytosis, phosphatidylinositol signaling system, choline metabolism in cancer and glycerolipid metabolism. Furthermore, to explore the potential biological effects of five screened ZRGs, we conducted GSEA analysis and found that they were related to intestinal immune network, natural killer cell-mediated cytotoxicity and cytokine receptor interaction, suggesting that the five Pr-ZRGs signature may be strongly correlated with the immune response process in the microenvironment of ccRCC.

In the two groups of ccRCC patients based on five-ZRG signature model classification, we found that immune infiltrating cells in the tumors of high-risk group were mainly T cells, tumor-infiltrating lymphocytes and macrophages, suggesting that chronic inflammation may promote tumor progression and reduce the survival time of patients. Further analysis of immune function in the two groups showed that CCR, check-point, cytolytic activity, para-inflammation and T cell co-stimulation were significantly increased in high-risk ccRCC patients. In the TME score, they also had significantly higher stromal score, immune score, and estimate score and lower tumor purity than that in low-risk patients. These results indicate that immune microenvironment of ccRCC plays a crucial role in the development of tumors. Immune cells and stromal cells in TME may promote the proliferation and migration of tumor cells, which resulted in the poor prognosis of ccRCC patients. Finally, we found that the different groups of patients

have different sensitivities to many antineoplastic drugs. For the 28 antineoplastic drugs identified, high-risk patients had a higher IC50, suggesting that tumor resistance in the high-risk patients was stronger. However, some limitations have existed in this research. First, this study is only a mining analysis of the database based on bioinformatics, multicenter and prospective studies are needed to further confirmed these findings accurately. Second, it is quite necessary to further explore the potential molecular mechanism of the proteins encoded by the five ZRGs in ccRCC.

5. Conclusions

Our study identified a novel prognosis-related five-ZRG signature with high accuracy, which is an independent prognostic factor for ccRCC. Furthermore, we integrated five-ZRG signature and other clinical features to construct a nomogram, which can predict the clinical outcome of ccRCC patients more accurately and reliably. Subsequently, we found that the five-ZRG signature were not only related to the prognosis of ccRCC, but related to tumor immune microenvironment and drug sensitivity. These findings may provide a new field of vision for immunotherapy and molecular targeted therapy in patients with ccRCC.

Abbreviations

ccRCC, clear cell renal cell carcinoma; TCGA, The Cancer Genome Atlas; ZNF, zinc finger; ZRGs, zinc finger-related genes; DE-ZRGs, differentially expressed zinc finger-related genes; LASSO, least absolute shrinkage and selection operator; ROC, receiver operating characteristic; KM, Kaplan–Meier; GSEA, Gene Set Enrichment Analysis; GDSC, Genomics of Drug Sensitivity in Cancer; ZNFs, zinc finger proteins; TME, tumor microenvironment; GO, gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; ssGSEA, single sample gene set enrichment analysis; IC50, half maximal inhibitory concentration; AUC, area under the curve.

Declarations

Ethical approval and consent to participate

Not applicable

Consent for publication

All authors have contributed to, read, and approved this submitted manuscript in its current form and consent for its publication.

Availability of supporting data

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

Not applicable

Authors' contributions

All authors made a substantial, direct, and intellectual contribution to the work. F.Z. and J.W. wrote the manuscript and designed the figures. J.Z., P.C. and Z.S. revised the manuscript. W.W. supervised the manuscript. F.Z. and J.W. contributed equally to the article.

Acknowledgments

We are grateful to the staff members of TCGA, Uniprot, GSEA and GDSC databases, which made the genomic data and clinical data of clear cell renal cell carcinoma available. We also thank Dr Hongpan Zhang for the constructive suggestions for the analysis.

Authors' information

¹ Department of Urology, Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020, China. ² Institute of Urology, Capital Medical University, Beijing, China.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68(6):394-424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin.* 2020;70(1):7-30.
3. Hsieh JJ, Purdue MP, Signoretti S, et al. Renal cell carcinoma. *Nat Rev Dis Primers.* 2017;3:17009.
4. Linehan WM. Genetic basis of kidney cancer: role of genomics for the development of disease-based therapeutics. *Genome Res.* 2012;22(11):2089-2100.
5. Wolff I, May M, Hoschke B, et al. Do we need new high-risk criteria for surgically treated renal cancer patients to improve the outcome of future clinical trials in the adjuvant setting? Results of a comprehensive analysis based on the multicenter CORONA database. *Eur J Surg Oncol.* 2016;42(5):744-750.
6. Wolfe SA, Nekludova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct.* 2000;29:183-212.

7. Persikov AV, Wetzel JL, Rowland EF, et al. A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. *Nucleic Acids Res.* 2015;43(3):1965-1984.
8. Vilas CK, Emery LE, Denchi EL, Miller KM. Caught with One's Zinc Fingers in the Genome Integrity Cookie Jar. *Trends Genet.* 2018;34(4):313-325.
9. Laity JH, Lee BM, Wright PE. Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol.* 2001;11(1):39-46.
10. Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP. Sticky fingers: zinc-fingers as protein-recognition motifs. *Trends Biochem Sci.* 2007;32(2):63-70.
11. Singh JK, van Attikum H. DNA double-strand break repair: Putting zinc fingers on the sore spot. *Semin Cell Dev Biol.* 2020.
12. Jen J, Wang YC. Zinc finger proteins in cancer progression. *J Biomed Sci.* 2016;23(1):53.
13. Henrion M, Frampton M, Scelo G, et al. Common variation at 2q22.3 (ZEB2) influences the risk of renal cancer. *Hum Mol Genet.* 2013;22(4):825-831.
14. Fang Y, Wei J, Cao J, et al. Protein expression of ZEB2 in renal cell carcinoma and its prognostic significance in patient survival. *PLoS One.* 2013;8(5):e62558.
15. Yu S, Dai J, Ma M, et al. RBCK1 promotes p53 degradation via ubiquitination in renal cell carcinoma. *Cell Death Dis.* 2019;10(4):254.
16. Li C, Li H, Zhang T, et al. ZFX is a Strong Predictor of Poor Prognosis in Renal Cell Carcinoma. *Med Sci Monit.* 2015;21:3380-3385.
17. Kang HW, Park H, Seo SP, et al. Methylation Signature for Prediction of Progression Free Survival in Surgically Treated Clear Cell Renal Cell Carcinoma. *J Korean Med Sci.* 2019;34(19):e144.
18. Kim B, Jung M, Moon KC. The Prognostic Significance of Protein Expression of CASZ1 in Clear Cell Renal Cell Carcinoma. *Dis Markers.* 2019;2019:1342161.
19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139-140.
20. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1-22.
21. Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biom J.* 2010;52(1):70-84.
22. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics.* 2000;56(2):337-344.
23. Eng KH, Schiller E, Morrell K. On representing the prognostic value of continuous gene expression biomarkers with the restricted mean survival curve. *Oncotarget.* 2015;6(34):36308-36318.
24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284-287.
25. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.*

