Mathematical Biosciences and Engineering

MBE, 20(6): 10694–10724.
DOI: 10.3934/mbe.2023474
Received: 10 February 2023
Revised: 15 March 2023
Accepted: 23 March 2023
Published: 17 April 2023

http://www.aimspress.com/journal/MBE

Research article

Identification of basement membrane-related prognostic signature for predicting prognosis, immune response and potential drug prediction in papillary renal cell carcinoma

Yujia Xi[1,†], Liying Song[2,†], Shuang Wang[2,†], Haonan Zhou[3], Jieying Ren[4], Ran Zhang[5], Feifan Fu[4], Qian Yang[4], Guosheng Duan[2] and Jingqi Wang[1,*]

[1] Department of Urology, The Second Hospital of Shanxi Medical University, Taiyuan, China
[2] Second School of Clinical Medicine, Shanxi Medical University, Taiyuan, China
[3] First School of Clinical Medicine, Shanxi Medical University, Taiyuan, China
[4] School of Basic Medicine, Shanxi Medical University, Taiyuan, China
[5] School of Public Health, Shanxi Medical University, Taiyuan, China

*** Correspondence:** Email: drwangjq@126.com.

† These three authors contributed equally.

Abstract: Papillary renal cell carcinoma (PRCC) is a malignant neoplasm of the kidney and is highly interesting due to its increasing incidence. Many studies have shown that the basement membrane (BM) plays an important role in the development of cancer, and structural and functional changes in the BM can be observed in most renal lesions. However, the role of BM in the malignant progression of PRCC and its impact on prognosis has not been fully studied. Therefore, this study aimed to explore the functional and prognostic value of basement membrane-associated genes (BMs) in PRCC patients. We identified differentially expressed BMs between PRCC tumor samples and normal tissue and systematically explored the relevance of BMs to immune infiltration. Moreover, we constructed a risk signature based on these differentially expressed genes (DEGs) using Lasso regression analysis and demonstrated their independence using Cox regression analysis. Finally, we predicted 9 small molecule drugs with the potential to treat PRCC and compared the differences in sensitivity to commonly used chemotherapeutic agents between high and low-risk groups to better target patients for more precise treatment planning. Taken together, our study suggested that BMs might

play a crucial role in the development of PRCC, and these results might provide new insights into the treatment of PRCC.

Keywords: papillary renal cell carcinoma; basement membrane; prediction model; signature; The Cancer Genome Atlas

1. Introduction

Renal cell carcinoma (RCC) is the 13th most common cancer worldwide, accounting for 2–3% of adult malignancies [1,2]. RCC can be divided into three main distinct histological subtypes: clear cell renal cell carcinoma (KIRC) in 75%, PRCC in 15–20%, and chromophobe renal cell carcinoma (KICH) in 5% [3]. PRCC can be divided into types I and II according to the molecular subtype classification. Compared to KIRC, type II PRCC is more aggressive and has a worse prognosis [4]. For localized PRCC, doctors usually choose radical nephrectomy or partial nephrectomy with wide surgical margin excisions. In the face of late-stage PRCC, MET inhibitors have been considered as a therapeutic strategy in advanced PRCC [5]. Targeted therapy is often considered for metastatic PRCC. Meanwhile, with the development of immune checkpoint inhibitors (ICI), combination therapy with tyrosine kinase inhibitors (anti-MET inhibitors) and ICI has potential to achieve good therapeutic effects [6]. Thus, refining the molecular and immune landscape of PRCC may be a crucial piece of work.

The BM is a special type of extracellular matrix (ECM) that contacts cells directly and is formed by two major ECM macromolecular proteins, collagen IV and laminin [7]. As part of the ECM BM isolates and interconnects the cell with the ECM of interest and serves as a bridge between the cell and the matrix microenvironment. In addition to providing a scaffold, BM also acts as a reservoir and regulator of growth factors, directing and fine-tuning cellular function [8]. Alterations in BM structure or function have been found in a variety of diseases, including cancers, autoimmune diseases and metabolic disorders [9]. Remodeling of the BM by cancer cells to invade the parenchyma is a defining event in the transition from dysplastic carcinoma in situ to invasive and malignant cancer [10].

As a complex macromolecular structure, BM undergoes critical transformations during the development of glomerular and renal tubules, and alterations in BM structure and function have been found in kidney diseases [11]. However, the effect of basement membrane genes (BMs) on PRCC has not been systematically studied. In this study, the interaction between BMs and PRCC was assessed systematically based on genomic information, and we constructed a prognostic model based on BMs to provide a reference for clinical practice.

2. Materials and methods

2.1. Data gathering and processing

RNA sequencing data from the experimental group of PRCC patients and healthy individuals was obtained in The Cancer Genome Atlas (TCGA) database (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga). Clinical information such as gender, age, pathological stage, TNM stage and the prognosis was also collected and collated. The R "limma"

package analyzed the control and experimental sequencing matrices with the critical threshold |logFC| > 0.585 and FDR < 0.05 screening DEGs between PRCC and normal, and the results were presented visually using the R "ggplot2" package. The DEGs were compared with over 200 BMs previously reported [12], of which the common genes were considered as the basis for subsequent studies of BM lesions in PRCC.

2.2. Development and verification of a BMs-based risk score model

Univariate Cox regression analysis was performed to screen genes associated with survival prognosis in the BMs of PRCC, followed by R "glmnet" package for LASSO Cox regression to choose prognostic signature genes and develop risk models to reflect the combined prognostic impact of BMs in PRCC. The resulting models were validated by 10-fold crossover, and the LASSO penalty parameter (λ) was determined. The model was constructed in the following manner (Eq (1)).

$$Risk\ score = \sum_{i=1}^{n}(Coef\ i * ExpGene\ i) \tag{1}$$

"ExpGene" represents the mean expression value in the prognostic signature gene sample, and "Coef" represents the model coefficient calculated from the Lasso regression of the prognostic signature gene. The risk values of the samples were estimated based on the created models, where the joint survival information of high and low-risk groups, divided by median, was used to construct Kaplan-Meier curves and ROC curves to reflect the prognoses of the samples and the ability of the models to evaluate them, respectively. The "ggsurvplot" function and R "survivalROC" package were used to visualize the plots therein.

2.3. Evaluation and validation of risk model

The relationships between the risk values of the model and various clinical data such as gender, age, pathological stage and survival time were presented in heat maps, box plots and survival curves in multiple dimensions to illustrate the clinical significance of the constructed model. Meanwhile, univariate and multivariate Cox regression analyses were performed to assess the independent predictive value of the risk score model for prognosis based on clinical characteristics data using the R "rms" package. Integrating the independent predictors obtained from the multivariate regression analysis using a 70% split sample of the dataset, column line plots and calibration curves were constructed to validate the predictive power of the risk models at 1, 3 and 5 years. The C-index of the line plot and the slope of the calibration curve reflect the strength of the predictive ability, and the closer the two parameters are to 1, the better the predictive accuracy.

