AC010973.2 Promotes Cell Proliferation and Is One of Six Stemness-Related Genes That Predict Overall Survival of Renal Clear Cell Carcinoma

Yingqing Liu

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Yuang Wei

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing

Xu Zhang

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing

Xiaohan Ren

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing

Jiawei Wang

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Lin Li

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Haibo Qin

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Wei Ding

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Xudong Shen

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Guangyao Li

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Zhongwen Lu

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Dong Zhang

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing

Chao Qin

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing

Lingsong Tao

Department of Urology, The Second People's Hospital of Wuhu, Anhui Province

Xinglin Chen (✉ nmuchenxinglin@163.com)

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing

Research Article

Abstract

Background

Extensive research has revealed that tumor stemness plays a central role in promoting tumor progression. However, the underlying involvement of stemness-related genes in renal clear cell carcinoma (ccRCC) remains controversial.

Methods

The data used for bioinformatics analysis were downloaded from The Cancer Genome Atlas database. The R software, SPSS and GraphPad Prism 8 were used for mapping and statistical analysis.

Results

We first quantified the stemness index of each patient through a machine learning algorithm. Then, we identified the differentially expressed genes between high and low stemness index as stemness-related genes. Based on these genes, we finally established a stable and effective prognosis model to predict patients' overall survival using a random forest algorithm (Training cohort; 1-year AUC: 0.67; 3-year AUC: 0.79; 5-year AUC: 0.73; Validation cohort; 1-year AUC: 0.66; 3-year AUC: 0.71; 5-year AUC: 0.7). The model genes include AC010973.2, RNU6-125P, AP001209.2, Z98885.1, KDM5C-IT1 and AL021368.3. The gene AC010973.2 was selected for further research for its highest importance. In vitro experiments demonstrated that AC010973.2 is highly expressed in ccRCC tissue and cell lines. Meanwhile, knockdown of AC010973.2 could significantly hamper the proliferation of ccRCC cells according to the colony formation and CCK8 assays.

Conclusion

In summary, our finding indicated that the stemness-related gene AC01097.3 is closely associated with patients' survival and could remarkably facilitate cell proliferation in ccRCC, making it potential to be a novel therapeutic target.

Introduction

Renal cell carcinoma (RCC) is the most common type of kidney cancer in human beings. The disease comprises > 10 histological and molecular subtypes, among which clear cell renal cell carcinoma (ccRCC) is the dominant form with a proportion of 75%-80% and an incidence of > 5%, along with frequent metastasis and poor prognosis [1]. The remaining subtypes, including papillary RCC, chromophobe RCC and unclassified RCC are called non-clear cell RCC (nccRCC). Nearly 92% of ccRCC carries the inactivation or mutation of the von Hippel-Lindau (VHL) gene located on human chromosome 3p,

resulting in highly vascularized and lipogenic characteristics of tumor tissues [2]. Early renal cancers efficiently benefit from partial nephrectomy (PN), while treatment for metastatic ccRCC is a challenge because the tumor is inherently insensitive to chemotherapy or radiotherapy, and few people show enough response to immunotherapy [3].

Cancer stem cell (CSC) is a kind of cell possessing the ability of continuously self-renewing and generating differentiated cells in malignant tumors. In the process of neoplasia, specific single cells in tissues grow hierarchically through a range of genetic and epigenetic modifications, generate the evolution of pre-malignant subpopulations, and finally display the malignant features as cancer cells [4,5]. It is believed that CSCs are necessary for tumor's aberrant proliferation, survival, invasion, and metastasis, for which reason CSCs have also been indicated as 'tumor-initiating cells' with reference to the result of Working Conference on CSCs [6]. In ccRCC, there is growing evidence of correlation between tumor progression and stemness-related pathways, as well as the pivotal role of CSCs in initiating the derivation of malignant subpopulations [7]. Prior studies have demonstrated that some biomarkers and chemokines were quite relevant to tumorigenesis and prognosis of renal cancer. Yu et al. concluded that patients with higher expression of NANOG and OCT4 in tumor tissues had lower survival rate [8]. Arezoo and colleagues also found that coexpression of OCT4 and NANOG in renal cancer was significantly associated with RCC subtypes and predicted poor prognosis of ccRCC patients, and additionally, they evaluated CXCR4 as a novel renal CSC marker in renal cancer using immunohistochemistry (IHC) [9,10]. Utilizing the same technique, Kyungeun et al. confirmed that CD133 was a favorable prognostic CSC biomarker for ccRCC [11]. However, in addition to histological methods to identify potential biomarkers, seldom integrative approaches have been utilized to explore how CSCs link to ccRCC.

In the present study, we performed comprehensive bioinformatic analysis and molecular research to investigate the bio function and prognostic value of stemness-associated genes in ccRCC. The results illuminated the prognostic and oncogenic role of AC010973.2 in ccRCC using multiple databases and ccRCC cell lines.

Methods

Public data acquisition and tumor stemness calculation

Expression profile and clinical information of ccRCC patients were downloaded from the TCGA database (TCGA-KIRC), including 72 normal tissue and 539 tumor tissue. All the data were preprocessed with the following steps: data merging and normalization, probe annotation, missing value completion. The one-class logistic regression machine learning algorithm (OCLR) was utilized to obtain gene expression-based stemness index (mRNAsi) and epigenetic regulation based-stemness index (EREG-mRNAsi). All patients were divided into high and low stemness groups according to the calculated stemness value. The limma package was used to identify differentially expressed genes (DEGs) between high and low stemness

groups with the threshold of |logFC > 1| and P.value < 0.05. These DEGs were defined as stemness-related genes.

Gene enrichment analysis and protein interaction

Gene ontology (GO) analysis was performed using clusterprofiler package to explore the underlying biological role of genes, consisting of three terms, biological process (BP), cellular component (CC) and molecular function (MF). Gene set enrichment analysis was conducted to compare the difference of biological pathways between two groups. The reference gene set was "Hallmark.7.4.symbol.gmt". Protein interaction of identified genes was performed in STRING online database and visualized by Cytoscape v3.7.2 software. Cytohubba plug-in was used to calculate the importance of all nodes and identify the key nodes.

Construction of prognostic prediction model based on stemness-related genes

All the patients were randomly divided into training and validation groups. Univariate cox regression analysis was firstly performed to identify prognosis-related genes. Then, dimensionality reduction was computed using random forest algorithm with ntree = 1000. Each gene was allocated a relative importance value. Multivariate cox regression analysis was conducted based on the top ten important genes to construct a prognostic prediction model. All the patients included in the prognosis model were assigned a riskscore value. Receiver operating characteristic (ROC) curve was used to assess the stability of the prognosis model. The area under roc curve (AUC) value > 0.7 was considered to have good predictive capabilities. Kaplan-Meier survival curves was used to compare the prognosis difference between two groups.

Nomogram, decision curve analysis (DCA) and calibration curves

A nomogram was established after combining the clinical features and our prognosis model using nomogramEx package in R software. DCA curve was plotted using rmda package to evaluate the performance of the nomogram. Calibration curve was used to display the difference between actual and predicted survival.

