Biochemistry and Biophysics Reports 35 (2023) 101506

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

journal homepage: www.elsevier.com/locate/bbrep

FAM72 family proteins as poor prognostic markers in clear cell renal carcinoma

Hui Gou [a,1], Ping Chen [c,1], Wenbing Wu [b,*]

[a] Department of Pharmacy, The Affiliated Hospital of Southwest Medical University, Luzhou, 646000, China
[b] Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Southwest Medical University, Luzhou, 646000, China
[c] Department of Pharmacy, Suining Central Hospital, Suining, 629000, China

ARTICLE INFO

Keywords:
FAM72
Clear cell renal carcinoma
TCGA
Mmune infiltration
Prognosis

ABSTRACT

Purpose: This study aimed to investigate the prognostic significance of the Family with Sequence Similarity 72 member (FAM72) gene family in clear cell renal carcinoma (ccRCC) using a bioinformatic approach.
Patients and methods: To investigate the association between FAM72 and ccRCC, we utilized various databases and analysis tools, including TCGA, GEPIA, Metscape, cBioPortal, and MethSurv. We conducted an analysis of FAM72 expression levels in ccRCC tissues compared to normal kidney tissues and performed univariate and multivariate Cox analysis to determine the relationship between FAM72 expression and patient prognosis. Furthermore, we carried out Gene Ontology (GO) and Gene Set Enrichment Analysis (GSEA) to identify enriched biological processes associated with FAM72 expression. Additionally, we analyzed immune cell infiltration and the level of methylation in ccRCC patients. Our bioinformatic analysis revealed that FAM72 expression levels were significantly higher in ccRCC tissues than in normal kidney tissues. High expression of FAM72 was associated with poor prognosis in ccRCC patients and was found to be an independent prognostic factor for ccRCC. GO and GSEA analyses indicated that FAM72 was enriched in biological processes related to mitosis, cell cycle, and DNA metabolism. Moreover, we found a significant correlation between FAM72 and immune cell infiltration and the level of methylation in ccRCC patients.
Conclusion: Our findings suggest that FAM72 could serve as an unfavorable prognostic molecular marker for ccRCC. A comprehensive understanding of FAM72 could provide crucial insights into tumor progression and prognosis.

1. Introduction

Renal cancer is a widespread urologic disease, with more than 400,000 new cases diagnosed annually [1]. Surgical resection is effective in treating early-stage renal cancer, but approximately 20%–30% of cases will relapse and metastasize post-surgery. Currently, metastatic renal cell carcinoma (RCC) is treated with tyrosine kinase inhibitors (TKIs), mTOR inhibitors, and immune checkpoint inhibitors (ICIs). Clear cell renal cell carcinoma (ccRCC), which is characterized by having an immunogenic tumor microenvironment (TME) with numerous immune cells [2], provides new opportunities for precision targeted therapy and immunotherapy as a novel approach for ccRCC treatment.

FAM72 is a gene family that is specific to neural stem cells. In humans, FAM72 is composed of four human-specific paralogs (A-D) that are associated with the Slit-Robo Rho GTPase activating protein 2 (SRGAP2) paralog on chromosome 1 [3]. FAM72 has been shown to have potential carcinogenic effects [4] and serves as an independent prognostic factor for several cancers, including glioblastoma multiforme [5,6], adrenocortical carcinoma [7], and lung adenocarcinoma [8]. FAM72A is associated with the occurrence, development, and prognosis of non-neural tissue tumors, such as colon cancer, breast cancer, and lung cancer, uterine cancer, ovarian cancer [9–11]. FAM72B is significantly linked to the prognosis and immune status of prostate cancer and lung adenocarcinoma [12–14]. FAM72C exhibits good predictive capability in the early diagnosis of various non-advanced precancerous lesions of inflammatory bowel disease and colorectal cancer [15,16]. FAM72D serves as a predictor of clear renal cell carcinoma and multiple myeloma and regulates cell proliferation in tumors [17,18]. These findings suggest that FAM72 could serve as a predictive biomarker for

* Corresponding author. No. 1, Section 1, Xianglin Road, Longma Tan District, Luzhou, Sichuan Province, China.
E-mail address: wuwenbing5@vip.163.com (W. Wu).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.bbrep.2023.101506
Received 22 April 2023; Received in revised form 8 June 2023; Accepted 24 June 2023

Abbreviations

FAM72	Family with sequence similarity 72 member
NSCs	neural stem cells; ccRCC, clear cell renal carcinoma
TKIs	tyrosine kinase inhibitors
ICIs	immune checkpoint inhibitors
RCC	renal cell carcinoma
TME	tumor microenvironment
SRGAP2	Slit-Robo Rho GTPase activating protein 2
TCGA	The Cancer Genome Atlas
KIRC	Kidney Clear Cell Carcinoma
TPM	transcripts per million reads
ROC	receiver operating characteristic curve
GBM	Glioblastoma multiforme
KM	Kaplan-Meier
DEG	differentially expressed genes
OS	overall survival
GO	Gene Ontology
ssGSEA	single-sample gene set enrichment analysis
GBM	Glioblastoma multiforme
CNS	central nervous system

Table 1
TCGA clear cell renal carcinoma patient characteristics.

Characteristic	Levels	Overall (539)	Percentage (%)
Gender	Female	186	34.5
	Male	353	65.5
Age	≤60	269	49.9
	>60	270	50.1
Race	Asian	8	1.5
	Black or African American	57	10.7
	White	467	87.8
T stage	T1	278	51.6
	T2	71	13.2
	T3	179	33.2
	T4	11	2
N stage	N0	241	93.8
	N1	16	6.2
M stage	M0	428	84.6
	M1	78	15.4
Pathologic stage	Stage I	272	50.7
	Stage II	59	11
	Stage III	123	22.9
	Stage IV	82	15.3
Histologic grade	G1	14	2.6
	G2	235	44.3
	G3	207	39
	G4	75	14.1
PFI event	Alive	378	70.1
	Dead	161	29.9
OS event	Alive	366	67.9
	Dead	173	32.1
FAM72A expression	Low	269	49.9
	High	270	50.1
FAM72B expression	Low	269	49.9
	High	270	50.1
FAM72C expression	Low	269	49.9
	High	270	50.1
FAM72D expression	Low	269	49.9
	High	270	50.1

cancer prognosis. However, the mechanism of FAM72 in tumorigenesis and development remains poorly understood, and it has not yet been identified as a prognostic factor for ccRCC.

This study is the first to investigate the function of FAM72 in ccRCC. We explored the role of FAM72 in evaluating the prognosis of ccRCC patients using survival analysis, univariate and multivariate Cox regression analysis, and a nomogram-based prognostic analysis, which demonstrated FAM72's good predictive performance. Gene ontology analysis and enrichment analysis provided insights into understanding the biological mechanisms of FAM72, while the correlation analysis between FAM72 and immune infiltration and methylation provided a theoretical basis for exploring the predictive value and potential mechanism of FAM72 in ccRCC prognosis.

2. Material and methods

2.1. Data source

The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) is a data portal that is part of the Cancer Genome Project, a large-scale initiative providing researchers with access to clinical and pathological information on 33 types of cancer. For this study, we acquired clinical and pathological data on patients with ccRCC, along with RNA-Seq expression data for 539 tumor samples and 72 paracancerous tissues, from the TCGA database. Using this data, we analyzed the expression of FAM72A-D and investigated the relationship between this gene and ccRCC prognosis and immune infiltration.

2.2. Expression analysis

To analyze the RNAseq data from the TCGA Kidney Clear Cell Carcinoma (KIRC) project, which was originally in the level 3 HTSeq-FPKM (Fragments Per Kilobase per Million) format, we converted it into TPM (Transcripts Per Million reads) format and applied a log2 transformation. We conducted statistical analysis using the Wilcoxon rank sum test, a non-parametric test suitable for analyzing small sample sizes, and visualized the results using the ggplot2 package of the R language (version 3.6.3), a widely used tool for producing high-quality graphics in scientific research.

2.3. Receiver operating characteristic curve analysis

We performed receiver operating characteristic (ROC) analysis to evaluate the predictive value of FAM72 expression for clinical outcomes. ROC analysis is a statistical method commonly used to assess the diagnostic accuracy of continuous variables based on their sensitivity and specificity. We used the pROC package (version 1.17.0.1) to conduct the analysis and the ggplot2 package to visualize the results.

2.4. Gene correlation analysis

To investigate the expression correlation between the FAM72A-D gene and other genes, we calculated the Pearson correlation coefficient between them using the R language (version 3.6.3). We corrected the p-values using the Benjamini-Hochberg (BH) method and used the resulting Pearson correlation coefficient values to create a correlation heat map. We utilized the ggplot2 package to generate the heat map.

2.5. Survival analysis

According to the average expression level of FAM72A-D genes, the patients were divided into FAM72 high expression group and FAM72 low expression group. To investigate the effect of FAM72 expression on the clinical prognosis of ccRCC patients, we used the R language survival package (version 3.2-10) and the Survminer package (version 0.4.9) to perform statistical analysis and run visualization on survival data. Plot

