



OPEN

Identification and validation of fatty acid metabolism-related lncRNA signatures as a novel prognostic model for clear cell renal cell carcinoma

Cheng Shen^{1,2,7}, Zhan Chen^{1,2,7}, Jie Jiang^{1,2,7}, Yong Zhang^{1,2}, Xinfeng Chen¹, Wei Xu^{1,2}, Rui Peng^{1,2}, Wenjing Zuo³, Qian Jiang⁴, Yihui Fan⁵, Xingxing Fang⁶✉ & Bing Zheng¹✉

Clear cell renal cell carcinoma (ccRCC) is a main subtype of renal cancer, and advanced ccRCC frequently has poor prognosis. Many studies have found that lipid metabolism influences tumor development and treatment. This study was to examine the prognostic and functional significance of genes associated with lipid metabolism in individuals with ccRCC. Using the database TCGA, differentially expressed genes (DEGs) associated with fatty acid metabolism (FAM) were identified. Prognostic risk score models for genes related to FAM were created using univariate and least absolute shrinkage and selection operator (LASSO) Cox regression analyses. Our findings demonstrate that the prognosis of patients with ccRCC correlate highly with the profiles of FAM-related lncRNAs (AC009166.1, LINC00605, LINC01615, HOXA-AS2, AC103706.1, AC009686.2, AL590094.1, AC093278.2). The prognostic signature can serve as an independent predictive predictor for patients with ccRCC. The predictive signature's diagnostic effectiveness was superior to individual clinicopathological factors. Between the low- and high-risk groups, immunity research revealed a startling difference in terms of cells, function, and checkpoint scores. Chemotherapeutic medications such lapatinib, AZD8055, and WIKI4 had better outcomes for patients in the high-risk group. Overall, the predictive signature can help with clinical selection of immunotherapeutic regimens and chemotherapeutic drugs, improving prognosis prediction for ccRCC patients.

Renal cell carcinoma (RCC) is one of the most prevalent kinds of urinary tract cancer, impacting approximately 430,000 individuals worldwide in 2020¹. Approximately 70% of all RCC cases are clear cell renal cell carcinoma (ccRCC), the most common and malignant subtype². Overall, the prognosis is uncertain due to tumor recurrence or metastasis during the course of the disease, with roughly one-third of patients presenting with metastases at the time of diagnosis³. Metastatic ccRCC is always associated with a high mortality rate^{4,5}. Patients with RCC that metastasizes have a dismal 12% OS rate at 5 years⁶⁻⁸. Despite improvements in diagnosis, screening, surgery, and pharmaceutical treatment, the clinical outcome of RCC remains poor. Hence, there is an urgent need to create more accurate prognostic models given the significant morbidity and mortality of ccRCC.

Reprogramming of energy metabolism is a characteristic of cancer development and stimulates cell growth and proliferation⁹. Increasing data suggest that fatty acid metabolism plays a critical role in metabolic reprogramming, influencing cell membrane formation, energy storage, and synthesis of signaling molecules^{10,11}. According to previous research, the rate-controlling enzyme ACLY, which is involved in the first stage of lipid production, can promote colon cancer spread¹². Increased lipolysis and fatty acid production in cervical cancer patients lead

¹Department of Urology, The Second Affiliated Hospital of Nantong University, Nantong 226001, China. ²Medical Research Center, The Second Affiliated Hospital of Nantong University, Nantong, China. ³Department of Orthopedics, The Second Affiliated Hospital of Nantong University, Nantong, China. ⁴Department of Paediatric, Chinese Medicine Hospital of Rudong, Nantong, China. ⁵Department of Pathogenic Biology, School of Medicine, Nantong University, Nantong, China. ⁶Department of Nephrology, The Second Affiliated Hospital of Nantong University, Nantong 226001, China. ⁷These authors contributed equally: Cheng Shen, Zhan Chen and Jie Jiang. ✉email: 48776592@qq.com; ntzb2008@163.com

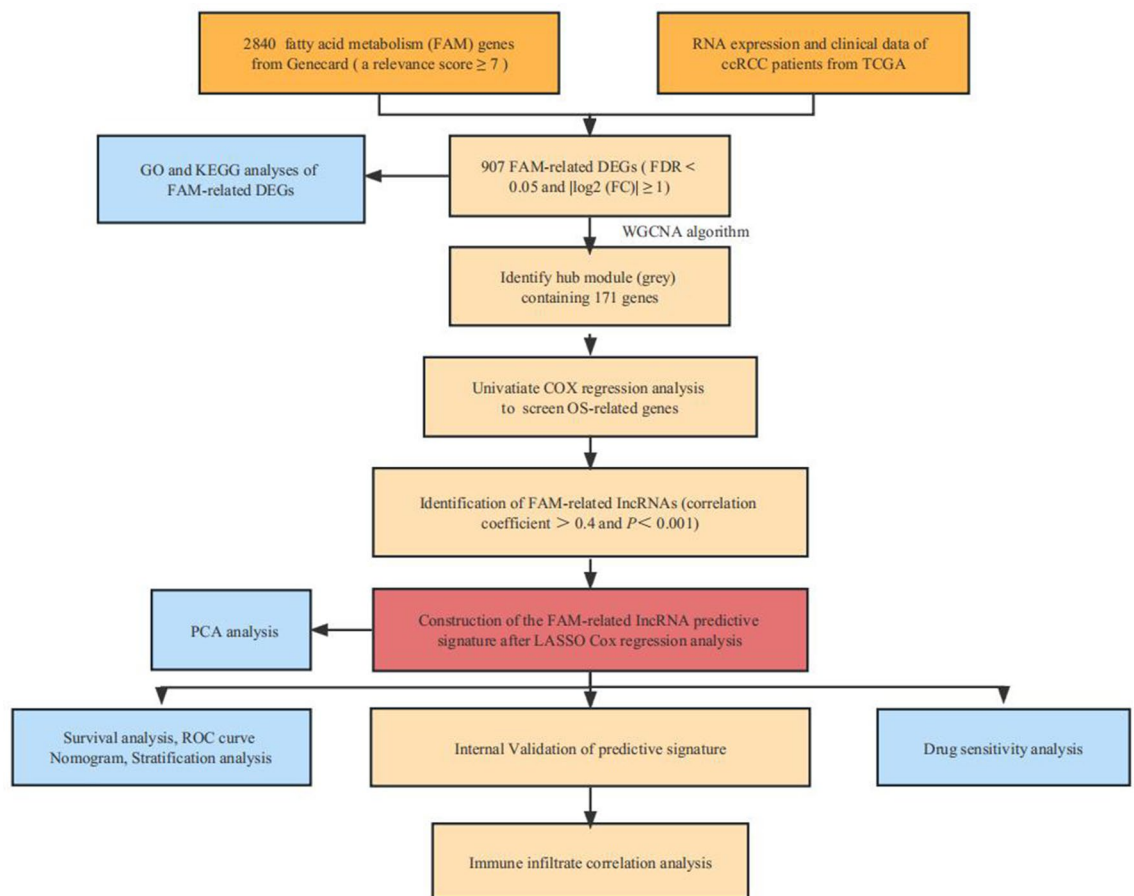


Figure 1. Flow chart of the study. ccRCC, renal clear cell carcinoma; TCGA, The Cancer Genome Atlas; DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; lncRNAs, long noncoding RNAs; ROC, receiver operating characteristic.

to lymphatic dissemination via activation of the nuclear factor κ B (NF- κ B) signaling pathway^{9,13}. Acute myeloid leukemia cells are more likely to survive when their fatty acid oxidation is activated¹⁴, and changes in fatty acid metabolism may affect the immune response and the effectiveness of chemotherapy and radiotherapy^{15,16}. Previous research has demonstrated that fatty acids can influence the phenotype and function of invading immune cells and may result in immunosuppression¹⁷. Recent research has demonstrated glycolytic reprogramming in RCC causes down-regulated METTL14 to accumulate BPTF and increase super-enhancers and distal lung metastasis¹⁸. Furthermore, RCC reprogrammes glutamine, tryptophan, and arginine metabolism to support tumor growth and carcinogenesis¹⁹. However, the prognostic and potential therapeutic implications of genes implicated in fatty acid metabolism in ccRCC patients, particularly in immunotherapy, are the subject of surprisingly few investigations.

Long noncoding RNAs (lncRNAs) play a crucial role in metabolic reprogramming in controlling metabolism-related pathways^{20–22}. For instance, lncRNA-GLCC1 was discovered to reprogramme glycolytic metabolism and was connected to colorectal cancer prognosis²³. In lung adenocarcinomas, ccat1/FABP5 facilitates tumor growth by regulating fatty acid metabolism²⁴. LINC01606 is an oncogene to encourage tumor cell stemness while defending colon cancer cells against iron-induced cell death²⁵. For several cancer types, CDKN2B-AS1 is seen as a prospective biomarker or therapeutic target²⁶.

As a result, we conducted this extensive, systematic study to identify genes involved in fatty acid metabolism using a weighted gene coexpression network analysis (WGCNA) approach and then created a lncRNA risk profile containing the eight FAM-related lncRNAs (AC009166.1, LINC00605, LINC01615, HOXA-AS2, AC103706.1, AC009686.2, AL590094.1, AC093278.2). We tested their value in predicting prognosis and response to chemotherapy and immunotherapy in patients with ccRCC. The findings of this study add to our understanding of how FAM affects ccRCC and will help ccRCC patients to receive more effective individualized care.

Results

Analysis of FAM-related gene enrichment. By following the steps illustrated in Fig. 1, we screened a total of 907 FAM-associated DEGs, 421 of which were upregulated and 486 downregulated (Fig. 2a and Supplementary File 1). Then, we determined the pathways in which the FAM-associated DEGs are involved using GO and KEGG enrichment analyses, including the PI3K-Akt signaling pathway, the HIF-1 signaling system, bile