Real-time (RT) qPCR of tissue and cells

The paired ccRCC and normal tissue were obtained from the first affiliated hospital of Nanjing Medical University. All patients signed informed consent. Four renal carcinoma cells (786-O, ACHN, Caki-1, Caki-2) and one normal kidney epithelial cells (HK-2) were purchased from Cell Bank of Chinese Academy of Sciences (Shanghai, China). Total RNA were extracted from tissue and cells using an RNA extraction kit according to the protocol. Total RNA was reversely transcribed to cDNA using a reverse transcription kit (TaqMan). RT-qPCR was then performed with SYBR Green methods (Applied Biosystems, Foster City, CA, USA). The primers used were as follows. AC010973.2, forward: 5'-TTCCGTTACAGAGCAAAACCT-3';

AC010973.2, reverse: 5'-ACCTCAGGGAACCTTGGATG-3'; GAPDH, forward: 5'-AC CACAGTCCATGCCATCAC-3'; and GAPDH, reverse: 5'-TCCACCACCCTG TTGCTGTA-3'.

Western blotting

Western blot was performed as previously described. Briefly, proteins were visualized by ECL solution and captured by Quantity One V4.31. The primary antibody of cleaved caspase-3, Bcl-2, Bax and GAPDH were purchased from Sigma.

RNA interference

RNA interference was performed as previously described [12]. The target sequence used for siRNA against AC010973.2 was: siRNA1, 5'-GCTGTCTACAGAATAAACT-3'; siRNA2, 5'-CAGCCCTTATCGTCCATGA-3'; siRNA3, 5'-AGGCAGACGGGAGGGCTTA-3'.

Clonogenic assay

Colony formation assay was performed as previously described [12]. Cells were seeded into six-well plates with 500 cells per well. Lastly, cells were fixed with formaldehyde and stained with crystal violet before counting the cells present.

CCK8 assay

CCK8 assay were performed using a CCK8 kit (Dojindo, Shanghai, China). Briefly, cells were plated into 96-well plates with 2000 cell per well and added 10μL CCK8. Cells were maintained for 1.5h in a humidified incubator at 37°C with 5% CO_2. The absorbance at 450nm was used to express the cell proliferation ability.

Statistical analysis

All the analysis were conducted in R software v4.0.0, SPSS v13.0 and GraphPad Prism 8. P-value was two-side and < 0.05 was regarded as statistically significant. Student T-test were used to compare the difference between the two groups.

Results

Identification of differentially expressed tumor stemness-related genes

By using the cutoff of |logFC|>1 and the P-value < 0.05, a total of 13729, 4718 and 3817 differentially expressed genes (DEGs) were obtained through differential expression analysis on high and low mRNAsi groups (Fig. 1A), high and low EREG-mRNAsi groups (Fig. 1B), tumor and normal groups (Fig. 1C), respectively. After intersecting DEGs from three ccRCC-related data sets, 732 tumor stemness-associated DEGs in ccRCC were finally identified (Fig. 1D).

Function and interaction analysis of tumor stemness-related DEGs

To explore the potential biofunction of 732 stemness-related DEGs in ccRCC, gene ontology (GO) enrichment analysis and protein-protein interaction (PPI) analysis were performed to uncover the intrinsic roles of these genes. The result of GO annotation analysis showed that DEGs were tightly associated with the construction of extracellular matrix (ECM) structure in respect of cellular component, molecular function and biological process (Fig. 2A). By calculating the score of nodes, the most significant modules made up of 15 nodes (Fig. 2C) and 30 nodes (Fig. 2D) from PPI network (Fig. 2B) then illustrated potential hub genes and interactions between them.

Construction, evaluation, and validation of the stemness indices-based 6-gene risk signature

To establish a prognosis prediction model for ccRCC patients based on the stemness indices (mRNAsi and EREG-mRNAsi), 529 ccRCC patients with complete follow-up information in the TCGA-KIRC database were randomly divided into a training cohort and validation cohort at a ratio of 1:1. Firstly, the univariate cox analysis was performed to identify prognosis-related genes with a P-value threshold < 0.05. Random Forest (RF) algorithm was then applied to reduce the dimension and calculate the relative importance score of these prognosis-related genes. As a result, each gene was given a score and the top 6 genes with the highest scores were identified (Fig. 3A) as most relevant to patient OS. Then, they were used in the establishment of a risk signature for the prognosis of ccRCC patients after multivariate cox analysis. In the 6-gene risk signature, the "$\sum \mathrm{coef} * \mathrm{Exp(genes)}$" formula to calculate the risk score for each patient was: Risk Score = 0.133 * AC010973.2 expression + 0.182 * AP001029.2 expression + (-0.271) * RNU6-125P expression + (-0.025) * Z98885.1+ (-0.114) * KDM5C-IT1 expression + (-0.175) * AL021368.3. According to the median risk score in the signature, patients from training (Fig. 3B) and validation (Fig. 3C) cohorts were equally divided into low-risk and high-risk groups. As successively shown in the figures were diagrams of risk score (left top), survival status (left medium), heat map of gene expression (left bottom), time-dependent ROC curves (right top) and Kaplan−Meier survival curves (right bottom). Time-dependent ROC curves showed AUC of the prognostic 6-gene signature was 0.67 at 1 year, 0.79 at 3 years and 0.73 at 5 years in training cohort (Fig. 3B), and a similar outcome of 0.66 at 1 year, 0.71 at 3 years and 0.7 at 5 years in validation cohort (Fig. 3C). Survival analysis revealed the huge advantage of patient's OS in the low-risk group compared to high-risk group both in training group (p < 0.0001, Fig. 3B) and validation group (p < 0.0001, Fig. 3C).

Clinical parameters-based model and comparison between two prognosis prediction models

By assigning OS as the outcome, a nomogram was constructed based on the clinical features and riskscore of ccRCC patients (Fig. 4A). As is shown in Fig. 4A, age (p < 0.001), stage (p < 0.001), T stage (p < 0.05), and riskscore (p < 0.001) showed significant associations with OS. To compare the ability of

clinical prediction model and 6-gene signature in prognosis prediction for ccRCC patients, decision curves analysis (DCA) was conducted by labeling model of risk signature as 'Riskscore', clinical prediction model as 'Clinical' and combination of two models as 'Combined' (Fig. 4B). There is a noticeable distinction of Net benefit (NB) between 'Riskscore' and 'Clinical' curves when threshold probability (Pt) ranges from 0.2 to 0.5, which suggested the superiority of gene signature over the clinical model bringing medical benefits for the whole cohort under prediction. By comparing the blue curve and yellow curve, it was proved that the predictive capacity of the combined model slightly preceded the gene signature alone. Assessment of the model for efficacy in predicting 1-year (Fig. 4C), 3-year (Fig. 4D) and 5-year (Fig. 4E) survival probability was conducted by measuring calibration through bootstrapping 1000 resamples (Gray = ideal).

Clinical relevance and potential biofunctional pathways of hub gene AC010973.2

Since AC010973.2 was the most significant gene in the risk signature (Fig. 3A), it was chosen for the following research. Clinical correlation analysis showed no meaningful relationships between AC010973.2 expression and age, gender, pathological grade, and M stage. However, the AC010973.2 value was significantly upregulated in patients' individual stage III-IV compared to stage I-II ($p < 0.05$), as well as in T stage III-IV compared to T stage I-II ($p < 0.05$) (Fig. 5A). The results revealed that AC010973.2 could take part in the proliferation and growth of the tumor as a pro-cancer gene.