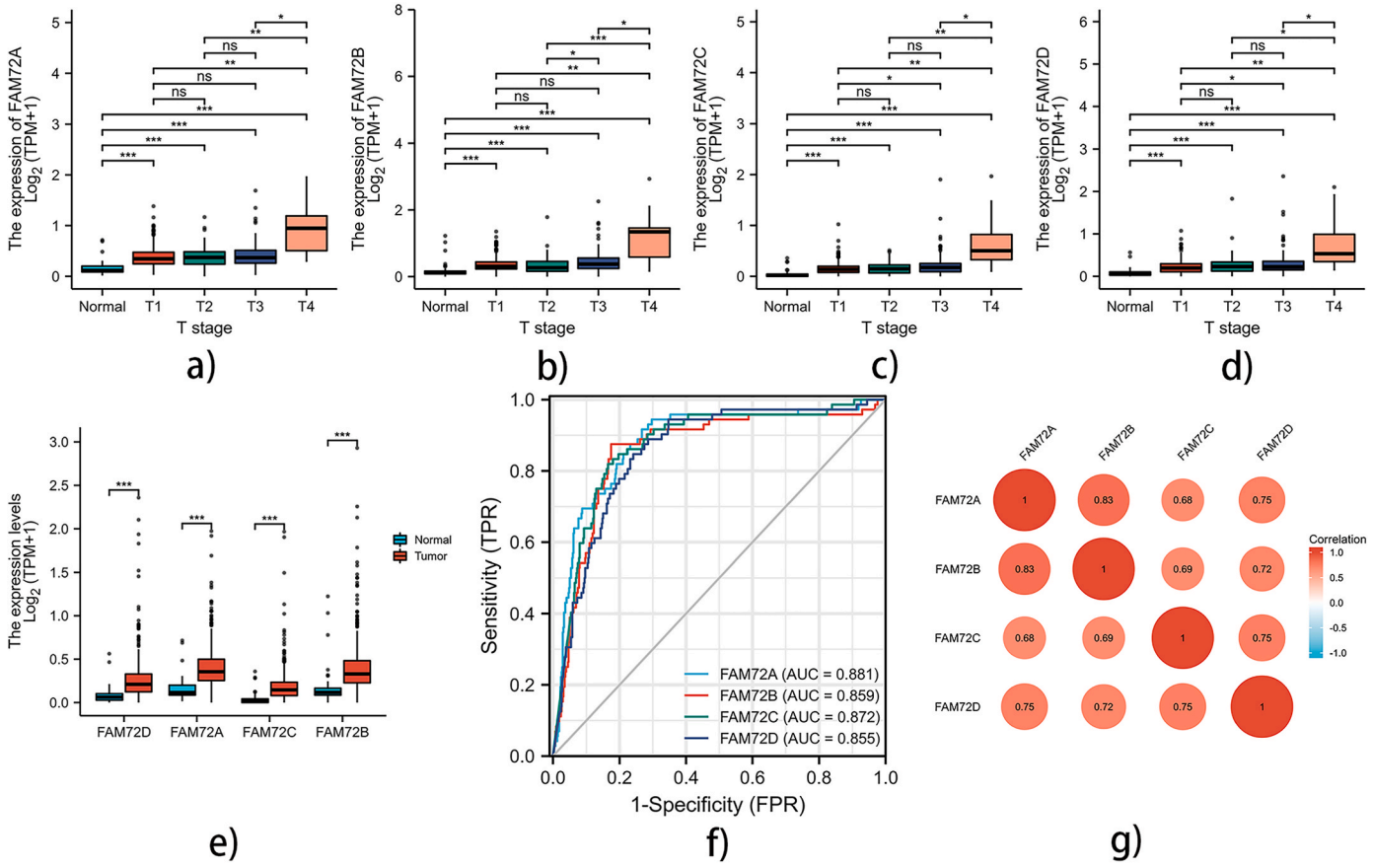

Fig. 1. Expression levels of FAM72A-D in ccRCC from TCGA data. (A–E) The expression levels of FAM72A-D in ccRCC and normal tissue. (F) Receiver operating characteristic analysis (ROC) of FAM72A-D in ccRCC. (G) Correlation among FAM72A-D members. *(*P < 0.05, **P < 0.01, ***P < 0.001)*.

Kaplan-Meier (KM) survival curves to compare differences.

2.6. Univariate and multivariate logit models and nomgram analysis

To further investigate the impact of FAM72 expression on ccRCC patient outcomes, we conducted univariate Cox regression analysis to explore the association between FAM72 gene expression levels and patient overall survival (OS). We also performed multivariate analysis to determine whether FAM72 was an independent prognostic factor for survival in ccRCC patients. We used the Cox regression module of the survival package (version 3.2-10) for statistical analysis of survival data, with FAM72 considered statistically significant when the p-value was less than 0.05. Moreover, we constructed nomograms predicting 1-year, 3-year, and 5-year survival probabilities by incorporating significant factors from the multivariate analysis, using the rms and survival packages in R software. To assess the goodness-of-fit of the Cox regression model against actual outcomes, we generated a prognosis calibration chart using the rms package (version 6.2-0) for both statistical analysis and visualization.

2.7. Function enrichment analysis

We extracted genes with correlation coefficient $|r| \geq 0.5$ and $p < 0.05$ with FAM72A-D expression, and performed functional enrichment analysis to find possible biological pathways related to FAM72A-D. R language (version 3.6.3) was used for statistical analysis and data visualization. We used Metscape (http://Metascape.org) for Gene Ontology (GO) enrichment analysis and the ClusterProfiler package (version 3.14.3) for KEGG enrichment analysis.

2.8. Gene set enrichment analysis

To divide the four FAM72 genes into high and low expression groups, we used the method described by Love et al. [19]. We then utilized the DESeq2 package in R language software to analyze the differential expression of individual genes and identify differentially expressed genes (DEGs) between the two groups. We subjected the resulting DEGs to enrichment analysis using the clusterProfiler package in R language software. For the enrichment analysis, we utilized predefined gene sets sourced from the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp).

2.9. Immune infiltration analysis

We estimated the level of immune infiltration in the tissue using transcriptome data and an algorithm that infers the fraction of immune cells present. For this analysis, we utilized the SVA package (version 1.34.0), with single-sample gene set enrichment analysis (ssGSEA) - the built-in algorithm of the GSVA package - as the immune infiltration algorithm. The classification of the 24 types of immune cells was based on the study by Bindea et al. [20]. Additionally, we investigated the correlation between FAM72 gene expression and immune checkpoint

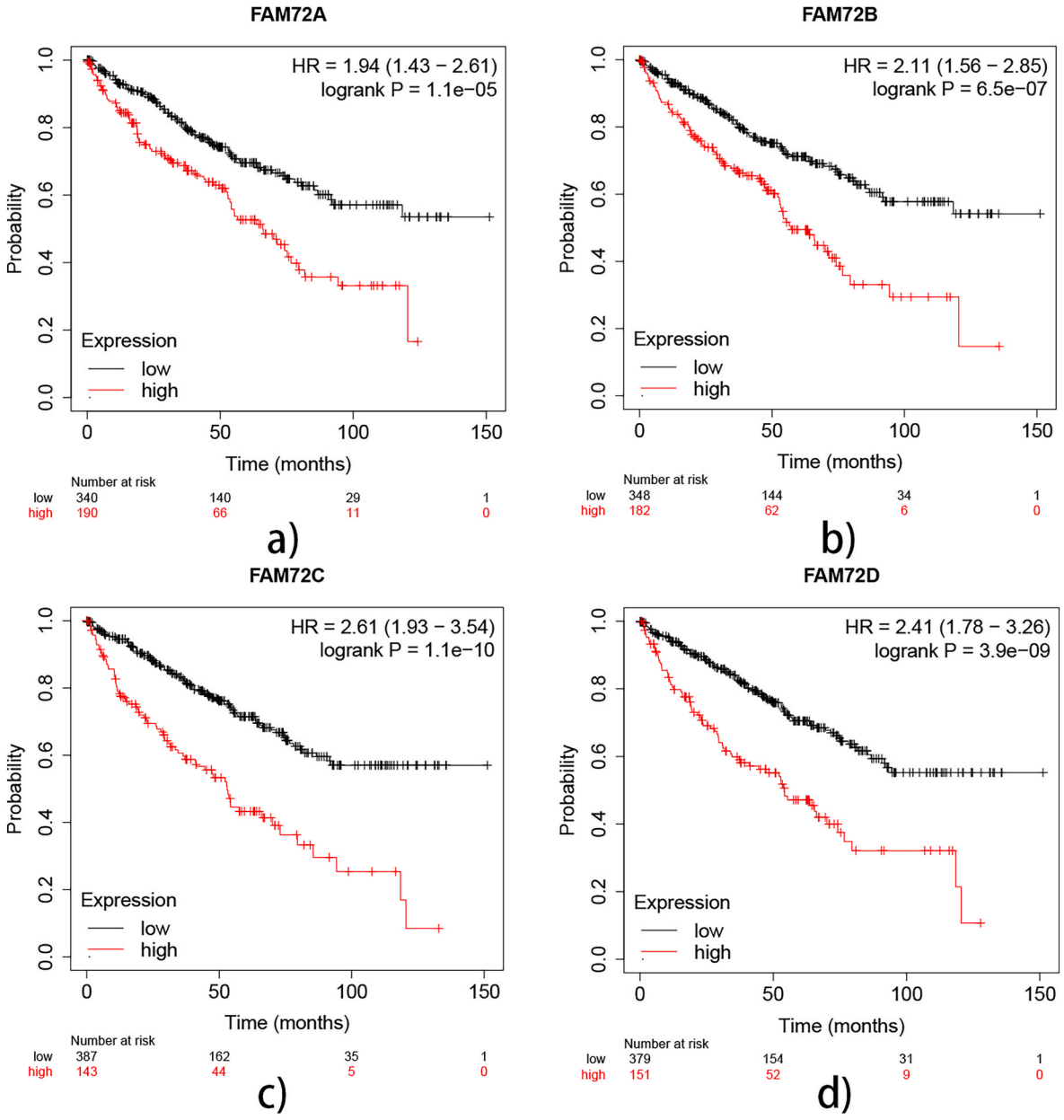

Fig. 2. The prognostic value of FAM72A-D expression in ccRCC. (A–D) Survival curves of OS from TCGA data (n = 530).

Table 2
Univariate and multivariate Cox regression model of prognosis for FAM72A-D in patients with clear cell renal carcinoma.

