Am J Clin Exp Urol 2024;12(6):352-366
www.ajceu.us /ISSN:2330-1910/AJCEU0158847

Original Article

Construction and evaluation of a prognostic model for metabolism-related genes in kidney renal clear cell carcinoma using TCGA database

Jingteng He, Mou Du, Xiaojun Bi, Peng Chen, Jian Li, Renli Tian, Lianhui Fan*, Qian Zhang*

Department of Urology, General Hospital of Northern Theater Command, Shenyang 110016, Liaoning, China.
**Equal contributors.*

Received June 29, 2024; Accepted October 15, 2024; Epub December 15, 2024; Published December 30, 2024

Abstracts: Objective: To investigate the expression of metabolism-related genes (MRGs) in kidney renal clear cell carcinoma (KIRC) and their association with patient prognosis, and to identify potential targets for intervention. Methods: Bioinformatics methods were employed to mine the KIRC transcriptome data in The Cancer Genome Atlas Program (TCGA) database in order to identify MRGs that are aberrantly expressed in cancerous tissues. Subsequently, a prognostic risk score model was constructed and its predictive capacity was evaluated. Finally, the expression of prognostically relevant MRGs was validated using external datasets and KIRC clinical samples. Results: A total of 789 differentially expressed MRGs associated with KIRC were screened, of which 465 genes were upregulated and 324 genes were downregulated, and finally 23 genes were screened to establish a risk score model. We found that the AUCs of the risk score model for predicting patients' 1-, 3- and 5-year overall survival (OS) were 0.804, 0.766 and 0.802, respectively. These findings suggest that the model has good predictive ability. A multifactorial Cox analysis revealed that 23 MRGs risk score was significantly associated with the overall survival of KIRC patients, and could therefore be used as an independent risk factor for the prognosis of KIRC patients (HR = 3.495, P < 0.001). Meanwhile, Kaplan-Meier analyses of the high-risk and low-risk groups indicated that the high-risk group exhibited a markedly inferior overall survival (OS) prognosis. The validation of clinical samples from KIRC patients and four external data sets (GSE36895, GSE40435, GSE53757 and GSE66272) demonstrated that KCNN4 and PLK1 were highly expressed in KIRC, whereas TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1 and UCN exhibited low expression in KIRC. Conclusion: Several MRGs are aberrantly expressed in KIRC, from which we screened 23 genes and constructed a MRGs prognostic risk model that can effectively predict the prognosis of KIRC patients and provide a new foundation for personalised diagnosis and treatment of KIRC.

Keywords: Kidney renal clear cell carcinoma, metabolism-related genes, prognosis model, risk score, cancer

Introduction

Kidney renal clear cell carcinoma (KIRC) represents the most prevalent form of renal cell carcinoma, accounting for 75%-85% of cases. Its aetiology is associated with heredity, smoking, obesity and hypertension, among other factors [1, 2]. Radical surgery represents the primary treatment option for early-stage KIRC. However, due to the lack of specific symptoms of early stage KIRC, patients often miss the opportunity for early diagnosis and treatment [3, 4]. Furthermore, patients with KIRC exhibit elevated rates of tumour recurrence and metastasis in comparison to patients with other renal cancer subtypes [5, 6]. Consequently, a comprehensive investigation of the potential molecular mechanisms governing the onset and progression of KIRC, coupled with the pursuit of novel, reliable prognostic assessment indicators and potential therapeutic targets for KIRC, is of paramount importance in accurately assessing the prognosis of patients and providing personalised treatment strategies.

An anomalous alteration in the metabolic profile is a distinctive attribute of the accelerated proliferation of tumour cells [7]. The high metabolic activity of the tumour results in a state of oxygen deprivation, low pH, electrolyte imbal-

https://doi.org/10.62347/XVZJ5704

ance, nutrient deficiency and a high state of oxidative stress [8, 9]. In KIRC, the tumour's aberrant growth characteristics result in an insufficient supply of nutrients, compelling the tumour cell to redirect its glucose, amino acid and lipid metabolism in order to meet the needs of its own synthesis and metabolism [10, 11]. The altered metabolism observed in renal cancer cells not only affects the phenotype of the tumour cells themselves, but also exerts an influence on the tumour microenvironment (TME), thereby influencing the growth of the tumour cells [10, 11]. KIRC is considered a metabolic disease and tumour cell metabolism is present throughout the interaction between tumour cells and TME [12, 13]. One of the features of KIRC is epigenetic changes in metabolism-related genes (MRGs), resulting in abnormal metabolic functions such as aerobic glycolysis, fatty acid metabolism and tryptophan-glutamine utilisation, involving a large number of metabolism-related genes such as ALDH6A1, FBP1, HAO2, TYMP, PSAT1, IL4I1, P4HA3, HK3, CPT1B and CYP26A1 [10-14]. Nevertheless, it remains uncertain whether aberrant MRGs expression can be employed for postoperative prognostic assessment of KIRC patients. In this study, we examined the correlation between MRGs expression and the prognosis of KIRC patients, and constructed a risk prognostic model to investigate its predictive value for the prognosis of KIRC patients, with the objective of providing a reference for the clinical diagnosis and treatment of KIRC.

Materials and methods

Data sources and acquisition

The Cancer Genome Atlas Program (TCGA) database (https://portal.gdc.cancer.gov/) was utilised to download access to the KIRC transcriptome RNA-seq (Workflow Type: HTSeq-counts) and corresponding patient clinical information. The data set comprises RNA-seq information for 537 KIRC and 71 paracancerous tissues. Genes encoding human metabolic enzymes and related transporters were obtained using Genecards [15], and a total of 2,169 MRGs-related genes were included using a relevance score greater than 20 as a screening condition. From these, MRGs with differential expression in KIRC tissues were compared and screened to construct a risk score model.

Identification of differentially expressed MRGs

The identification of differentially expressed genes (DEGs) were screened using the R software package Deseq2 for TCGA data KIRC data normalisation and difference analysis according to the thresholds of $|\log2 \text{ fold change}| \geq 1$ and $P < 0.05$ [16]. Intersections with MRGs were then performed to obtain differentially expressed MRGs, and corresponding heat maps and volcano maps were generated.

Enrichment analysis

Gene ontology (GO) enrichment analysis of the differentially expressed MRGs was conducted using the clusterProfiler package [17], and the results were visualised using the GOplot package [18].

Prognosis-related MRGs screening and prognostic risk score model construction

A LASSO regression analysis was conducted using the R language package glmnet with the objective of identifying prognostically relevant MRGs. Subsequently, a stepwise multifactorial Cox regression was employed to calculate the correlation coefficient (β) for each target gene. The 789 genes were obtained from the intersection of the differentially expressed genes (DEGs) and modular genes (MGRs). The FPKM matrices of the 789 genes were selected from the TCGA-KIRC dataset. The data were subjected to preprocessing, including the handling of missing values and data normalization. The dependent variables were the survival time and survival status of the patients, while the independent variable was the expression matrix of the 789 genes of the patients. A 10-fold cross-validation was conducted using the cv.glmnet function in the glmnet package to identify the optimal regularisation parameter λ ($\log\lambda = -2.036$ in this study, corresponding to the best_lambda value obtained from the lasso_cv$lambda.min). The coefficients of the model were extracted using the coef function. The genes were screened using regularisation, and the feature genes were those corresponding to the non-zero coefficients in the LASSO model. These coefficients were then multiplied and summed with the gene expression and normalised to obtain a risk score (RISK SCORE) for each sample, thereby constructing a risk score model [19]. Risk scores were calculated for

each patient, and patients were stratified into low-risk and high-risk groups based on the median threshold. Survival curves were plotted by Kaplan-Meier survival analysis for both subgroups, and the area under the curve (AUC) of the receiver operating characteristic (ROC) curve for 1-, 3- and 5-year overall survival was calculated using the R language survival and time ROC packages.