- 2005;102(43):15545-15550.
26. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013;39(4):782-795.
 27. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
 28. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013;41(Database issue):D955-961.
 29. Diaz-Montero CM, Rini BI, Finke JH. The immunology of renal cell carcinoma. *Nat Rev Nephrol*. 2020;16(12):721-735.
 30. Jonasch E, Walker CL, Rathmell WK. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat Rev Nephrol*. 2020.
 31. Li R, Weng L, Liu B, et al. TRIM59 predicts poor prognosis and promotes pancreatic cancer progression via the PI3K/AKT/mTOR-glycolysis signaling axis. *J Cell Biochem*. 2020;121(2):1986-1997.
 32. Hao L, Du B, Xi X. TRIM59 is a novel potential prognostic biomarker in patients with non-small cell lung cancer: A research based on bioinformatics analysis. *Oncol Lett*. 2017;14(2):2153-2164.
 33. Tan P, Ye Y, He L, et al. TRIM59 promotes breast cancer motility by suppressing p62-selective autophagic degradation of PDCD10. *PLoS Biol*. 2018;16(11):e3000051.
 34. Sang Y, Li Y, Song L, et al. TRIM59 Promotes Gliomagenesis by Inhibiting TC45 Dephosphorylation of STAT3. *Cancer Res*. 2018;78(7):1792-1804.
 35. Lin WY, Wang H, Song X, et al. Knockdown of tripartite motif 59 (TRIM59) inhibits tumor growth in prostate cancer. *Eur Rev Med Pharmacol Sci*. 2016;20(23):4864-4873.
 36. Movilla N, Bustelo XR. Biological and regulatory properties of Vav-3, a new member of the Vav family of oncoproteins. *Mol Cell Biol*. 1999;19(11):7870-7885.
 37. Uen YH, Fang CL, Hseu YC, et al. VAV3 oncogene expression in colorectal cancer: clinical aspects and functional characterization. *Sci Rep*. 2015;5:9360.
 38. Tan B, Li Y, Shi X, et al. Expression of Vav3 protein and its prognostic value in patients with gastric cancer. *Pathol Res Pract*. 2017;213(5):435-440.
 39. Liu Y, Wu X, Dong Z, Lu S. The molecular mechanism of Vav3 oncogene on upregulation of androgen receptor activity in prostate cancer cells. *Int J Oncol*. 2010;36(3):623-633.
 40. Rao S, Lyons LS, Fahrenholtz CD, et al. A novel nuclear role for the Vav3 nucleotide exchange factor in androgen receptor coactivation in prostate cancer. *Oncogene*. 2012;31(6):716-727.
 41. Schwab KR, Patterson LT, Hartman HA, et al. Pygo1 and Pygo2 roles in Wnt signaling in mammalian kidney development. *BMC Biol*. 2007;5:15.
 42. Thompson B, Townsley F, Rosin-Arbesfeld R, Musisi H, Bienz M. A new nuclear component of the Wnt signalling pathway. *Nat Cell Biol*. 2002;4(5):367-373.

43. Chinigo G, Fiorio Pla A, Gkika D. TRP Channels and Small GTPases Interplay in the Main Hallmarks of Metastatic Cancer. *Front Pharmacol.* 2020;11:581455.
44. Yang YL, Zhang Y, Li DD, et al. RNF144A functions as a tumor suppressor in breast cancer through ubiquitin ligase activity-dependent regulation of stability and oncogenic functions of HSPA2. *Cell Death Differ.* 2020;27(3):1105-1118.

Figures

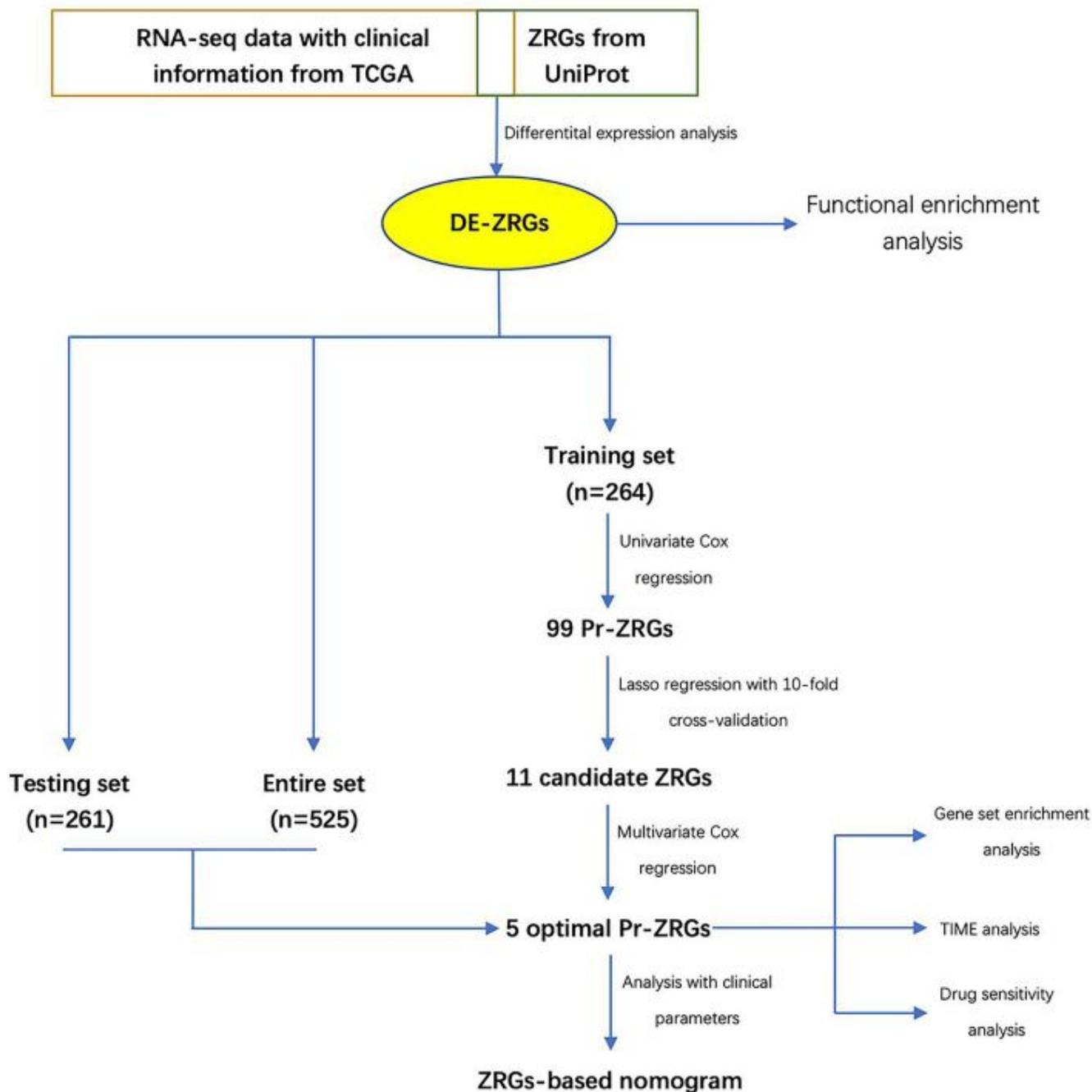


Figure 1

The flowchart of this study. TCGA: The Cancer Genome Atlas; ZRGs: Zinc finger(ZNF)-related-genes; DE-ZRGs: differentially expressed ZNF-related-genes; Pr-ZRGs: prognosis-related ZNF-related-genes; Lasso: Least absolute shrinkage and selection operator. TIME: tumor immune microenvironment.