Characteristics	Total(N)	HR(95% CI)	P value
FAM72A			
Univariate analysis			
Gender (Male vs. Female)	539	0.930 (0.682-1.268)	0.648
Age (>60 vs.≤60)	539	1.765 (1.298-2.398)	<0.001****
Race (Other vs. White)	532	1.222 (0.678-2.201)	0.505
T stage (T2/T3/T4 vs. T1)	539	2.917 (2.095-4.061)	<0.001****
N stage (N1 vs. N0)	257	3.453 (1.832-6.508)	<0.001****
M stage (M1 vs. M0)	506	4.389 (3.212-5.999)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	3.299 (2.342-4.648)	<0.001****
FAM72A (High vs. Low)	539	1.343 (0.995-1.812)	0.054
Multivariate analysis			
Age (>60 vs.≤60)	539	1.949 (1.271-2.990)	0.002***
T stage (T2/T3/T4 vs. T1)	539	0.671 (0.188-2.401)	0.54
N stage (N1 vs. N0)	257	2.208 (1.089-4.477)	0.028*
M stage (M1 vs. M0)	506	3.243 (2.002-5.254)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	2.573 (0.670-9.881)	0.169
FAM72A (High vs. Low)	539	1.518 (0.989-2.331)	0.056
FAM72B			
Univariate analysis			
Gender (Male vs. Female)	539	0.930 (0.682-1.268)	0.648
Age (>60 vs.≤60)	539	1.765 (1.298-2.398)	<0.001****
Race (Other vs. White)	532	1.222 (0.678-2.201)	0.505
T stage (T2/T3/T4 vs. T1)	539	2.917 (2.095-4.061)	<0.001****
N stage (N1 vs. N0)	257	3.453 (1.832-6.508)	<0.001****
M stage (M1 vs. M0)	506	4.389 (3.212-5.999)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	3.299 (2.342-4.648)	<0.001****
FAM72B (High vs. Low)	539	1.724 (1.269-2.342)	<0.001****
Multivariate analysis			
Age (>60 vs.≤60)	539	1.901 (1.242-2.909)	0.003***
T stage (T2/T3/T4 vs. T1)	539	0.585 (0.168-2.046)	0.402
N stage (N1 vs. N0)	257	1.810 (0.894-3.666)	0.099
M stage (M1 vs. M0)	506	3.078 (1.896-4.998)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	2.977 (0.797-11.113)	0.105
FAM72B (High vs. Low)	539	1.764 (1.132-2.749)	0.012*
FAM72C			
Univariate analysis			
Gender (Male vs. Female)	539	0.930 (0.682-1.268)	0.648
Age (>60 vs.≤60)	539	1.765 (1.298-2.398)	<0.001****
Race (Other vs. White)	532	1.222 (0.678-2.201)	0.505
T stage (T2/T3/T4 vs. T1)	539	2.917 (2.095-4.061)	<0.001****
N stage (N1 vs. N0)	257	3.453 (1.832-6.508)	<0.001****
M stage (M1 vs. M0)	506	4.389 (3.212-5.999)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	3.299 (2.342-4.648)	<0.001****
FAM72C (High vs. Low)	539	1.601 (1.181-2.170)	0.002***
Multivariate analysis			
Age (>60 vs.≤60)	539	1.915 (1.250-2.934)	0.003***
T stage (T2/T3/T4 vs. T1)	539	0.574 (0.163-2.024)	0.388
N stage (N1 vs. N0)	257	2.083 (1.032-4.205)	0.041*
M stage (M1 vs. M0)	506	3.282 (2.021-5.329)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	3.056 (0.809-11.540)	0.099
FAM72C (High vs. Low)	539	1.196 (0.790-1.812)	0.397
FAM72D			
Univariate analysis			
Gender (Male vs. Female)	539	0.930 (0.682-1.268)	0.648
Age (>60 vs.≤60)	539	1.765 (1.298-2.398)	<0.001****
Race (Other vs. White)	532	1.222 (0.678-2.201)	0.505
T stage (T2/T3/T4 vs. T1)	539	2.917 (2.095-4.061)	<0.001****
N stage (N1 vs. N0)	257	3.453 (1.832-6.508)	<0.001****
M stage (M1 vs. M0)	506	4.389 (3.212-5.999)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	3.299 (2.342-4.648)	<0.001****
FAM72D (High vs. Low)	539	1.627 (1.200-2.205)	0.002***
Multivariate analysis			
Age (>60 vs.≤60)	539	1.953 (1.275-2.992)	0.002***
T stage (T2/T3/T4 vs. T1)	539	0.657 (0.187-2.306)	0.512
N stage (N1 vs. N0)	257	2.071 (1.035-4.144)	0.04*
M stage (M1 vs. M0)	506	3.257 (2.016-5.264)	<0.001****
Pathologic stage (Stage II&III&IV vs. I)	536	2.715 (0.722-10.217)	0.14
FAM72D (High vs. Low)	539	1.500 (0.988-2.279)	0.057

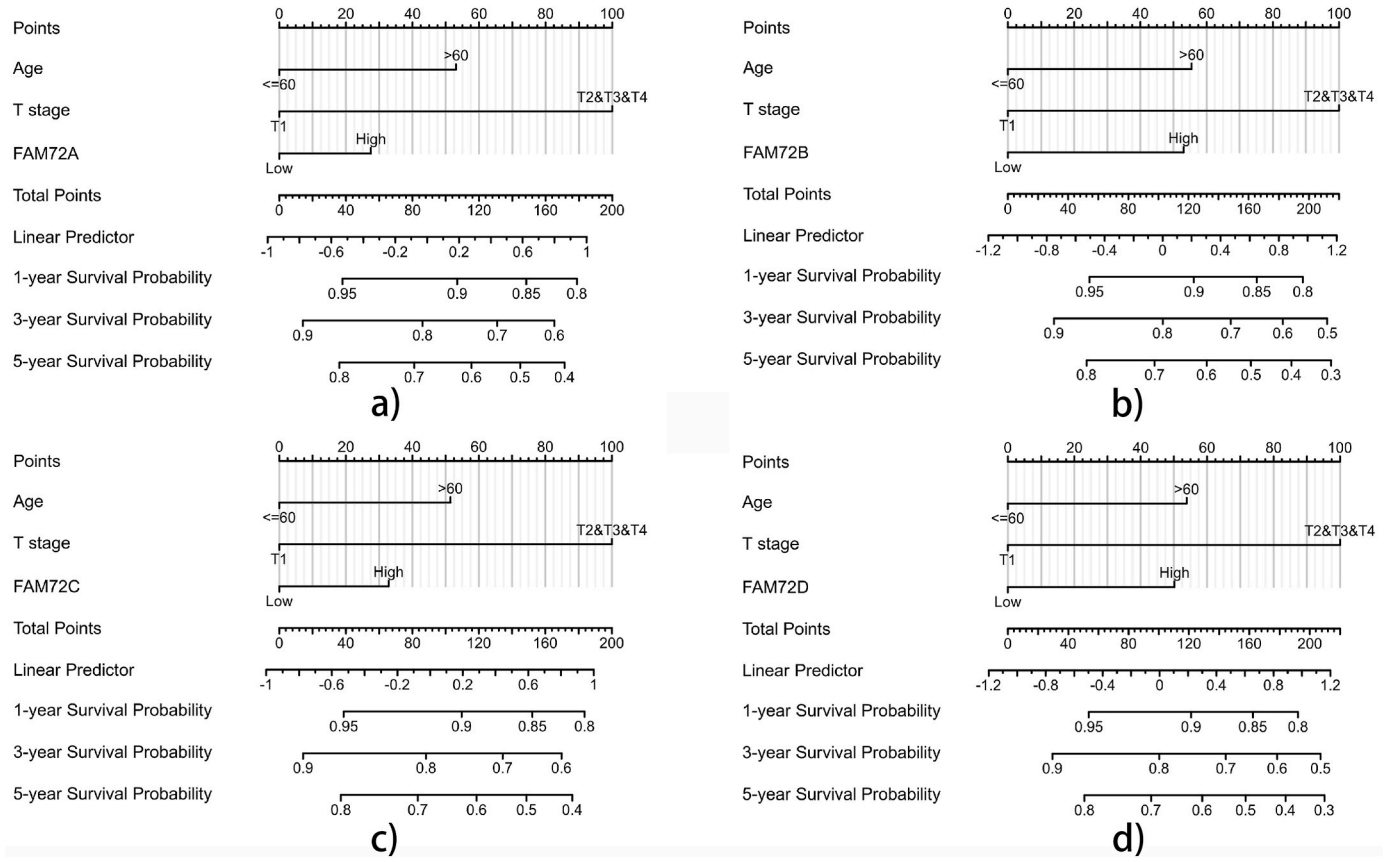

Fig. 3. Nomogram for predicting the probability of 1-, 3- and 5-year OS of ccRCC patients. (A–D) A nomogram that integrates FAM72A-D and other prognostic factors in ccRCC from TCGA data.

molecules PDCD1 (PD-1), CD274 (PD-L1), CTLA4, and CD80 (LAG-3) using the Spearman correlation coefficient method. We conducted statistical analysis and data visualization using the ggplot2 package in the R language.

2.10. FAM72 methylation analysis

Previous studies have suggested a potential correlation between FAM72 expression and methylation, with evidence indicating that the methylation status of the FAM72 promoter may influence gene expression in tumor tissues such as Glioblastoma multiforme (GBM), breast cancer, and liver cancer [6]. To investigate the relationship between FAM72 gene expression and methylation, we conducted an analysis using the MethSurv platform (https://biit.cs.ut.ee/MethSurv/). This platform enables methylation analysis of tumor samples in the TCGA database and allows for survival analysis of individual CpG methylation sites [21].

3. Results

3.1. Clinical features

We acquired gene expression and clinical data from a cohort of 539 ccRCC patients in the TCGA database, which included information on age, sex, race, tumor stage, and other relevant clinical features as shown in Table 1.

3.2. Expression of FAM72 in ccRCC tissues

We found a significant positive correlation between the expression levels of FAM72A-D and various clinical features of ccRCC, including tumor volume, depth of invasion, and extent of involvement of adjacent tissues (Fig. 1A–E). Moreover, we observed a significant increase in the expression of all four FAM72 genes in tumor tissues compared to normal tissues, with corresponding AUC values of 0.881 (CI: 837–0.926), 0.859 (CI: 0.809-0.908), 0.872 (CI: 0.828-0.916), and 0.855 (CI: 0.811-0.899) (Fig. 1F). Additionally, we identified a significant positive correlation between FAM72A-D gene expression levels, with FAM72A and FAM72B showing the highest correlation coefficient ($r = 0.83$) (Fig. 1G).

3.3. Relationship between FAM72 and OS

Patients with high expression of FAM72A-D exhibited a significantly lower overall survival rate than those with low expression ($p < 0.001$), according to Fig. 2A–D. The hazard ratios (HR) of all four FAM72 genes are more than 1, which suggests that they might act as ccRCC risk factors. To learn more about the predictive significance of FAM72 expression levels, we employed univariate Cox regression analysis. The findings revealed that higher expression levels of FAM72B, C, and D were substantially linked with lower overall survival ($p < 0.001$, $p = 0.002$, and $p = 0.002$, respectively) while FAM72A showed marginal significance ($p = 0.054$) (Table 2). In a multivariate analysis utilizing Cox regression models, FAM72B, age, and distant metastasis (M stage) were independently associated with overall survival in ccRCC patients.