2.4. Protein-protein interaction network analysis and recognition of transcription factors with signature

Proteins enter the cell with a similar protein association formed by a protein-protein interaction (PPI) network, which indicated the protein mechanism. In cell and systems biology, the assessment and analysis of PPI networks and their function was the fundamental goal in explaining and gaining an understanding of the workings of cellular mechanisms. A PPI network from PRCC patient

DEGs was constructed using the STRING database (https://string-db.org/) (version 11.5) [13,14]. Then, the PPI network was imported into Cytoscape (v.3.9.1) for visualization and further analysis of PPI network modules. GeneMANIA (http://genemania.org/) was used to explore the 20 genes most associated with the signature genes and their interactions [15]. Transcription factors (TFs) were proteins that attach to specific genes and control the rate of transcription of genetic information and were critical to understanding the molecular level. The NetworkAnalyst platform was used to identify topologically plausible TFs from the JASPAR database tending to bind to our common signature genes. NetworkAnalyst (https://www.networkanalyst.ca/) was a database with a comprehensive range of gene expression profiling and network visualization analysis [16].

2.5. Functional enrichment analyses and gene set enrichment analysis

The R "clusterProfiler" and "GSVA" packages were respectively utilized for GO functional enrichment, KEGG pathway enrichment analysis and gene set enrichment analysis (GSEA) of BMs in PRCC. The enrichment analyses were visualized using the R "ggplot2" package after being filtered by $P < 0.05$ and FDR < 25% criteria.

2.6. Promotion and suppression of tumor immune responses

Immune cell infiltration and immune checkpoint expression could essentially reflect the immune promotion or suppression of the tumor microenvironment and provided a reference for the immunotherapy of PRCC. In this regard, differences in immune cell infiltration and immune checkpoints [17] between high and low-risk groups were assessed based on the CIBERSORT, TIMER, QUANTISEQ, CIBERSORT-ABS, XCELL, MCP counter and EPIC algorithms. The results of immune analyses were presented as heat maps and box plots. The "ssGSEA" was used to determine the scores of immune cells and analyze the activity of immune-related pathways using the "gsva" program. Additionally, the TIMER database (https://cistrome.shinyapps.io/timer/) was used to determine the relationship between immune cells and the 10 BMs, improving our understanding of the role of BMs in PRCC. The TIMER web server was a comprehensive resource for the systematic analysis of immune infiltrates in different cancer types. The TIMER algorithm allowed the estimation of the abundance of six immune infiltrates (B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages and dendritic cells) [18,19].

2.7. Drug prediction and sensitivity analysis

The DsigDB database (https://maayanlab.cloud/Enrichr/) [20–22] predicted small molecule drugs targeting genes characteristic of the BMs of PRCC, and the top 9 drugs were selected as potential drugs to be developed for the treatment of PRCC development using p-value ranking. Based on the differences in prognostic characteristics and immune infiltration between high and low-risk groups, the sensitive drugs were predicted to be different for different risk groups. Accordingly, the R "PRrophic" package analyzed 50% of the maximum inhibitory concentration (IC50) of commonly used renal papillary carcinoma drugs for pharmacovigilance analysis to provide a reference indication for clinical therapeutic use.

2.8. Analysis of gene expression and methylation levels

The University of ALabama at Birmingham CANcer data analysis Portal (UALCAN) was a comprehensive, user-friendly interactive web resource for the analysis of cancer databases [23,24]. It focused on the analysis of data from TCGA database and provides charts and graphs. UALCAN allowed users to identify biomarkers or perform computer validation of potential genes of interest and assess epigenetic regulation of gene expression through promoter methylation. UALCAN also provided additional information on selected genes/targets through links to HPRD, GeneCards, Pubmed, TargetScan, the Human Protein Atlas, DRUGBANK, Open Targets and GTEx.

2.9. Verification of gene expression

GSE15641, which contained 23 normal tissue samples and 11 PRCC tissue samples, was selected as a verification set to verify the gene expression of BMs. Verification data was from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). The GEO database was a publicly available genomics data repository, which contains array and sequence data.

2.10. Statistics analysis

Wilcoxon test and Kaplan-Meier analysis were used to compare survival differences between high and low-risk groups. Single and multiple regression analyses were performed to determine independent prognostic factors for the models. Meanwhile, box-line plots, column plots and heat maps were used for multi-dimensional responses to assess the ability and predictive value of the models. All statistical tests in this study were analyzed by R software (4.2.0), and all results were within the range of statistical significance ($p < 0.05$).

3. Results

3.1. Establishing and validating a signature based on BMs

A total of 98 differentially expressed BMs were selected between normal and PRCC samples, including 41 up-regulated genes and 57 down-regulated genes (Figure 1(a)). The prognostic value of the 98 DEGs was investigated by univariate Cox regression analysis. The results showed that only 23 BMs had prognostic values (Figure 1(b)). The 23 BMs were analyzed by LASSO Cox regression, and a prognostic risk model featuring 10 BMs (Figure 1(c) and (d)) was constructed as follows: risk score = (0.8059 × expression level of ADAMTS6) + (7.9349 × expression level of CD44) + (−0.0004 × expression level of CTSD) + (0.0233 × expression level of DDR2) + (0.0047 × expression level of FBLN1) + (0.0027 × expression level of HSPG2) + (0.0099 × expression level of LAMA1) + (0.0003 × expression level of LUM) + (0.0052 × expression level of THBS2) + (0.0004 × expression level of TIMP1). The samples were divided into high and low-risk groups based on the median risk model score. PCA analysis indicated different dimensions between high-risk and low-risk groups (Figure 2(a)). There was a significant difference ($p < 0.001$) in survival and mortality rates between the high and low-risk groups based on the median risk mode, which suggested a negative correlation between the risk scores and survival prognosis (Figure 2(b) and (d)). The result of time-dependent ROC analysis

showed that the prognostic accuracy in TCGA dataset was 0.781 at 1 year, 0.741 at 3 years and 0.668 at 5 years (Figure 2(c)). A random sample of 70% of TCGA dataset was taken to validate the risk model again, showing that validation group survival curve p-value = 0.004, and AUCs were 0.874 at 1 year, 0.816 at 3 years and 0.675 at 5 years (Figure 3(a)–(c)). Heatmap showed the differences of 10 BMs between high and low-risk patients in TCGA and validation set (Figure 2(e) and Figure 3(d)).