To explore the associated pathways AC010973.2 involved in, GSVA analysis was used to detect signatures differentiation between high and low expression groups. Several representative cancer pathways were found tightly related to AC010973.2, such as E2F targets, IL6-JAK-STAT3 signaling, apoptosis signaling, hypoxia signaling and fatty acid metabolism (Fig. 5B). Then, AC010973.2 was applied as an independent predictor for patient's OS in ccRCC, and the ROC curves showed the AUC was 0.645, 0.616 and 0.677 for survival probability at one year, three years and five years (Fig. 5C).

AC010973.2 expression was upregulated in ccRCC tissues and cell lines

Based on five paired ccRCC tissue, the results of qPCR revealed prominent upregulation of AC010973.2 mRNA expression in tumoral tissues ($p < 0.01$, Fig. 6A), which was also validated in cell level (Fig. 6B). ACHN and Caki-1 were transfected with siRNA-AC010973.2 for further research due to their highest expression level of AC010973.2. The knockdown of AC010973.2 was verified by mRNA expression detection (Fig. 6C).

AC010973.2 promoted tumor proliferation through apoptosis signaling pathway in ccRCC

In the result of GSVA, apoptosis signaling expression showed significant up-regulation in the high AC010973.2 group (Fig. 5B). It was known that Caspase-3 serves as one of the most important terminal

operators of apoptosis in cells, and Bcl-2 and Bax are pivotal regulators of Cas-3. After silencing AC010973.2 in ACHN and Caki-1 (Fig. 6C), expression of the three proteins was detected by Western Blot. The protein expression of Cas3 and Bax exhibited an obvious increase while Bcl-2 was distinctly downregulated, both in si-ACHN and si-Caki-1 cell lines, compared to ACHN and Caki-1 cell lines (Fig. 7A-B).

The results of colony formation assays showed the colonies of si-AC010973.2 group were significantly less than the control group in both ACHN and Caki-1 cell lines (Fig. 7C), suggesting that AC010973.2 knockdown weakened the ability of tumor cells. The CCK8 assays also demonstrated that the knockdown of AC010973.2 prominently inhibited the proliferation capacity of ccRCC cell lines (Fig. 7D). Altogether, these results indicated that AC010973.2 might promote tumor growth and survival through negatively regulating apoptosis in ccRCC.

Discussion

ccRCC, especially metastatic ccRCC, demonstrated high morbidity and mortality in renal cell carcinoma patients [13]. Tumor stemness-associated genes, a small copies of tumor genes, were contributed to tumor initiation, invasion, metastasis and chemical resistance [14]. Many studies have pointed out that tumor stemness-associated cells were characterized by unlimited proliferation tendency and multiple differentiation potential and played an essential role in developing multi-drug resistance and tumor recurrence [15]. Therefore, sensitive and reliable prognostic biomarkers were needed to identify patients with poor prognosis and who can benefit from treatment. In this study, we explored the association between tumor stemness-associated genes and CCRCC, especially metastatic ccRCC. A prognostic signature based on stemness-associated DEGs in ccRCC was developed and successfully validated in independent datasets using the Cox regression model. Further study demonstrated that the prognostic model constructed by 732 intersection genes showed more advantages than clinical traits in predicting patient survival convenience. Finally, the results of colony formation assays showed that the knockdown of AC010973.2 prominently inhibited the proliferation capacity of ccRCC cell lines. As a new method for predicting tumor prognosis, the predictive effect of stemness-associated genes had been confirmed in our studies.

In a recent study, Chuan Liu et al figured out the relationship between the immune-related gene pairs and cancer [16]. However, there is few studies focused on the predictive value of stemness-associated genes. In the past few years, stemness-associated genes had been proposed as a novel sub-category gene, and were proven to play a very important role in the treatment of tumors [17]. A very important malignant feature of many tumors which showed important therapeutic significance is that the presence of malignant cells with stemness-associated [18]. These stemness-associated cells can self-renew and maintain tumor growth and were seemed to be one of the main reasons for tumor resistance, recurrence and metastasis [19]. Interestingly, tumor stemness-associated genes have also been found to promote tumor metastasis in different tumor models, and have been widely involved in the ability to regulate

proliferation behavior in various tumors [20]. Besides, a study found that the stemness-associated cell behavior and malignant characteristics are linked through epithelial-mesenchymal transition [21]. In addition to promote stem cell differentiation, stemness-associated genes were also related to tumor invasion and metastasis [22]. In our research, stemness-associated genes not only promoted the progression of ccRCC, but also promote the formation of metastasis of ccRCC, and finally resulted in poor prognosis.

In this study, we integrated the transcriptome profiling from the TCGA database and stemness-associated genes. Next, we conducted GO&KEGG enrichment analysis and a Protein-protein interaction network. Based on 732 intersection genes, random forest algorithms were used to model the prognosis of all 732 intersection genes, and we finally selected the top 6 important genes for modeling. The results indicated that prognostic markers could be independently used to predict OS in ccRCC patients. In the univariate analysis, risk, grade, stage, pT, pM, and pN are related to the OS of ccRCC patients. Multivariate analysis showed that pT, and risk were associated with OS. A nomogram is established to predict the probability of survival at 1, 3 and 5 years through combining important clinical parameters and risk scores,. AC010973.2, as the most critical model gene, were selected for further research. The time-dependent ROC curve showed the predictive value of AC010973.2. GSVA analysis demonstrated that some representative cancer pathways were closely related to AC010973.2, such as E2F target, IL6-JAK-STAT3 signaling, apoptosis signaling, hypoxia signaling, and fatty acid metabolism. In the validation part, qPCR results showed that the expression of AC010973.2 mRNA in tumor tissues was significantly up-regulated. AC010973.2 knockdown weakens the ability of tumor cells. Besides, CCK8 analysis also showed that the knockdown AC010973.2 could significantly inhibit the proliferation of ccRCC cell lines.

Conclusion

In summary, we explored the biological functions and prognostic value of stemness-associated genes in ccRCC. Bioinformatics analysis showed that stemness-associated genes might regulate the apoptosis signaling pathway to influence tumor progression. The prognostic signature of stemness-associated genes identified in our study might serve as promising diagnostic and prognostic biomarkers in ccRCC. Furthermore, AC010973.2 was identified for further research. In vitro experiments showed that AC010973.2 could significantly promote cell proliferation of ccRCC and might be a potential prognostic marker in ccRCC.

Declarations

Acknowledgments

Funding

Author information

Affiliations

Department of Urology,The Second People's Hospital of Wuhu, Anhui Province 241000,China

Yingqing Liu, Jiawei Wang, Lin Li, Haibo Qin, Wei Ding, Xudong Shen, Guangyao Li, Zhongwen Lu, Lingsong Tao

Department of Urology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Yuang Wei, Xu Zhang, Xiaohan Ren, Dong Zhang, Chao Qin, Xinglin Chen

Contributions

YL, JW, and WD designed the experiments and offered direction for the project. LL, HQ, XS and ZL conducted the experiments, analyzed and visualized the results. XR drafted the manuscript. LT and XC reviewed the manuscript and made revisions. All authors participated and approved the final version of the manuscript.