External validation of differential expression of model genes

The 23 genes were constructed to model differential expression in the KIRC and normal groups. Differential expression analyses were performed on four datasets (GSE36895, GSE40435, GSE53757 and GSE66272) and the results were visualised using the ggpubr package.

Clinical sample collection and quantitative real-time polymerase chain reaction analysis

Tissues exhibiting malignant and pre-malignant characteristics from 11 pairs of patients diagnosed with kidney renal clear cell carcinoma (KIRC) were procured from the General Hospital of Northern Theater Command and the Second Affiliated Hospital of the Army Military Medical University. This study was approved by the Ethics Committee of the General Hospital of Northern Theater Command and conducted following the Helsinki Declaration. Inclusion Criteria: The specimens were obtained from patients who underwent radical surgery for renal cancer and the pathological nature of the patients was clear cell carcinoma of the kidney. Exclusion Criteria: Pathological findings of non-renal clear cell carcinoma; patients with other metabolic abnormalities; patients with other neoplastic conditions. Total RNA was extracted from all tissues using Trizol (Invitrogen, 15596026). The SYBR Green One-Step qRT-PCR kit (Invitrogen, 11736059) was employed to quantify total RNA (100 ng). PrimeScript RT Master Mix (Shenggong, Shanghai, China) was used for cDNA synthesis. The PCR primers are listed below.

KCNN4, forward: 5'-GTGCCTCAGAGCAAAAGTCC-3'; KCNN4, reverse: 5'-CTACTTGGACTGCTGGCTGGG-3'; TEK, forward: 5'-TCCAGGCAACTTGACTTCGG-3'; TEK, reverse: 5'CCTTGAACCTTGTAACGGATAG-3'; GAPDH, forward: 5'-GGAGCGAGATCCCTCCAAAAT-3'; GAPDH, reverse: 5'-GGCTGTTGTCATACTTCTCATGG-3'; PLG, forward: 5'-GCGCCAGCACAGAGCTCTGCTCAAC-3'; PLG, reverse: 5'-GCTTGCTACTTGTAAGAACTAATAACTTC-3'; ANGPTL3, forward: 5'-CACCAATGTTTCCCCCAAT-3'; ANGPTL3, reverse: 5'-AAGATACCCTTTTTTACGCTCCTG-3'; TFAP2A, forward: 5'-GCTGCCTCACCAGCTGTCGGTATTGT-3'; TFAP2A, reverse: 5'-ACAGGGACACGGGGCCTTTCTCCC-3'; ANK3, forward: 5'-CAGGGGCTGTATTCTAGCAACT-3'; ANK3, reverse: 5'-CCCCTCCGTCTCACACTATTTT-3'; ATP1A1, forward: 5'-GGCCTTTAAGGTTGGACGTG-3'; ATP1A1, reverse: 5'-CACAGTAACATTGAGAACCCCC-3'; UCN, forward: 5'-CGAGCAGAACCGCATCATATT-3'; UCN, reverse: 5'-ACAGTGCCCTGGTGGCTCT-3'; PLK1, forward: 5'-AAAGAGATCCCGGAGGTCCTA-3'; PLK1, reverse: 5'-GGCTGCGGTGAATGGATATTTC-3'.

The 2-ΔΔCt method was conducted to calculate the RNA expression levels. A Student's t-test was used to compare the expression levels of each RNA between the different groups.

Statistical analysis

The data were expressed as mean ± SE. Comparisons between the two groups were performed by t-test, the relationship between metabolism-related genes and the prognosis of KIRC patients was analysed by Cox regression analysis, and the relationship between risk scores and prognosis of KIRC patients was analysed by log-rank test with survival analysis. $P < 0.05$ indicated that the difference was statistically significant.

Results

Screening for differentially expressed MRGs

By differential analysis with the R language Deseq2 package (**Figure 1A** and **1B**), 789 MRGs were differentially expressed between KIRC and paracancerous tissues (**Figure 1C**), of which 465 and 324 MRGs were expressed up-regulated and down-regulated, respectively, in KIRC cancerous tissues.

Enrichment analysis of differential MRGs

The GO enrichment analysis revealed that 465 up-regulated differentially expressed MRGs were mainly enriched in biological processes such as leukocyte proliferation, positive regula-

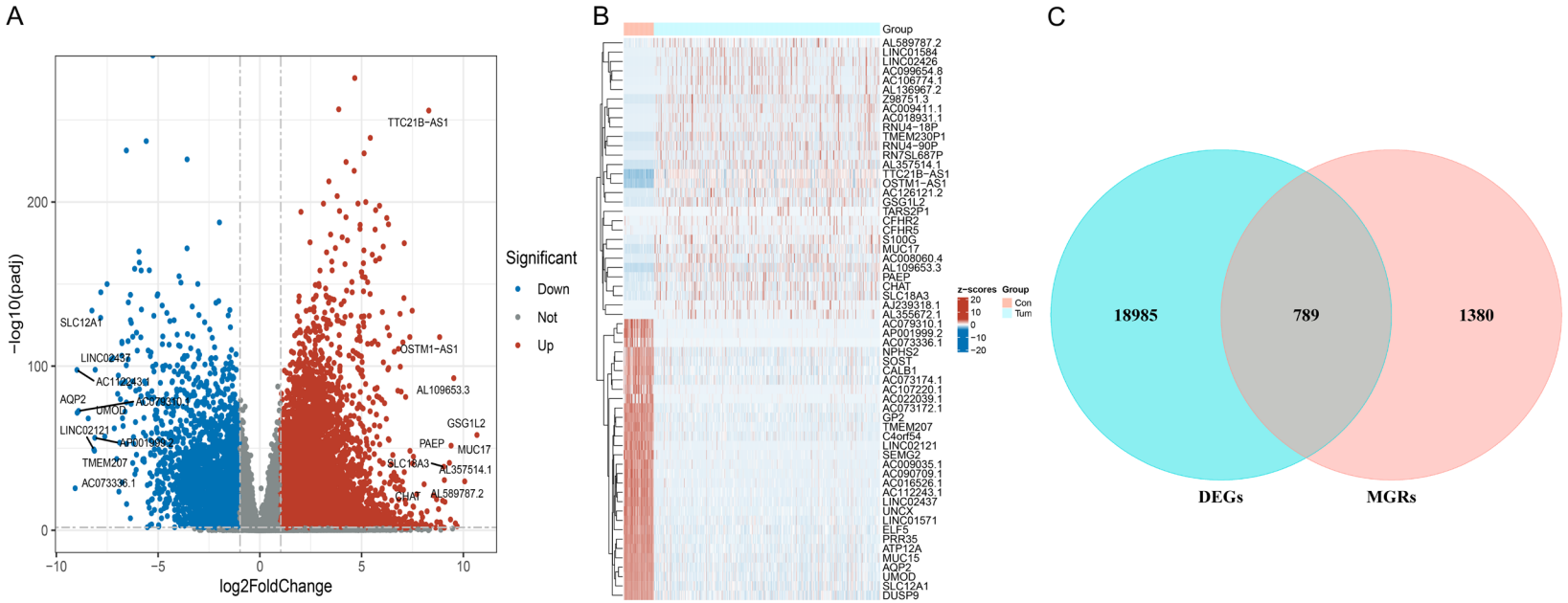


Figure 1. Differential expression of metabolism-related genes in cancer and paracancerous tissues. A. Volcano plot of differentially expressed genes. B. Heat map of differentially expressed genes. C. Intersection analysis of differentially expressed genes and MRGs genes.

tion of cytokine production, positive regulation of leukocyte activation, positive regulation of cellular activation (BP), and cellular components such as plasma membrane, stromal membrane and endocytic vesicles. With regard to molecular function (MF), the analysis indicated that these genes were primarily involved in the regulation of transmembrane proteins and cytokines (**Figure 2A** and **2C**). In contrast, the 324 down-regulated differentially expressed MRGs were predominantly associated with biological processes, including sodium and anion transport, organic acid fatty acid and carboxylic acid transport, and cellular components such as the basolateral membrane, apical plasma membrane and mitochondrial matrix. With regard to molecular function, the primary involvement was in the regulation of metal ion and organic acid transmembrane transport proteins (**Figure 2D** and **2F**). The KEGG enrichment analysis revealed that the upregulated differential genes (**Figure 2B**) were predominantly associated with tumour-related pathways, such as PI3K-Akt, TNF and HIF1. While the down-regulated differential genes (**Figure 2E**) were mainly enriched in the pathways of carbon metabolism, glucose and amino acid metabolism.