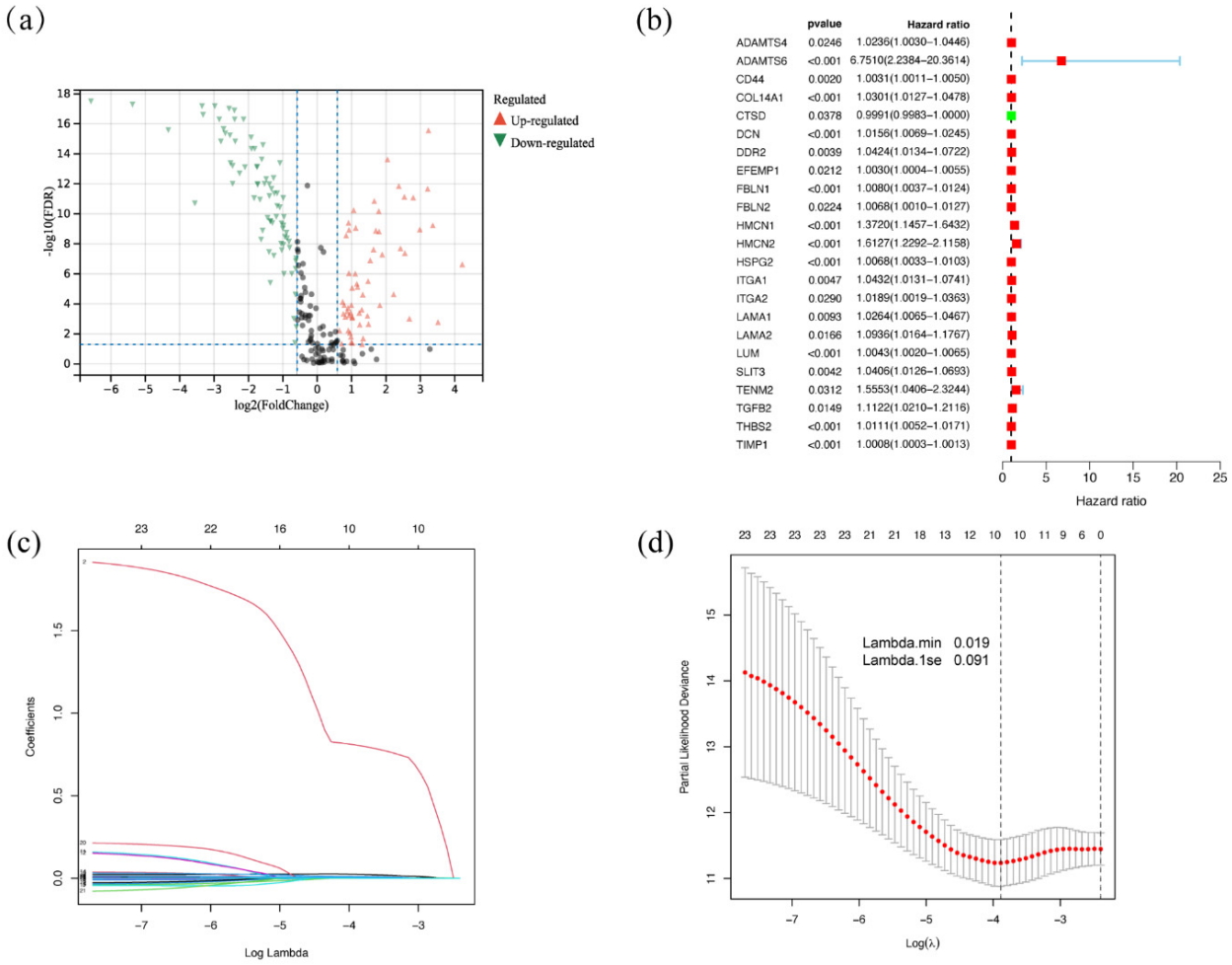

Figure 1. Differential expression of BMs. (a) Volcano map of differential BMs, green for down-regulated genes and red for up-regulated genes. (b) Forest map of univariate Cox regression analysis to identify prognosis-related BMs. (c) Selection of the optimal parameter (lambda) in the least absolute shrinkage and selection operator (LASSO) model. Dotted vertical lines were drawn at the optimal values using the minimum criteria. (d) Selection of predictors using the LASSO regression analysis with 10-fold cross-validation.