Corresponding author

Correspondence to Xinglin Chen.

E-mail: nmuchenxinglin@163.com

Ethics declarations

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethical approval

The studies involving human participants were reviewed and approved by the Ethics Committee of Jiangsu Province Hospital. The patients/participants provided their written informed consent to participate in this study. Only the data was collected from the database about the patients and all authors have declared that they agree to publish. all methods were performed in accordance with relevant guidelines and regulations.

Informed consent

Informed consent was obtained from all individual participants included in the study.

References

1. Cohen, H. T. & McGovern, F. J. Renal-cell carcinoma. *The New England journal of medicine*, **353**, 2477–2490 https://doi.org/10.1056/NEJMra043172 (2005).

2. Koul, H. *et al.* Molecular aspects of renal cell carcinoma: a review. *American journal of cancer research*, **1**, 240–254 (2011).

3. Sánchez-Gastaldo, A., Kempf, E., Del González, A. & Duran, I. Systemic treatment of renal cell cancer: A comprehensive review. *Cancer treatment reviews*, **60**, 77–89 https://doi.org/10.1016/j.ctrv.2017.08.010 (2017).

4. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome - biological and translational implications. *Nature reviews. Cancer*, **11**, 726–734 https://doi.org/10.1038/nrc3130 (2011).

5. Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature*, **481**, 306–313 https://doi.org/10.1038/nature10762 (2012).

6. Valent, P. *et al.* Cancer stem cell definitions and terminology: the devil is in the details. *Nature reviews. Cancer*, **12**, 767–775 https://doi.org/10.1038/nrc3368 (2012).

7. Lytle, N. K., Barber, A. G. & Reya, T. Stem cell fate in cancer growth, progression and therapy resistance. *Nature reviews. Cancer*, **18**, 669–680 https://doi.org/10.1038/s41568-018-0056-x (2018).

8. Yu, B. *et al.* Expressions of stem cell transcription factors Nanog and Oct4 in renal cell carcinoma tissues and clinical significance. *Artificial cells, nanomedicine, and biotechnology*, **44**, 1818–1823 https://doi.org/10.3109/21691401.2015.1105238 (2016).

9. Rasti, A. *et al.* Co-expression of Cancer Stem Cell Markers OCT4 and NANOG Predicts Poor Prognosis in Renal Cell Carcinomas. *Scientific reports*, **8**, 11739 https://doi.org/10.1038/s41598-018-30168-4 (2018).

10. Rasti, A. *et al.* Reduced expression of CXCR4, a novel renal cancer stem cell marker, is associated with high-grade renal cell carcinoma. *Journal of cancer research and clinical oncology*, **143**, 95–104 https://doi.org/10.1007/s00432-016-2239-8 (2017).

11. Kim, K., Ihm, H., Ro, J. Y. & Cho, Y. M. High-level expression of stem cell marker CD133 in clear cell renal cell carcinoma with favorable prognosis. *Oncol Lett*, **2**, 1095–1100 https://doi.org/10.3892/ol.2011.365 (2011).

12. Ren, X. *et al.* COL5A2 Promotes Proliferation and Invasion in Prostate Cancer and Is One of Seven Gleason-Related Genes That Predict Recurrence-Free Survival. *Frontiers in oncology*, **11**, 583083 https://doi.org/10.3389/fonc.2021.583083 (2021).

13. Zhu, Z. *et al.* Biological functions and prognostic value of RNA Binding Proteins in clear cell Renal Cell Carcinoma. *Journal of Cancer*, **11**, 6591–6600 https://doi.org/10.7150/jca.49175 (2020).

14. Piao, L. *et al.* SETD8 promotes stemness characteristics and is a potential prognostic biomarker of gastric adenocarcinoma. *Experimental and molecular pathology*, **117**, 104560 https://doi.org/10.1016/j.yexmp.2020.104560 (2020).

15. Piao, L. *et al.* SETD8 is a prognostic biomarker that contributes to stem-like cell properties in non-small cell lung cancer. *Pathology, research and practice*, **216**, 153258 https://doi.org/10.1016/j.prp.2020.153258 (2020).

16. Liu, C. *et al.* Comprehensive analysis of a 14 immune-related gene pair signature to predict the prognosis and immune features of gastric cancer. *International immunopharmacology*, **89**, 107074 https://doi.org/10.1016/j.intimp.2020.107074 (2020).

17. Federer-Gsponer, J. R. *et al.* Patterns of stemness-associated markers in the development of castration-resistant prostate cancer., **80**, 1108–1117 https://doi.org/10.1002/pros.24039 (2020).

18. Wang, N. *et al.* SIPA1 enhances SMAD2/3 expression to maintain stem cell features in breast cancer cells. *Stem cell research*, **49**, 102099 https://doi.org/10.1016/j.scr.2020.102099 (2020).

19. Rodrigues, F. S. *et al.* IKKβ Kinase Promotes Stemness, Migration, and Invasion in KRAS-Driven Lung Adenocarcinoma Cells. *International journal of molecular sciences*, **21**, https://doi.org/10.3390/ijms21165806 (2020).

20. Pucci, M., Gomes Ferreira, I., Malagolini, N., Ferracin, M. & Dall'Olio, F. The Sd(a) Synthase B4GALNT2 Reduces Malignancy and Stemness in Colon Cancer Cell Lines Independently of Sialyl Lewis X Inhibition. *International journal of molecular sciences*, **21**, https://doi.org/10.3390/ijms21186558 (2020).

21. Tataranni, T. *et al.* Dichloroacetate Affects Mitochondrial Function and Stemness-Associated Properties in Pancreatic Cancer Cell Lines., **8**, https://doi.org/10.3390/cells8050478 (2019).