Construction of clinical prediction models

The screened MRGs were re-filtered by LASSO regression (**Figure 3A**) and cross-validation (**Figure 3B**). Subsequently, 23 risk genes associated with patient prognosis were subjected to further filtration by stepwise multifactorial Cox regression analysis, with the objective of constructing a risk score prediction model. The results of the one-way COX analysis of the model genes are presented in **Figure 3C**. The risk score calculation formula is provided in **Table 1**.

Evaluation of prognostic models

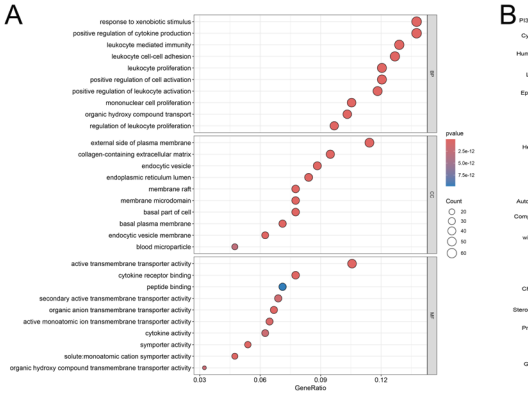
To verify whether the above risk score could be employed as an independent factor to predict the prognosis of KIRC patients, Cox regression analysis was performed by combining the patients' clinical characteristics and risk score. The results of univariate Cox analysis showed that patient tumour grade, TM stage, Age, platelet and risk score were all significantly correlated with the overall survival of KIRC patients ($P < 0.05$, **Figure 4A**). However, the results of the multifactorial Cox analysis showed that only age, tumour T-stage, stage and risk score were significantly associated with the overall survival of KIRC patients. These findings suggest that risk score may serve as an independent risk factor for the prognosis of patients with KIRC (**Figure 4B**). Nevertheless, risk score was more closely associated with a poor prognosis (HR = 3.495, $P < 0.001$). The risk score model was employed to calculate the risk score of patients with kidney renal clear cell carcinoma (KIRC) using the median risk score as the cut-off value, resulting in the division of patients into high-risk and low-risk groups (**Figure 4E**). The AUC was employed to assess the predictive capacity of the risk scoring model (**Figure 4C**). The AUC for this risk scoring model in predicting the patients' 1-, 3-, and 5-year overall survival rates were 0.804, 0.766, and 0.802, respectively, indicating that the model exhibited robust predictive ability. A Kaplan-Meier analysis of the two groups of patients using the log-rank test demonstrated a significant difference in overall survival prognosis between the two groups ($P < 0.05$, **Figure 4D**), with the high-risk group exhibiting a markedly inferior overall survival prognosis.

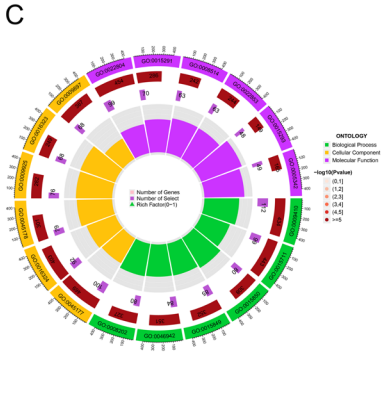
Construction of nomograms

Nomograms were constructed for the variables that were statistically significant in the results of the multifactorial COX analysis, and the results are shown in **Figure 5A**, which showed that T stage, age, risk score and stage could be used as independent prognostic indicators for predicting the 1-, 3- and 5-year survival of KIRC patients. The calibration curve results showed that the nomogram was around the ideal line, indicating its excellent ability to predict the prognosis of KIRC patients (**Figure 5B**). The time-dependent ROC curves demonstrated that the nomogram exhibited high accuracy in predicting patient survival at one, three, and five years, with an area under the curve (AUC) of approximately 0.9 (**Figure 5C**). The one-, three-, and five-year decision curve analysis (DCA) curves of the patients indicated that the nomogram was the most effective variable among all included variables, further confirming the accuracy of the nomogram in predicting the prognostic level of KIRC patients (**Figure 5D-F**).

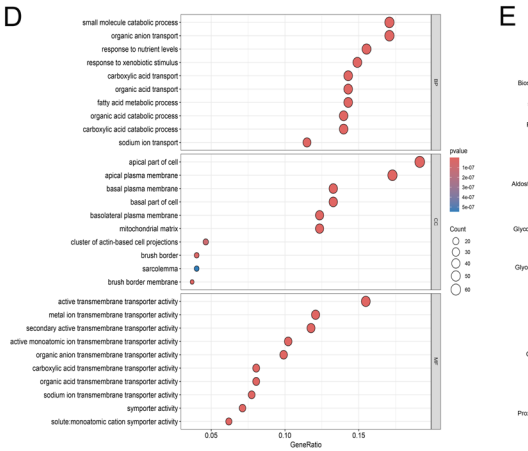
External validation of model gene differential expression

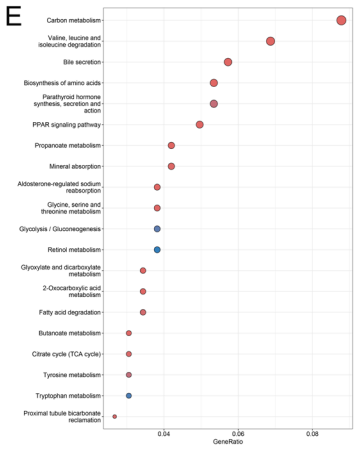
The four external datasets (GSE36895, GSE40435, GSE53757 and GSE66272) were anal-

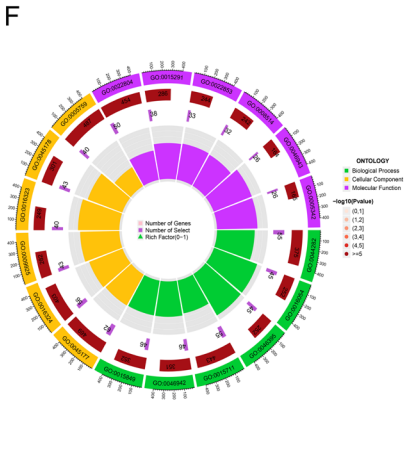




GO Term	Description
GO:0009410	response to xenobiotic stimulus
GO:0070661	leukocyte proliferation
GO:0007159	leukocyte cell-cell adhesion
GO:0001819	positive regulation of cytokine production
GO:0002696	positive regulation of leukocyte activation
GO:0050867	positive regulation of cell activation
GO:0009897	external side of plasma membrane
GO:0005788	endoplasmic reticulum lumen
GO:0030139	endocytic vesicle
GO:0062023	collagen-containing extracellular matrix
GO:0045121	membrane raft
GO:0098857	membrane microdomain
GO:0022804	active transmembrane transporter activity
GO:0005126	cytokine receptor binding
GO:0015293	symporter activity
GO:0008514	organic anion transmembrane transporter activity
GO:0015294	solute:monoatomic cation symporter activity
GO:0022853	active monoatomic ion transmembrane transporter activity