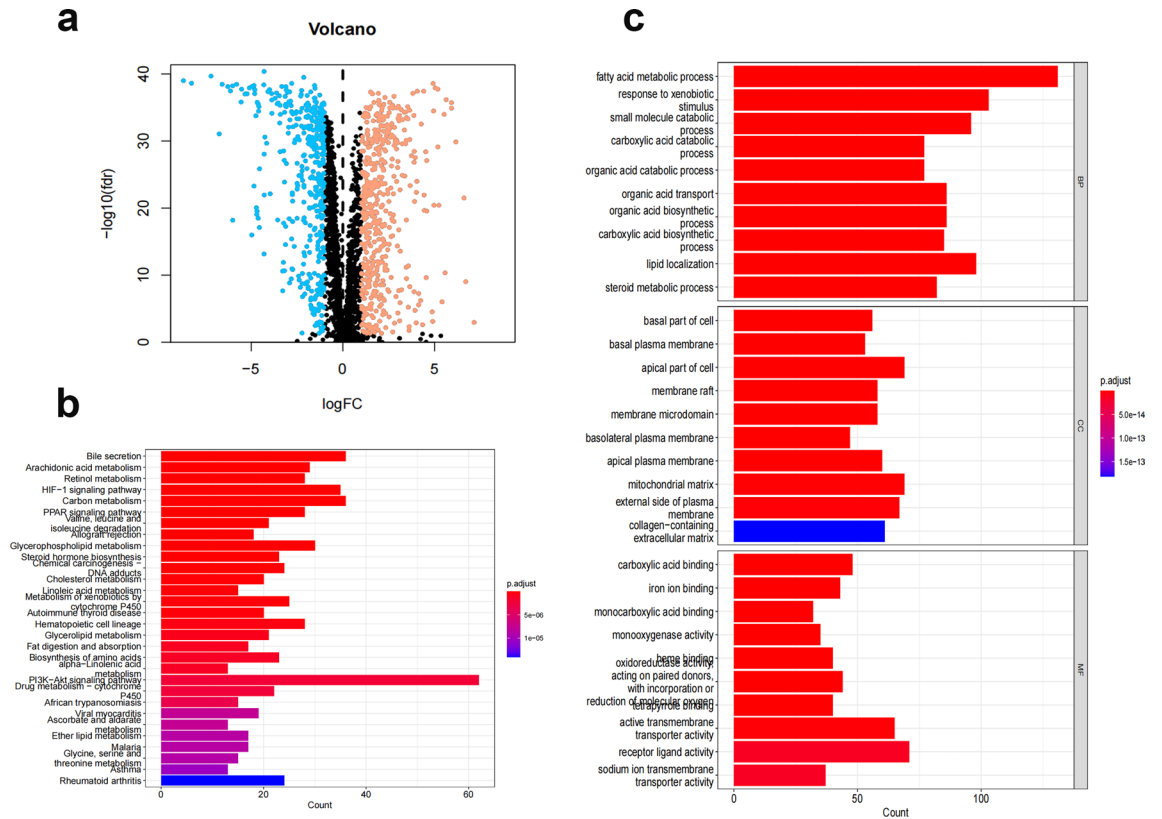


Figure 2. GO and KEGG analysis of FAM-associated DEGs in cancerous and adjacent tissues. **(a)** The 907 FAM-related genes in ccRCC. Yellow dots indicate upregulated genes, and blue dots indicate downregulated genes. **(b)** KEGG analysis of FAM-associated DEGs. **(c)** GO analysis of FAM-associated DEGs. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes; FC, fold change; FDR: false discovery rate; BP, biological process; CC, cellular component; MF, molecular function.

secretion, and the carbon metabolism-related signaling pathway in cancer. KEGG analysis is mainly the Metabolic Signaling pathway (Fig. 2b). Analysis of DEG enrichment in biological processes found that these genes play a significant role in FAM, response to environmental stimuli, and lipid localization. In the cellular component category, the DEGs are primarily abundant in the mitochondrial matrix near the apex of the cell. The DEGs are primarily enriched in receptor–ligand activity, active transmembrane transporter activity, and carboxylic acid binding GO activities within the molecular function category (Fig. 2c). Most of the genes upregulated affect lipid metabolism.

Web module mining WGCNA. WGCNA was applied to the DEGs filtered from TCGA. The connection between genes in the gene network satisfied the scale-free network distribution when the soft threshold power $\beta = 9$ (scale-free $R^2 = 0.9$) was initially estimated using the scale-free topology criterion (Fig. 3a,b). Then, a phylogenetic tree was constructed to mine coexpression modules (cut height ≥ 0.25) (Fig. 3c). Using hierarchical clustering, modules were analyzed, and modules on the same branch displayed similar gene expression patterns (Fig. 3d). Four coexpression modules were created by combining the related gene modules (Fig. 3e). Among them, the tumor score and the gray module showed a strong positive correlation ($\text{cor} = 0.53$; $P = 9e-14$) (Fig. 3f). Consequently, 171 important genes from the gray modules were chosen for additional study (Supplementary File 2).

Differentially expressed lncRNAs associated with FAM genes used to develop predictive characteristics. We examined 1258 lncRNAs linked to differentially expressed FAM genes. There were 555 lncRNAs (Supplementary File 3) related to prognosis in ccRCC patients, according to single-variable Cox regression analysis. We created a predictive model utilizing LASSO Cox regression and 1000-fold cross-validation to predict the prognosis of patients with ccRCC (Fig. 4a,b). Eight lncRNAs (AC009166.1, LINC00605, LINC01615, HOXA-AS2, AC103706.1, AC009686.2, AL590094.1, AC093278.2) related to FAM were used to create predictive characteristics. The expression levels of the eight FAM-related lncRNAs in ccRCC patients are displayed in Fig. 4c (Supplementary File 4). We used the R packages ggalluvial and Cytoscape to further visualize lncRNAs. Thirty-eight lncRNA–mRNA pairs were used for the coexpression network (Fig. 4d). LINC01615, HOXA-AS2, AC103706.1, AC009686.2, and AL590094.1 were found to be risk factors and AC009166.1, AC093278.2, and LINC00605 protective factors (Fig. 4e). The risk score was calculated as follows: Risk score = $(-0.53 \times \text{AC009166.1}$

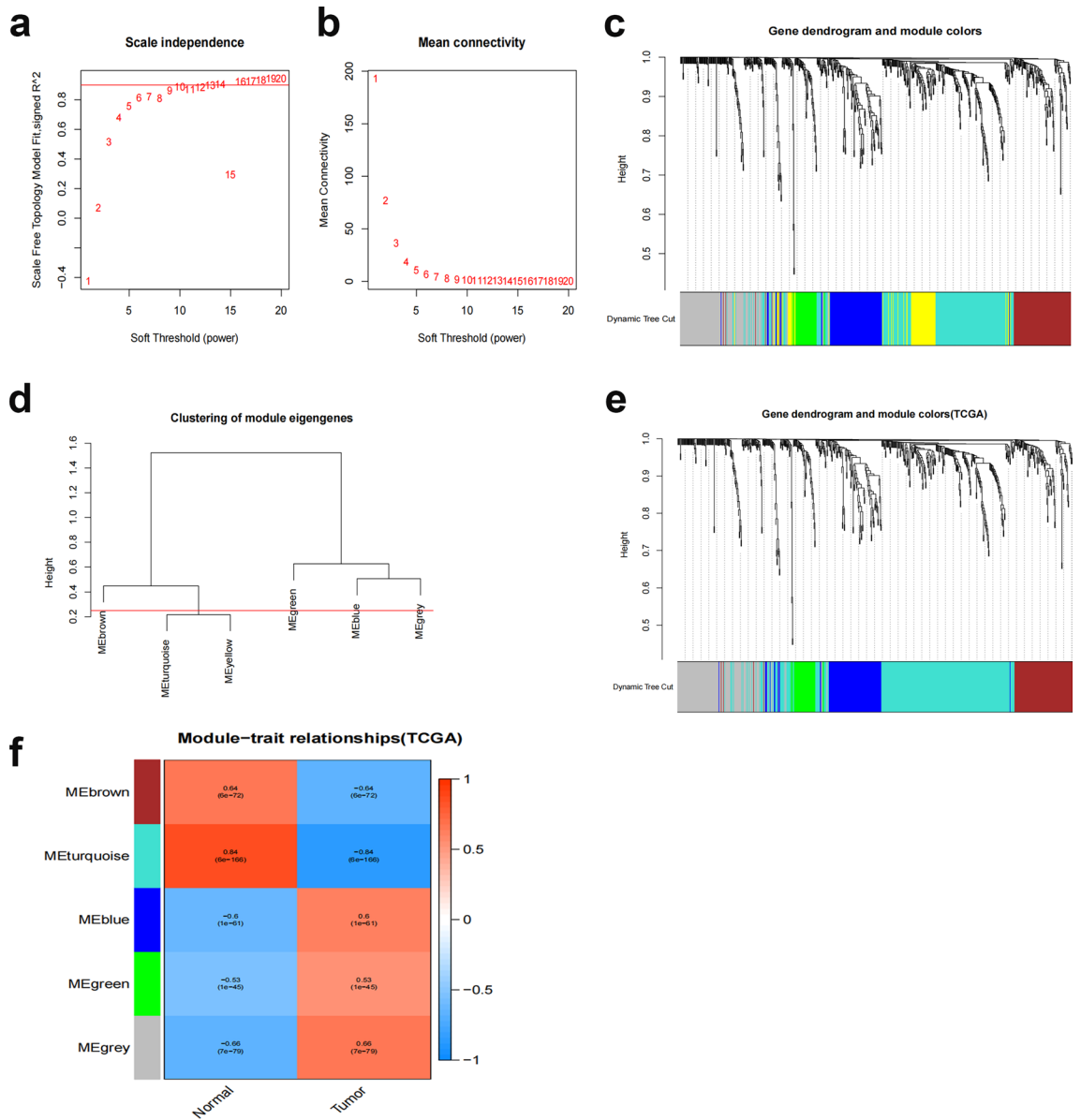


Figure 3. WGCNA network module mining. **(a,b)** Determination of the optimal soft threshold by network topology analysis. The scale-free topology model fit, signed R^2 is satisfied when $\beta=9$. **(c)** Gene dendrogram and nodal color of WGCNA. **(d)** Hierarchical clustering analysis of the WGCNA module. **(e)** Gene dendrograms are based on clustering. **(f)** Correlation between modules.

expression) + (- 0.64 × LINC00605 expression) + (0.35 × LINC01615 expression) + (0.53 × HOXA-AS2 expression) + (0.55 × AC103706.1 expression) + (0.71 × AC009686.2 expression) + (0.71 × AC009686.2 expression) + (0.41 × AL590094.1 expression) + (- 0.36 × AC093278.2 expression).

Analysis of the relationship between prognosis and prognostic traits of individuals with ccRC. The risk score of each patient was determined according to the algorithm, and the patients were then separated into high-risk and low-risk groups based on the median value. Kaplan–Meier analysis of overall survival times revealed that the low-risk group’s OS was significantly lower than that of the high-risk group (Fig. 5a, $p < 0.001$). The difference in risk scores is depicted in Fig. 5b, where it can be seen that as risk scores rise, more deaths occur (Fig. 5c). Age, grading, staging, M-staging, and risk score were significantly associated with OS in ccRCC patients according to one-way Cox regression analysis, and age, grading, staging, and risk score were independent predictors of OS in ccRCC patients according to multiway Cox regression analysis (Fig. 5d), indicating that the risk characteristics were independent risk factors for ccRCC prognosis (Fig. 5e). The predictive value of other clinicopathological variables was lower than the AUC of the risk score (AUC = 0.776) (Fig. 5f). The AUCs for 1-, 3-, and 5-year survival were 0.796, 0.773, and 0.785, respectively, demonstrating good predictive performance (Fig. 5g). We examined the differences in clinicopathological characteristics between the high-