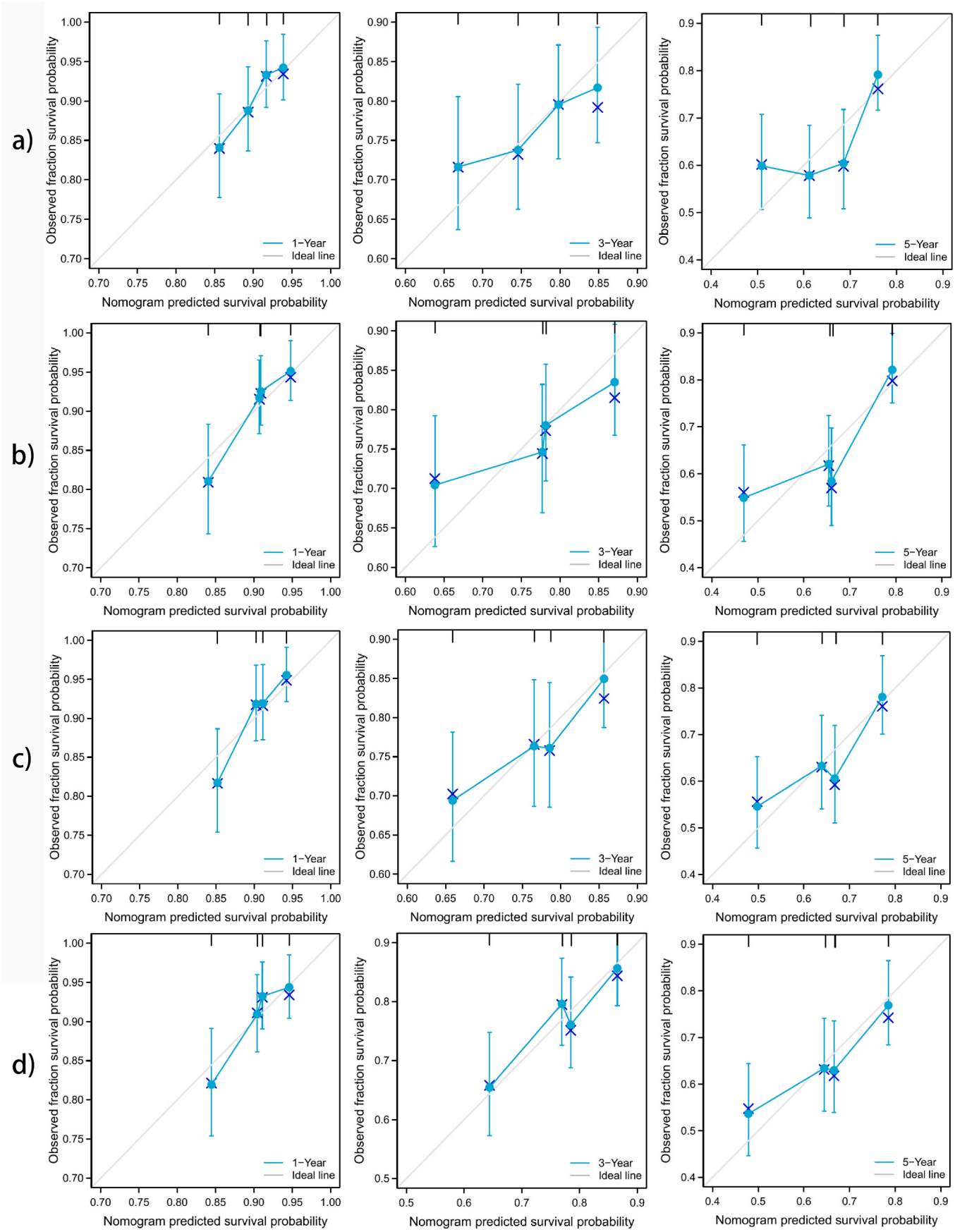

Fig. 4. Calibration curve for predicting the probability of 1-, 3- and 5-year OS for ccRCC patients. (A–D) The calibration curve of the nomogram in ccRCC from TCGA data.

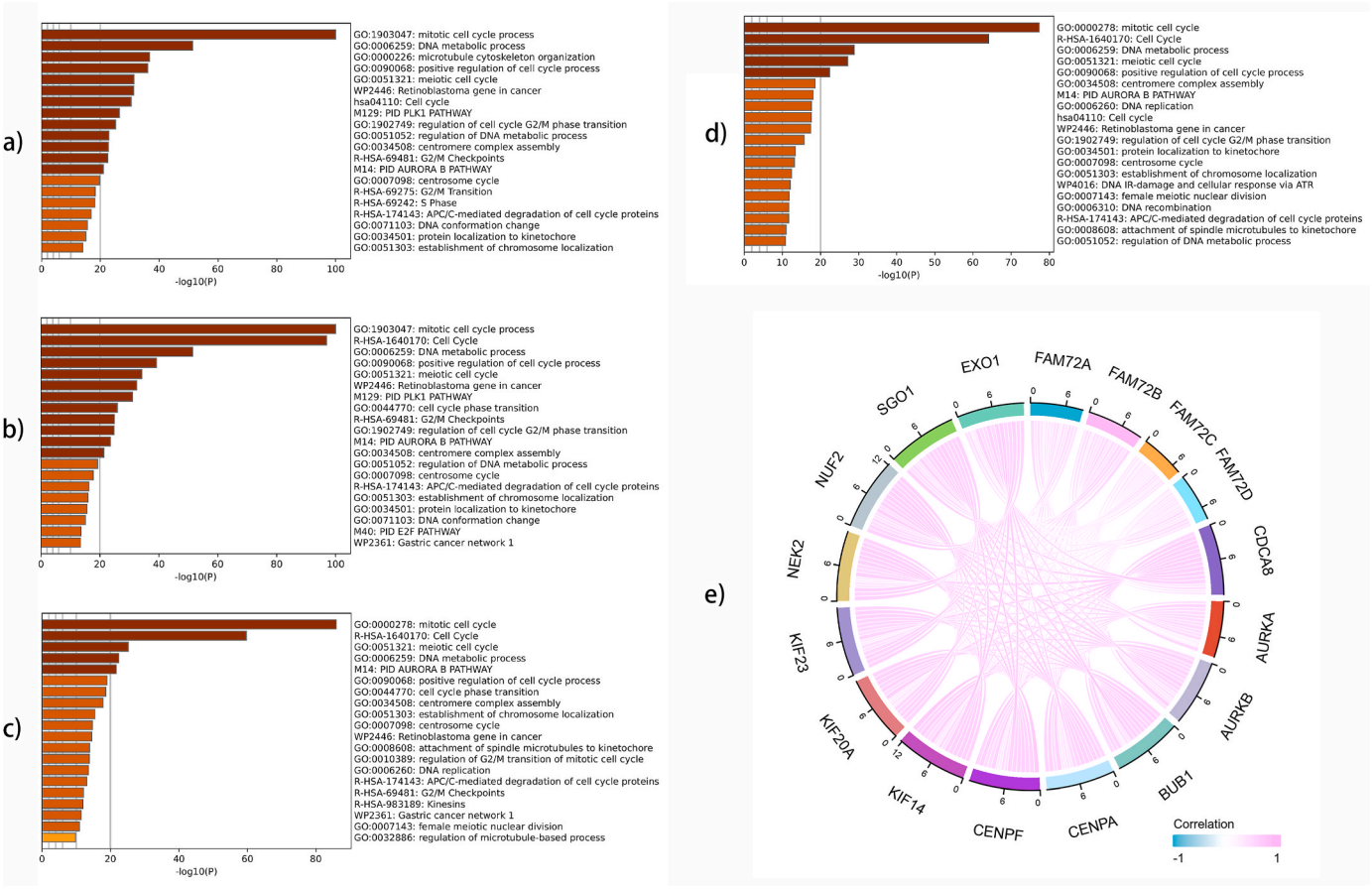

Fig. 5. Functional enrichment of FAM72A-D in ccRCC. (A–D) Gene ontology (GO) enrichment analysis of FAM72A-D and its co-expression genes in Metascape. The GO enriched terms are colored by *p*-value, where terms containing more genes tend to have more significant *p*-value. (E) Correlation between FAM72 and mitosis-related genes.

Our study shows that the overall survival rates of ccRCC patients with high expression of FAM72 are significantly lower.

3.4. Analysis of prognostic model based on FAM72 and clinical case factors

We created a nomogram (Fig. 3A–D) that combines the expression levels of FAM72A-D with separate clinical risk factors such age and pathological stage. The nomogram makes it possible to calculate total points, greater scores denote a worse prognosis. The constancy of the 45° line across several circumstances in the calibration diagram showed a high agreement between the anticipated probability and the actual probability. According to the concordance values (C-index) for FAM72A-D, which were 0.587, 0.6, 0.597, and 0.602, respectively, the model does a respectable job of accurately forecasting actual results (Fig. 4A–D).

3.5. Functional enrichment analysis of FAM72-related genes

The overall number of genes we examined included 438 for FAM72A, 413 for FAM72B, 185 for FAM72C, and 254 for FAM72D. Using GO analysis, we discovered that these genes are predominantly involved in mitotic cell cycle processes, DNA metabolic process organization, microtubule cytoskeleton, and other associated processes (Fig. 5A–D). Since the FAM72 gene is highly expressed during mitosis in renal clear cell carcinoma cells, we further analyzed the relationship between the FAM72 gene and a group of genes involved in cell division, such as genes related to chromosome segregation CDCA8, AURKA, BUB1, NUF2, and SGO1, AURKB, NEK2, centromere formation-related genes CENPA, CENPF, spindle formation-related genes KIF14, KIF20A, KIF23, DNA damage and repair-related gene EXO1 [39–42]. The results showed that these genes had a strong expression correlation with members of the FAM72 gene family (Fig. 5E). Further, KEGG pathway analysis revealed that genes associated with the FAM72 gene family were considerably enriched in a number of signaling pathways, including the cell cycle, oocyte meiosis, progesterone-mediated oocyte maturation, the Fanconi anemia pathway, and homologous recombination (Fig. 6A–D).

3.6. FAM72-related signaling pathways based on GSEA

We conducted logFC-based gene set enrichment analysis (GSEA) on more than 50,000 differentially expressed genes after single-gene differential expression analysis. According to our research, the cell cycle checkpoints and pathways for cytokine-cytokine receptor interaction were considerably enriched in the genes of FAM72A that were differently expressed. Differentially expressed FAM72B genes were linked to extracellular matrix structure and leishmaniasis. On the other hand, FAM72C's differentially expressed genes were predominately involved