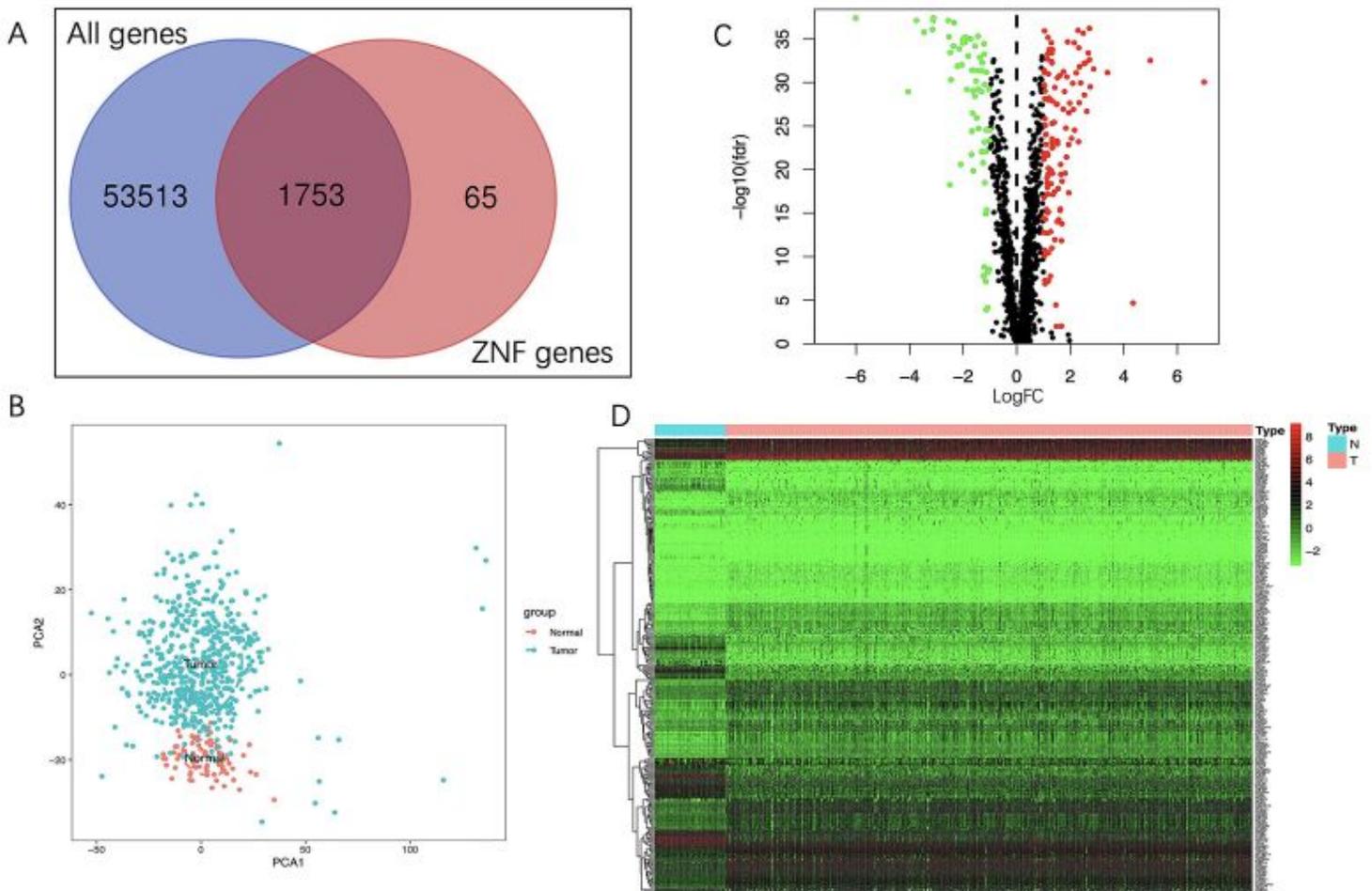


Figure 2

Screening of the DE-ZRGs. (A) The ZRGs expressed in tumor specimens and adjacent nontumor renal specimens. (B) Principal components analysis (PCA) plot of the expression profile of ZRGs between tumor and normal tissue. (C) Volcano plot of DE-ZRGs between tumor and normal tissue. (D) Heatmap of the DE-ZRGs.

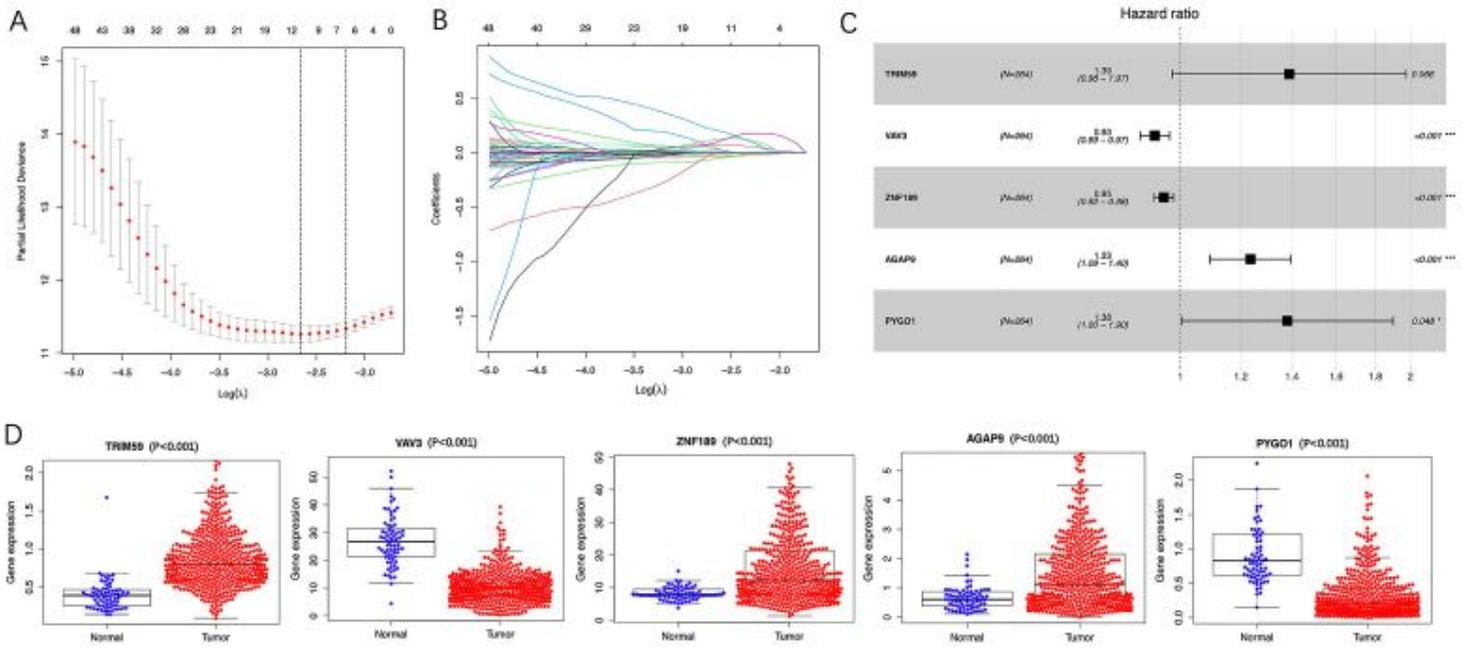


Figure 3

Identification of 5 optimal Pr-ZRGs and their expression levels in ccRCC. (A) LASSO coefficients profiles of 99 Pr-ZRGs. (B) 11 candidate ZRGs obtained by LASSO regression with 10-fold cross-validation using minimum λ value. (C) 5 optimal Pr-ZRGs got by Multivariate Cox regression analysis. (D) Expression pattern of the 5 optimal Pr-ZRGs between tumor and normal kidney tissue.