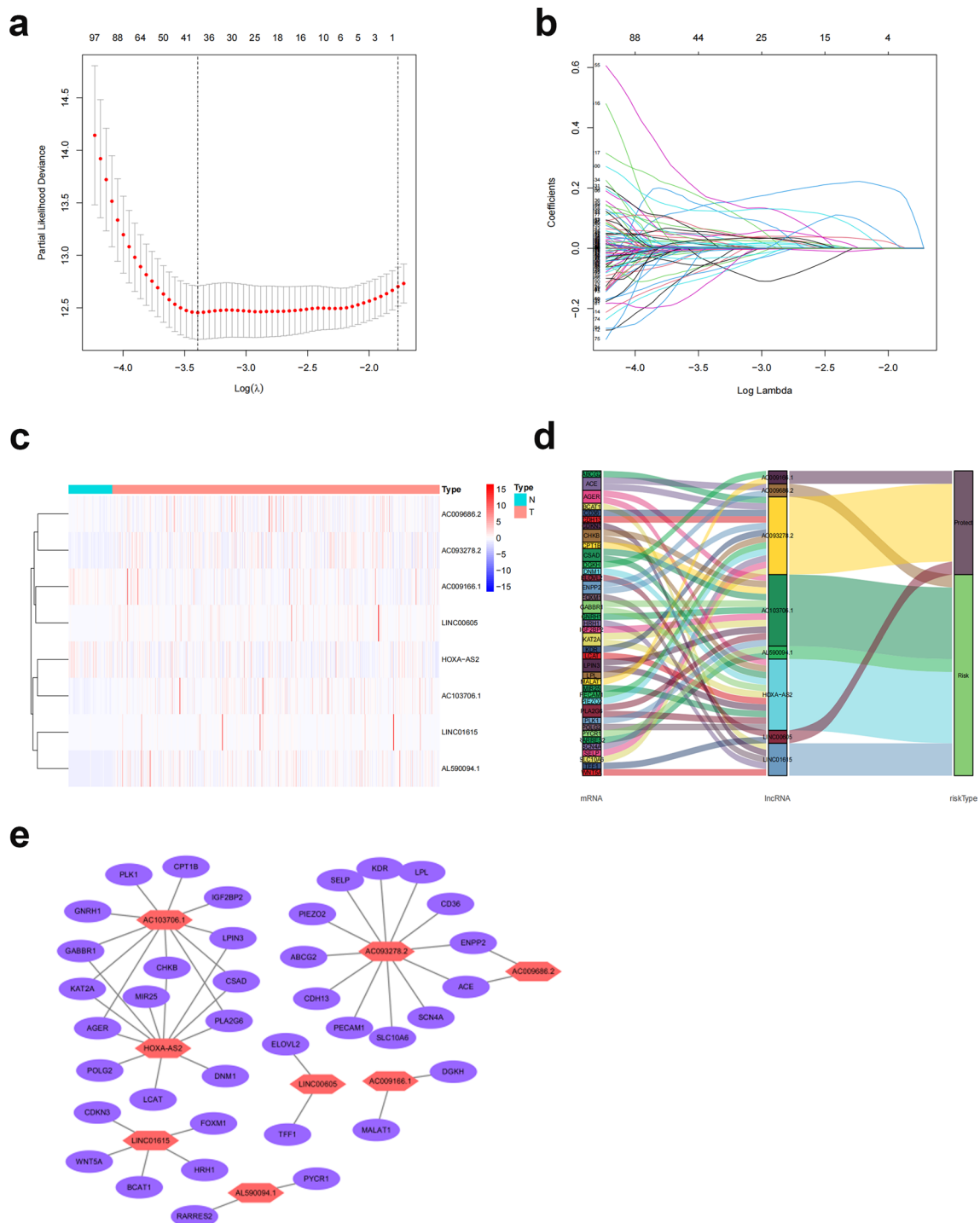


Figure 4. Analysis of the expected signal, including screening, expression levels, and lncRNA–mRNA networks for the eight lncRNAs related to FAM. **(a)** Tenfold cross-validation error rate plot. **(b)** LASSO coefficient profiles of FAM-associated lncRNAs. **(c)** Heatmap (red represents high expression, whereas blue represents low expression) of the eight FAM-associated lncRNAs between the normal (brilliant blue) and the tumor tissues (red) (R software (version 4.2.1, URL: <http://www.r-project.org>)). **(d)** Prognostic FAM-associated lncRNA coexpression network. **(e)** A Sankey diagram of lncRNAs linked with prognostic FAM. Long non-coding RNAs, lncRNAs; renal clear cell carcinoma, ccRCC.

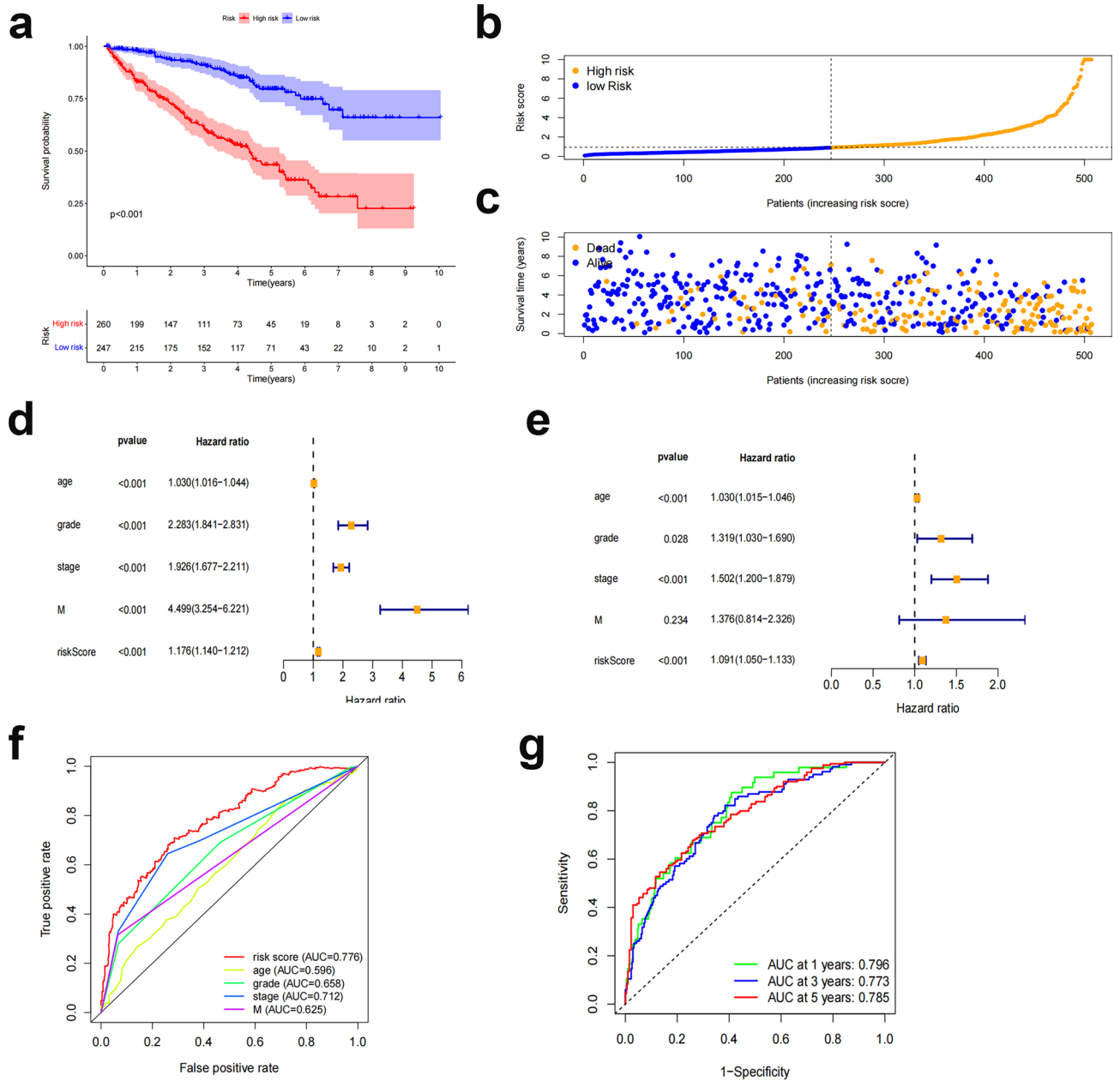


Figure 5. Correlation between prognostic features and prognosis for ccRCC patients. **(a)** Kaplan–Meier analysis of OS rates for patients with ccRCC in high- and low-risk groups. **(b)** Distribution of risk scores among patients with ccRCC. **(c)** Percentage of patients who died and survived for each risk score. Red denotes the number of fatalities, whereas blue denotes the number of survivors. **(d)** Forest plot of a Cox regression analysis with one variable. **(e)** Multivariate Cox regression analysis forest plot. **(f)** Risk scores versus clinicopathological factors using ROC curves. **(g)** ROC curves and AUCs for 1-year, 3-year, and 5-year survival predictions. AUC, area under the curve; T, tumor; M, metastasis; ROC, receiver operating characteristic; ccRCC, renal clear cell carcinoma.

risk and low-risk groups to rule out the influence of these factors. However, there was no discernible difference between the groups (Fig. 6).

To further predict the prognosis of ccRCC patients, we constructed nomogram prediction plots integrating clinicopathological variables and risk scores to predict prognosis at 1, 3, and 5 years (Fig. 7a). Following calibration, the findings revealed that the anticipated survival and actual OS rates were in good agreement (Fig. 7b–d).

Relationship between predictive characteristics of various clinicopathological indicators and ccRCC patient prognosis.

To examine the association between prognostic indicators and predictive markers in ccRCC patients under various clinicopathological variable classifications, subgroup analysis of survival was conducted among patients with varying ages, grades, stages and M stages and both sexes. OS was poorer in

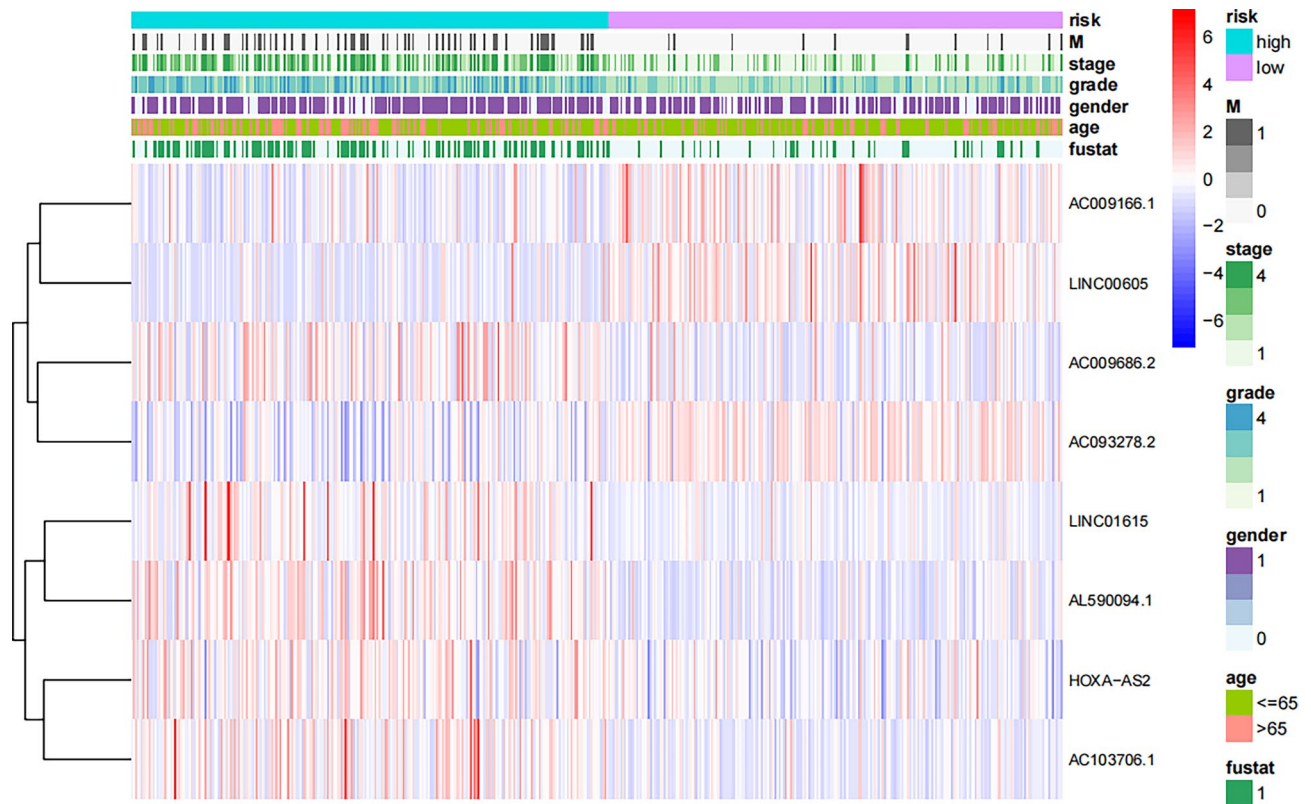


Figure 6. Heatmap (Purple: high expression; brilliant blue: low expression) for the connections between, clinicopathologic features and the risk groups (R software (version 4.2.1, URL: <http://www.r-project.org>)). lncRNAs, long noncoding RNAs; T, tumor; M, metastasis.