22. Chandimali, N., Jeong, D. K. & Kwon, T. Peroxiredoxin II Regulates Cancer Stem Cells and Stemness-Associated Properties of Cancers. *Cancers*, **10**, https://doi.org/10.3390/cancers10090305 (2018).
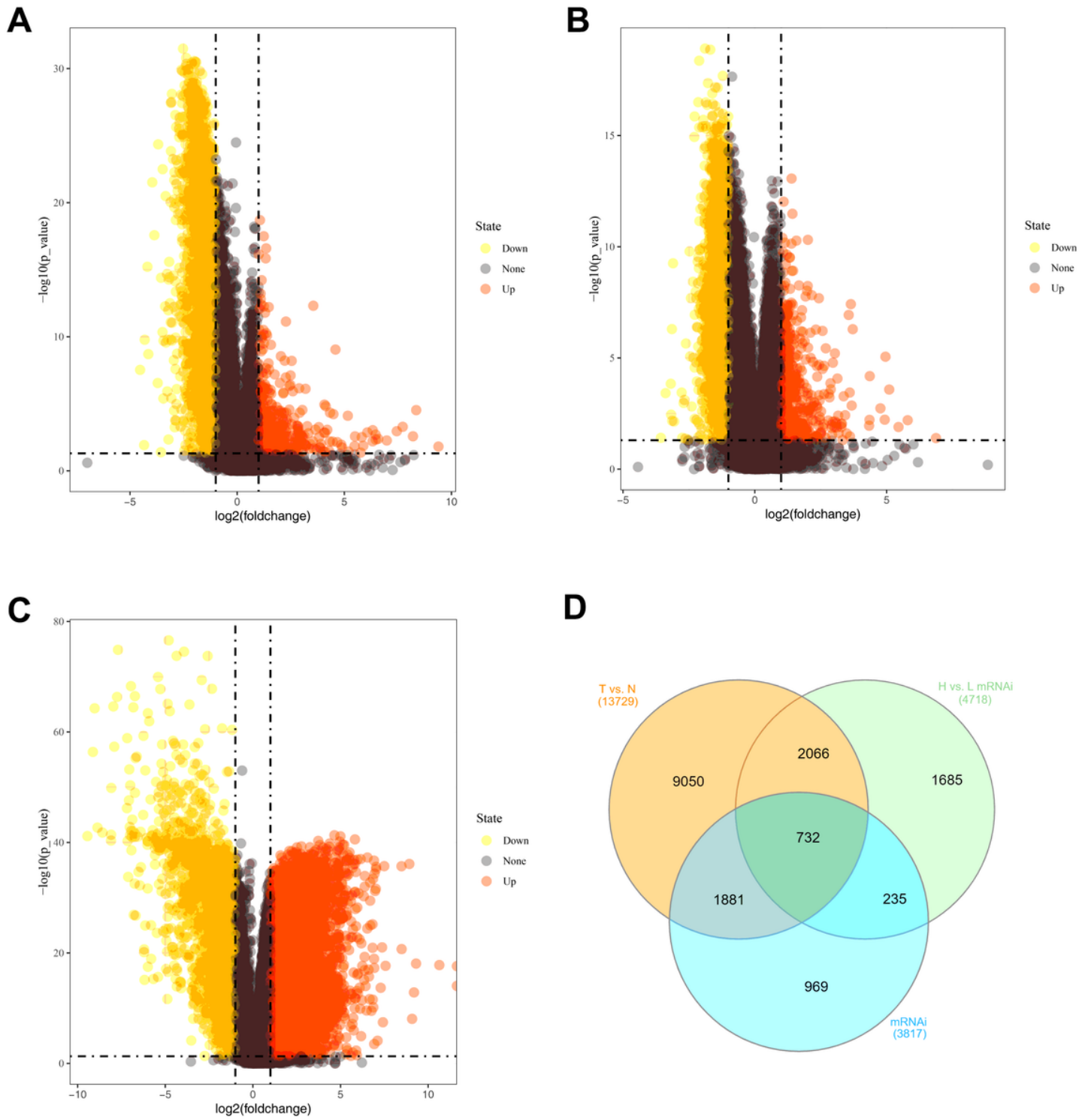
Figures

Figure 1

Identification of stemness indices-related DEGs. Notes: Volcano plots for DEGs in high-mRNAsi and low-mRNAsi ccRCC samples from (A), high-EREG-mRNAsi and low-EREG-mRNAsi ccRCC samples from (B), and tumor and normal tissues (C). (D) Venn diagram for DEGs from 3 different data sets. DEGs: differentially expressed genes; ccRCC: clear cell renal cell carcinoma
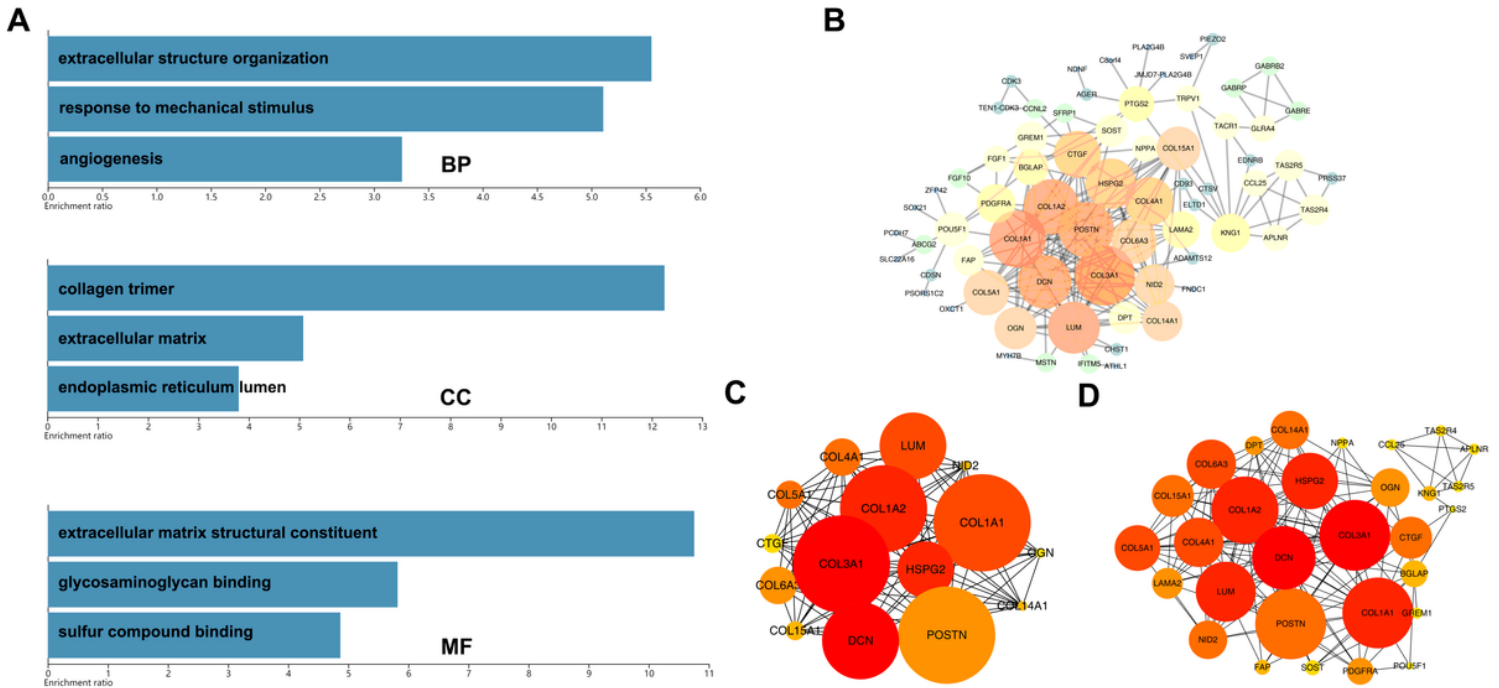
Figure 2