GO Term	Description
GO:0044282	small molecule catabolic process
GO:0016054	organic acid catabolic process
GO:0046395	carboxylic acid catabolic process
GO:0015711	organic anion transport
GO:0046942	carboxylic acid transport
GO:0015849	organic acid transport
GO:0045177	apical part of cell
GO:0016324	apical plasma membrane
GO:0009925	basal plasma membrane
GO:0016323	basolateral plasma membrane
GO:0045178	basal part of cell
GO:0005759	mitochondrial matrix
GO:0022804	active transmembrane transporter activity
GO:0015291	secondary active transmembrane transporter activity
GO:0022853	active monoatomic ion transmembrane transporter activity
GO:0008514	organic anion transmembrane transporter activity
GO:0046943	carboxylic acid transmembrane transporter activity
GO:0005342	organic acid transmembrane transporter activity

Figure 2. GO enrichment analysis of the up- and down-regulated differential metabolism genes. A. Bubble plot of GO enrichment analysis of up-regulated differential metabolism genes. B. Bubble plot of KEGG enrichment analysis of up-regulated genes. C. Circle plot of GO enrichment analysis. D. Bubble and circle plots of GO enrichment analysis of down-regulated differential metabolism genes. E. Bubble plot of KEGG enrichment analysis of down-regulated genes. F. Circle plot of GO enrichment analysis.

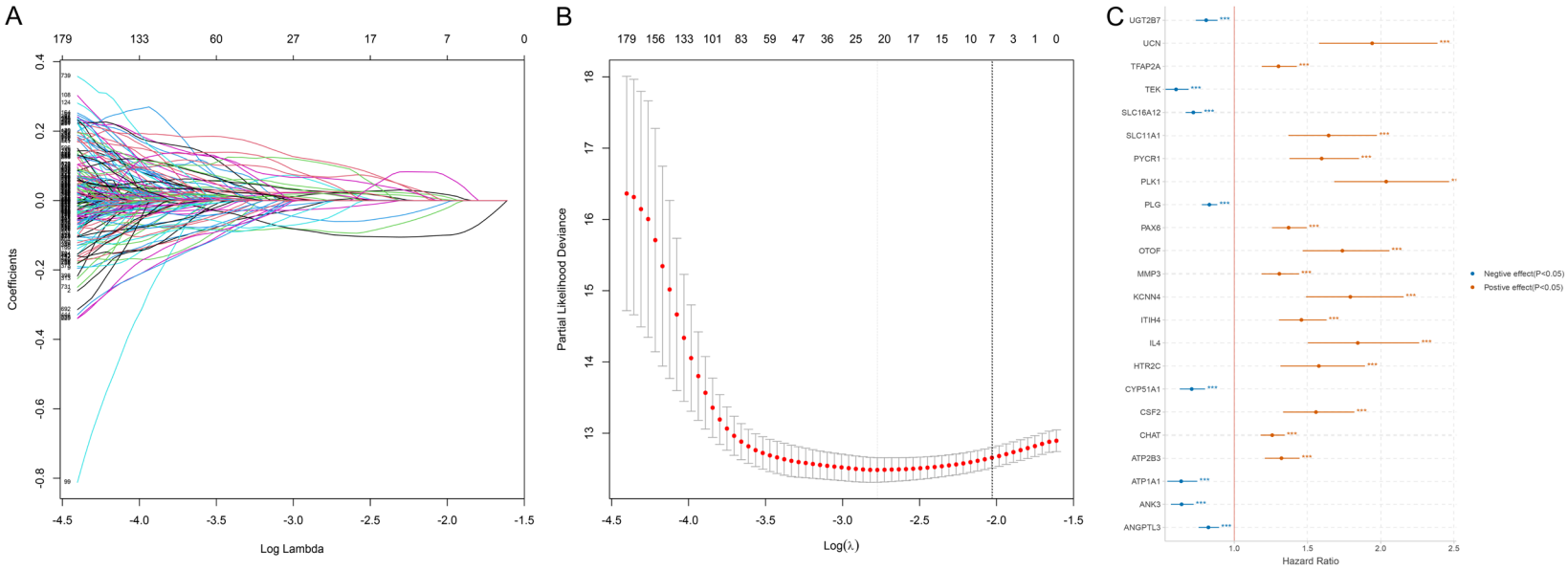


Figure 3. Construction of the LASSO+COX model for MRGs genes. A. LASSO screening for prognostically relevant MRGs. B. LASSO regression cross-validation results. C. One-way COX forest plot for model genes.

Table 1. Calculation of risk scores for 23 KIRC prognosis-related genes

Number	Gene	Coef
1	CYP51A1	-0.000949064872992064
2	PAX6	0.0212413870083898
3	SLC11A1	0.0428904900148218
4	ITIH4	0.0319399371104166
5	ATP2B3	0.0254556142474471
6	CHAT	0.019832964172135
7	KCNN4	0.00102261747196768
8	IL4	0.110231658841011
9	OTOF	0.0636953177955828
10	TEK	-0.0901203786992055
11	PLG	-0.0232749048089913
12	ANGPTL3	-0.00271334734211268
13	TFAP2A	0.0191583970019194
14	HTR2C	0.0521174704075536
15	MMP3	0.0251749248446669
16	ANK3	-0.0589841575482183
17	SLC16A12	-0.0967919455847867
18	ATP1A1	-0.0503935669089723
19	UCN	0.0673926761924249
20	CSF2	0.000629953213422847
21	PLK1	0.103793892512404
22	UGT2B7	-0.0052510267958739
23	PYCR1	0.0181619050092819

ysed to observe the differential expression of model genes. The samples were divided into two groups, KIRC and Normal, and the results demonstrated that KCNN4, TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1, UCN and PLK1 exhibited significant differential expression profiles across all four datasets. Among the genes examined, KCNN4 and PLK1 exhibited high expression levels in ccRCC, while TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1, and UCN demonstrated high expression levels in normal tissue (**Figure 6A-D**). Furthermore, we corroborated the aforementioned observations by analysing clinical samples from KIRC patients. This analysis confirmed that KCNN4 and PLK1 were highly expressed in KIRC, while TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1 and UCN were low expressed in KIRC (**Figure 7**).

Discussion

The accurate prediction of the prognosis and risk grading of cancer patients is directly related to the choice of treatment options and has a significant impact on treatment outcomes. KIRC is a less malignant form of kidney cancer, however, the likelihood of KIRC metastasising via the bloodstream is extremely high, with metastatic potential in approximately 60% of patients [20, 21]. Conversely, KIRC is insensitive to radiation and chemotherapy, and early individualised treatment of patients is imperative to enhance survival rates [20-22]. The prognosis of KIRC is influenced by a variety of factors, including tumour stage, histological type and treatment regimen, but due to the highly heterogeneous nature of KIRC, these commonly used clinical metrics are not sufficient for personalised and accurate prognostic assessment of patients [23, 24]. The metabolic profile of KIRC tumour cells is typically altered in order to adapt to the energy and material requirements of the tumour. Furthermore, the initiation and progression of tumours are highly correlated with the metabolic reprogramming of cancer cells [10-14]. The abnormal expression of genes involved in tumour metabolism, as well as the presence of mutations and ectopics, directly impacts the proliferation, invasion, angiogenesis and immune escape of KIRC tumour cells [10-14]. Alterations in MRGs in KIRC tumours play an important role in tumour initiation, development and metastasis. Assessing the characteristics and prognosis of KIRC patients by studying the alterations in MRGs has been increasingly emphasised.