all high-risk groups than in low-risk groups (Fig. 8), suggesting that predictive characteristics can predict ccRCC prognosis under diverse conditions.

Internal validation of predictive characteristics. We randomly divided ccRCC patients into two groups to verify the applicability of the prognostic prediction signature based on the whole TCGA dataset; the demographic characteristics are displayed in Table 1. The high-risk group of the training cohort had worse survival than the low-risk group, in line with the findings for the overall dataset (Fig. 9a, $p = 1.22e-12$). In addition, the OS rate was lower in the high-risk group than in the low-risk group in the validation cohort (Fig. 9b, $p = 4.17e-06$). ROC curves for both groups provided an indicator of the patient's clinical performance. AUCs for the training group's 1-, 3-, and 5-year survival rates were 0.835, 0.808, and 0.814, respectively (Fig. 9c); those for the validation group's 1-, 3-, and 5-year survival rates were 0.752, 0.737, and 0.769, respectively (Fig. 9d).

Infiltration of immune cells and functional analysis. To illustrate the spatial distribution of high- and low-risk samples, we visualized the distribution of patients based on the complete genome, FAM-associated gene sets, FAM-associated lncRNAs, and predictive characteristics using principal component analysis (PCA). High- and low-risk patients grouped into separate quadrants based on predictive factors (Fig. 10). As TME composition is an essential indicator of the development of cancer, we evaluated the distribution of several immune cell subtypes in the two risk groups using several genetic markers. The results showed that patients in the high- and low-risk groups had significantly different levels of activated dendritic cells (aDCs), CD8+ T cells, immature dendritic cells (iDCs), mast cells, macrophages, T helper cells, T follicular helper (Tfh) cells, T helper type 1 (Th1) cells, T helper type 2 (Th2) cells, and tumor-infiltrating lymphocytes (TILs) (Fig. 11a). Moreover, the high-risk group had higher immune function scores for chemokine receptors (CCR), checkpoints, lysis activity, inflammation promotion, part inflammation, T-cell coinhibition, T-cell costimulation, and type I IFN response (Fig. 11b), indicating an immunologically active state. Furthermore, immune checkpoint expression differed between the groups (Fig. 11c).

Predictive characteristics and ccRCC therapy correlation. We utilized the OncoPredict package to predict drug sensitivity scores in the high- and low-risk groups, and sensitivity ratings were favorably associated with chemotherapeutic medication IC50 values. The high-risk population showed strong sensitivity to lapatinib, AZD8055, and WIKI4 (Fig. 12a–c). In the low-risk group, sensitivity to axitinib, cediranib, and osimertinib was high (Fig. 12d–f). Overall, differentiating between high- and low-risk patients can help in developing a personalized treatment plan.

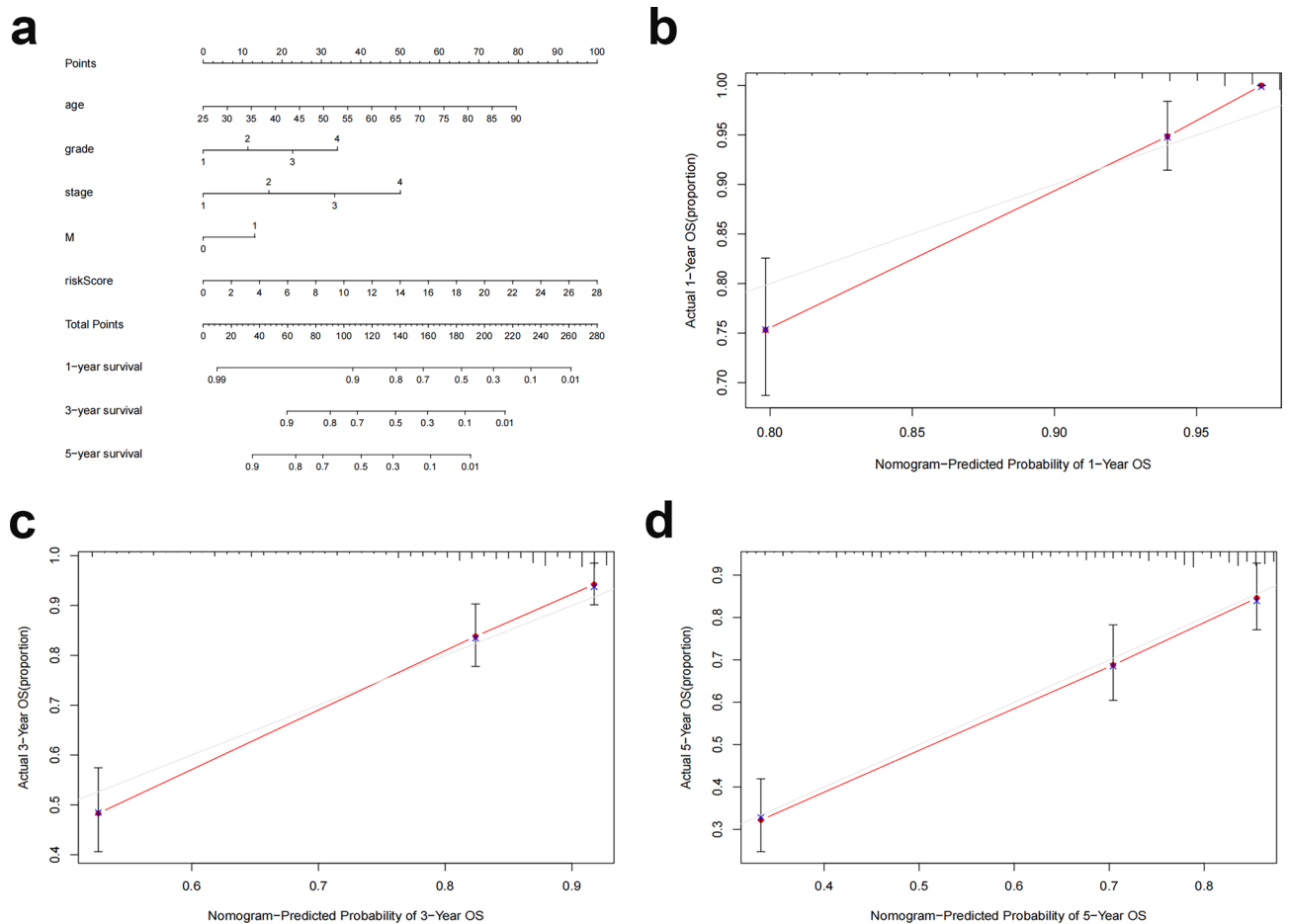


Figure 7. Creation and validation of a nomogram. (a) Clinicopathological factors and risk scores combined with a nomogram to predict survival in ccRCC patients at 1, 3, and 5 years. (b–d) Calibration curves to assess whether expected survival at 1, 3, and 5 years agrees with actual OS rates. M, metastasis; OS, overall survival; ccRCC, clear cell renal cell carcinoma.

Correlation between risk scores/FAM genes and clinical variables. We examined the relationship between the above clinical factors and risk ratings of the eight FAM genes based on gene expression and associated clinical data we collected from TCGA. The results demonstrated that AC009686.2, AL590094.1, and risk scores were connected with age, tumor stage, and tumor grade and that AC093278.2 correlated with tumor stage and tumor grade (Fig. 13).

Discussion

Renal cancer is a heterogeneous disease. According to the latest research, it is thought that the majority of ccRCC cases are caused by a combination of mechanisms, including dysregulation of hypoxia-inducible factor (HIF) signaling, mutations in vital histone and chromatin-modifying enzymes, and metabolic reprogramming of cellular metabolism^{27,28}. Despite improvements in targeted therapy and diagnostic methods in recent years, metastasis and invasion in malignancies such as ccRCC can result in a gravely unfavorable prognosis²⁹. Sunitinib, a broad-spectrum small-molecule inhibitor of receptor tyrosine kinases (RTKs), is currently the standard treatment for first-line treatment of advanced clear cell renal cell carcinoma (ccRCC). However, the majority of patients develop resistance and experience disease progression^{30,31}.