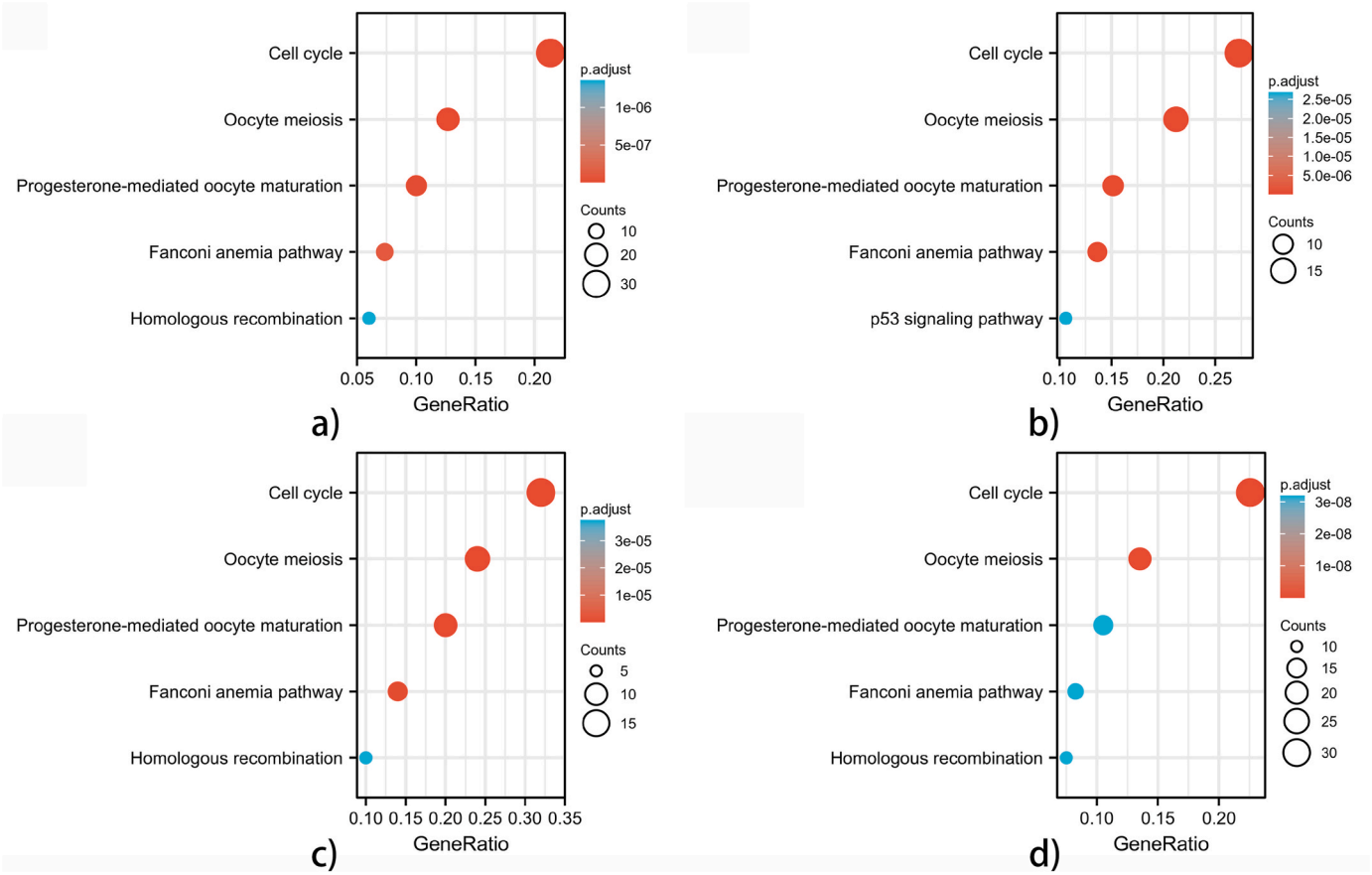

Fig. 6. The top 5 pathways were differentially enriched according to the level of DEGs in FAM72A-D related ccRCC. (A–D) The enrichment plot was obtained from the gene set enrichment analysis (GSEA).

in ion channel transport and SLC-mediated transmembrane transport pathways. Finally, differentially expressed FAM72D genes were mostly implicated in cell cycle checkpoints and an anti-inflammatory response that facilitates Leishmania parasite infection (Fig. 7A–D). Recent studie suggests that FOXM1 can control the transcriptional level of FAM72A gene expression, which in turn can influence cell cycle progression and apoptotic signal transduction [43]. We also examined the relationship between FOXM1 and the FMA72A-D gene. The findings revealed that FOXM1 was poorly connected with the expression of FAM72C ($r = 0.55$) and FAM72D ($r = 0.42$) and moderately correlated with the expression of FAM72A ($r = 0.55$) and FAM72B ($r = 0.59$) (Fig. 7E).

3.7. Correlation between FAM72 expression and immune infiltration

The FAM72 gene family significantly positively correlated with 11 different immune cell types, including T helper cells, T central memory (Tcm), T helper 1 (Th1) cells, T effector memory (Tem), Th2 cells, T cells, activated DCs (aDC), T follicular Helper cells (TFH), CD8 T cells, regulatory T cells (TReg), and cytotoxic cells, according to our immune infiltration analysis (Fig. 8A–D). The expression of immunological checkpoint molecules such as PDCD1 (PD-1), CD274 (PD-L1), CTLA4 (cytotoxic T lymphocyte antigen-4), and CD80 (lymphocyte activation gene 3, LAG-3) showed a substantial positive connection with the expression of the FAM72 gene (Fig. 9A–H).

3.8. Correlation between FAM72 gene expression and methylation

We found that the methylation level of FAM72 was significantly low, as shown in Fig. 10, and that increased gene expression may be connected with this low methylation level (Fig. 10A–C). We examined the relationship between the methylation of each CpG island and patient survival rates in order to learn more about the effect of FAM72 methylation on ccRCC patient prognosis. Our results showed that hypomethylation at various locations of the FAM72 gene was significantly linked with worse prognosis, with the exception of cg04968835 ($p = 0.4$) and cg07344025 ($p = 0.15$) (Fig. 10D).

4. Discussion

The second most common malignant tumor of the urinary system, after bladder cancer, is renal cell carcinoma. Kidney clear cell carcinoma is the most prevalent renal cell carcinoma and is hence frequently referred to as RCC. The main form of treatment for localized RCC is still surgical excision because conventional radiation and chemotherapy are ineffective against RCC. However, after surgical excision, up to 40% of localized kidney malignancies may subsequently form metastatic tumors. Due to its resistance to chemotherapy and radiotherapy as well as the absence of efficient therapeutic alternatives, metastatic RCC has a very bad prognosis [22]. The median survival time of RCC patients is still shorter than expected despite the widespread use of numerous targeted medications in their care. Therefore, it is crucial to find

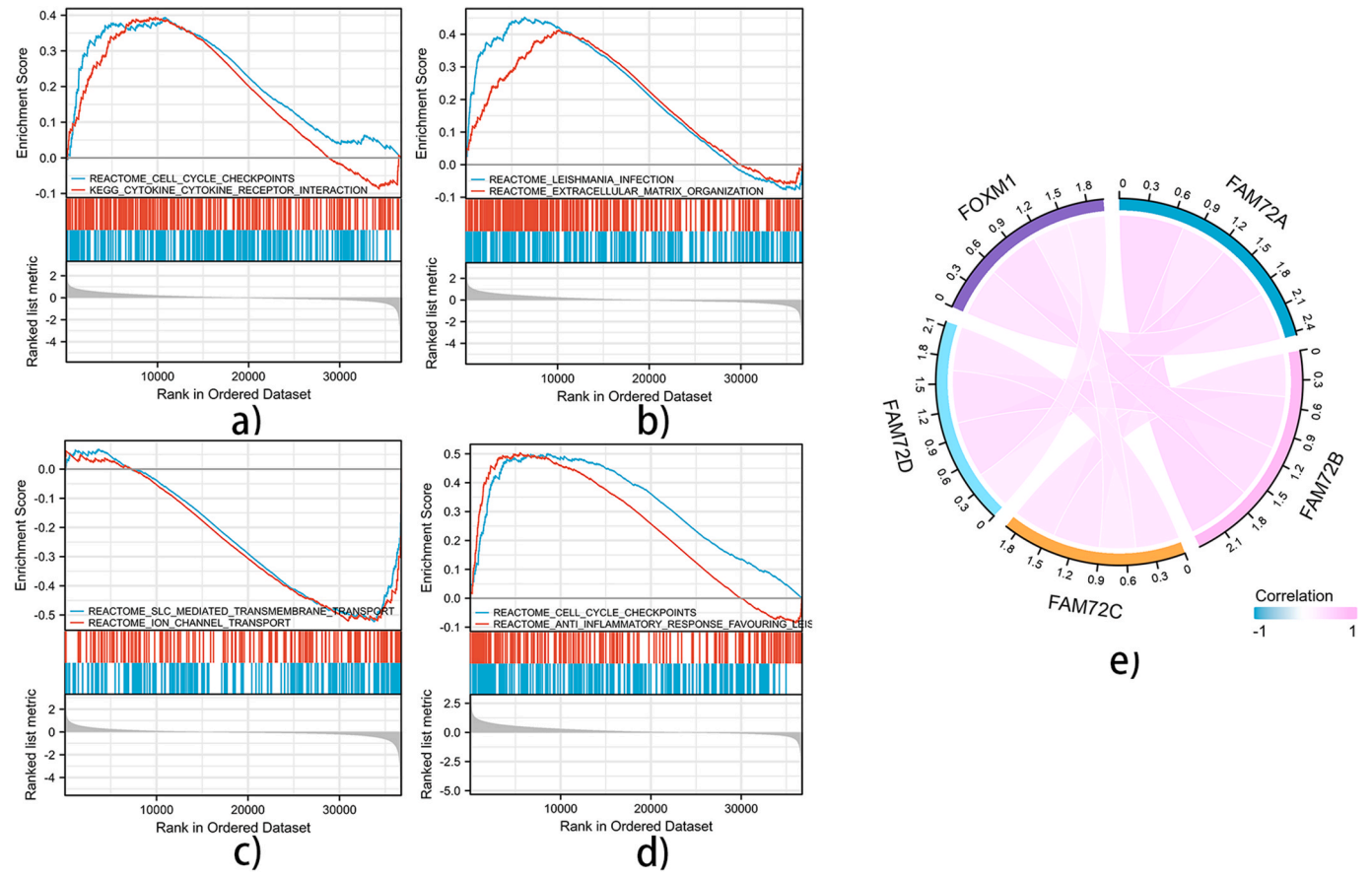

Fig. 7. The functional pathways of FAM72A-D in ccRCC. (A–D) Several pathways were enriched in FAM72A-D related ccRCC. (E) Chord chart of correlation test between FOXM1 and FAM72A-D.

treatment targets for RCC as well as markers for early identification [23].

According to our findings, FAM72 may be a useful predictive biomarker for ccRCC, with higher levels of FAM72 expression in patients generally being associated with worse prognoses, most likely as a result of hypomethylation at their CpG islands. Using the comprehensive human clinical cancer research database cBioPortal, which contains the Cancer Genome Atlas (TCGA), Rahane et al. [5] carried out a thorough investigation of FAM72 (A-D) expression and somatic mutation data in 31 tumors, including glioblastoma multiforme (GBM). They obtained information on human gene mutations from TCGA using a computerized clinical data analysis approach, and discovered that the transcription of the mitotic cell cycle gene FAM72 was associated with the expression of the proliferation marker MKI67. It was discovered that neural stem cells (NSCs) may turn into cancer stem cells, giving rise to brain tumor cells responsible for brain tumors such as GBM, if the gene transcriptional control unit, which is the intergenic region of two subgene units SRGAP2 and FAM72, is out of control. Ho et al. [6] also covered the monitoring of the |-SRGAP2-FAM72-| master gene, its function in GBM, and the prospective application of FAM72 for the diagnosis of other cancers beyond the central nervous system (CNS). These earlier discoveries imply that FAM72 might serve as an oncogene. Our research strengthened the evidence for the overexpression of the FAM72 gene in ccRCC, confirming its potential contribution to the disease's development and laying the groundwork for the use of the FAM72 gene family in the diagnosis and treatment of the disease.

Our GO analysis showed that genes associated with FAM72 are primarily involved in cell cycle processes, such as cell mitosis, meiosis, DNA synthesis, centrosome assembly, and particularly cell cycle checkpoints in the G2/M phase, which are essential for controlling cell proliferation. Retinoblastoma genes in cancers, the PID PLK1 signaling pathway, and the PID AURORA B signaling pathway were further enriched terms (Fig. 5). However, additional research is required to validate these results. Genes connected to FAM72 expression were significantly enriched in the cell cycle, homologous recombination, and other pathways strongly related to tumor initiation and progression, according to KEGG and GSEA analysis. According to Wang et al., FAM72A may control cell growth by modifying the metabolism of cellular reactive oxygen species, especially in tumors caused by the Epstein-Barr virus [24]. This suggests that FAM72A and members of its family may also encourage tumor growth and a bad prognosis through related mechanisms.