In this study, we screened 789 differentially expressed genes associated with KIRC metabolism, of which 465 genes were up-regulated and 324 genes were down-regulated. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses revealed that these MRGs were closely associated with tumour-related signals such as leukocyte proliferation, cytokine production, glucose and amino acid metabolism, PI3K-Akt, TNF, HIF1 and others. In order to construct a clinical prediction model, we finally screened 23 genes to establish a risk score model by Cox regression combined with LASSO regression analysis. Additionally, multifactorial Cox analysis revealed that only age, tumour T-stage, tumour stage and 23 metabolism-related genes risk score were significantly associated with overall survival in KIRC patients. These could be considered as independent risk factors for the prognosis of KIRC patients, although the risk

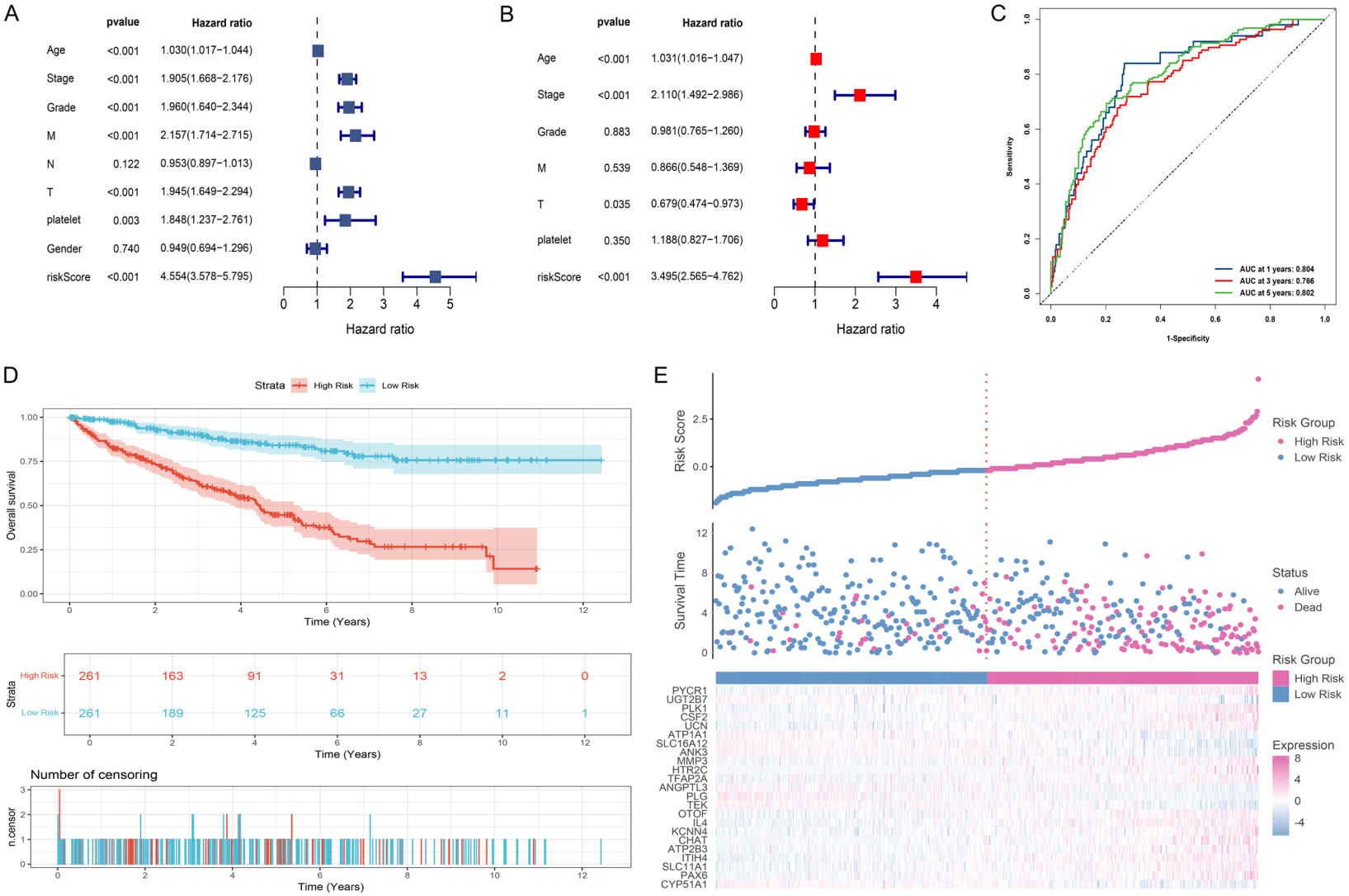


Figure 4. Construction and validation of the clinical significance of the riskScore. A and B. Single- and multi-factor COX forest plots. C. Receiver operating characteristic (ROC) curve of the riskScore for predicting patient survival at 1, 3, and 5-year. D. Kaplan-Meier survival curve of the model in the TCGA-KIRC dataset. E. Diagram of the risk factors associated with the model.