Accumulating data show that angiogenesis, growth, proliferation, and aggressiveness of cancer cells all strongly correlate with metabolic instability^{32,33}. According to preliminary data, the physiological behavior of human malignancies is thought to be related to aberrant glycolytic metabolism³⁴. In colorectal cancer, abnormal anaerobic metabolic pathways are crucial for development of tumor stem cells (CSCs), which promote rapid tumor development, progression, and medication resistance^{35,36}. In addition, cellular energy production, membrane synthesis, and signaling pathways linked to tumor formation are all impacted by FAM^{32,37}. Fatty acid anabolism/catabolism dysregulation may promote development of cancerous cells¹⁰. Previous research has shown that the glutathione redox system prevents the ccRCC iron mortality induced by poor lipid metabolism³⁸, and a recent study revealed that FE2F1 activates SREBP1-dependent fatty acid production to increase proliferation and metastasis of ccRCC³⁹. Although several researchers have concentrated on the role that FAM plays in different tumors, it is still unclear how this issue applies to ccRCC. It may be possible to gain new knowledge about

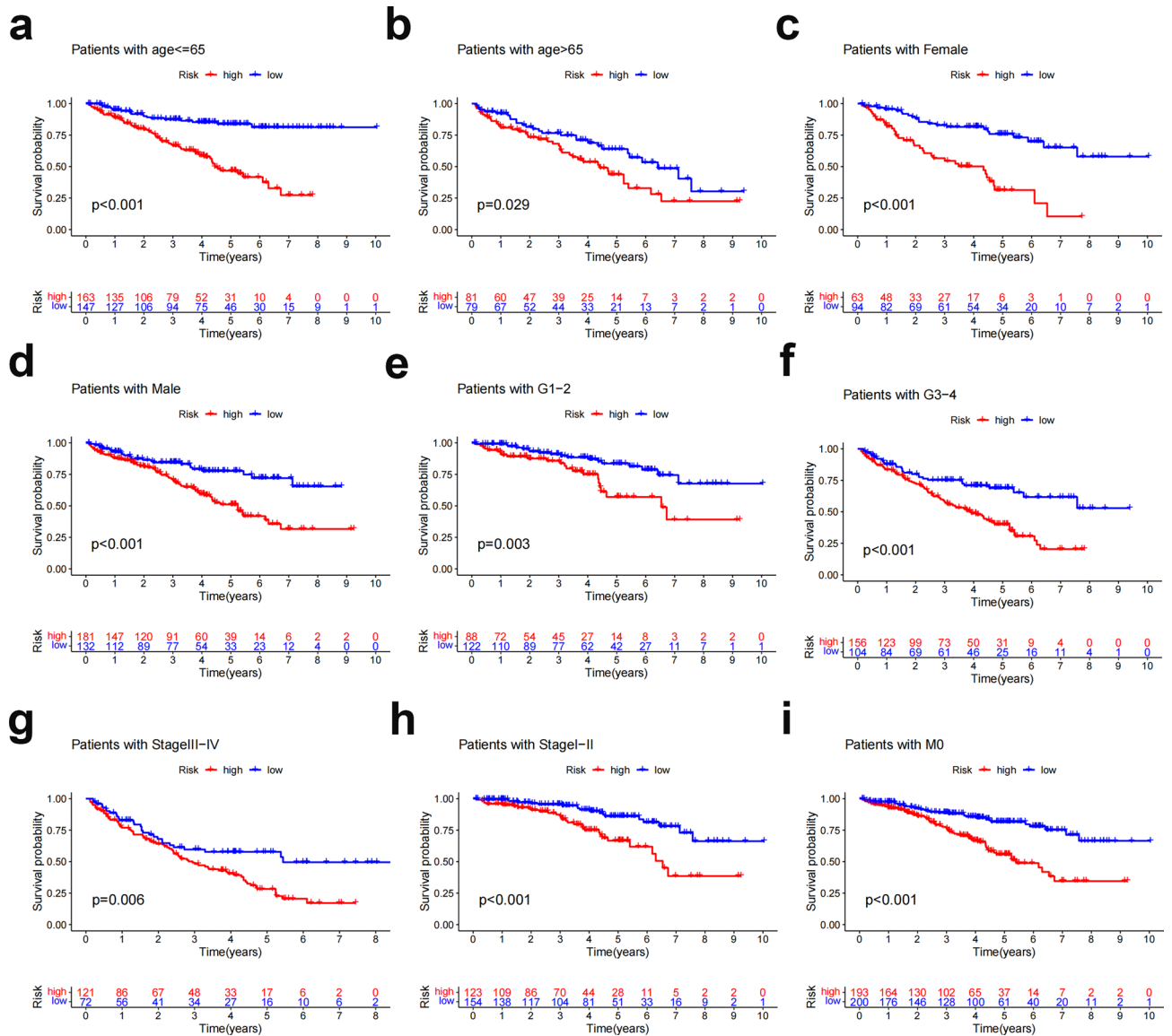


Figure 8. Kaplan–Meier survival curves for high- and low-risk patients based on clinicopathological factors. (a,b) Age. (c,d) Sex. (e,f) Grading. (g,h) Staging. (i) M staging. M, distant metastases.

the biological behavior of these tumors and identify novel and more potent therapeutic approaches by revealing important molecular markers linked to FAM and investigating their role in the development of ccRCC.

In the current work, we thoroughly investigated the function of FAM genes in ccRCC. First, the PI3K-Akt, HIF-1, and other tumor-related signaling pathways were found to be primarily enriched among DEGs; thus, FAM genes with differential expression are linked to tumor growth. Then, we chose the genes in the gray modules that were highly related to tumors for ensuing analysis after using WGCNA to identify four modules. We uncovered eight important FAM genes that may serve as predictive biomarkers in clinical management of ccRCC patients, as based on LASSO and multivariate Cox regression analysis. Through univariate and multivariate Cox regression analyses, a prognostic risk score model that divide patients into high-risk and low-risk categories was produced, with the prognostic risk score as an independent prognostic factor. The model’s prediction ability was validated with risk assessment nomogram clinical characteristics (age, sex, and tumor TNM stage). Here, we present a risk model that may be used to identify ccRCC patients with poor prognosis and to manage this disease.

Based on a literature review, these eight genes are somewhat related to cancers, and five genes linked explicitly to ccRCC were found among them according to PubMed. For instance, it has been revealed that the tumor-infiltrating lymphocyte-associated long noncoding RNA LINC01615 is involved in predicting ccRCC⁴⁰. By controlling several lncRNAs, sorafenib plays a role in the response to sorafenib in several cancer cell types⁴¹. Among three additional genes, AC103706.1 may be able to determine the prognosis of individuals with renal subhyalinosis⁴², and the ccRCC immune microenvironment and prognosis can be predicted by the ferroptosis-related long non-coding RNA AL590094.1⁴³. Ac093278.2, an immune-related long noncoding RNA, is used to predict prognosis in ccRCC⁴⁴. A PubMed search for "GENE and ccRCC" revealed that the three genes AC009166.1, LINC00605, and AC009686.2 have not been investigated in ccRCC. Additional fundamental and clinical studies are necessary to

Variables	Entire TCGA dataset (n = 507)	Internal validation cohort	
		First cohort (n = 255)	Second cohort (n = 252)
Age (%)			
≤ 65	336 (66.3)	165 (64.7)	171 (67.9)
> 65	171 (33.7)	90 (35.3)	81 (32.1)
Gender (%)			
Female	174 (34.3)	91 (35.7)	83 (32.9)
Male	333 (65.7)	164 (64.3)	169 (67.1)
Grade (%)			
G1 + 2	227 (44.8)	115 (45.1)	112 (44.4)
G3 + 4	272 (53.6)	138 (54.1)	134 (53.2)
GX + unknown	8 (1.6)	2 (0.8)	6 (2.4)
Stage (%)			
I + II	306 (60.3)	151 (59.2)	155 (61.5)
III + IV	198 (39.1)	101 (39.6)	97 (38.5)
TX + unknown	3 (0.6)	3 (1.2)	0 (0)
T (%)			
T1 + 2	324 (63.9)	160 (62.7)	164 (65.1)
T3 + 4	183 (36.1)	95 (37.3)	88 (34.9)
M (%)			
M0	392 (77.3)	202 (79.2)	190 (75.4)
M1	73 (14.4)	37 (14.5)	36 (14.3)
MX + unknown	42 (8.3)	16 (6.3)	26 (10.3)
N (%)			
NO	16 (3.2)	10 (4.0)	6 (2.4)
N1	266 (52.4)	135 (52.9)	131 (52.0)
NX	225 (44.4)	110 (43.1)	115 (45.6)

Table 1. Clinical characteristics of the different cohorts of patients. T, tumor; M, metastasis; N, lymph node.

confirm the mechanisms underlying the predictive value of the eight genes on which our observations and risk models are based, as well as to develop new therapeutic targets to improve OS in ccRCC patients.

Many researchers have reported the relationship between FAM and the TME. For instance, under hypoxic conditions, Tregs in glioblastoma depend on fatty acid oxidation (FAO) for energy⁴⁵. Significant CD8+ T-cell infiltration has been linked to clinical outcomes and immunological response in ESCA, according to earlier research^{46,47}. An elevated CD8+ T-cell to Treg ratio has been linked to poor prognosis in RCC⁴⁸. In samples from individuals with ccRCC in the high-risk category, we noted a rise in infiltrating CD8+ T cells, indicating poor prognosis. Furthermore, poor prognosis is linked to advanced thyroid carcinoma and high tumor-associated macrophage infiltration⁴⁹. The high-risk group also had greater type I IFN response scores and less antitumor immunity in addition to increased tumor immune cell infiltration. IFN production is a major regulator of PD-L1 expression in tumor cells and host cells, and it may enhance the benefit of anti-PD-1 therapy⁵⁰. Therefore, a diminished antitumor immune response may portend worse outcome for those at risk. In general, patients with advanced cancers have a better chance of surviving with checkpoint inhibitor-based immunotherapy⁵¹.

To our knowledge, this work is the first investigation of prognostic FAM-related lncRNAs in ccRCC. The signature, which is based on eight FAM-related lncRNAs, offers a practical way to predict clinical outcome and offers some guidance for choosing immunotherapeutic and chemotherapeutic treatments. However, there are a number of limitations that should be considered. First, we solely used data from TCGA for internal validation, and external validation requires data from other databases to confirm the applicability of the prediction signature. Second, we used information from public databases as retrospective data. Therefore, additional real-world data in the future are required to confirm therapeutic applicability. In addition, our research lacks validation by in vivo and in vitro experiments. The next step is to confirm how FAM-related lncRNAs function in ccRCC.

Conclusion

In conclusion, our FAM-associated lncRNA signature can independently predict prognosis of patients with ccRCC. Furthermore, it offers a potentially fruitful route for selecting antitumor immunotherapy and chemotherapeutic drugs.

Methods

Data acquisition and processing. Through TCGA (<https://portal.gdc.cancer.gov/>), we accessed RNA-seq data with FPKM normalization for renal clear cell carcinoma (TCGA- KIRC) together with related clinical and prognostic data. lncRNA expression and survival time data were gathered for 611 patients. Patients with