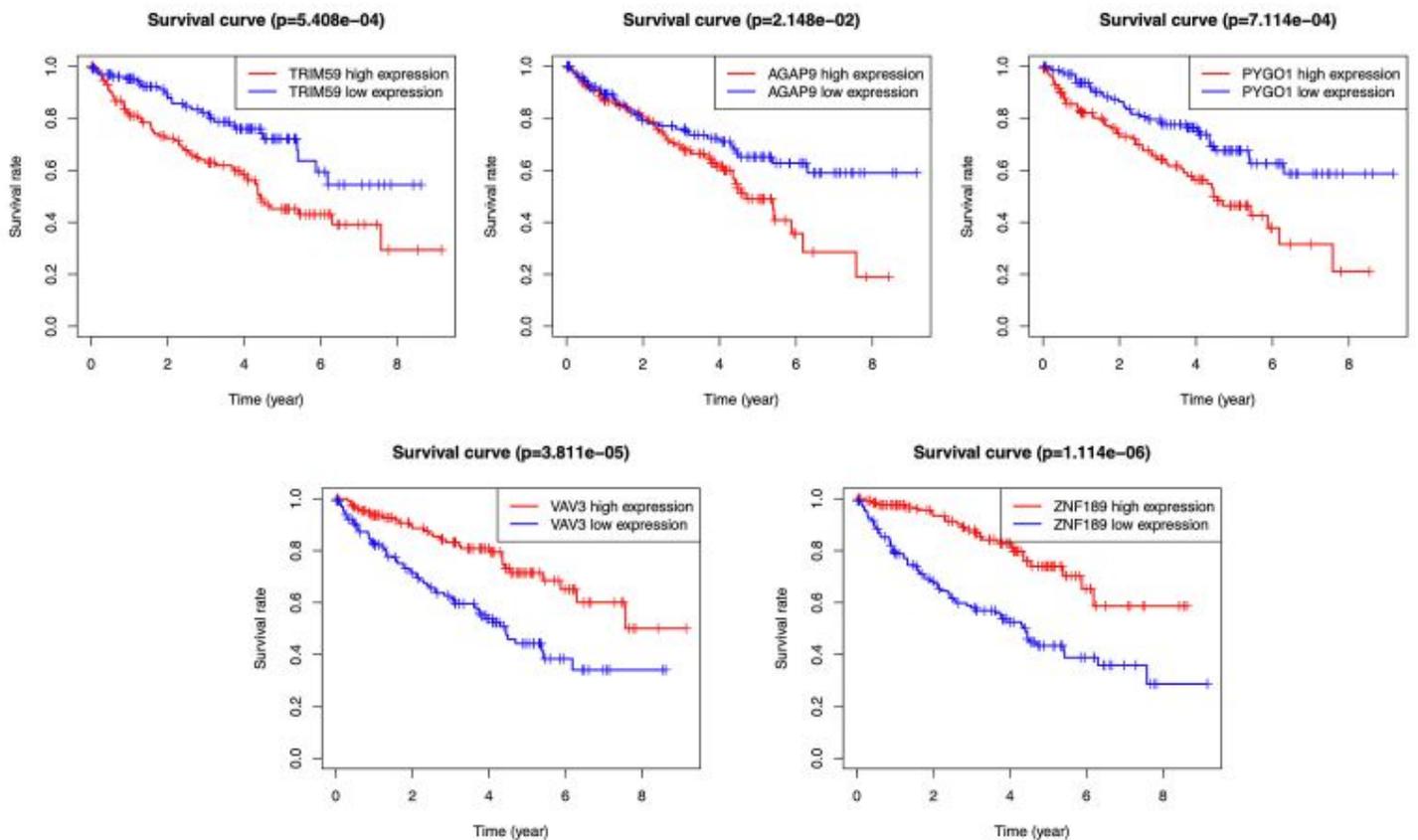


Figure 4

Kaplan–Meier (KM) survival analysis of TRIM59, AGAP9, PYGO1, VAV3 and ZNF189.