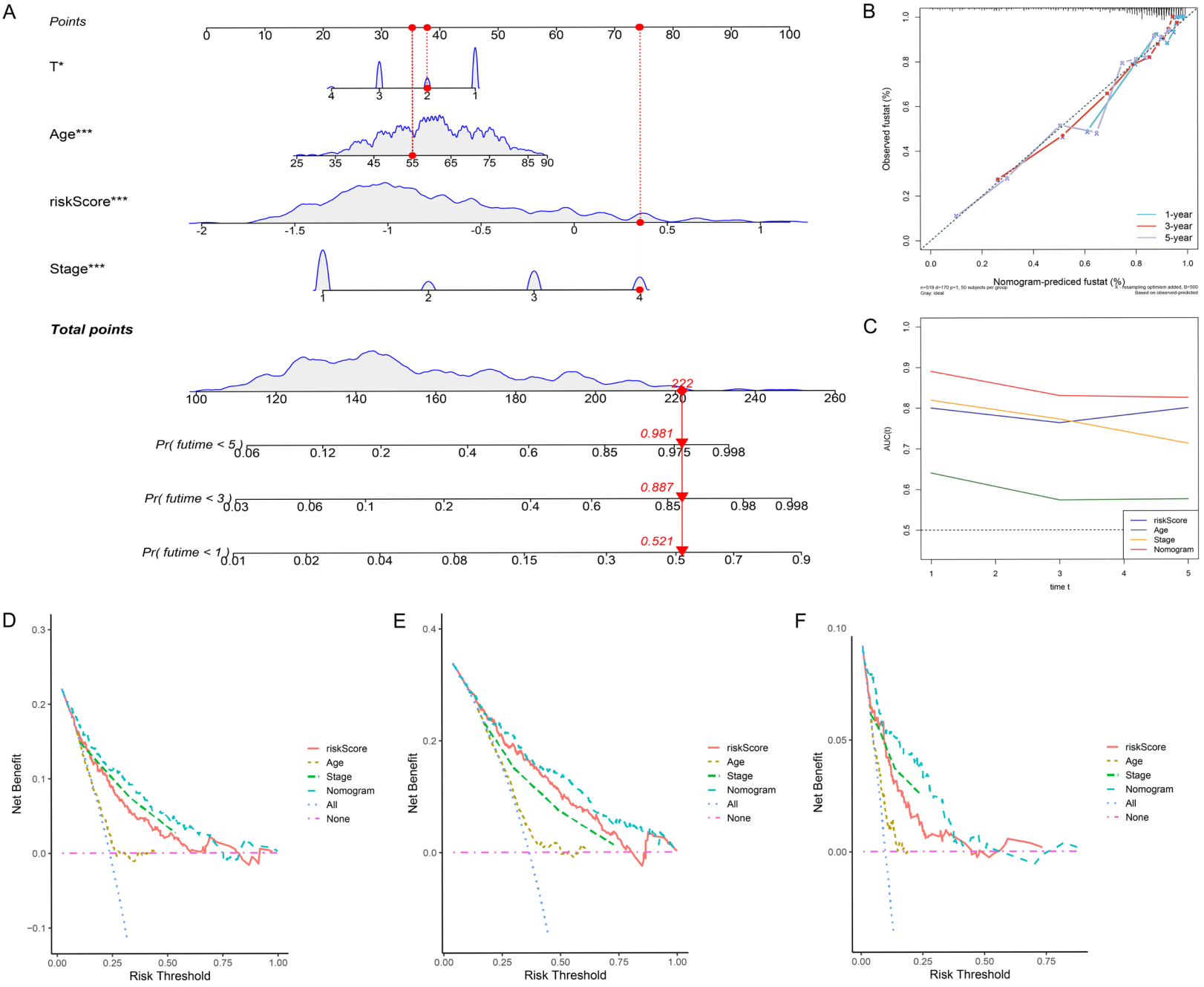

A
Points
T*
Age***
riskScore***
Stage***
Total points
Pr(futime < 5)
Pr(futime < 3)
Pr(futime < 1)
222
0.981
0.887
0.521
B
Observed fustat (%)
Nomogram-prediced fustat (%)
1-year
3-year
5-year
C
AUC(t)
time t
riskScore
Age
Stage
Nomogram
D
E
F
Net Benefit
Risk Threshold
riskScore
Age
Stage
Nomogram
All
None

Figure 5. Nomogram construction. A. Nomogram of P < 0.05 variables for multifactorial COX analysis. B. Calibration curves of the nomogram. C. Time dependent ROC curves for variables. D-F. DCA curves at 1, 3 and 5-year for variables.

score was more closely associated with a poor prognosis (HR = 3.495, P < 0.001). Further analysis revealed that the area under the curve (AUC) of the risk score model for predicting patients' 1-, 3- and 5-year overall survival were 0.804, 0.766 and 0.802, respectively, suggesting that the model had good predictive ability. Additionally, the Kaplan-Meier (K-M) analysis of the high-risk and low-risk groups demonstrated that the high-risk group had a significantly worse overall survival prognosis.

In recent years, the development of molecular biology has enabled researchers to identify that ccRCC is not simply a tumour disease, but rather that its development is closely related to the metabolic reprogramming of renal cells [25]. The metabolic reprogramming process involves various aspects of lipid metabolism, glucose metabolism, and amino acid metabolism, which influence the rate of tumour growth, invasiveness, and responsiveness to therapy [10-14]. One of the key mechanisms regulating KIRC metabolic reprogramming is alteration in gene expression related to metabolism [10-14]. In this study, we identified 23 metabolic risk genes associated with KIRC prognosis. The findings were validated using four external data sets (GSE36895, GSE40435, GSE53757, and GSE66272) and found that KCNN4, TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1, UCN, and PLK1 had significant differential expression profiles in the four data sets. Among the identified genes, KCNN4 and PLK1 exhibited high expression levels in KIRC, whereas TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1, and UCN displayed low expression levels in KIRC. Furthermore, we validated the expression profiles of KCNN4 and PLK1 in KIRC samples, confirming their high expression levels, while TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1, and UCN exhibited low expression levels.

The results demonstrated that KCNN4 expression was higher in KIRC samples compared to normal controls and was negatively correlated with prognosis and KCNN4 may affect KIRC prognosis by influencing the TME immune status [26]. In pan-cancer analyses, PLK1 was identified as a significant prognostic factor and was associated with tumour immunity across all cancers, particularly KIRC [27]. The expression of TEK is low in KIRC, and Kaplan-Meier curve analysis has demonstrated that the downregulation of TEK expression was associated with a poor prognosis for patients with KIRC [28]. Furthermore, the expression of PLG, ANGPTL3 and UCN is significantly reduced in KIRC, and this is an important biological prognostic factor in KIRC [29-31]. TFAP2A has been identified as a prognostic marker in KIRC, exhibiting a correlation with tumour immune infiltration and drug responsiveness [32]. The expression of ANK3 was found to be significantly decreased in KIRC, with patients exhibiting low ANK3 expression demonstrating reduced survival rates. Furthermore, a significant correlation was observed between ANK3 expression in KIRC and the infiltration levels of B cells, CD8+ T cells, macrophages and neutrophils [33]. The expression of ATP1A1 is markedly reduced in human KIRC tissues. In vitro and in vivo data of ATP1A1 inhibitory roles in KIRC progression suggest that ATP1A1 is a potential novel suppressor protein for KIRC [34]. The roles of KCNN4, TEK, PLG, ANGPTL3, TFAP2A, ANK3, ATP1A1, UCN, and PLK1 in the progression of KIRC and on the prognosis of KIRC patients have been reported, but none of these studies have discussed in-depth molecular mechanisms. Consequently, further investigation of the molecular mechanisms underlying the role of these genes in KIRC development is essential to identify molecular targets for precision treatment of KIRC.

Limitations and prospects of the study: (1) only TCGA-KIRC data were included in this study, and too few samples may affect the stability and reproducibility of the results, (2) although LASSO and Cox regression models perform well in feature selection and survival analysis, they have some algorithmic limitations, (3) there is a lack of further in-depth studies in cellular and animal models to identify key therapeutic targets. Future studies should increase the sample size to improve the reliability and reproducibility of the results. Meanwhile, future studies could attempt to combine multiple machine learning algorithms to improve the accuracy of