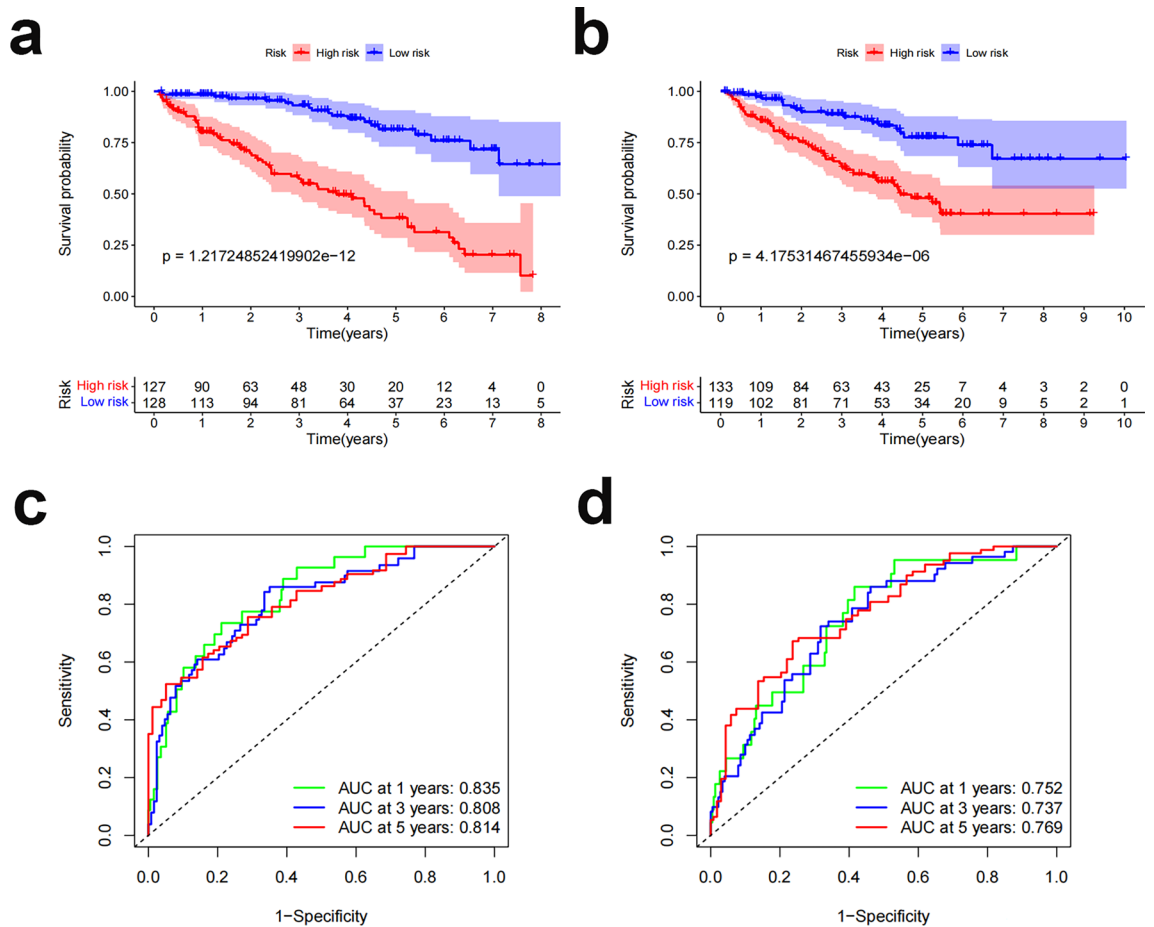


Figure 9. Internal validation of OS prediction signatures based on internal datasets. (a) Training group Kaplan–Meier survival curves. (b) Test group Kaplan–Meier survival curves. (c) ROC curves and AUCs for 1-, 3-, and 5-year survival in the training group. (d) ROC curves and AUCs for 1-, 3-, and 5-year survival in the test group. ROC, receiver operating characteristic; AUC, area under the curve; OS, overall survival.

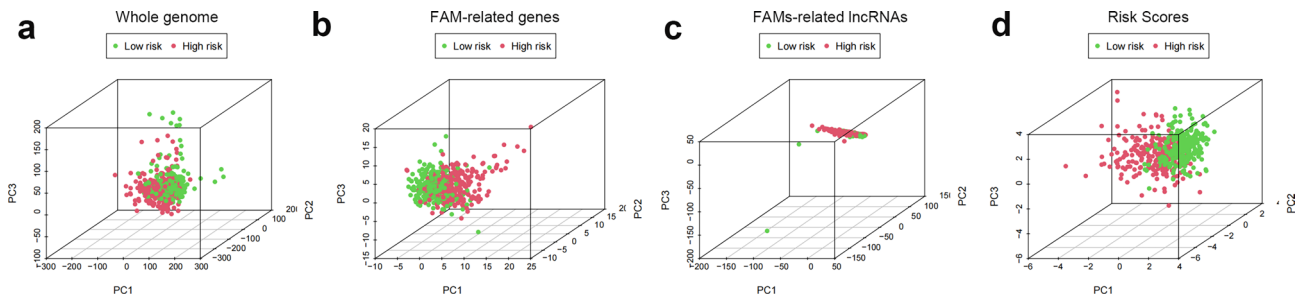


Figure 10. PCA profiles showing patient distribution based on (a) genome-wide; (b) zinc finger-associated genes; (c) zinc finger-associated lncRNAs; and (d) risk scores. The separation between red and green dots is stronger in the high- than in the low-risk group.

less than 30 days of follow-up were excluded; 507 patients met the inclusion requirements. The patients were randomly assigned in a 1:1 ratio to a training group (n = 255) and test group (n = 252). GeneCards (<https://www.genecards.org>) provided download access to 9368 genes involved in FAM.

Screening for DEGs. First, FAM genes in GeneCards with correlation values below 7 were disregarded. Then, DEGs were obtained by comparing normal and ccRCC tissues in TCGA and the filter criteria using a false discovery rate (FDR) of < 0.05 and $|\log_2\text{-fold change (FC)}| \geq 1$ ⁵². Nine hundred-seven differentially expressed FAM-related genes were subsequently identified for further investigation.

Functional enrichment analysis. Using the "ggplot2" package, we carried out analyses using the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. We found mRNAs that had

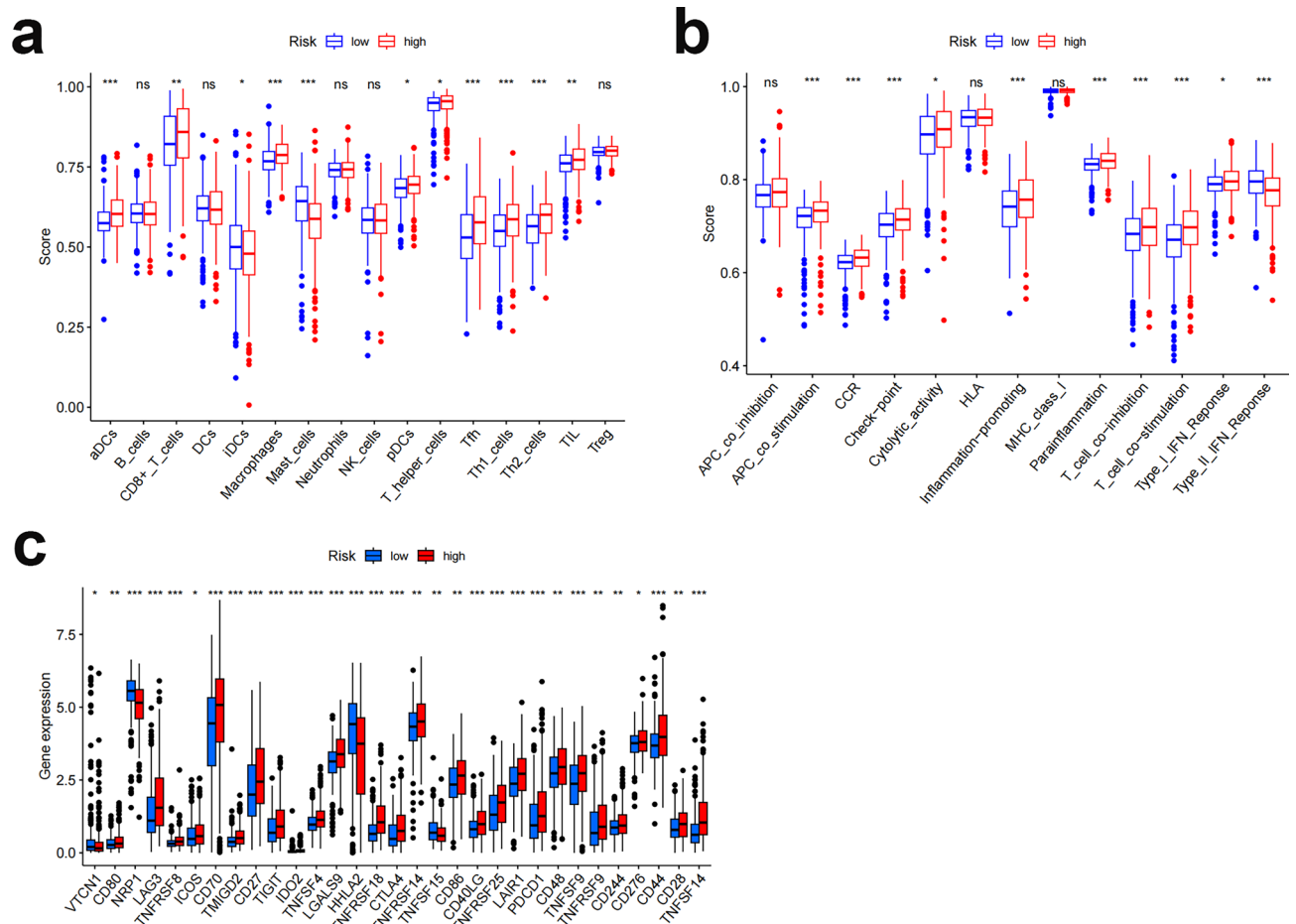


Figure 11. Immuno-infiltration analysis. Results of ssGSEA scoring. **(a)** Sixteen immune cell scores. **(b)** Thirteen immune-related function scores. **(c)** Expression of immune checkpoints in high- and low-risk populations. ssGSEA, single-sample gene set enrichment analysis; aDCs, activated dendritic cells; iDCs, immature dendritic cells; NKs, natural killer cells; pDCs, plasmacytoid dendritic cells; Tfh, T follicle helper cell; Th1, T helper factor type 1; Th2, T helper cell type 2; TIL, tumor-infiltrating lymphocyte; Treg, T-regulatory cell; APC, antigen-presenting cell; CCR, chemokine receptor; HLA, human leukocyte antigen; MHC, major histocompatibility complex; IFN, interferon. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns, not significant.

strong correlations with the 8 FAM-related lncRNAs mentioned above while assessing the biological function of these lncRNAs in ccRCC. Five coexpression networks of lncRNAs and mRNAs were created and visualized with Sankey plots. Correlation coefficient criteria were set at > 0.4 or -0.4 , and P values < 0.001 were deemed statistically significant.

WGCNA. According to an earlier description⁵³, ccRCC-related modules were identified using the WGCNA (version 1.61) package. To summarize, soft thresholds were calculated using a scale-free topology criterion. The minimal module size was set to 50 genes, and the ideal soft threshold was selected. Dynamic tree cuts with the MEDissThres option set to 0.25 were used to identify modules.