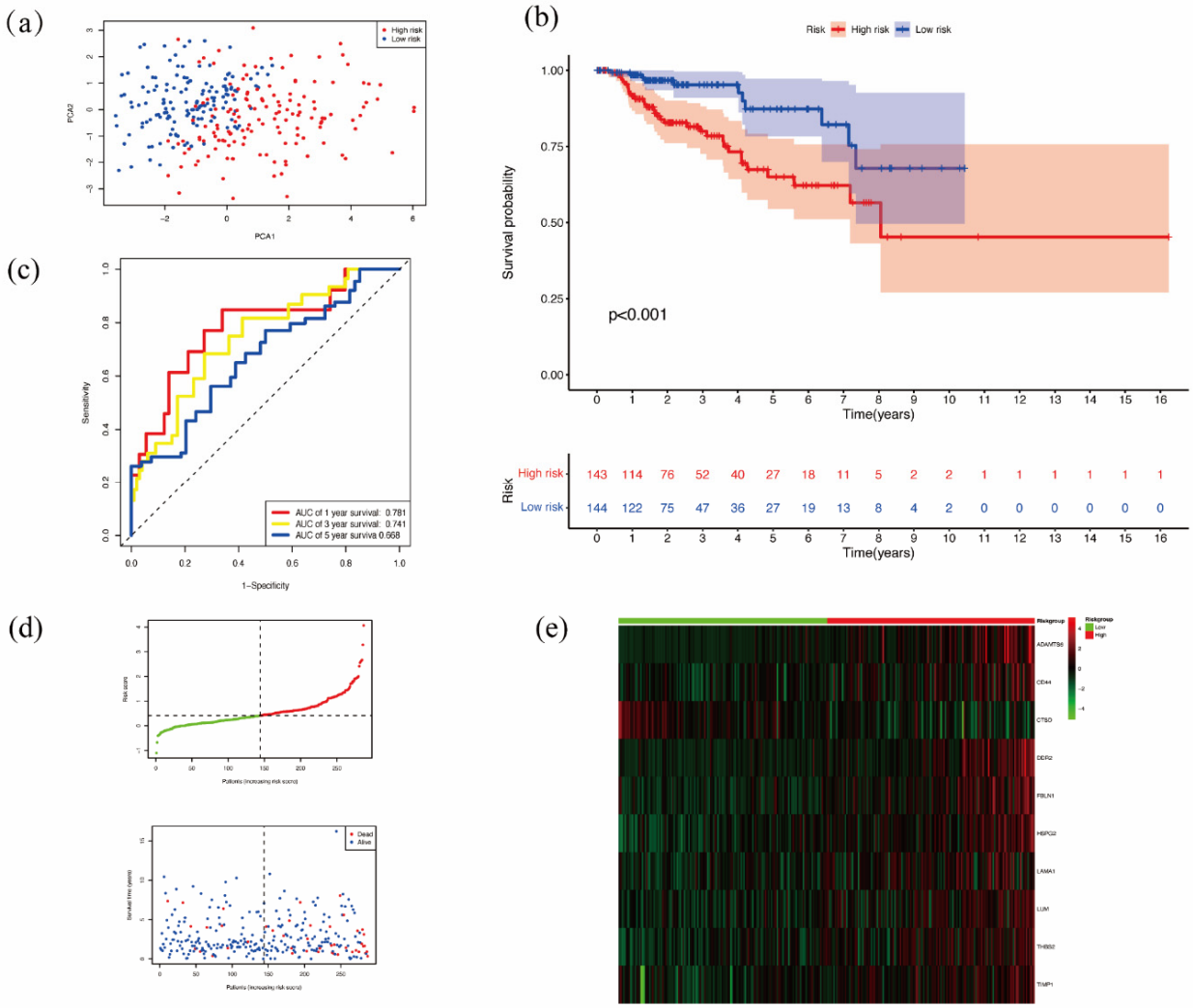

Figure 2. Construction of BM-based prognostic features in TCGA set. (a) PCA analysis in TCGA set. (b) Kaplan-Meier survival analysis of PRCC patients between high and low-risk groups in TCGA. (c) Time-independent receiver operating characteristic (ROC) analysis of risk scores predicting overall survival in TCGA. (d) Distribution of survival status based on median risk scores in TCGA. (e) Heat map showing differences in 10 BM-associated genes between high-risk and low-risk patients in TCGA.

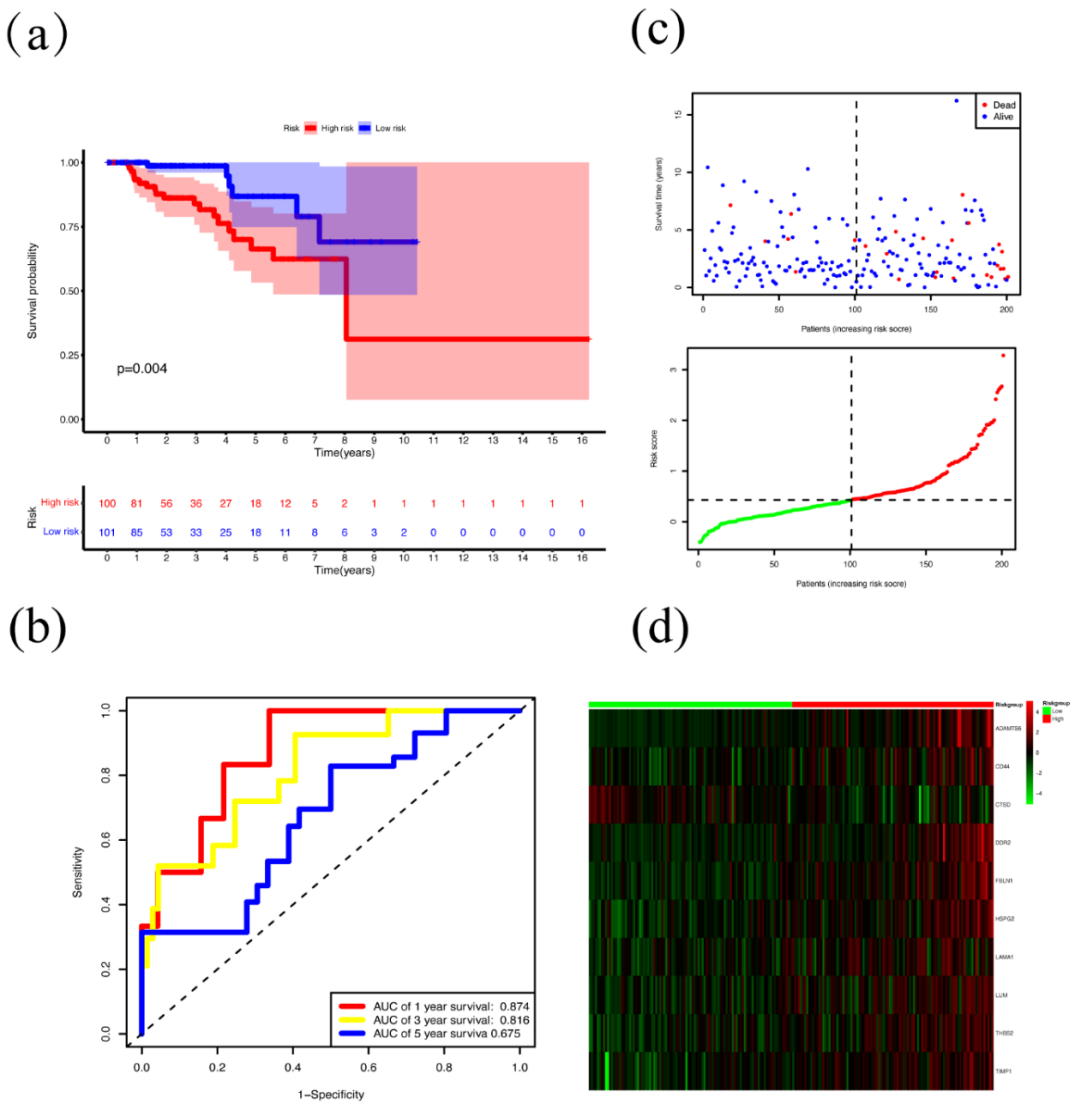

Figure 3. 70% of TCGA cohort was randomly selected as the validation set to validate BM-based prognostic signature. (a) Kaplan-Meier survival analysis of PRCC patients between the high and low-risk groups in the validation set; (b) time-independent receiver operating characteristic (ROC) analysis of risk scores predicting overall survival in the validation set; (c) distribution of survival status based on median risk scores in the validation set. (d) Heat map showing differences in 10 BM-associated genes between high and low-risk patients in the validation set.

3.2. Association between clinical characteristics and the signature was an independent indicator of the prognosis of PRCC patients

The univariate and multivariate Cox analyses were executed to demonstrate whether these features could be independent prognostic indicators. Univariate analysis indicated that risk score, TNM stage, age and gender were significantly associated with survival in PRCC patients ($p < 0.001$) (Figure 4(a)). Multivariate analysis showed that risk score, N stage and age remained significantly associated with prognosis ($p < 0.05$) (Figure 4(b)). These results suggested that BMs-based

characteristics were independent prognostic indicators for PRCC patients.