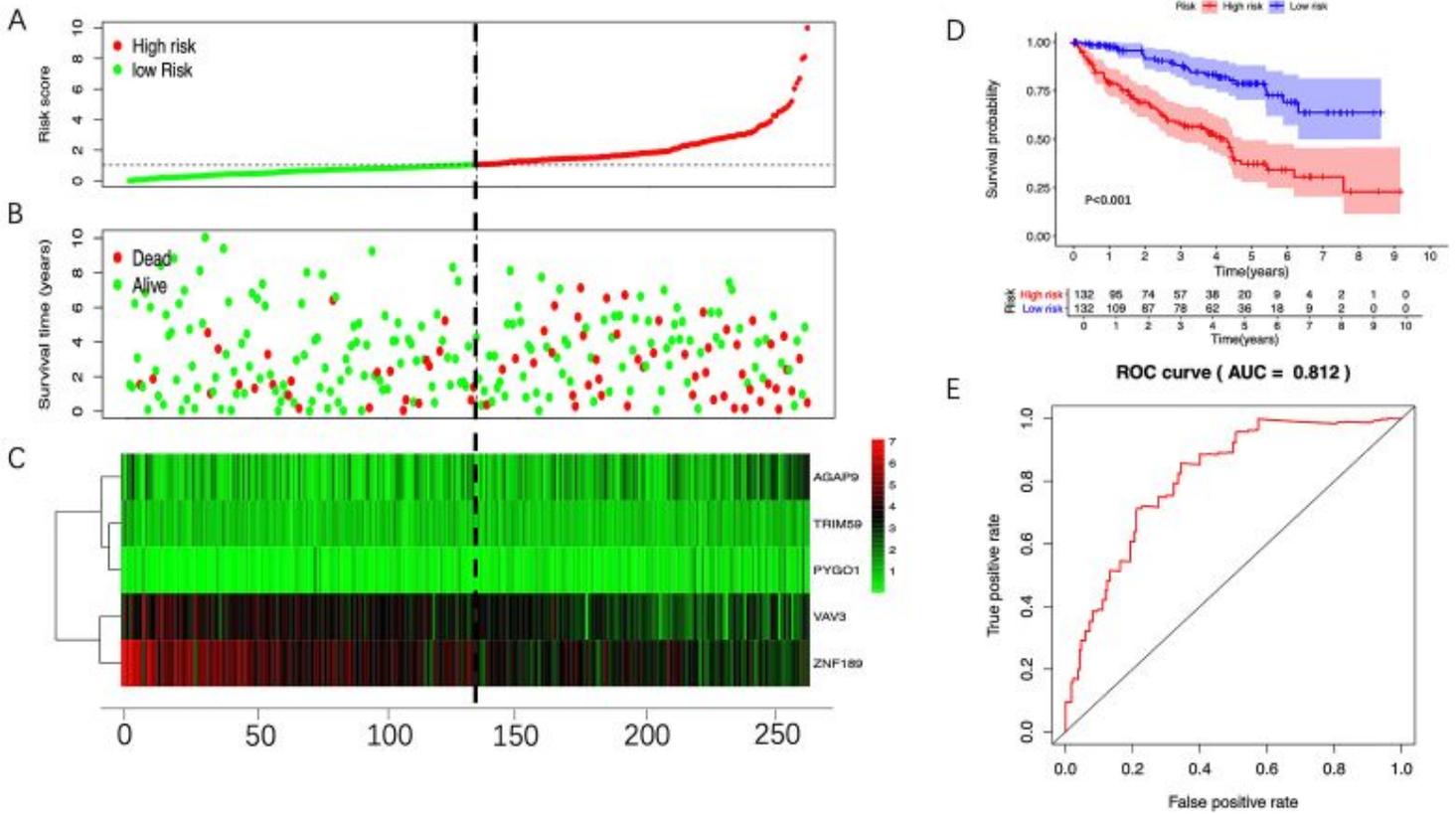


Figure 5

Prognostic analysis of five-ZRG signature in the training set. The dotted line represented the median risk score and stratified the ccRCC patients into low- and high-risk group. (A) The curve of risk score. (B) Survival status of the ccRCC patients. (C) Heatmap of the expression levels of the five optimal Pr-ZRGs in low- and high-risk group. (D) KM survival analysis of the five-ZRG signature. (E) Time-dependent ROC curve analysis of the five-ZRG signature.

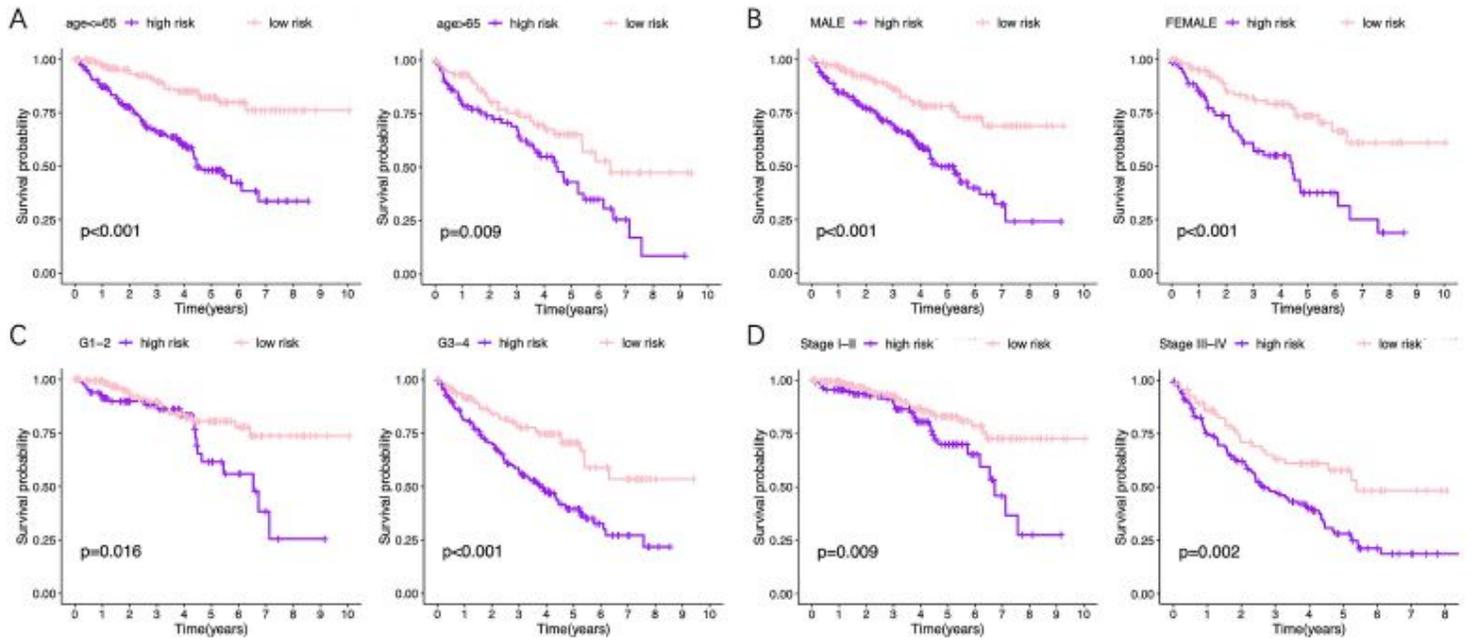


Figure 6

KM survival analysis of the five-ZRG signature in different subgroups including (A) younger than 65 years old, older than 65 years old; (B) male, female; (C) grade 1/2, grade 3/4; (D) stage I-II, stage III-IV.