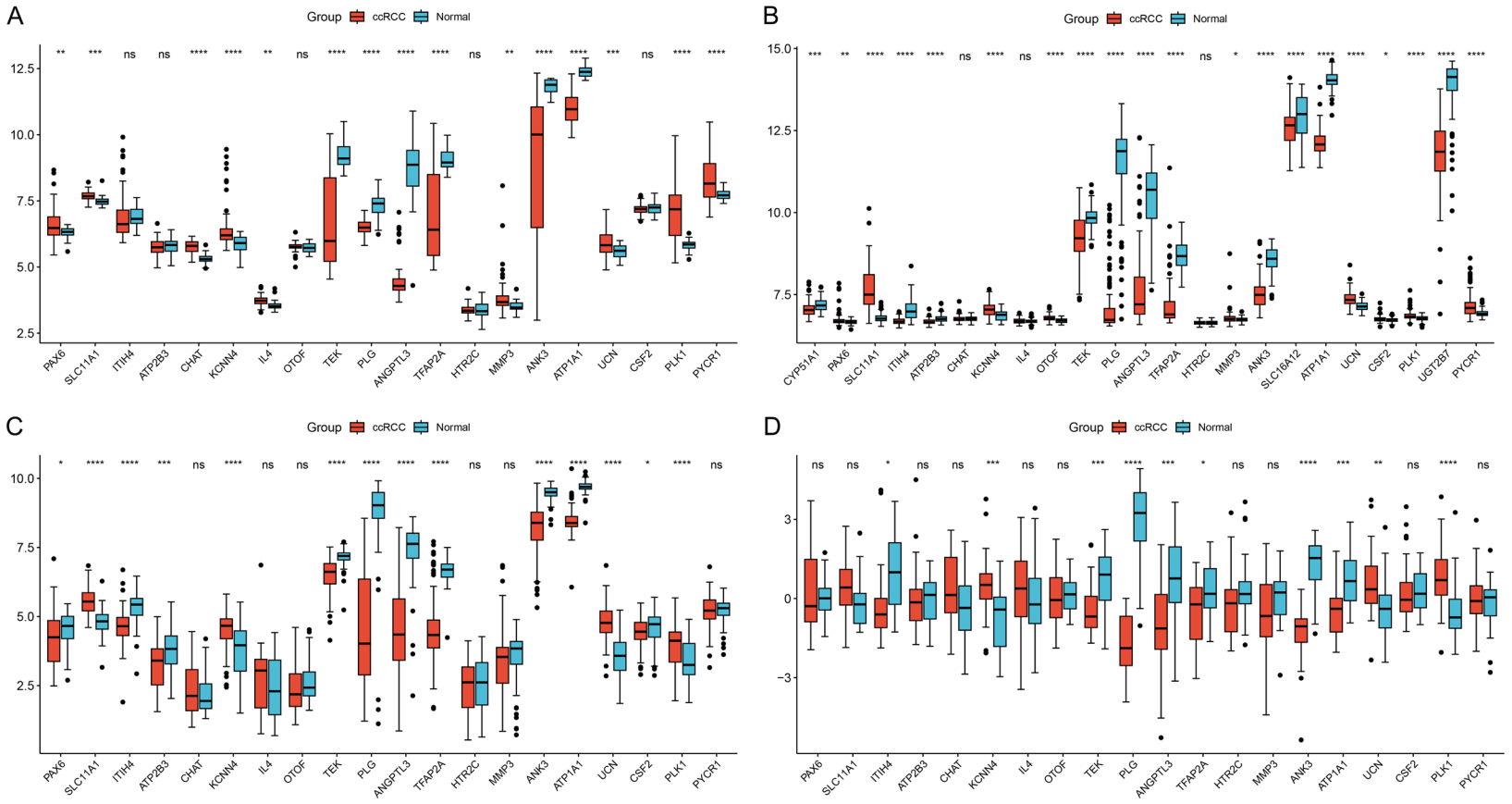


Figure 6. Differential expression of externally validated model genes. A. Differential expression of model genes in the GSE36895 dataset. B. Differential expression of model genes in the GSE40435 dataset. C. Differential expression of model genes in the GSE53757 dataset. D. Differential expression of model genes in the GSE66272 dataset.

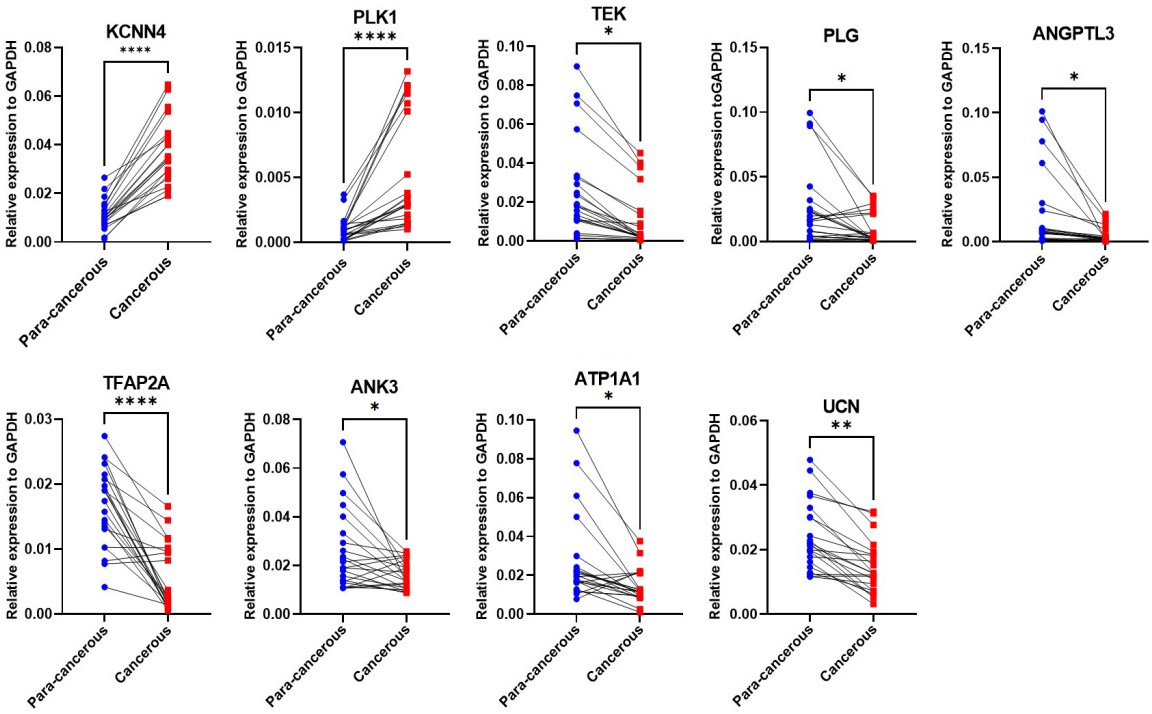


Figure 7. Analysis of the differential expression of model genes in clinical samples from the KIRC.

feature selection and prediction. In addition, multi-centre studies using data from different regions and different populations should be conducted to verify the generalisability of the results. Finally, further in-depth studies in cellular and animal models will be carried out to identify targets and drugs for treatment and intervention in kidney cancer.

In conclusion, this study employed a bioinformatics approach based on the TCGA database to analyse the impact of MRGs on the prognosis of KIRC patients. It is potentially informative to study the role and mechanism of MRGs in the progression of KIRC. Concurrently, we devised a 23-gene signature-based risk score prediction model with favourable predictive efficacy, which can be employed as an autonomous indicator for the prognostic assessment of KIRC patients and provide a reference for the prognostic assessment of KIRC patients and the development of personalised treatment plans. Nevertheless, the precise mechanism through which these risk genes and associated pathways contribute to the pathogenesis of KIRC remains unclear and necessitates further investigation and validation.

Disclosure of conflict of interest

None.

Address correspondence to: Qian Zhang and Lianhui Fan, Department of Urology, General Hospital of Northern Theater Command, Shenyang 110016, Liaoning, China. E-mail: mengqiqian@outlook.com (QZ); fanlianhui2099@163.com (LHF)

References

[1] Bahadoram S, Davoodi M, Hassanzadeh S, Bahadoram M, Barahman M and Mafakher L. Renal cell carcinoma: an overview of the epidemiology, diagnosis, and treatment. G Ital Nefrol 2022; 39: 2022-vol3.

[2] Hsieh JJ, Purdue MP, Signoretti S, Swanton C, Albiges L, Schmidinger M, Heng DY, Larkin J and Ficarra V. Renal cell carcinoma. Nat Rev Dis Primers 2017; 3: 17009.

[3] Chowdhury N and Drake CG. Kidney cancer: an overview of current therapeutic approaches. Urol Clin North Am 2020; 47: 419-431.

[4] Young M, Jackson-Spence F, Beltran L, Day E, Suarez C, Bex A, Powles T and Szabados B. Renal cell carcinoma. Lancet 2024; 404: 476-491.