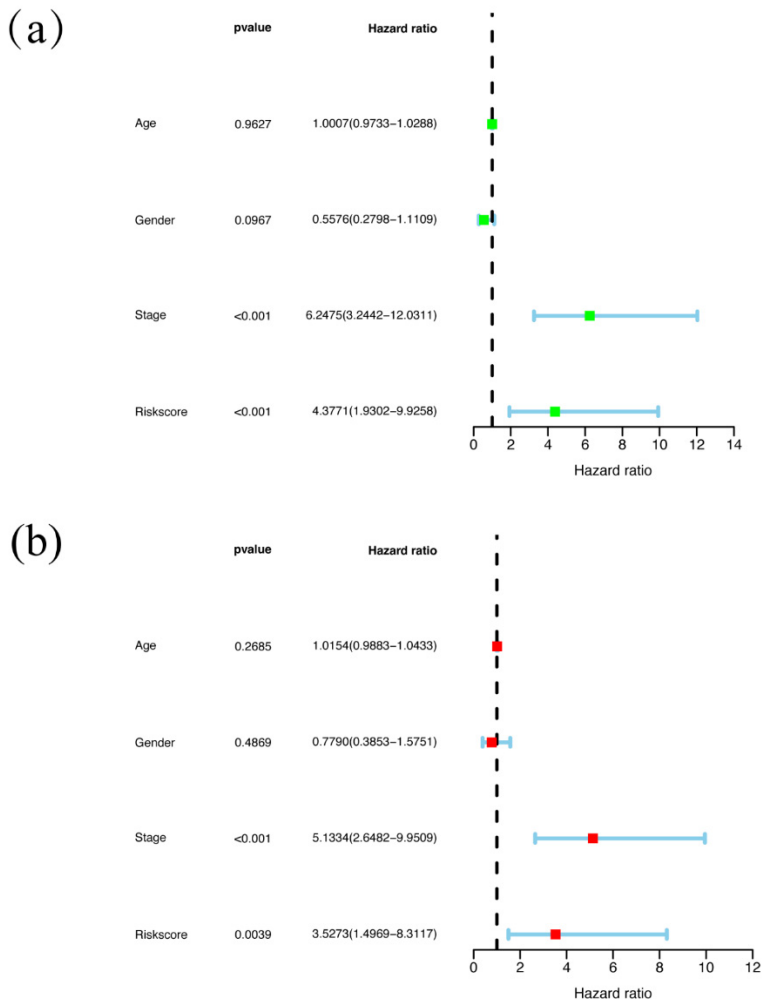

Figure 4. The signature was an independent prognostic factor for PRCC in TCGA cohort study. (a) Correlation between risk scores for OS and clinicopathological factors by univariate Cox regression analysis. (b) Correlation between risk scores for OS and clinicopathological factors by multivariate Cox regression analysis.

The Chi-square test was used to investigate whether the prognostic signature was involved in the development and progression of PRCC. Results (Figure 5(a)) revealed that there were significant sex differences ($p < 0.01$), TNM stage ($p < 0.05$) and T stage ($p < 0.01$) between high-risk group and low-risk group, except for age ($p > 0.05$). Wilcoxon test results (Figure 5(b)) showed that age, gender, T stage and TNM stage were significantly related to risk score ($p < 0.05$). In addition, stratified analyses were further performed to investigate the prognostic significance of the signature in subgroups. Our study suggested that BMs-based signatures had a good prognostic predictive ability in age > 65 years ($p = 0.014$), ≤ 65 years ($p = 0.002$), males ($p = 0.002$), Stage III, IV ($p = 0.009$) and T3, T4 ($p = 0.018$). However, BM-based signals did not perform well in predicting female, Stage I, II and T1, T2 prognosis ($p > 0.05$) (Figure 6).

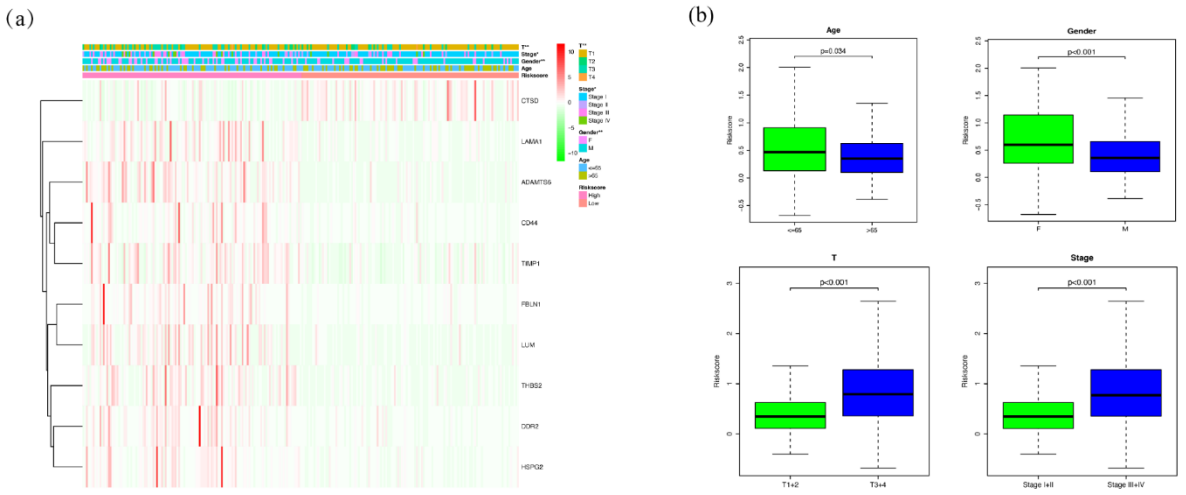

Figure 5. Correlation between risk score and clinical factors. (a) Heat map showing the correlation between risk groups and clinical characteristics. (b) Boxplot showing a correlation between risk scores and clinical factors.