T-cell immune infiltration in tumors has been linked to a better prognosis and a better response to cancer treatments [25,26]. In addition to attracting other immune cells to the tumor location, T cells are essential for identifying and killing tumor cells [27]. Increased T-cell infiltration in tumors has been associated with better survival and treatment response in a number of human or animal studies in melanoma [28], osteosarcoma, and small cell lung cancer [29]. Our results suggest that the expression of FAM72A-D can reflect the degree of

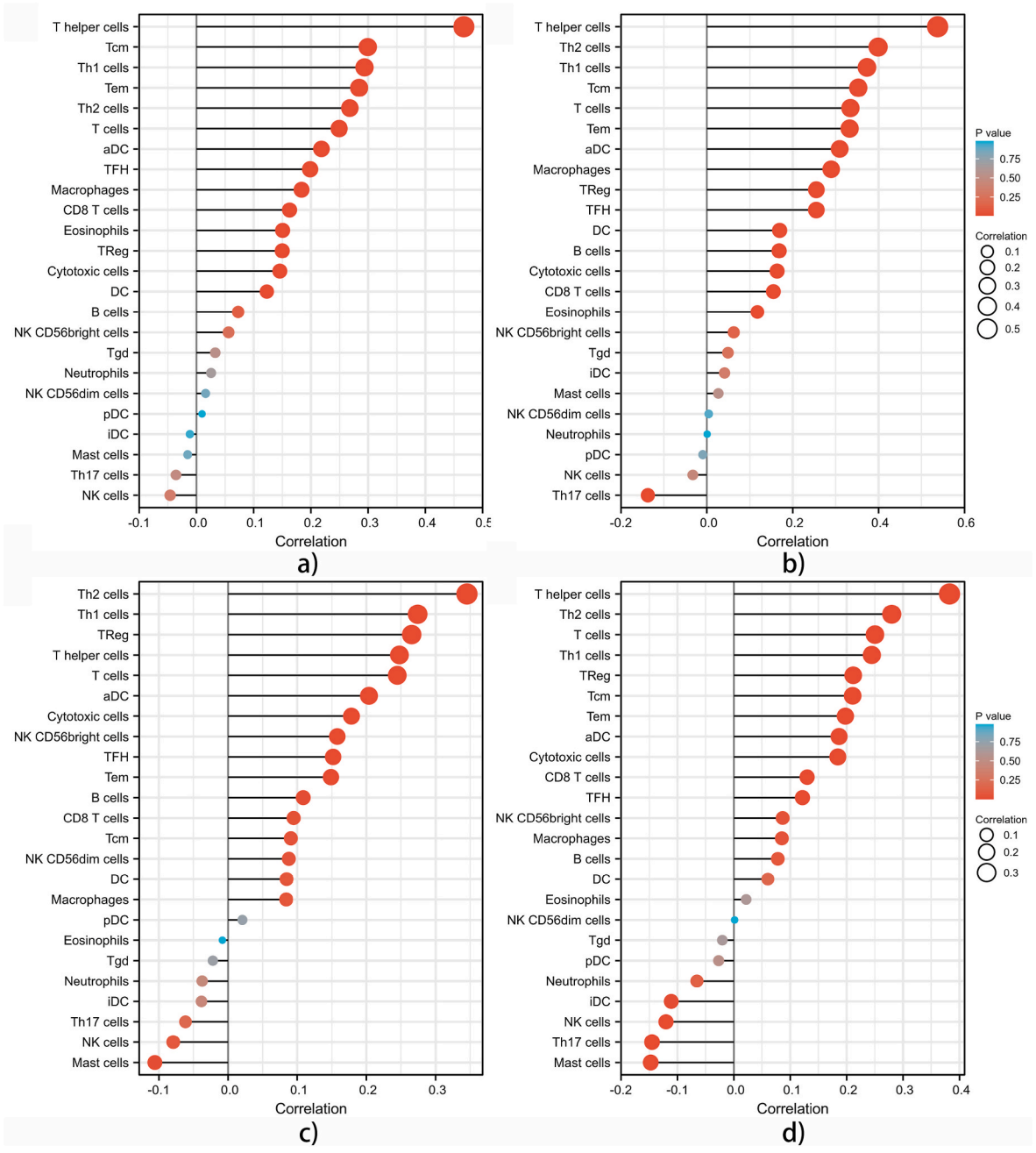

Fig. 8. FAM72A-D was positively correlated with the infiltration of most immune cells. (A–D) The association between the expression level of FAM72A-D and the immune infiltration in the tumor microenvironment.

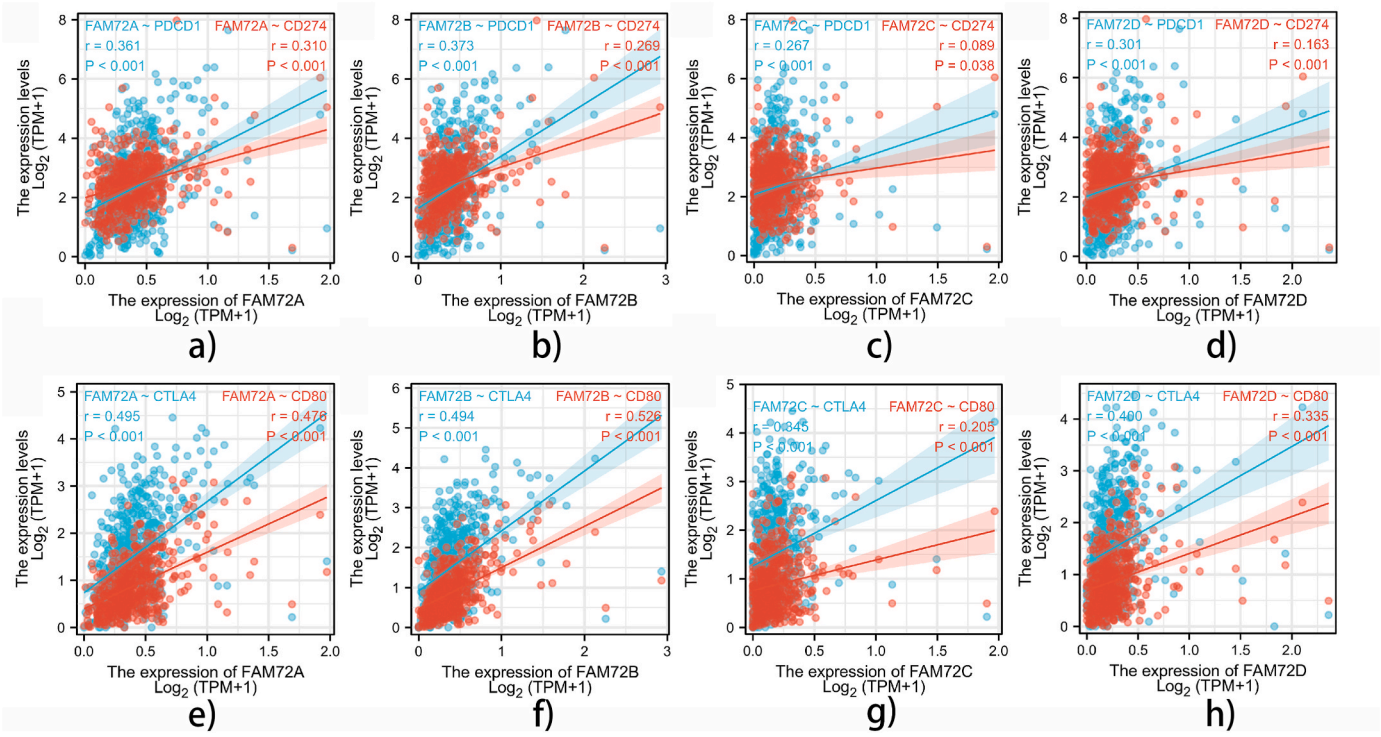

Fig. 9. FAM72A-D is positively correlated with immune checkpoint molecules. Correlation between the expression of FAM72A-D and immune checkpoint molecules PDCD1/CD274 (A–D), CTLA4/CD80 (E–H).

immune infiltration in ccRCC and offer insight into immunotherapy for ccRCC. Immune infiltration includes T cells such as T helper cells, Tcm, Tem, Treg, TFH, cytotoxic cells, CD8 T cells, and dendritic cells such as aDC. Regulatory T cells (Treg) can compromise immune surveillance in healthy people and lessen the anti-tumor immune response in tumor-bearing hosts, which is significant to keep in mind [30]. The expression of PDCD1 (PD-1), CD274 (PD-L1), CTLA4 and CD80 (LAG-3) in ccRCC also showed a favorable connection with FAM72A-D. The immune response is modified by their binding because PD-L1 is a ligand for PD-1. Recent years have seen a major increase in the attention on the PD-1/PD-L1 axis in tumor immunotherapy [31]. PD-L1+ tumor-infiltrating immune cells and high stromal $CD8^+$ TIL density are linked to high stromal PD-1+ tumor-infiltrating lymphocyte density, which frequently results in a poor prognosis [32]. There is a strong correlation between high CTLA4 levels and a poor prognosis in individuals with nasopharyngeal cancer [33] and thymoma [34]. This is because CTLA4+ tumor cells might impair DC maturation and activity. LAG-3 has been studied in a number of malignancies and has been shown to work in concert with the PD-1/PD-L1 axis [35]. High LAG-3 expression in tumor tissue has been linked in studies to a poor prognosis for HCC [36]. Our findings imply that FAM72 may be a promising predictive biomarker for the efficacy of immunotherapy.

A significant epigenetic alteration known as DNA methylation regulates gene expression throughout the onset and spread of cancer. Researchers can more fully comprehend the regulatory role of DNA methylation and make accurate prognostic predictions for tumor patients by thoroughly analyzing DNA methylation and gene expression data [37]. The evaluation of prognostic value can be improved by using multi-gene or gene family biomarkers [38]. In this study, we created a predictive model based on three FAM72 A/B/D genes that are methylation-driven. We discovered that in tumors from ccRCC patients, hypomethylation of these genes was substantially related with increased expression and a poor prognosis. As a result, the FAM72 gene-based epigenetic regulatory pattern of methylation may be used as an additional predictive reference for ccRCC patients.

Although FAM72 A-D and the prognosis of ccRCC patients have been linked by bioinformatic analysis utilizing open-access databases, additional study is required to establish its accuracy assessment and practical application. Through cell and animal investigations, we intend to explore the mechanism of the FAM72 gene family in ccRCC in our upcoming research, offering fresh insights into its potential application as a therapeutic and predictive biomarker.