Function analysis on DEGs. Notes: (A) GO enrichment analysis of 732 DEGs, BP: biological process; CC: cell component; MF: molecular function. (B) PPI network of 732 DEGs after removing nodes without edges. (C) Top 30 nodes of PPI network. (D) Top 15 nodes of PPI network. GO: gene ontology; PPI: protein-protein interaction
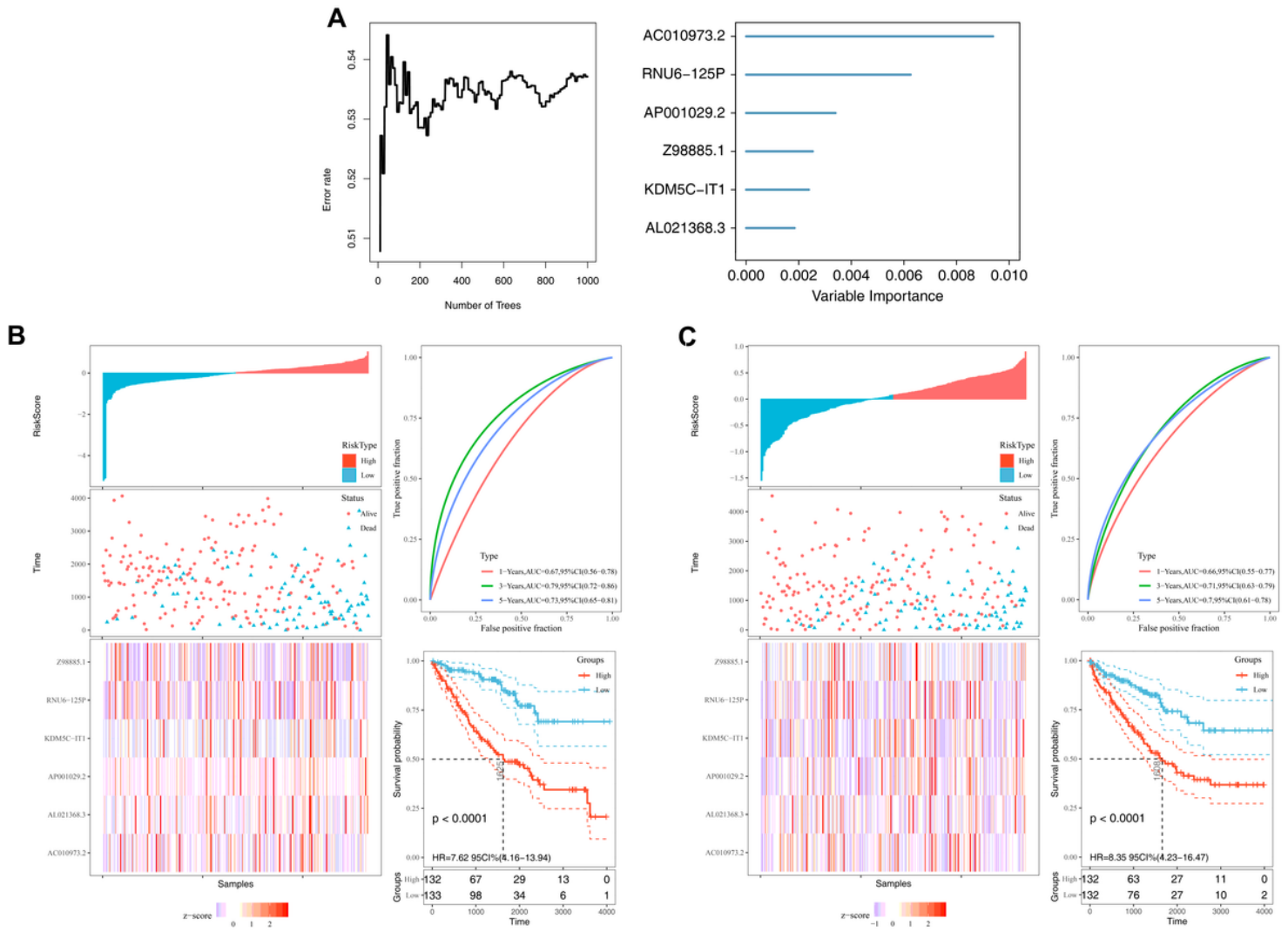
Figure 3

Establishment and assessment of the risk signature for ccRCC. Notes: (A) Calculation of relative importance of each DEG in the prognosis of ccRCC patients using Random Forest algorithm. In (B) Training cohort and (C) Validation cohort, diagrams are risk score map (left top), survival status map (left medium), gene expression heat map (left bottom), time-dependent ROC curves (right top) and Kaplan–Meier survival curves (right bottom). ROC: receiver operating characteristics
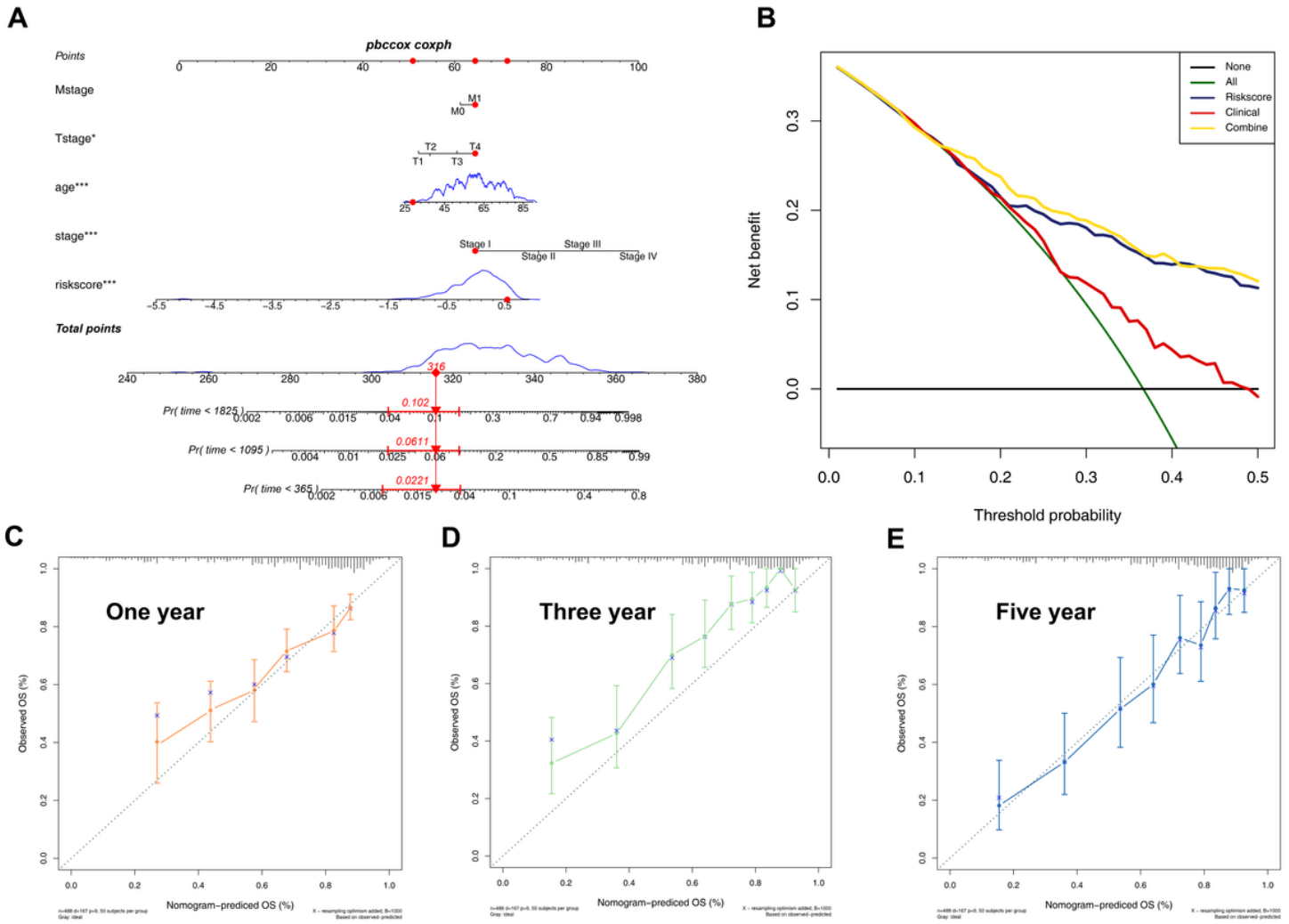
Figure 4