[5] Dong D, Mu Z, Wei N, Sun M, Wang W, Xin N, Shao Y and Zhao C. Long non-coding RNA

ZFAS1 promotes proliferation and metastasis of clear cell renal cell carcinoma via targeting miR-10a/SKA1 pathway. Biomed Pharmacother 2019; 111: 917-925.
[6] Ljungberg B, Albiges L, Abu-Ghanem Y, Bedke J, Capitanio U, Dabestani S, Fernández-Pello S, Giles RH, Hofmann F, Hora M, Klatte T, Kuusk T, Lam TB, Marconi L, Powles T, Tahbaz R, Volpe A and Bex A. European Association of Urology Guidelines on renal cell carcinoma: the 2022 update. Eur Urol 2022; 82: 399-410.
[7] Park JH, Pyun WY and Park HW. Cancer metabolism: phenotype, signaling and therapeutic targets. Cells 2020; 9: 2308.
[8] Bader JE, Voss K and Rathmell JC. Targeting metabolism to improve the tumor microenvironment for cancer immunotherapy. Mol Cell 2020; 78: 1019-1033.
[9] Lyssiotis CA and Kimmelman AC. Metabolic interactions in the tumor microenvironment. Trends Cell Biol 2017; 27: 863-875.
[10] Liu M, Pan Q, Xiao R, Yu Y, Lu W and Wang L. A cluster of metabolism-related genes predict prognosis and progression of clear cell renal cell carcinoma. Sci Rep 2020; 10: 12949.
[11] Wettersten HI, Aboud OA, Lara PN Jr and Weiss RH. Metabolic reprogramming in clear cell renal cell carcinoma. Nat Rev Nephrol 2017; 13: 410-419.
[12] Linehan WM, Schmidt LS, Crooks DR, Wei D, Srinivasan R, Lang M and Ricketts CJ. The metabolic basis of kidney cancer. Cancer Discov 2019; 9: 1006-1021.
[13] Tan SK, Hougen HY, Merchan JR, Gonzalgo ML and Welford SM. Fatty acid metabolism reprogramming in ccRCC: mechanisms and potential targets. Nat Rev Urol 2023; 20: 48-60.
[14] Ricketts CJ, De Cubas AA, Fan H, Smith CC, Lang M, Reznik E, Bowlby R, Gibb EA, Akbani R, Beroukhim R, Bottaro DP, Choueiri TK, Gibbs RA, Godwin AK, Haake S, Hakimi AA, Henske EP, Hsieh JJ, Ho TH, Kanchi RS, Krishnan B, Kwiatkowski DJ, Liu W, Merino MJ, Mills GB, Myers J, Nickerson ML, Reuter VE, Schmidt LS, Shelley CS, Shen H, Shuch B, Signoretti S, Srinivasan R, Tamboli P, Thomas G, Vincent BG, Vocke CD, Wheeler DA, Yang L, Kim WY, Robertson AG; Cancer Genome Atlas Research Network, Spellman PT, Rathmell WK and Linehan WM. The cancer genome atlas comprehensive molecular characterization of renal cell carcinoma. Cell Rep 2018; 23: 313-326, e5.
[15] Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, Sirota-Madi A, Olender T, Golan Y, Stelzer G, Harel A and Lancet D. GeneCards version 3: the human gene integrator. Database (Oxford) 2010; 2010: baq020.
[16] Love MI, Huber W and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014; 15: 550.
[17] Yu G, Wang LG, Han Y and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 2012; 16: 284-7.
[18] Walter W, Sánchez-Cabo F and Ricote M. GOplot: an R package for visually combining expression data with functional analysis. Bioinformatics 2015; 31: 2912-4.
[19] Yang L, Qu Q, Hao Z, Sha K, Li Z and Li S. Powerful identification of large quantitative trait loci using genome-wide R/glmnet-based regression. J Hered 2022; 113: 472-478.
[20] Mohammadi M, Mansouri K, Mohammadi P, Pournazari M and Najafi H. Exosomes in renal cell carcinoma: challenges and opportunities. Mol Biol Rep 2024; 51: 443.
[21] Lyskjær I, Iisager L, Axelsen CT, Nielsen TK, Dyrskjøt L and Fristrup N. Management of renal cell carcinoma: promising biomarkers and the challenges to reach the clinic. Clin Cancer Res 2024; 30: 663-672.
[22] Chen YW, Wang L, Panian J, Dhanji S, Derweesh I, Rose B, Bagrodia A and McKay RR. Treatment landscape of renal cell carcinoma. Curr Treat Options Oncol 2023; 24: 1889-1916.
[23] Rosellini M, Marchetti A, Mollica V, Rizzo A, Santoni M and Massari F. Prognostic and predictive biomarkers for immunotherapy in advanced renal cell carcinoma. Nat Rev Urol 2023; 20: 133-157.
[24] Pezzicoli G, Ciciriello F, Musci V, Salonne F, Ragno A and Rizzo M. Genomic profiling and molecular characterization of clear cell renal cell carcinoma. Curr Oncol 2023; 30: 9276-9290.
[25] Hu J, Wang SG, Hou Y, Chen Z, Liu L, Li R, Li N, Zhou L, Yang Y, Wang L, Wang L, Yang X, Lei Y, Deng C, Li Y, Deng Z, Ding Y, Kuang Y, Yao Z, Xun Y, Li F, Li H, Hu J, Liu Z, Wang T, Hao Y, Jiao X, Guan W, Tao Z, Ren S and Chen K. Multiomic profiling of clear cell renal cell carcinoma identifies metabolic reprogramming associated with disease progression. Nat Genet 2024; 56: 442-457.
[26] Chen S, Wang C, Su X, Dai X, Li S and Mo Z. KCNN4 is a potential prognostic marker and critical factor affecting the immune status of the tumor microenvironment in kidney renal clear cell carcinoma. Transl Androl Urol 2021; 10: 2454-2470
[27] Qian Y, Li Y, Chen K, Liu N, Hong X, Wu D, Xu Z, Zhou L, Xu L, Jia R and Ge YZ. Pan-cancer transcriptomic analysis identifies PLK1 crucial for the tumorigenesis of clear cell renal cell carcinoma. J Inflamm Res 2022: 15: 1099-1116.

[28] Ha M, Son YR, Kim J, Park SM, Hong CM, Choi D, Kang W, Kim JH, Lee KJ, Park D, Han ME, Oh SO, Lee D and Kim YH. TEK is a novel prognostic marker for clear cell renal cell carcinoma. Eur Rev Med Pharmacol Sci 2019; 23: 1451-1458.

[29] Iqbal S, Begum F, Nyamai DW, Jalal N and Shaw P. An integrated computational analysis of high-risk SNPs in angiopoietin-like proteins (ANGPTL3 and ANGPTL8) reveals perturbed protein dynamics associated with cancer. Molecules 2023; 28: 4648.

[30] Pei D, Xu C, Wang D, Shi X, Zhang Y, Liu Y, Guo J, Liu N and Zhu H. A novel prognostic signature associated with the tumor microenvironment in kidney renal clear cell carcinoma. Front Oncol 2022: 12: 912155.

[31] He Z, Gu Y, Yang H, Fu Q, Zhao M, Xie Y, Liu Y and Du W. Identification and verification of a novel anoikis-related gene signature with prognostic significance in clear cell renal cell carcinoma. J Cancer Res Clin Oncol 2023; 149: 11661-11678.

[32] Lin D, Hu B, Zhu S and Wu Y. Exploring a ferroptosis and oxidative stress-based prognostic model for clear cell renal cell carcinoma. Front Oncol 2023; 13: 1131473.

[33] Somsuan K and Aluksanasuwan S. Bioinformatic analyses reveal the prognostic significance and potential role of ankyrin 3 (ANK3) in kidney renal clear cell carcinoma. Genomics Inform 2023; 21: e22.

[34] Zhang D, Zhang P, Yang P, He Y, Wang X, Yang Y, Zhu H, Xu N and Liang S. Downregulation of ATP1A1 promotes cancer development in renal cell carcinoma. Clin Proteomics 2017; 14: 15.