Taken together, these results suggest that the high expression of FAM72 in ccRCC may be caused by the hypomethylation modification of the gene and the action of the transcription factor FOXM1, while FAM72 can affect immune cell infiltration, expression of immune checkpoint molecules and cell cycle promote the progression of clear cell renal cell carcinoma. The FAM72 family may be a potential poor prognostic molecular marker in ccRCC, and a comprehensive understanding of it can provide important insights into tumor progression and prognosis (Fig. 11).

Funding

This study was supported by Southwest Medical University (00170015) and (00160127)

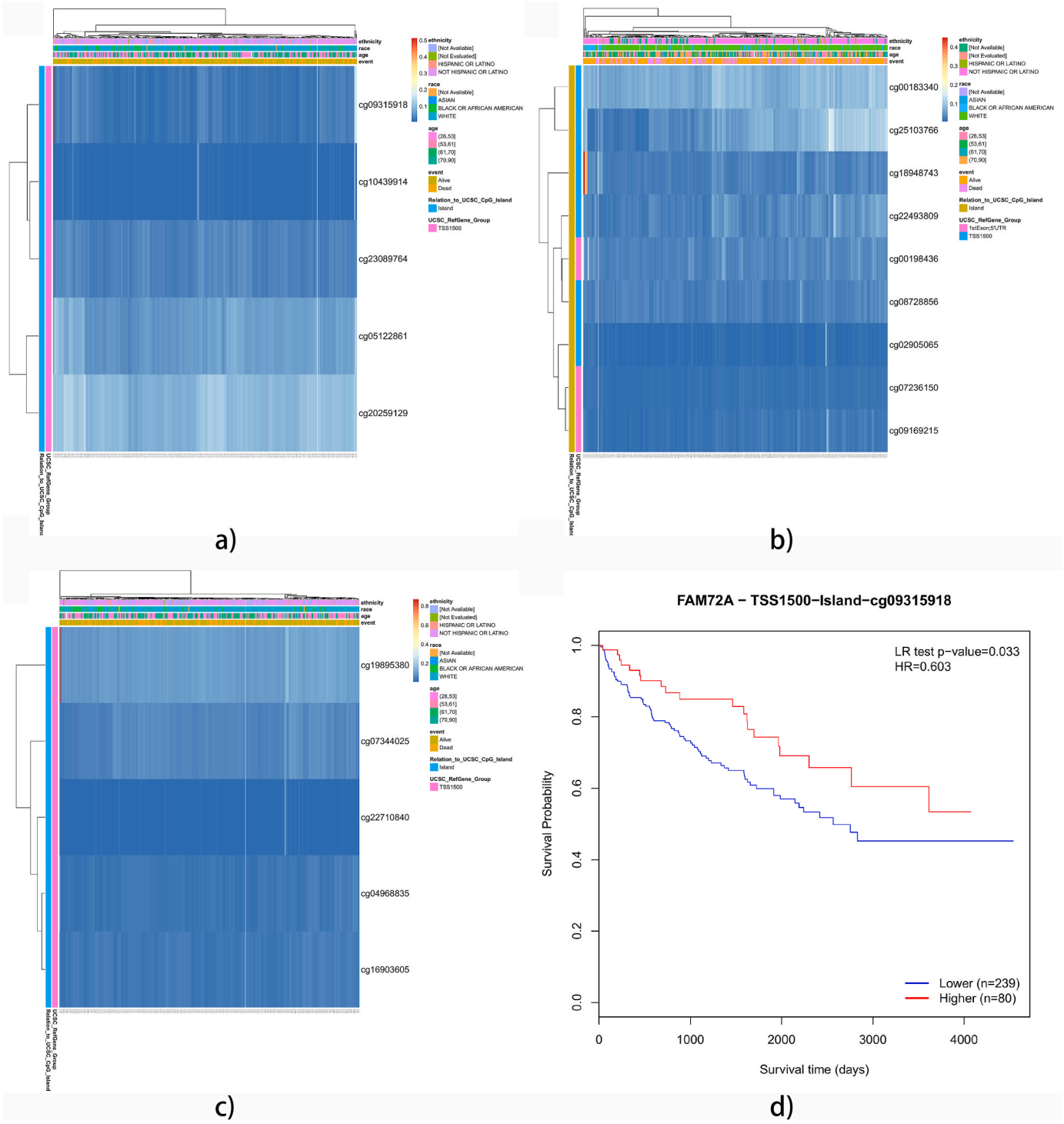

Fig. 10. The methylation of FAM72 A/B/D in ccRCC. (A–C) The visualization between the methylation level and the FAM72 A/B/D expression. (D) The Kaplan-Meier survival of the promoter methylation of FAM72A.

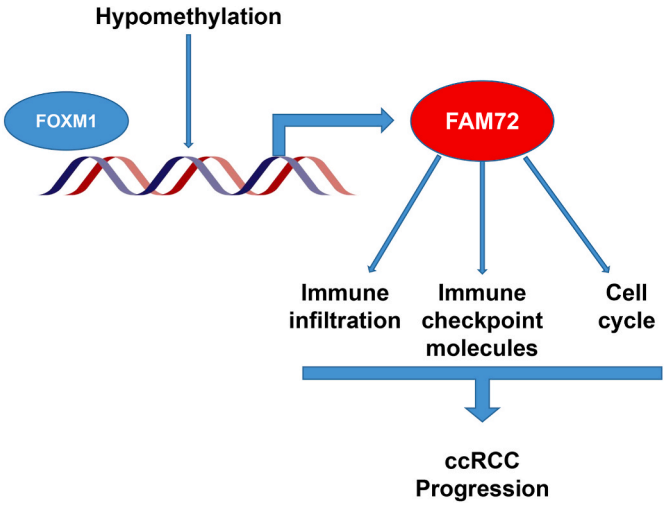

Fig. 11. Overview of the relationship between FAM72 and ccRCC progression.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

[1] J. Huang, D.K.-W. Leung, E.O.-T. Chan, V. Lok, S. Leung, I. Wong, X.-Q. Lao, Z.-J. Zheng, P.K.-F. Chiu, C.-F. Ng, J.H.-M. Wong, A. Volpe, A.S. Merseburger, T. Powles, J.Y.-C. Teoh, M.C.S. Wong, A global trend analysis of kidney cancer incidence and mortality and their associations with smoking, alcohol consumption, and metabolic syndrome, Eur. Urol. Focus. 8 (2022) 200–209, https://doi.org/10.1016/j.euf.2020.12.020.

[2] S. Allison, New insights into the ccRCC microenvironment, Nat. Rev. Nephrol. 17 (2021), https://doi.org/10.1038/s41581-021-00457-9, 511–511.

[3] A. Kutzner, S. Pramanik, P.-S. Kim, K. Heese, All-or-(N)One – an epistemological characterization of the human tumorigenic neuronal paralogous FAM72 gene loci, Genomics 106 (2015) 278–285, https://doi.org/10.1016/j.ygeno.2015.07.003.

[4] Y. Wang, J. Lin, K. Yan, J. Wang, Identification of a robust five-gene risk model in prostate cancer: a robust likelihood-based survival analysis, Int. J. Genomics. 2020 (2020), 1097602, https://doi.org/10.1155/2020/1097602.

[5] C.S. Rahane, A. Kutzner, K. Heese, A cancer tissue-specific FAM72 expression profile defines a novel glioblastoma multiform (GBM) gene-mutation signature, J. Neuro Oncol. 141 (2019) 57–70, https://doi.org/10.1007/s11060-018-03029-3.

[6] N.T.T. Ho, C.S. Rahane, S. Pramanik, P.-S. Kim, A. Kutzner, K. Heese, FAM72, glioblastoma multiforme (GBM) and beyond, Cancers 13 (2021) 1025, https://doi.org/10.3390/cancers13051025.

[7] C.S. Rahane, A. Kutzner, K. Heese, Establishing a human adrenocortical carcinoma (ACC)-specific gene mutation signature, Cancer Genet 230 (2019) 1–12, https://doi.org/10.1016/j.cancergen.2018.10.005.

[8] Y. Yu, Z. Wang, Q. Zheng, J. Li, FAM72 serves as a biomarker of poor prognosis in human lung adenocarcinoma, Aging 13 (2021) 8155–8176, https://doi.org/10.18632/aging.202625.

[9] K. Heese, The protein p17 signaling pathways in cancer, Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med. 34 (2013) 4081–4087, https://doi.org/10.1007/s13277-013-0999-1.

[10] S. Pramanik, A. Kutzner, K. Heese, Lead discovery and in silico 3D structure modeling of tumorigenic FAM72A (p17), Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med. 36 (2015) 239–249, https://doi.org/10.1007/s13277-014-2620-7.

[11] C. Guo, X. Zhang, S.P. Fink, P. Platzer, K. Wilson, J.K.V. Willson, Z. Wang, S. D. Markowitz, Ugene, a newly identified protein that is commonly overexpressed in cancer and binds uracil DNA glycosylase, Cancer Res. 68 (2008) 6118–6126, https://doi.org/10.1158/0008-5472.CAN-08-1259.

[12] B. Cheng, Q. He, Y. Cheng, H. Yang, L. Pei, Q. Deng, H. Long, L. Zhu, R. Jiang, A three-gene classifier associated with MicroRNA-mediated regulation predicts prostate cancer recurrence after radical prostatectomy, Front. Genet. 10 (2019) 1402, https://doi.org/10.3389/fgene.2019.01402.

[13] P. Rajan, J. Stockley, I.M. Sudbery, J.T. Fleming, A. Hedley, G. Kalna, D. Sims, C. P. Ponting, A. Heger, C.N. Robson, R.M. McMenemin, I.D. Pedley, H.Y. Leung, Identification of a candidate prognostic gene signature by transcriptome analysis of matched pre- and post-treatment prostatic biopsies from patients with advanced prostate cancer, BMC Cancer 14 (2014) 977, https://doi.org/10.1186/1471-2407-14-977.

[14] X. Li, X. Chen, Y. Yuan, R. Zhai, W.C. Cho, X. Jiang, Identifying immune cell infiltration and effective diagnostic biomarkers in lung adenocarcinoma by comprehensive bioinformatics analysis and in vitro study, Front. Oncol. 12 (2022), 916947, https://doi.org/10.3389/fonc.2022.916947.

[15] X. Zhang, S. Wan, Y. Yu, W. Ruan, H. Wang, L. Xu, C. Wang, S. Chen, T. Cao, Q. Peng, S. Li, T. Hu, Z. Jiang, Z. Chen, J.-B. Fan, Identifying potential DNA methylation markers in early-stage colorectal Cancer, Genomics 112 (2020) 3365–3373, https://doi.org/10.1016/j.ygeno.2020.06.007.