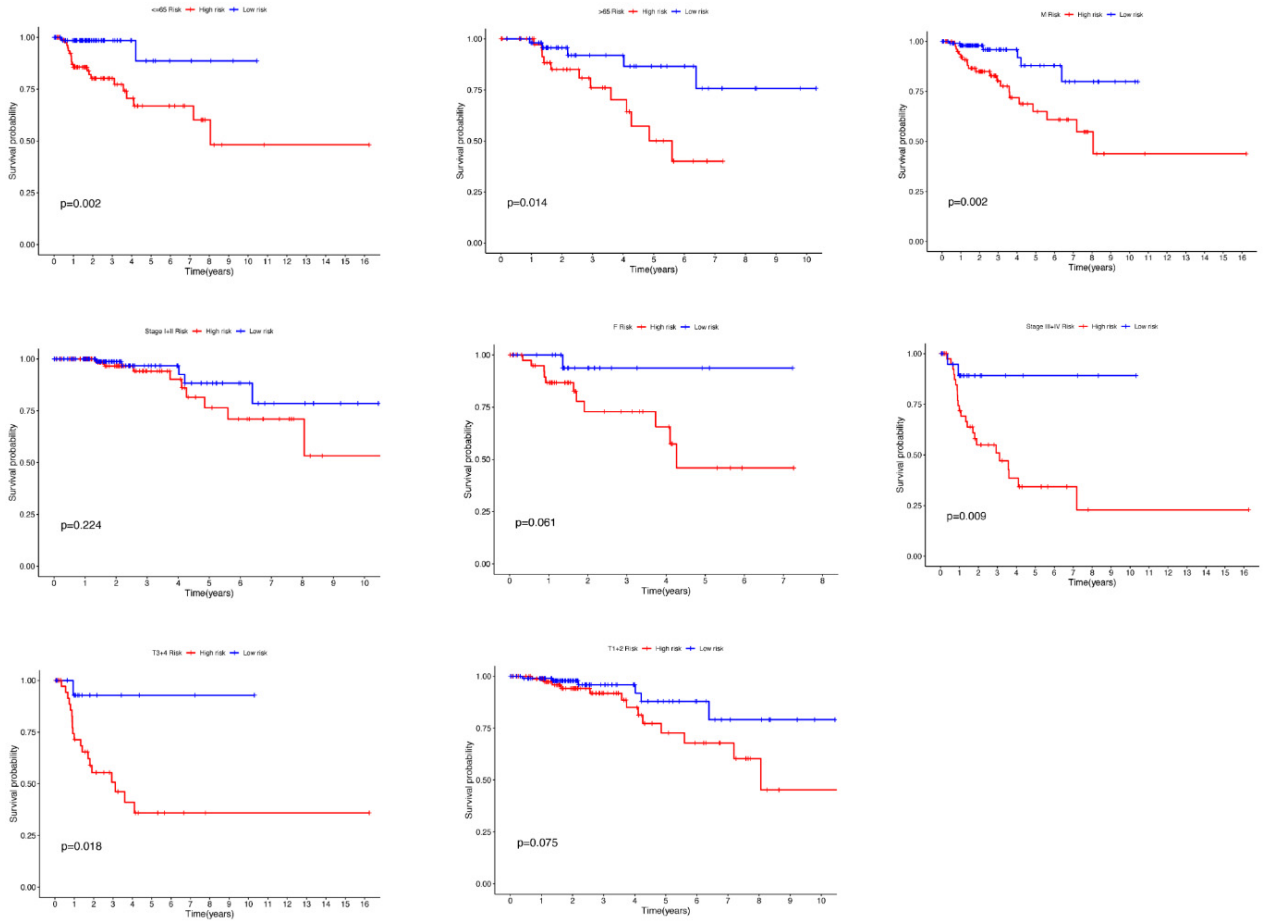

Figure 6. Kaplan-Meier plots of OS differences between high and low-risk groups for TCGA stratified by sex, age, T stage and TNM stage.

3.3. Construction of a nomogram

A nomogram combined various prognostic indicators to assess individual survival probabilities graphically. To further predict survival in patients with PRCC, we constructed a nomogram consisting of age, gender, risk score and stage. Nomography predicted 1, 3 and 5-year survival in PRCC patients (Figure 7(a)). The calibration curve showed that the actual survival of patients met the predicted value (Figure 7(b)). C-index and DCA methods were used to test the accuracy and stability of the model, which confirmed the good predictive ability of the nomogram (Figures 7(c) and (d)).

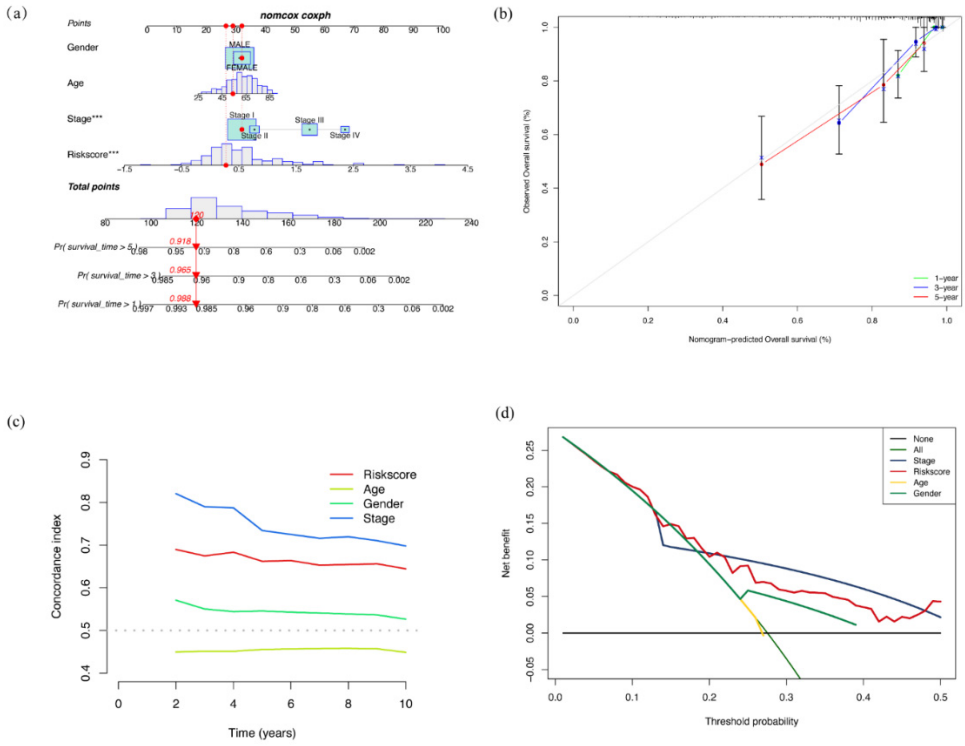

Figure 7. (a) A nomogram consisting of the risk score and other clinical indicators for predicting 1, 3, and 5-year OS of PRCC. (b) Calibration plot for nomogram predicted and observed 1, 3 and 5-year overall survival rates. Nomogram-predicted survival and actual survival were plotted on the x-axis and y-axis, respectively. The white dotted line represents the best prediction, and the blue, red and green solid lines represent the nomogram prediction. The vertical bars represent a 95% confidence interval. (c) Coherence index (C-index) values indicate the prognostic predictive power of clinical factors and risk score. (d) DCA curve based on Cox regression model.

3.4. Mutation analysis in high and low-risk groups and drug sensitivity analysis

To determine the differences in PRCC-associated gene mutations between the high and low-risk groups, the gene mutations in the high and low-risk groups were calculated and combined with clinical traits. General information on representative mutations in both groups is presented in Figure 8(a) and (b). The TTN, KMT2C, MET, MUC16 and SETD2 genes had the top five mutation frequencies in the high-

risk group, while the top five genes TTN, MUC4, FAT1, KIAA1109 and WDFY3 had the highest mutation frequencies in the low-risk group.

We further investigated the difference in sensitivity to commonly used chemotherapeutic drugs for PRCC between the two groups (Sunitinib, Pazopanib, Axitinib, Sorafenib, Temsirolimus and Erlotinib). The results of the GDSC database analysis showed that the IC50 values of Axitinib, Pazopanib and Erlotinib in the high-risk group were significantly lower than those in the low-risk group, suggesting that patients in the high-risk group were more sensitive to axitinib, pazopanib and erlotinib (Figure 8(c)–(e)).