Decision Curves Analysis (DCA) of clinical model and risk signature. Notes: (A) Nomogram of clinical parameters-based prediction model for prognosis of patients. (B) Decision curves for different predicting models. Horizontal line: assume all patients as low risk (PiPr); Red curve: clinical prediction model; blue curve: gene signature; Yellow curve: combination use of two models. Calibration curves for nomogram in predicting survival probability at 1 year, 3 years and 5 years. The calibration measurement was conducted through bootstrapping 1000 resamples in TCGA-KIRC database (C-E).
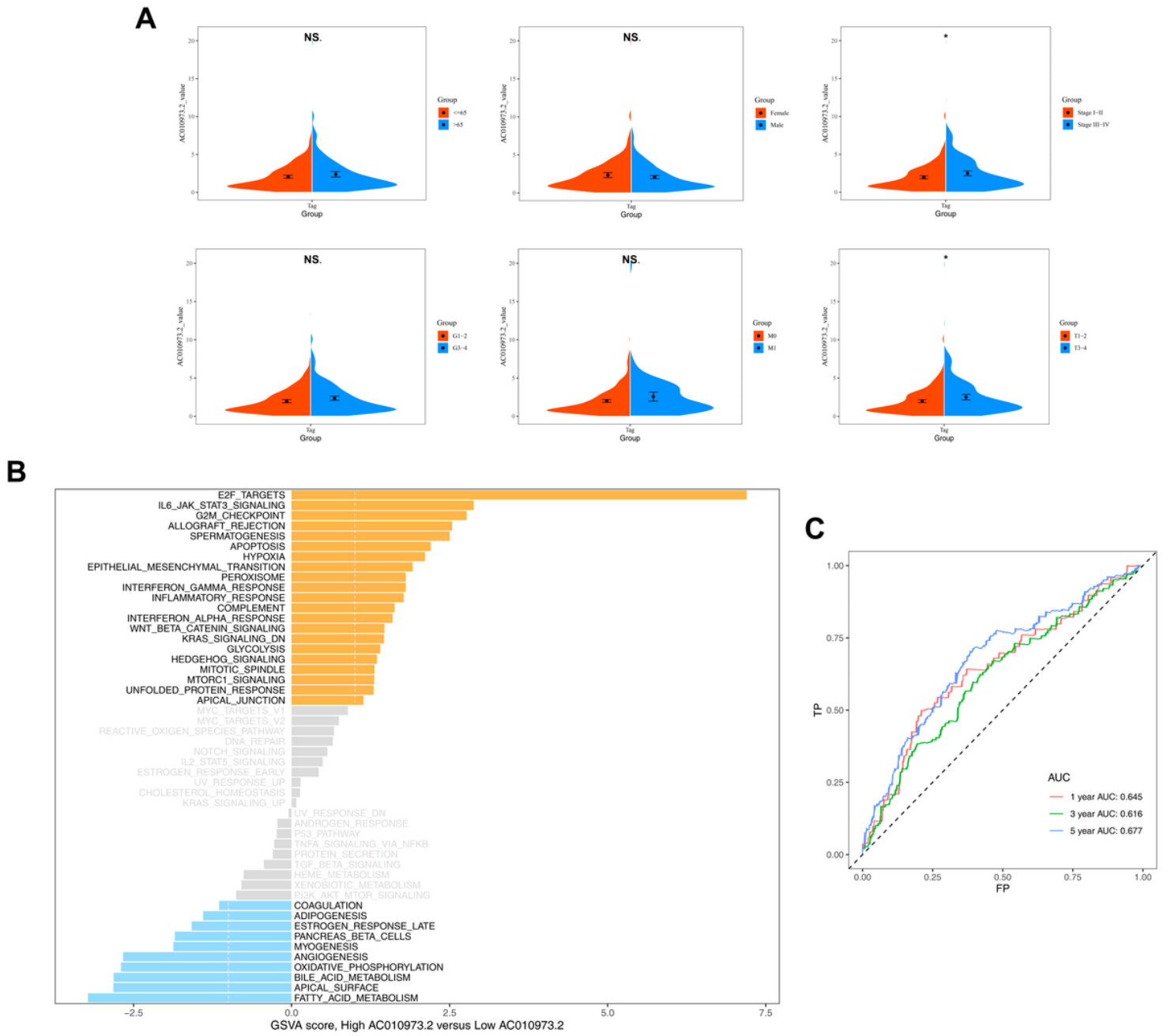
Figure 5