Prediction characteristics of FAM-related lncRNAs. Using "limma" software, we determined the association between FAM-related genes and lncRNAs. Applying the filters $|R2| > 0.4$ and $p < 0.001$, a total of 1258 FAM-related lncRNAs were retrieved along with their associated expression data. We initially performed one-way Cox regression analysis to obtain FAM-related lncRNAs linked to prognosis in ccRCC patients. Then, we used LASSO Cox regression analysis to obtain FAM-related lncRNAs to create predictive features. Risk score = $\beta_i \times \text{expression}_i$, where β_i is the coefficient and i is the lncRNA, was applied to define the prognostic model. In both datasets, risk scores were computed for each sample. The samples were split into high-risk and low-risk groups based on the median risk score, which served as the cutoff point. The survivor and survminer R packages were employed to conduct Kaplan–Meier survival analysis to compare the survival rates of the two groups. Clustering of risk characteristics was visualized using principal component analysis (PCA) utilizing "limma" and "scatterplot3d"^{54,55}. Additionally, univariate and multivariate Cox regression analyses were carried out to evaluate the association between clinical variables and risk scores. The timeROC package in R was utilized

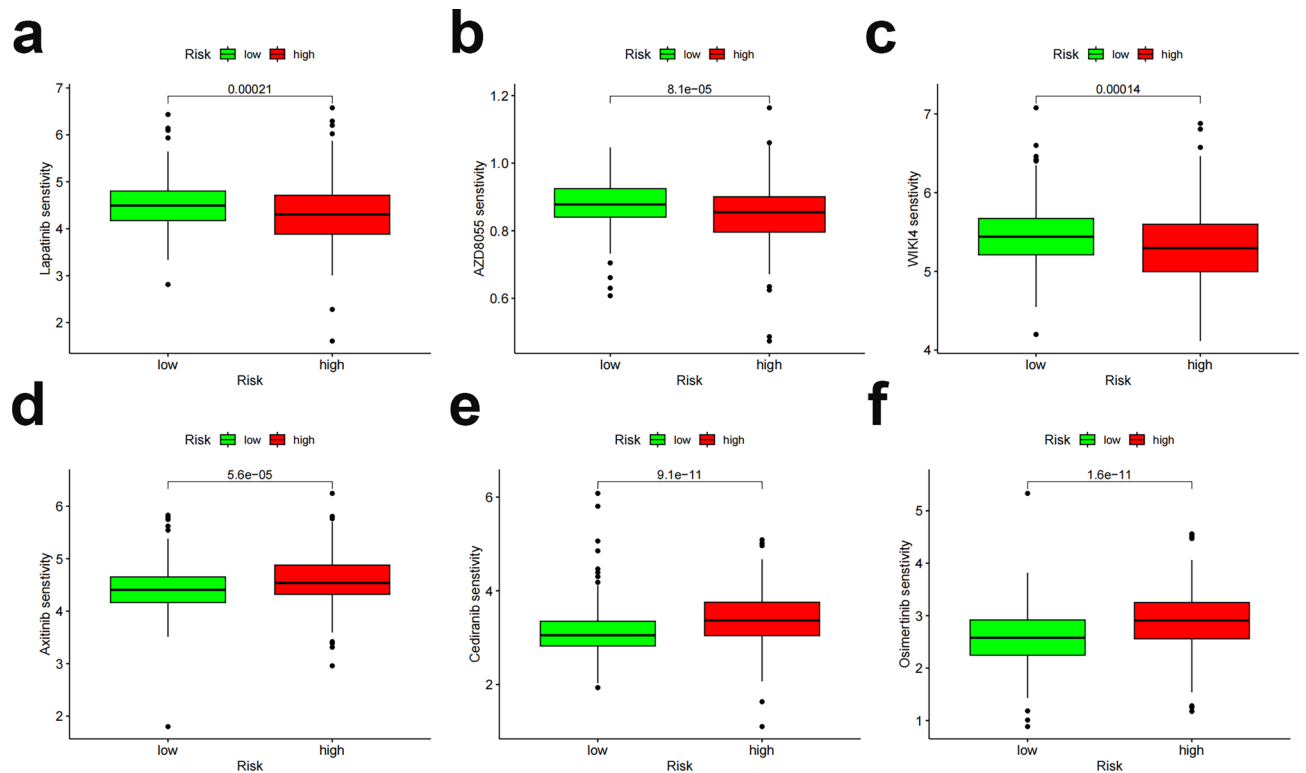


Figure 12. Drug sensitivity analysis. (a–c) Predicted sensitivity scores for lapatinib, AZD8055, and WIKI4 as chemotherapy candidates for patients with high FAM-related lncRNA scores. (d–f) Axitinib, cediranib, and osimertinib are candidates for patients with low scores. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.0001$.

to validate the accuracy and capacity of prediction⁵⁶. All approaches were included in an internal validation cohort to confirm the applicability of the predictive signature.

Creating column line diagrams. We created nomogram survival charts that predicted 1-, 2-, 3-, and 5-year survival in patients with ccRCC by combining risk scores with age, sex, stage, and stage M clinicopathological features. Calibration curves were then applied to determine whether the predicted survival is compatible with actual survival.

Drug sensitivity analysis with predictive signatures. We assessed the contribution of predictive features in predicting response to ccRCC treatment using the Genomics of Drug Sensitivity in Cancer (GDSC) database. This public dataset collects information on therapeutic sensitivity and molecular indicators of drug response in cancer cells⁵⁷. OncoPredict was used to download GDSC2 gene expression profiles and related drug response information⁵⁸. We used sensitivity ratings to estimate the half-maximal inhibitory concentration (IC50) of each medicine in ccRCC patients.

Quantification of immune infiltration levels. The "gsa" program was utilized to compare immune cells and pathways between the two groups utilizing ssGSEA. Using the "GSVA" program, single-sample gene set enrichment analysis (ssGSEA) was employed to determine the infiltration score of 16 immune cells and activation of 13 immune-related pathways⁵⁹. Finally, we investigated the correlation between risk scores and immunological checkpoints by contrasting gene expression levels between high-risk and low-risk groups.

Statistical analysis. R software (version 4.2.1, <http://www.r-project.org>) was used to conduct all statistical analyses. Expression levels of DEGs related to FAM in healthy and malignant tissues were compared using the Wilcoxon test. FAM-related lncRNAs were discovered to be an independent prognostic risk factor for ccRCC based on univariate and multivariate Cox models. The area under the curve (AUC) of the receiver operating characteristic (ROC) curve was used to determine how accurately the FAM-related lncRNA model predicts ccRCC. All statistical tests were two-sided, with $p < 0.05$ regarded as statistically significant. Statistical analyses were carried out using the R environment and the Bioconductor package (version 3.5.5).

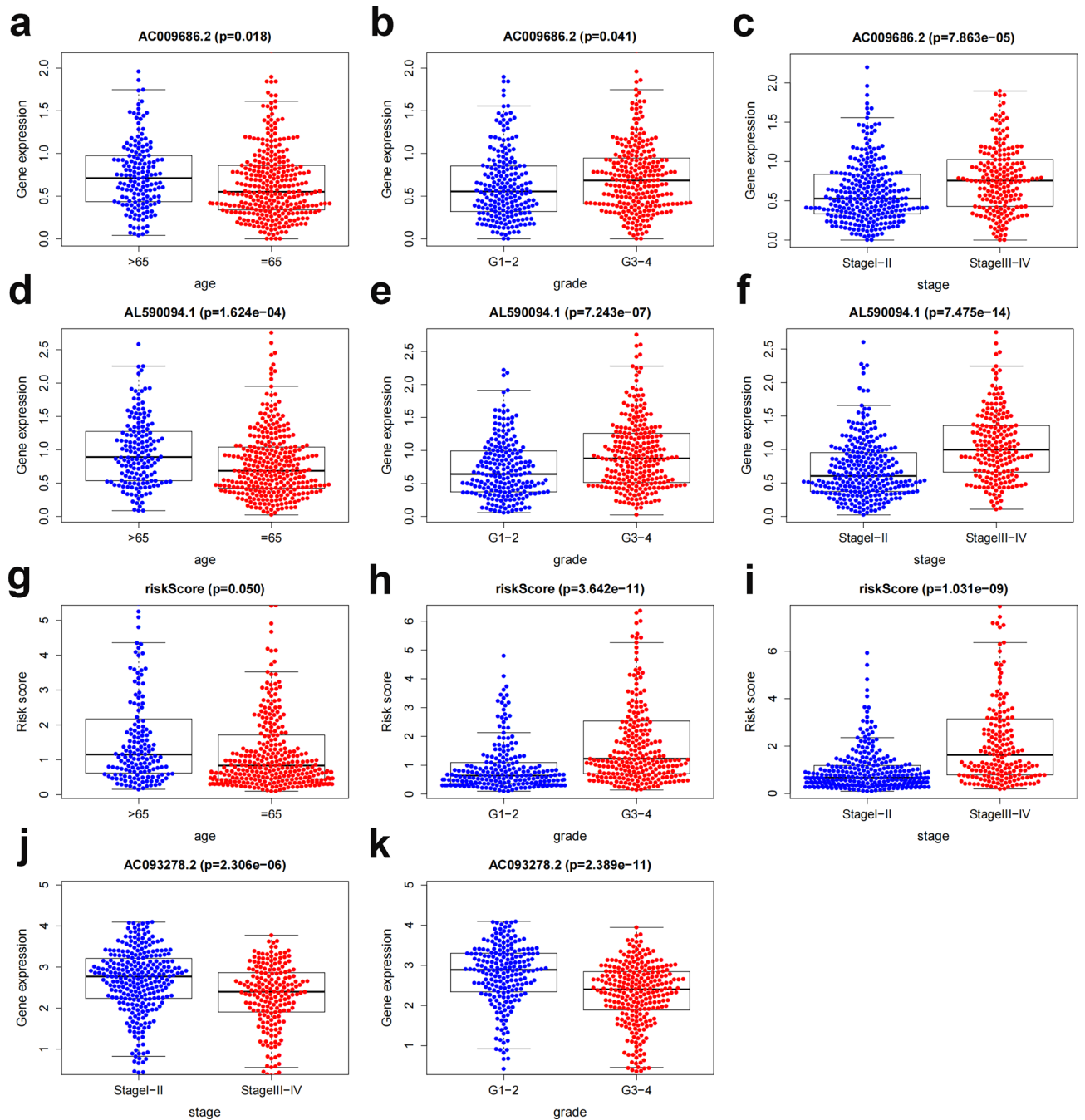


Figure 13. Correlation analysis between model genes and clinical features. (a–c) Correlation of AC009686.2 expression levels with age, grade, and stage. (d–f) Correlation of AL590094.1 expression levels with age, grade, and stage. (g–i) Correlation of risk score with age, grade, and stage. (j,k) AC093278.2 Correlation of expression level with tumor grade and stage.

Data availability

The corresponding author will gladly provide access to the datasets used and analyzed during the current work upon reasonable request.

Received: 19 November 2022; Accepted: 22 April 2023

Published online: 29 April 2023

References

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71, 209–249 (2021).
2. Jonasch, E., Gao, J. & Rathmell, W. K. Renal cell carcinoma. *BMJ* 349, g4797 (2014).