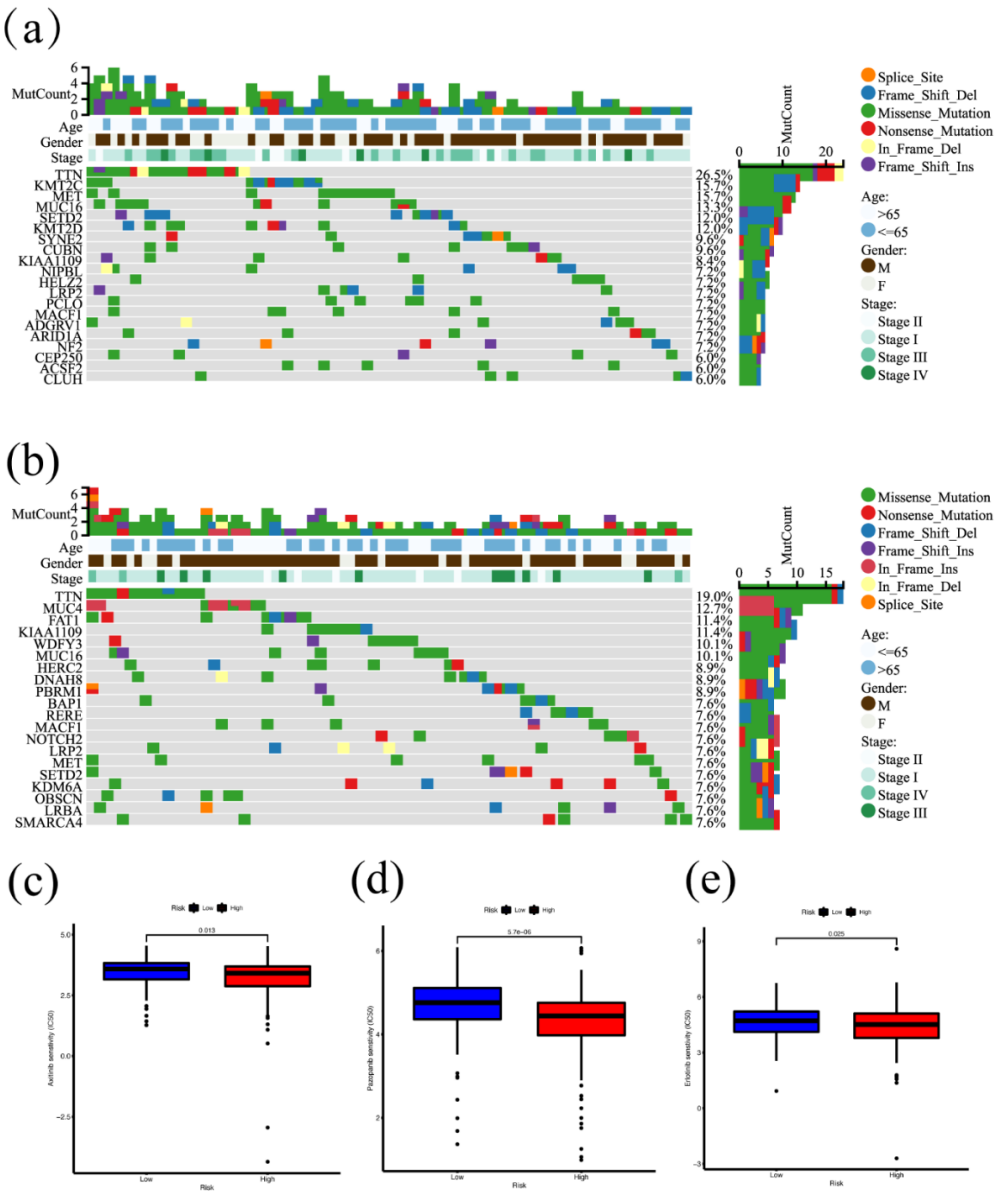

Figure 8. (a) A waterfall plot of mutation profiles in the high-risk group. (b) Waterfall plot of mutation profiles in the low-risk group. (c)–(e) Boxplots depicting the difference in drug sensitivity analysis of axitinib, pazopanib and erlotinib between the high-risk and low-risk groups.

3.5. Functional enrichment analyses and gene set enrichment analysis

GO and KEGG enrichment analyses were used to explore potential functions of differentially expressed BMs. Bioprocess results showed that 98 BMs were significantly involved in an ECM organization, extracellular structure organization, external encapsulating structure organization, cell-substrate adhesion and integrin-mediated signaling pathway. Analysis of cellular components showed that collagen-containing ECM, BM, integrin complex, a protein complex involved in cell adhesion and collagen trimer were significantly enriched. Molecular function analysis indicated that BMs were mainly located in ECM structural constituent, integrin binding, glycosaminoglycan binding, collagen binding and ECM binding (Figure 9(a)).

The KEGG results suggested that these BMs were mainly involved in ECM-receptor interaction, focal adhesion, PI3K-Akt signaling pathway, human papillomavirus infection, amoebiasis, hypertrophic cardiomyopathy and so on. (Figure 9(b)).

To further elucidate the molecular mechanisms underlying BM-based signatures, GSEA analysis was performed. The results of GSEA analysis showed that the main enrichment areas, such as alcoholism, herpes simplex virus 1 infection, human papillomavirus infection, influenza A, and PI3K-Akt signaling pathway, were observed in the KEGG gene sets (Figure 9(c)).