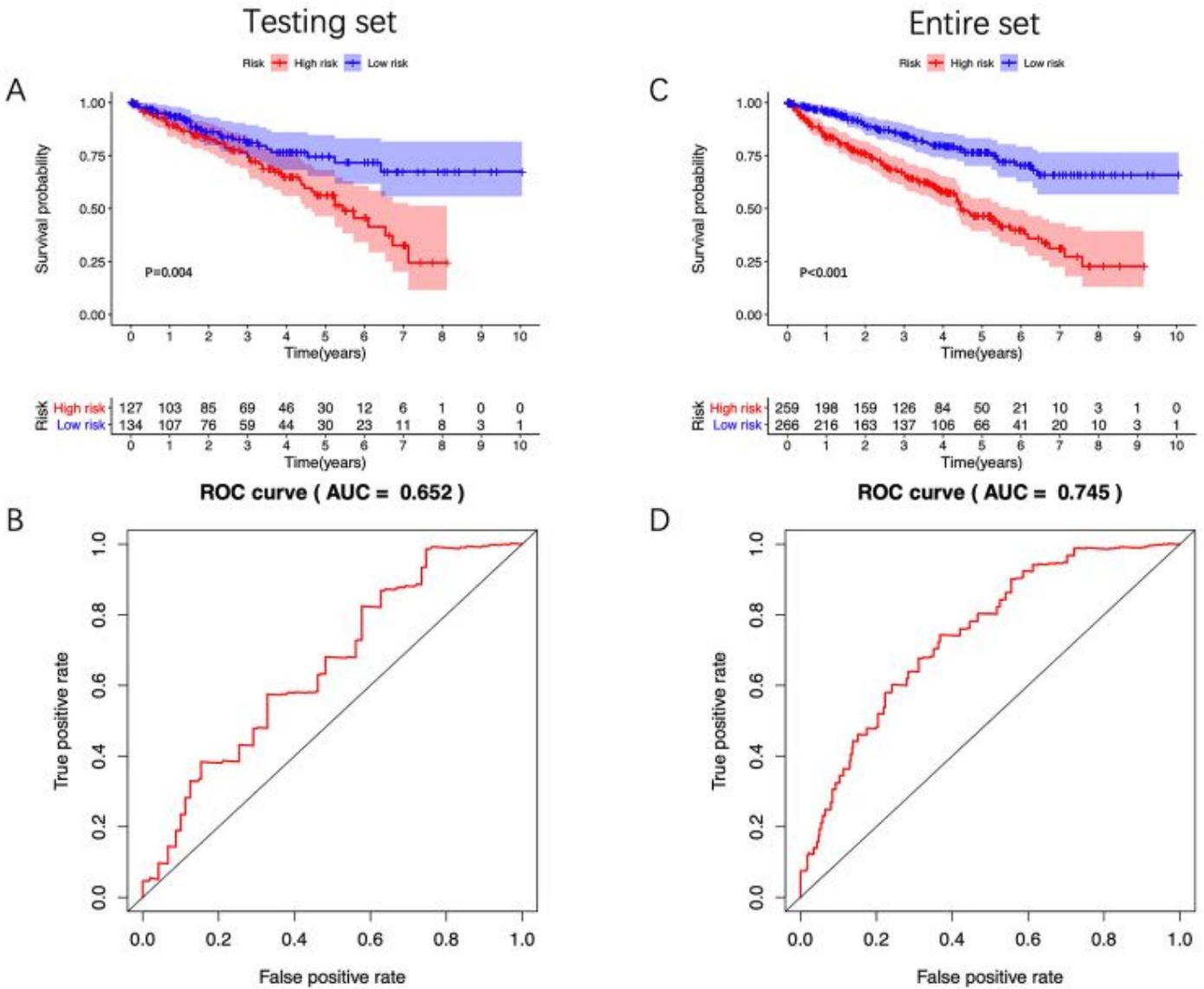


Figure 7

Validation of the five-ZRG signature. KM survival analysis of the five-ZRG signature in the testing set (A) and entire set (C). Time-dependent ROC curve analysis of the five-ZRG signature in the testing set (B) and entire set (D).

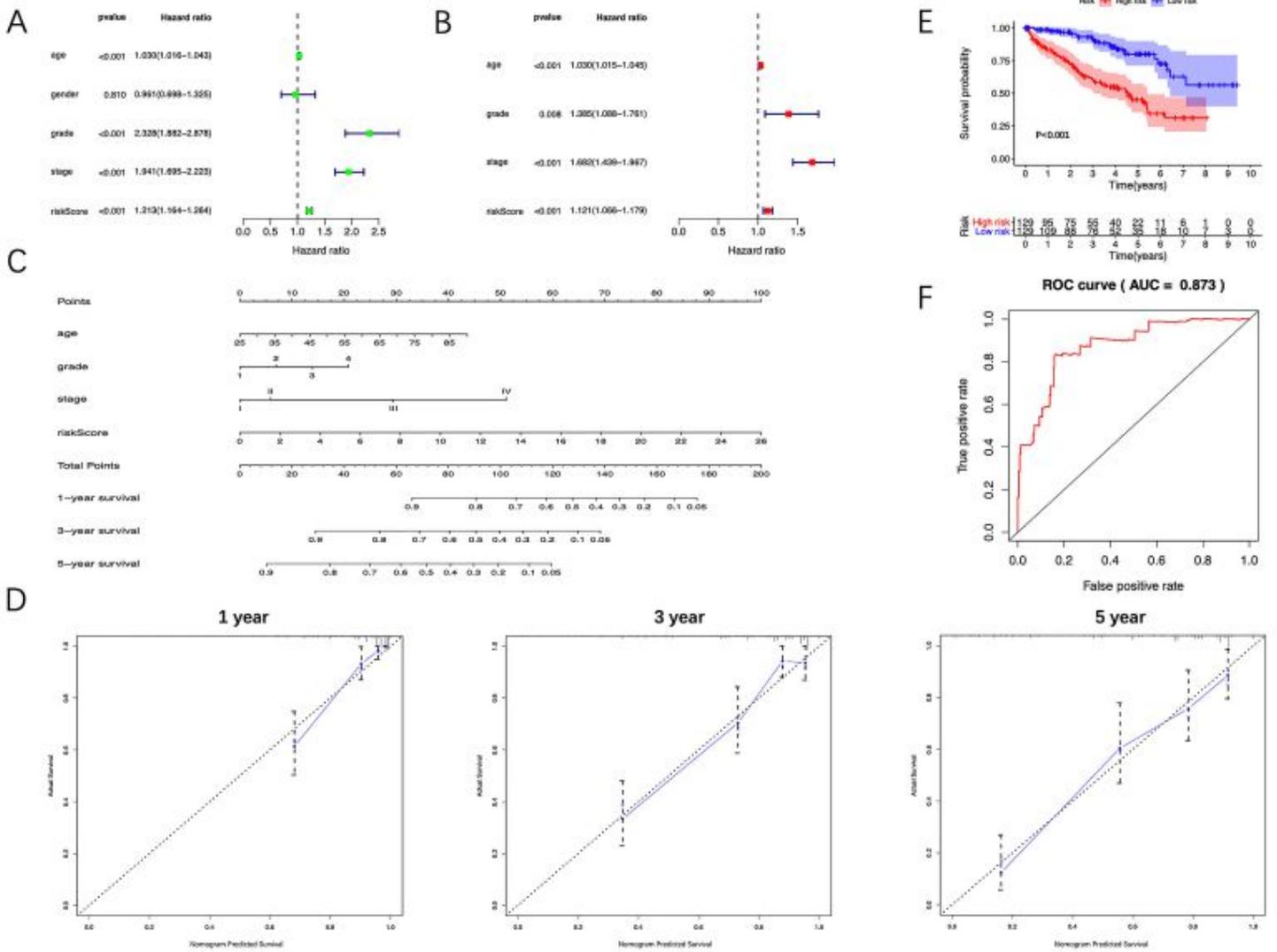


Figure 8

Identifying the independent prognostic factors and construction of ZRG-based nomogram. (A) Forrest plot of univariate Cox regression analysis in ccRCC. (B) Forrest plot of multivariate Cox regression analysis in ccRCC. (C) Nomogram integrated five ZGR-based risk score, age, grade and stage. (D) The calibration plot of the nomogram for agreement test between 1-, 3- and 5-year OS prediction and the actual outcome. (E) OS of the high-risk group was significantly worse than that of the low-risk group. (F) Time-dependent ROC curves of the nomogram.