3. Thoenes, W., Störkel, S. & Rumpelt, H. J. Histopathology and classification of renal cell tumors (adenomas, oncocyomas and carcinomas). The basic cytological and histopathological elements and their use for diagnostics. *Pathol. Res. Pract.* **181**, 125–143 (1986).
4. Lalani, A.-K.A. *et al.* Systemic treatment of metastatic clear cell renal cell carcinoma in 2018: Current paradigms, use of immunotherapy, and future directions. *Eur. Urol.* **75**, 100–110 (2019).
5. Jonasch, E., Walker, C. L. & Rathmell, W. K. Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat. Rev. Nephrol.* **17**, 245–261 (2021).
6. Garcia, J. A. & Rini, B. I. Recent progress in the management of advanced renal cell carcinoma. *CA Cancer J. Clin.* **57**, 112–125 (2007).
7. Barata, P. C. & Rini, B. I. Treatment of renal cell carcinoma: Current status and future directions: BRCA testing, ovarian cancer, prevention, therapy, screening. *CA Cancer J. Clin.* **67**, 507–524 (2017).
8. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2018: Cancer statistics, 2018. *CA Cancer J. Clin.* **68**, 7–30 (2018).
9. Ding, C. *et al.* Characterization of the fatty acid metabolism in colorectal cancer to guide clinical therapy. *Mol. Ther. Oncolytics* **20**, 532–544 (2021).
10. Li, Z. & Zhang, H. Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell Mol. Life Sci.* **73**, 377–392 (2016).
11. Carracedo, A., Cantley, L. C. & Pandolfi, P. P. Cancer metabolism: Fatty acid oxidation in the limelight. *Nat. Rev. Cancer* **13**, 227–232 (2013).
12. Wen, J. *et al.* ACLY facilitates colon cancer cell metastasis by CTNNB1. *J. Exp. Clin. Cancer Res.* **38**, 401 (2019).
13. Zhang, C. *et al.* FABP5 promotes lymph node metastasis in cervical cancer by reprogramming fatty acid metabolism. *Theranostics* **10**, 6561–6580 (2020).
14. Tabe, Y., Konopleva, M. & Andreeff, M. Fatty acid metabolism, bone marrow adipocytes, and AML. *Front. Oncol.* **10**, 155 (2020).
15. Han, S. *et al.* CPT1A/2-mediated FAO enhancement—a metabolic target in radioresistant breast cancer. *Front. Oncol.* **9**, 1201 (2019).
16. Wu, Y., Fabritius, M. & Ip, C. Chemotherapeutic sensitization by endoplasmic reticulum stress: Increasing the efficacy of taxane against prostate cancer. *Cancer Biol. Ther.* **8**, 146–152 (2009).
17. Corn, K. C., Windham, M. A. & Rafat, M. Lipids in the tumor microenvironment: From cancer progression to treatment. *Prog. Lipid Res.* **80**, 101055 (2020).
18. Zhang, C. *et al.* Downregulated METTL14 accumulates BPTF that reinforces super-enhancers and distal lung metastasis via glycolytic reprogramming in renal cell carcinoma. *Theranostics* **11**, 3676–3693 (2021).
19. Chakraborty, S., Balan, M., Sabarwal, A., Choueiri, T. K. & Pal, S. Metabolic reprogramming in renal cancer: Events of a metabolic disease. *Biochim. Biophys. Acta Rev. Cancer* **1876**, 188559 (2021).
20. Liu, H., Luo, J., Luan, S., He, C. & Li, Z. Long non-coding RNAs involved in cancer metabolic reprogramming. *Cell Mol. Life Sci.* **76**, 495–504 (2019).
21. Ghafouri-Fard, S., Shoorei, H. & Taheri, M. The role of long non-coding RNAs in cancer metabolism: A concise review. *Front. Oncol.* **10**, 555825 (2020).
22. Sobhani, N. *et al.* Predictive and prognostic value of non-coding RNA in breast cancer. *Cancers* **14**, 2952 (2022).
23. Tang, J. *et al.* LncRNA GLCC1 promotes colorectal carcinogenesis and glucose metabolism by stabilizing c-Myc. *Nat. Commun.* **10**, 3499 (2019).
24. Chen, J., Alduais, Y., Zhang, K., Zhu, X. & Chen, B. CCAT1/FABP5 promotes tumour progression through mediating fatty acid metabolism and stabilizing PI3K/AKT/mTOR signalling in lung adenocarcinoma. *J. Cell Mol. Med.* **25**, 9199–9213 (2021).
25. Luo, Y. *et al.* Long noncoding RNA LINC01606 protects colon cancer cells from ferroptotic cell death and promotes stemness by SCD1-Wnt/ β -catenin-TFE3 feedback loop signalling. *Clin. Transl. Med.* **12**, e752 (2022).
26. Song, C., Qi, Y., Zhang, J., Guo, C. & Yuan, C. CDKN2B-AS1: An indispensable long non-coding RNA in multiple diseases. *Curr. Pharm. Des.* **26**, 5335–5346 (2020).
27. Wettersten, H. I., Aboud, O. A., Lara, P. N. & Weiss, R. H. Metabolic reprogramming in clear cell renal cell carcinoma. *Nat. Rev. Nephrol.* **13**, 410–419 (2017).
28. Xie, M. *et al.* The long intergenic non-protein coding RNA 707 promotes proliferation and metastasis of gastric cancer by interacting with mRNA stabilizing protein HuR. *Cancer Lett.* **443**, 67–79 (2019).
29. Chaffer, C. L. & Weinberg, R. A. A Perspective on Cancer Cell Metastasis. *Science* **331**, 1559–1564 (2011).
30. Huang, D. *et al.* Sunitinib Acts Primarily on Tumor Endothelium rather than Tumor Cells to Inhibit the Growth of Renal Cell Carcinoma. *Cancer Res.* **70**, 1053–1062 (2010).
31. Bergers, G. & Hanahan, D. Modes of resistance to anti-angiogenic therapy. *Nat Rev Cancer* **8**, 592–603 (2008).
32. He, D. *et al.* Prognostic value of fatty acid metabolism-related genes in patients with hepatocellular carcinoma. *Aging (Albany NY)* **13**, 17847–17863 (2021).
33. Xiao, D., Wu, J., Zhao, H., Jiang, X. & Nie, C. RPP25 as a Prognostic-Related Biomarker That Correlates With Tumor Metabolism in Glioblastoma. *Front Oncol.* **11**, 714904 (2021).
34. Wei, X. *et al.* DT-13 inhibited the proliferation of colorectal cancer via glycolytic metabolism and AMPK/mTOR signaling pathway. *Phytomedicine* **54**, 120–131 (2019).
35. Vincent, Z., Urakami, K., Maruyama, K., Yamaguchi, K. & Kusuvara, M. CD133-positive cancer stem cells from Colo205 human colon adenocarcinoma cell line show resistance to chemotherapy and display a specific metabolomic profile. *Genes Cancer* **5**, 250–260 (2014).
36. Zeki, S. S., Graham, T. A. & Wright, N. A. Stem cells and their implications for colorectal cancer. *Nat. Rev. Gastroenterol Hepatol.* **8**, 90–100 (2011).
37. Currie, E., Schulze, A., Zechner, R., Walther, T. C. & Farese, R. V. Cellular fatty acid metabolism and cancer. *Cell Metab.* **18**, 153–161 (2013).
38. Miess, H. *et al.* The glutathione redox system is essential to prevent ferroptosis caused by impaired lipid metabolism in clear cell renal cell carcinoma. *Oncogene* **37**, 5435–5450 (2018).
39. Shen, D. *et al.* E2F1 promotes proliferation and metastasis of clear cell renal cell carcinoma via activation of SREBP1-dependent fatty acid biosynthesis. *Cancer Lett.* **514**, 48–62 (2021).
40. Deng, Y. *et al.* Identification and experimental validation of a tumor-infiltrating lymphocytes-related long noncoding RNA signature for prognosis of clear cell renal cell carcinoma. *Front Immunol.* **13**, 1046790 (2022).
41. Faranda, T. *et al.* Differential expression profiling of long non-coding RNA GAS5 and miR-126-3p in human cancer cells in response to sorafenib. *Sci. Rep.* **9**, 9118 (2019).
42. Xue, S. *et al.* Identification of a combined lncRNA prognostic signature and knockdown of lncRNA MANCR to inhibit progression of clear cell renal cell carcinoma by bioinformatics analysis. *Transl. Androl. Urol.* **11**, 1304–1317 (2022).
43. Zhou, Z. *et al.* Identification and Validation of a Ferroptosis-Related Long Non-Coding RNA (FRlncRNA) Signature to Predict Survival Outcomes and the Immune Microenvironment in Patients With Clear Cell Renal Cell Carcinoma. *Front Genet.* **13**, 787884 (2022).
44. Sun, Z., Jing, C., Xiao, C. & Li, T. Long Non-Coding RNA Profile Study Identifies an Immune-Related lncRNA Prognostic Signature for Kidney Renal Clear Cell Carcinoma. *Front Oncol.* **10**, 1430 (2020).

45. Miska, J. *et al.* HIF-1 α Is a Metabolic Switch between Glycolytic-Driven Migration and Oxidative Phosphorylation-Driven Immunosuppression of Tregs in Glioblastoma. *Cell Rep.* **27**, 226–237.e4 (2019).
46. Duan, J. *et al.* A nomogram-based immunoprofile predicts overall survival for previously untreated patients with esophageal squamous cell carcinoma after esophagectomy. *J. Immunother Cancer* **6**, 100 (2018).
47. Gao, Y. *et al.* Prognostic value of tumor-infiltrating lymphocytes in esophageal cancer: an updated meta-analysis of 30 studies with 5,122 patients. *Ann. Transl. Med.* **8**, 822 (2020).
48. Murakami, T. *et al.* Multiplexed single-cell pathology reveals the association of CD8 T-cell heterogeneity with prognostic outcomes in renal cell carcinoma. *Cancer Immunol. Immunother.* **70**, 3001–3013 (2021).
49. Ryder, M., Ghossein, R. A., Ricarte-Filho, J. C. M., Knauf, J. A. & Fagin, J. A. Increased density of tumor-associated macrophages is associated with decreased survival in advanced thyroid cancer. *Endocr. Relat. Cancer* **15**, 1069–1074 (2008).
50. Ayers, M. *et al.* IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017).
51. Hellmann, M. D. *et al.* Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell* **33**, 843–852.e4 (2018).
52. Li, W., Gao, L.-N., Song, P.-P. & You, C.-G. Development and validation of a RNA binding protein-associated prognostic model for lung adenocarcinoma. *Aging (Albany NY)* **12**, 3558–3573 (2020).
53. Lou, Y. *et al.* Characterization of transcriptional modules related to fibrosing-NAFLD progression. *Sci. Rep.* **7**, 4748 (2017).
54. Kim, S., Kang, D., Huo, Z., Park, Y. & Tseng, G. C. Meta-analytic principal component analysis in integrative omics application. *Bioinformatics* **34**, 1321–1328 (2018).
55. Li, Z., Safo, S. E. & Long, Q. Incorporating biological information in sparse principal component analysis with application to genomic data. *BMC Bioinformatics* **18**, 332 (2017).
56. Heagerty, P. J. & Zheng, Y. Survival model predictive accuracy and ROC curves. *Biometrics* **61**, 92–105 (2005).
57. Iorio, F. *et al.* A Landscape of Pharmacogenomic Interactions in Cancer. *Cell* **166**, 740–754 (2016).
58. Maeser, D., Gruener, R. F. & Huang, R. S. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief Bioinform.* **22**, bbab260 (2021).
59. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. *Cell* **160**, 48–61 (2015).

Acknowledgements

We are grateful to the Second Affiliated Hospital of Nantong University for their support of this work and the public TCGA database. The present study was supported by the Natural Science Foundation of Jiangsu (BE2017682), Medical Research Project of Nantong Health Commission (MS2022017), Jiangsu Health Commission (Z2022014), Nantong Science and Technology Bureau (MS22019009), Youth Project of Health Commission of Nantong City (QN2022017), and Basic Research and Social Minsheng Plan Project (JC12022008).

Author contributions

C.S. and B.Z. envisioned the study. W.X., J.J., R.P. collected the data. X.F.C., Z.C., Q.J., and Y.Z. checked the data. J.J., Y.H.F. and W.J.Z. analyzed the data. C.S. wrote the manuscript. B.Z. and X.X.F. reviewed and revised the manuscript. All authors read and approved the final manuscript. The authors declare that they consent to publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-34027-9>.

Correspondence and requests for materials should be addressed to X.F. or B.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023