[16] Y. Zhang, Q. Wu, L. Xu, H. Wang, X. Liu, S. Li, T. Hu, Y. Liu, Q. Peng, Z. Chen, X. Wu, J.-B. Fan, Sensitive detection of colorectal cancer in peripheral blood by a novel methylation assay, Clin. Epigenet. 13 (2021) 90, https://doi.org/10.1186/s13148-021-01076-8.

[17] F. Hu, W. Zeng, X. Liu, A gene signature of survival prediction for kidney renal cell carcinoma by multi-omic data analysis, Int. J. Mol. Sci. 20 (2019) 5720, https://doi.org/10.3390/ijms20225720.

[18] F. Chatonnet, A. Pignarre, A.A. Sérandour, G. Caron, S. Avner, N. Robert, A. Kassambara, A. Laurent, M. Bizot, X. Agirre, F. Prosper, J.I. Martin-Subero, J. Moreaux, T. Fest, G. Salbert, The hydroxymethylome of multiple myeloma identifies FAM72D as a 1q21 marker linked to proliferation, Haematologica 105 (2020) 774–783, https://doi.org/10.3324/haematol.2019.222133.

[19] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550, https://doi.org/10.1186/s13059-014-0550-8.

[20] G. Bindea, B. Mlecnik, M. Tosolini, A. Kirilovsky, M. Waldner, A.C. Obenauf, H. Angell, T. Fredriksen, L. Lafontaine, A. Berger, P. Bruneval, W.H. Fridman, C. Becker, F. Pagès, M.R. Speicher, Z. Trajanoski, J. Galon, Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer, Immunity 39 (2013) 782–795, https://doi.org/10.1016/j.immuni.2013.10.003.

[21] V. Modhukur, T. Iljasenko, T. Metsalu, K. Lokk, T. Laisk-Podar, J. Vilo, MethSurv: a web tool to perform multivariable survival analysis using DNA methylation data, Epigenomics 10 (2018) 277–288, https://doi.org/10.2217/epi-2017-0118.

[22] N. Kroeger, W. Xie, J.-L. Lee, G.A. Bjarnason, J.J. Knox, M.J. Mackenzie, L. Wood, S. Srinivas, U.N. Vaishamayan, S.-Y. Rha, S.K. Pal, T. Yuasa, F. Donskov, N. Agarwal, C.K. Kollmannsberger, M.-H. Tan, S.A. North, B.I. Rini, T.K. Choueiri, D.Y.C. Heng, Metastatic non-clear cell renal cell carcinoma treated with targeted therapy agents: characterization of survival outcome and application of the International mRCC Database Consortium criteria, Cancer 119 (2013) 2999–3006, https://doi.org/10.1002/cncr.28151.

[23] P. Wu, T. Xiang, J. Wang, R. Lv, G. Wu, TYROBP is a potential prognostic biomarker of clear cell renal cell carcinoma, FEBS Open Bio 10 (2020) 2588–2604, https://doi.org/10.1002/2211-5463.12993.

[24] L.-T. Wang, C.-S. Lin, C.-Y. Chai, K.-Y. Liu, J.-Y. Chen, S.-H. Hsu, Functional interaction of Ugene and EBV infection mediates tumorigenic effects, Oncogene 30 (2011) 2921–2932, https://doi.org/10.1038/onc.2011.16.

[25] A. Kumar, R. Watkins, A.E. Vilgelm, Cell therapy with TILs: training and taming T cells to fight cancer, Front. Immunol. 12 (2021), 690499, https://doi.org/10.3389/fimmu.2021.690499.

[26] H.R. Ali, E. Provenzano, S.-J. Dawson, F.M. Blows, B. Liu, M. Shah, H.M. Earl, C. J. Poole, L. Hiller, J.A. Dunn, S.J. Bowden, C. Twelves, J.M.S. Bartlett, S.M. A. Mahmoud, E. Rakha, I.O. Ellis, S. Liu, D. Gao, T.O. Nielsen, P.D.P. Pharoah, C. Caldas, Association between CD8+ T-cell infiltration and breast cancer survival in 12,439 patients, Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. 25 (2014) 1536–1543, https://doi.org/10.1093/annonc/mdu191.

[27] S. Turcotte, A. Gros, E. Tran, C.-C.R. Lee, J.R. Wunderlich, P.F. Robbins, S. A. Rosenberg, Tumor-reactive CD8+ T cells in metastatic gastrointestinal cancer refractory to chemotherapy, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 20 (2014) 331–343, https://doi.org/10.1158/1078-0432.CCR-13-1736.

[28] M. Yan, J. Hu, Y. Ping, L. Xu, G. Liao, Z. Jiang, B. Pang, S. Sun, Y. Zhang, Y. Xiao, X. Li, Single-cell transcriptomic analysis reveals a tumor-reactive T cell signature associated with clinical outcome and immunotherapy response in melanoma, Front. Immunol. 12 (2021), 758288, https://doi.org/10.3389/fimmu.2021.758288.

[29] J.A. Park, L. Wang, N.-K.V. Cheung, Modulating tumor infiltrating myeloid cells to enhance bispecific antibody-driven T cell infiltration and anti-tumor response, J. Hematol. Oncol.J Hematol Oncol. 14 (2021) 142, https://doi.org/10.1186/s13045-021-01156-5.

[30] H. Nishikawa, S. Koyama, Mechanisms of regulatory T cell infiltration in tumors: implications for innovative immune precision therapies, J. Immunother. Cancer 9 (2021), e002591, https://doi.org/10.1136/jitc-2021-002591.

[31] S. Groeger, F. Wu, F. Wagenlehner, T. Dansranjav, S. Ruf, F. Denter, J. Meyle, PD-L1 up-regulation in prostate cancer cells by porphyromonas gingivalis, Front. Cell. Infect. Microbiol. 12 (2022), 935806, https://doi.org/10.3389/fcimb.2022.935806.

[32] B. Wang, W. Pan, M. Yang, W. Yang, W. He, X. Chen, J. Bi, N. Jiang, J. Huang, T. Lin, Programmed death ligand-1 is associated with tumor infiltrating lymphocytes and poorer survival in urothelial cell carcinoma of the bladder, Cancer Sci. 110 (2019) 489–498, https://doi.org/10.1111/cas.13887.

[33] P.-Y. Huang, S.-S. Guo, Y. Zhang, J.-B. Lu, Q.-Y. Chen, L.-Q. Tang, L. Zhang, L.-T. Liu, L. Zhang, H.-Q. Mai, Tumor CTLA-4 overexpression predicts poor survival in patients with nasopharyngeal carcinoma, Oncotarget 7 (2016) 13060–13068, https://doi.org/10.18632/oncotarget.7421.

[34] G. Santoni, C. Amantini, M.B. Morelli, D. Tomassoni, M. Santoni, O. Marinelli, M. Nabissi, C. Cardinali, V. Paolucci, M. Torniai, S. Rinaldi, F. Morgese, G. Bernardini, R. Berardi, High CTLA-4 expression correlates with poor prognosis in thymoma patients, Oncotarget 9 (2018) 16665–16677, https://doi.org/10.18632/oncotarget.24645.

[35] W.J. Lee, Y.J. Lee, M.E. Choi, K.A. Yun, C.H. Won, M.W. Lee, J.H. Choi, S.E. Chang, Expression of lymphocyte-activating gene 3 and T-cell immunoreceptor with immunoglobulin and ITIM domains in cutaneous melanoma and their correlation with programmed cell death 1 expression in tumor-infiltrating lymphocytes, J. Am. Acad. Dermatol. 81 (2019) 219–227, https://doi.org/10.1016/j.jaad.2019.03.012.

[36] M. Guo, F. Yuan, F. Qi, J. Sun, Q. Rao, Z. Zhao, P. Huang, T. Fang, B. Yang, J. Xia, Expression and clinical significance of LAG-3, FGL1, PD-L1 and CD8+T cells in hepatocellular carcinoma using multiplex quantitative analysis, J. Transl. Med. 18 (2020) 306, https://doi.org/10.1186/s12967-020-02469-8.

[37] Y. Liu, Y. Liu, R. Huang, W. Song, J. Wang, Z. Xiao, S. Dong, Y. Yang, X. Yang, Dependency of the cancer-specific transcriptional regulation circuitry on the promoter DNA methylome, Cell Rep. 26 (2019) 3461–3474.e5, https://doi.org/10.1016/j.celrep.2019.02.084.

[38] S.W.K. Ng, A. Mitchell, J.A. Kennedy, W.C. Chen, J. McLeod, N. Ibrahimova, A. Arruda, A. Popescu, V. Gupta, A.D. Schimmer, A.C. Schuh, K.W. Yee, L. Bullinger, T. Herold, D. Görlich, T. Büchner, W. Hiddemann, W.E. Berdel, B. Wörmann, M. Cheok, C. Preudhomme, H. Dombret, K. Metzeler, C. Buske, B. Löwenberg, P.J.M. Valk, P.W. Zandstra, M.D. Minden, J.E. Dick, J.C.Y. Wang, A 17-gene stemness score for rapid determination of risk in acute leukaemia, Nature 540 (2016) 433–437, https://doi.org/10.1038/nature20598.

[39] X. Li, W. Huang, W. Huang, T. Wei, W. Zhu, G. Chen, J. Zhang, Kinesin family members KIF2C/4A/10/11/14/18B/20A/23 predict poor prognosis and promote cell proliferation in hepatocellular carcinoma, Am. J. Transl. Res. 12 (2020) 1614–1639.

[40] K. Gupta, J. Levinsohn, G. Linderman, D. Chen, T.Y. Sun, D. Dong, M.M. Taketo, M. Bosenberg, Y. Kluger, K. Choate, P. Myung, Single-cell analysis reveals a hair follicle dermal niche molecular differentiation trajectory that begins prior to morphogenesis, Dev. Cell 48 (2019) 17–31.e6, https://doi.org/10.1016/j.devcel.2018.11.032.

[41] H.-S. Lee, Z. Lin, S. Chae, Y.-S. Yoo, B.-G. Kim, Y. Lee, J.L. Johnson, Y.-S. Kim, L. C. Cantley, C.-W. Lee, H. Yu, H. Cho, The chromatin remodeler RSF1 controls centromeric histone modifications to coordinate chromosome segregation, Nat. Commun. 9 (2018) 3848, https://doi.org/10.1038/s41467-018-06377-w.

[42] S. Stroik, K. Kurtz, K. Lin, S. Karachenets, C.L. Myers, A.-K. Bielinsky, E. A. Hendrickson, EXO1 resection at G-quadruplex structures facilitates resolution and replication, Nucleic Acids Res. 48 (2020) 4960–4975, https://doi.org/10.1093/nar/gkaa199.

[43] Y. Fu, X. Jia, J. Yuan, Y. Yang, T. Zhang, Q. Yu, J. Zhou, T. Wang, Fam72a functions as a cell-cycle-controlled gene during proliferation and antagonizes apoptosis through reprogramming PP2A substrates, Dev. Cell 58 (2023) 398–415.e7, https://doi.org/10.1016/j.devcel.2023.02.006.