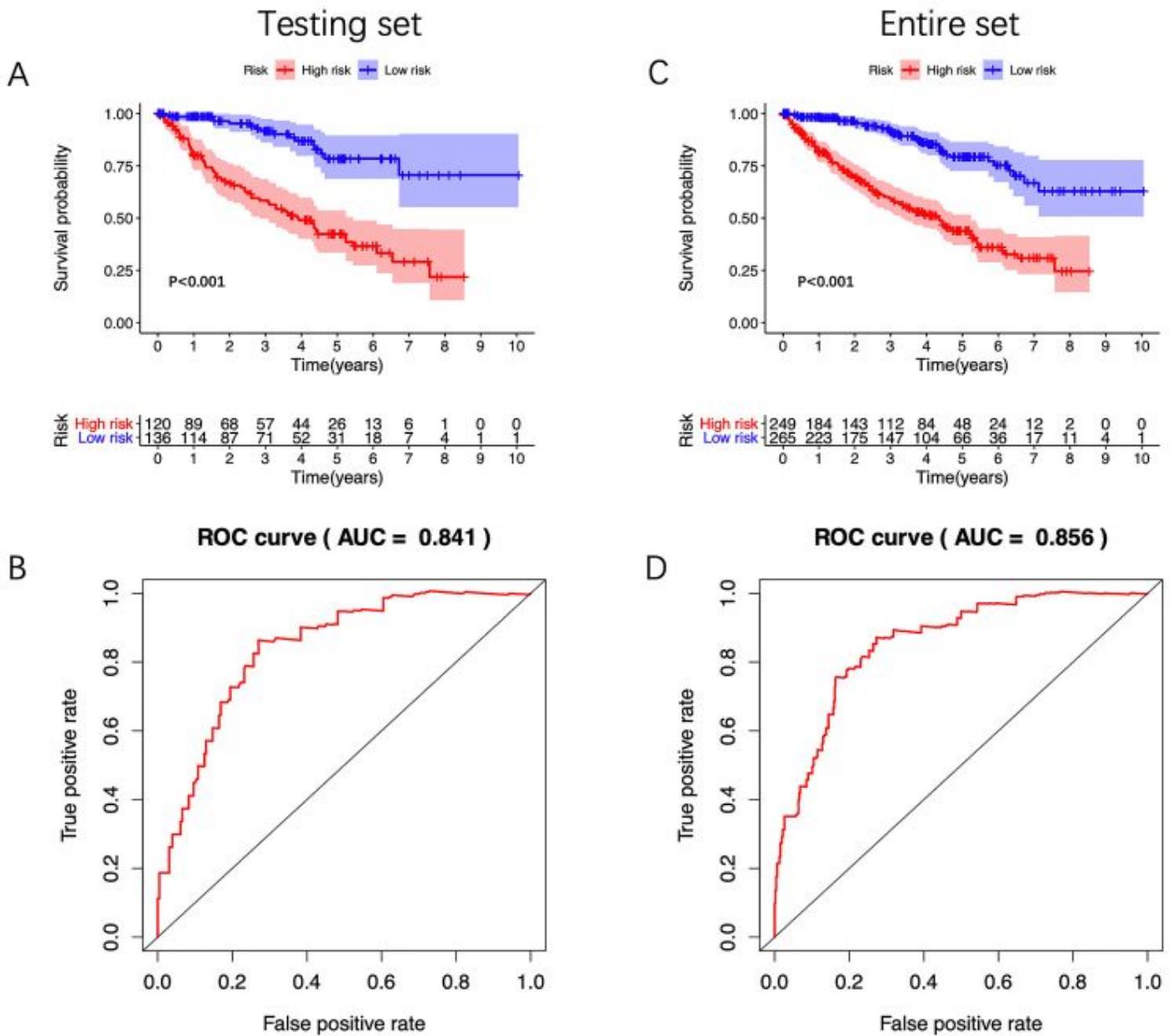


Figure 9

Validation of the nomogram in the testing and entire set. KM survival analysis of the nomogram in the testing set (A) and entire set (C). Time-dependent ROC curve analysis of the nomogram in the testing set (B) and entire set (D).

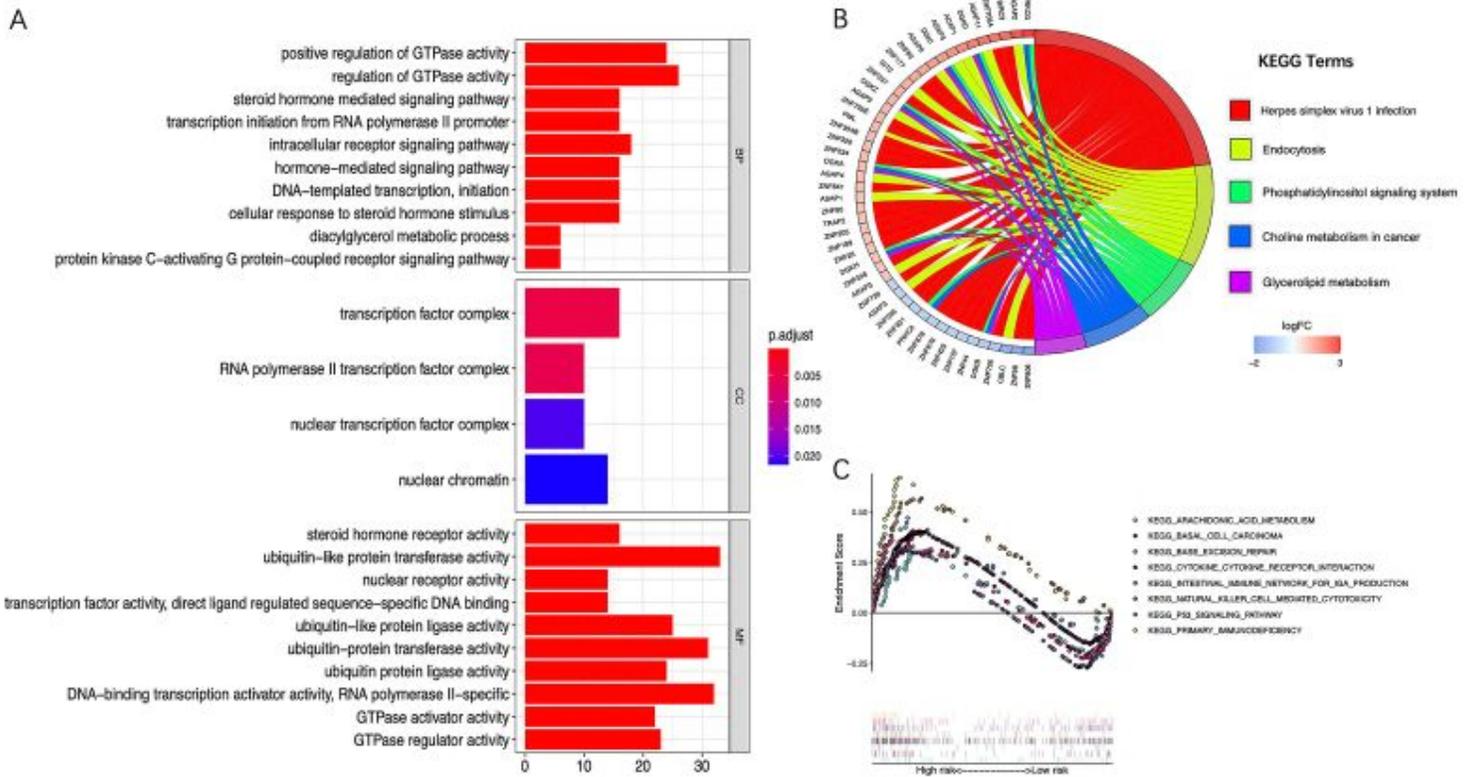


Figure 10

Functional enrichment analysis. (A) GO enrichment analysis of the DE-ZRGs. (B) KEGG enrichment analysis of the DE-ZRGs. (C) GSEA plot for significant enriched biological processes associated with the five optimal Pr-ZRGs.