Clinical correlation analysis and GSVA of AC010973.2. Notes: (A) Violin plots for distribution of AC010973.2 value in different groups of diverse clinicopathological features in XXX patients from KIRC database. (B) Diagram of GSVA on AC010973.2. The collection used was H-hallmark gene sets. Y-axis represented diverse hallmark gene sets. X-axis represented gene sets expression level in high (yellow bars) and low (blue bars) AC010973.2 groups. (C) ROC curves for AC010973.2 alone to predict 1-year, 3-year and 5-year OS of ccRCC patients. AUC for them were exhibited in the plot.
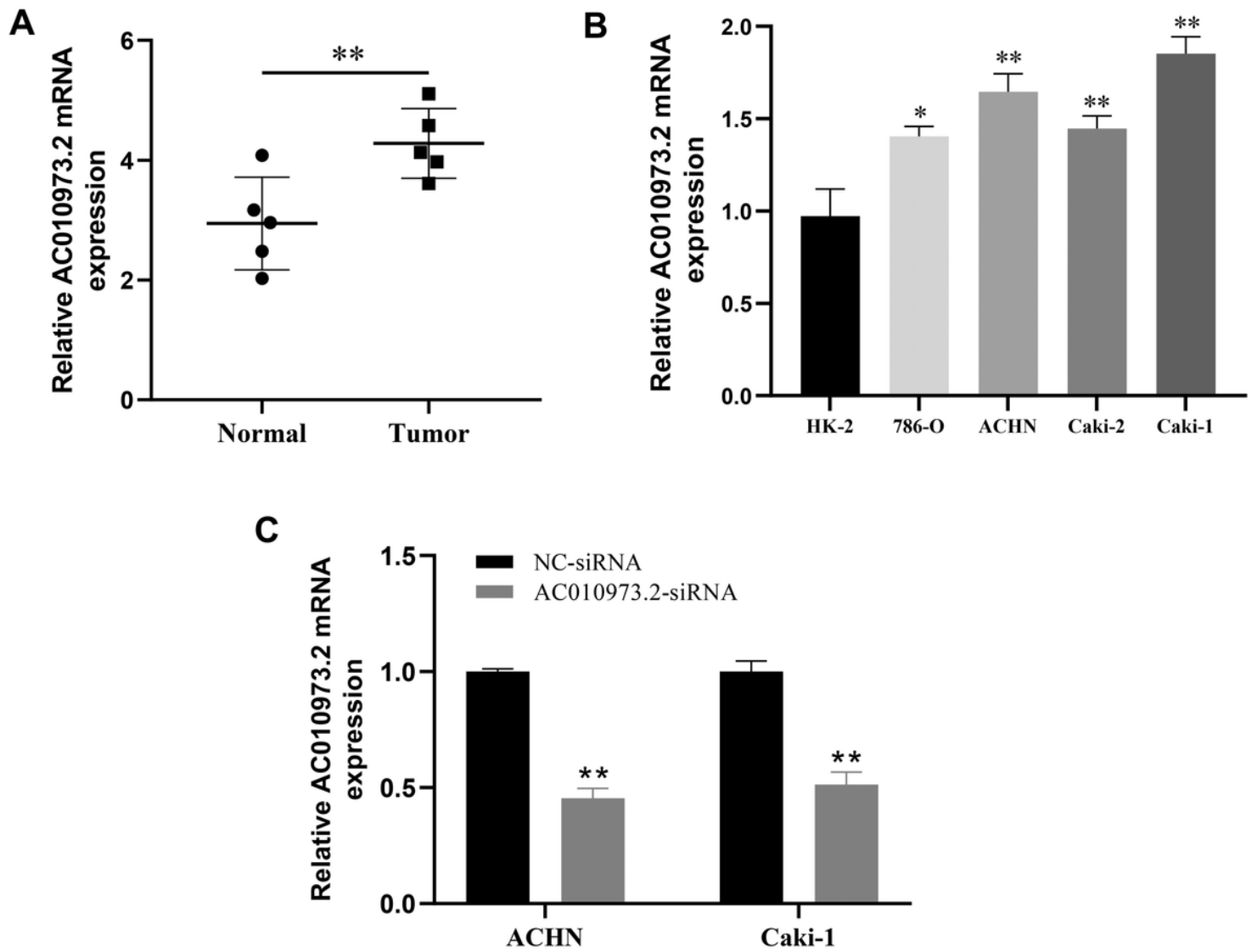
Figure 6

Upregulation of AC010973.2 mRNA expression in ccRCC. Notes: (A) AC010973.2 mRNA expression in tumors and non-malignant tissues from ccRCC patients. (B) AC010973.2 mRNA expression in ccRCC cell lines and normal renal tubular epithelial cells. (C) mRNA expression in AC010973.2 knockdown cell lines.
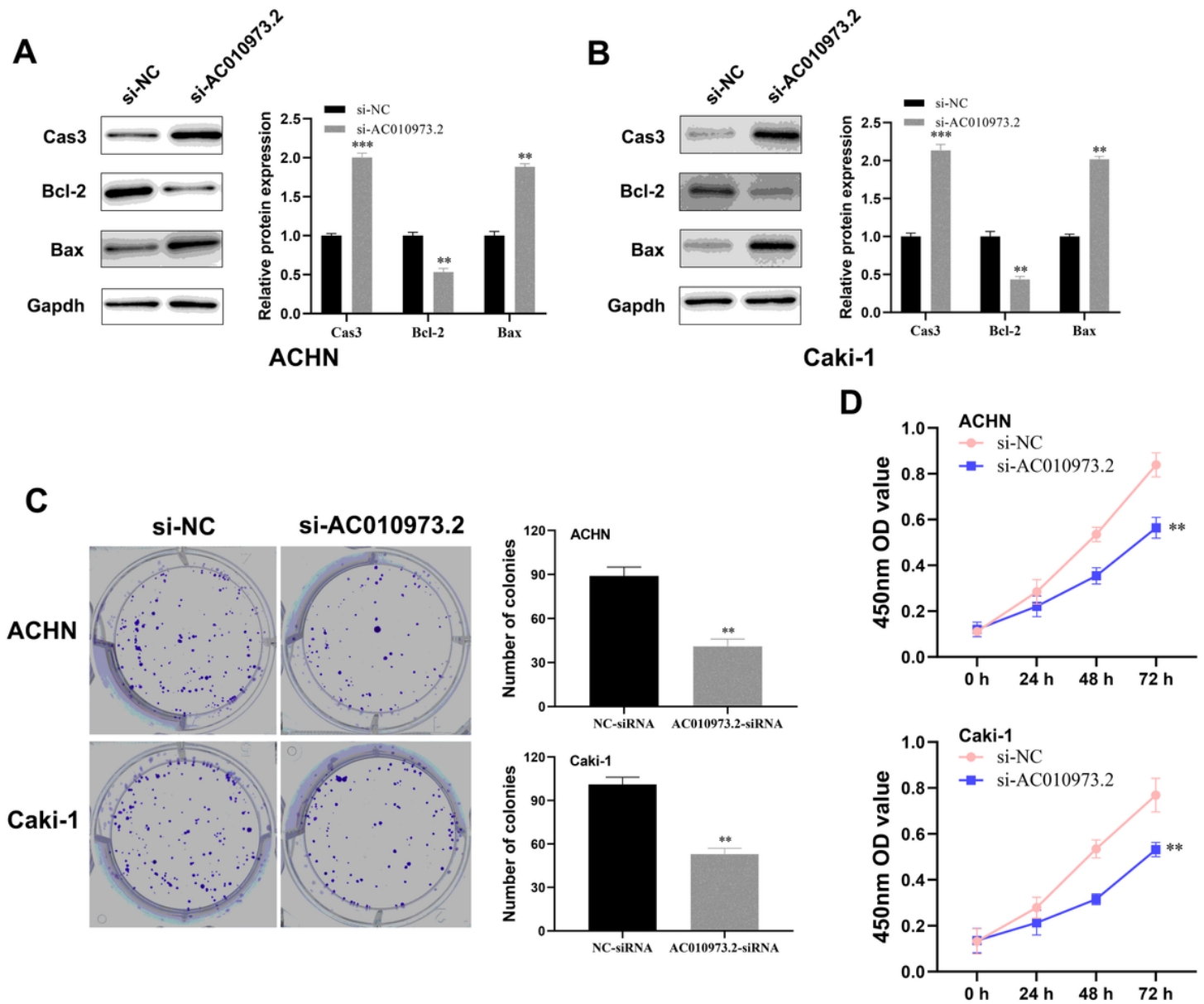
Figure 7

AC010973.2 regulated apoptosis signaling pathway and promoted cell proliferation. Notes: After the transfection with siRNA-AC010973.2, protein levels of Caspase-3, Bax and Bcl-2 in ACHN (A) and Caki-1 (B) cell lines through Western Blot. (C) Colony formation assay of ACHN and Caki-1 after the knockdown of AC010973.2. (D) CCK8 assay of ACHN and Caki-1 after the knockdown of AC010973.2.