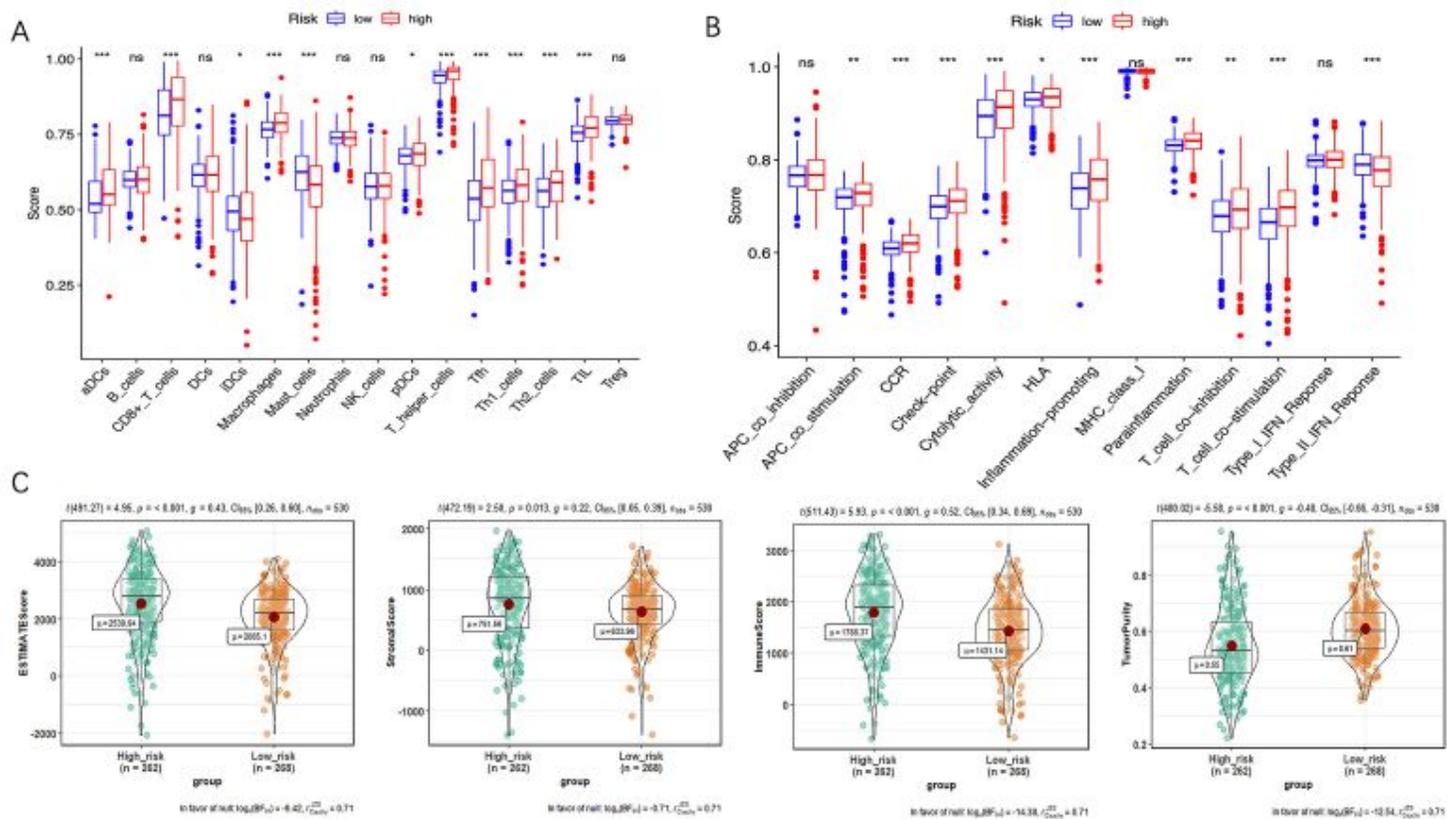


Figure 11

The landscape of immune infiltration and tumor microenvironment in ccRCC. (A) The infiltrating levels of immune cell in high- and low-risk group. (B) The immune function between high- and low-risk group. (C) The tumor microenvironment between high- and low-risk group.

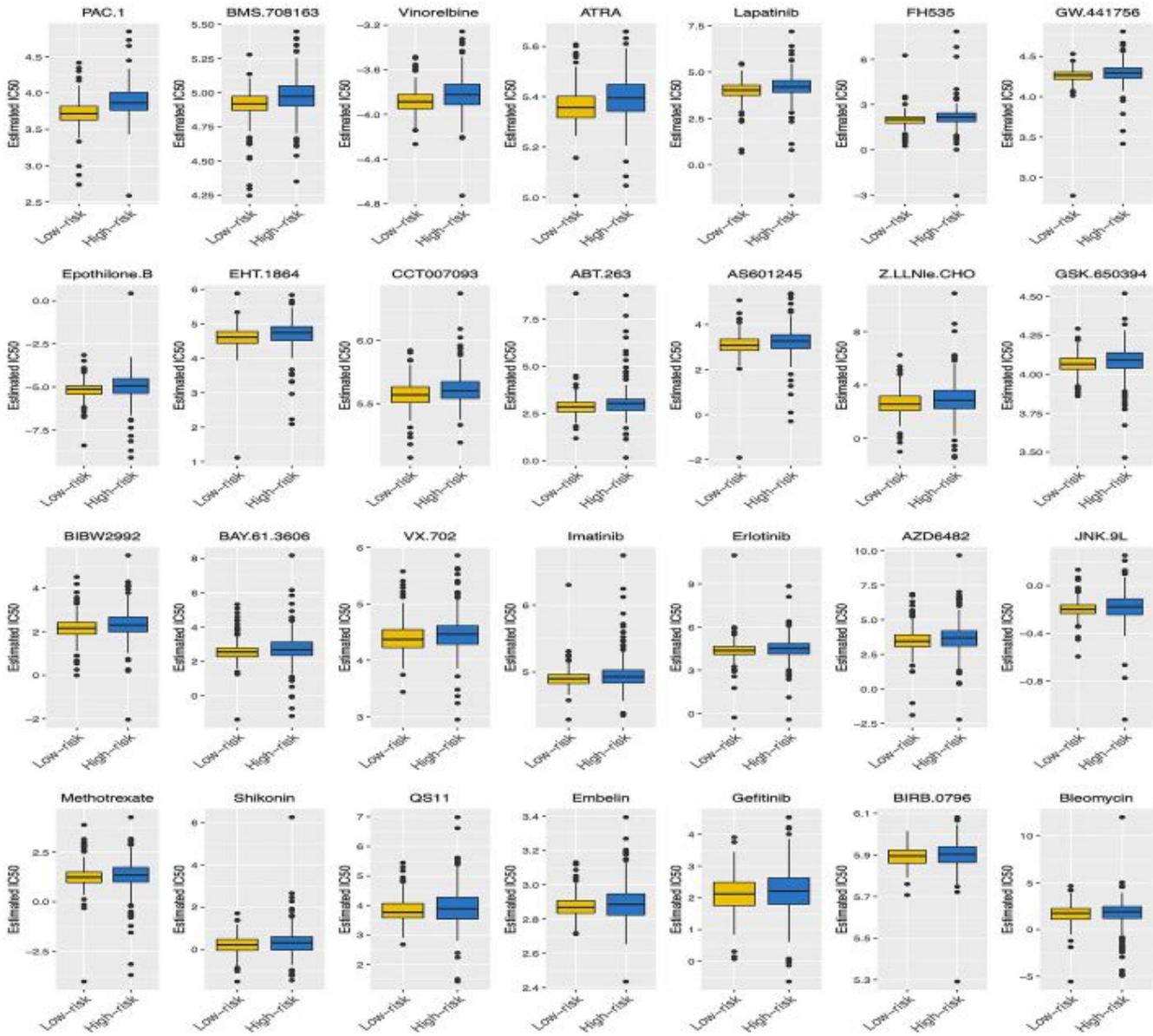


Figure 12

Drug sensitivity analysis between the high-and low-risk group.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile4.xlsx](#)
- [Additionalfile5.xlsx](#)

- [Descriptionofadditionalfiles.docx](#)