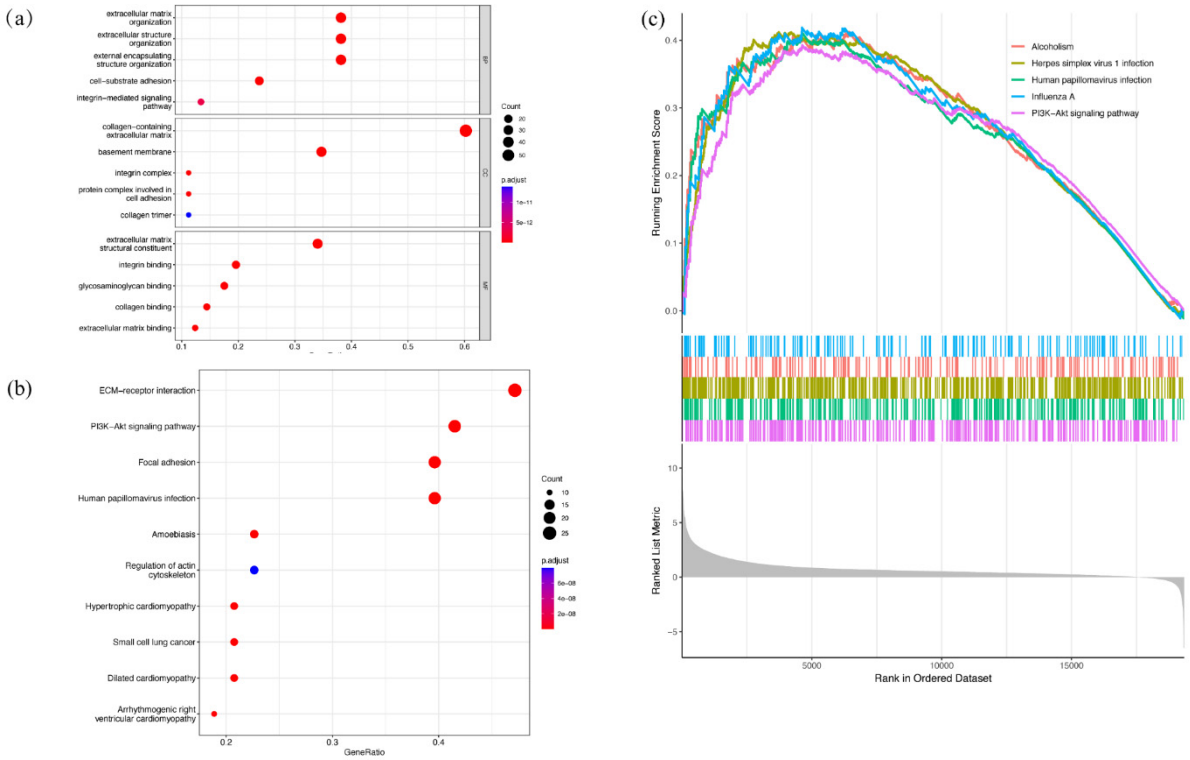

Figure 9. GO, KEGG and GSEA enrichment analyses of differentially expressed BMs. (a) Bubble plot of GO functional enrichment analysis; BP, biological process; CC, cellular component; MF, molecular function. (b) Bubble plot of KEGG pathway enrichment analysis. Results with a P-value < 0.05 were considered statistically significant. (c) GSEA analysis for differentially expressed BMs. Results with P-value < 0.05 were considered statistically significant.

3.6. Protein-protein interaction and transcription factor prediction

The STRING database analysis showed that the PPI network of differentially expressed BMs consisted of 94 nodes and 637 edges (Figure 10(a)). The most meaningful module consists of 30 BMs including 30 nodes and 213 edges (Figure 10(b)).

GeneMANIA analysis indicated that the protein linkage correspondence among the 10 signature genes was mainly Co-expression (64.21%), Physical interactions (15.73%), predicted protein interactions (9.18%), shared localization domains (2.65%) and pathway (2.44%). Enrichment analysis of genes in the network graph was performed, mainly enriched on collagen-containing ECM, ECM structural lumen, aminoglycan catabolic process, glycosaminoglycan catabolic process, vacuolar and ECM organization (Figure 10(c))

Transcription factor (TF) interactions with signatures were presented in Figure 10(d). From the TF-gene interaction network analysis, 73 TFs have been identified, which basically indicates a strong association between them.

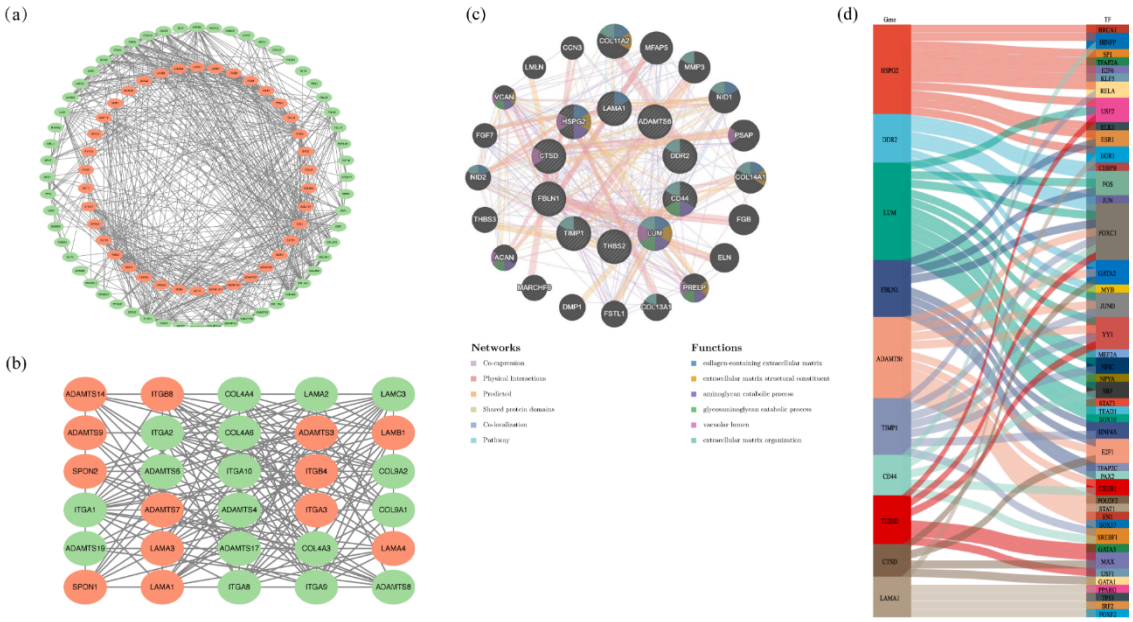

Figure 10. (a) The PPIs of differentially expressed BMs. (b) The core module of the PPI network. (c) The network graph was made to present the 20 most associated proteins, the protein linkage correspondence and the enrichment analysis of 10-signature genes. (d) A Sankey diagram was performed to present the relationship between the 10-signature genes and TFs.

3.7. Immune infiltration level analysis of the BM-based signature

Analysis of TIMER, CIBERSORT, CIBERSORT-ABS, XCELL, QUANTISEQ, EPIC and MCP-COUNTER presented the relationships between features and immuno-penetration on a heatmap (Figure 11(a)). The details of the 7 algorithms were shown in Table 1−7 and supplement Figure S1. Based on the functional analyses, we further compared the enrichment scores of 16 types of immune

cells and the activity of 13 immune-related pathways between the low- and high-risk groups in TCGA cohort by employing the single-sample gene set enrichment analysis (ssGSEA). The high-risk subgroup had higher levels of infiltration of immune cells, including B cells, iDCs, Tfh cells and Treg cells, than the low-risk subgroup, whereas macrophages and pDCs were higher in the low-risk group (Figure 11(b)). The results meant that immunotherapy might work better for patients in the high-risk group. Additionally, the high-risk group showed higher scores in Type II IFN, while the low-risk group showed higher scores in APC co-inhibition, HLA and T cell co-inhibition (Figure 11(c)). The correlation of signature with immune cells and immune pathways was presented in Figure 11(d). The results suggested that CD44 had the highest correlation with most immune cells and immune pathways, which had a positive correlation. In contrast, LAMA1 was negatively associated with most immune cells and pathways. Given the importance of checkpoint inhibition immunotherapy, we also investigated the association between risk scores and 47 common immune checkpoints. We found that 24 checkpoints including BTLA, CD27, CD40, PDCD1, PDCD1LG2, TIGIT and so on had significant differences in the expression between the two groups of patients. Moreover, most of checkpoints were elevated in the high-risk group, suggesting immunosuppression and exhaustion phenotypes in the high-risk group (Figure 12 and Supplement Figure S2).

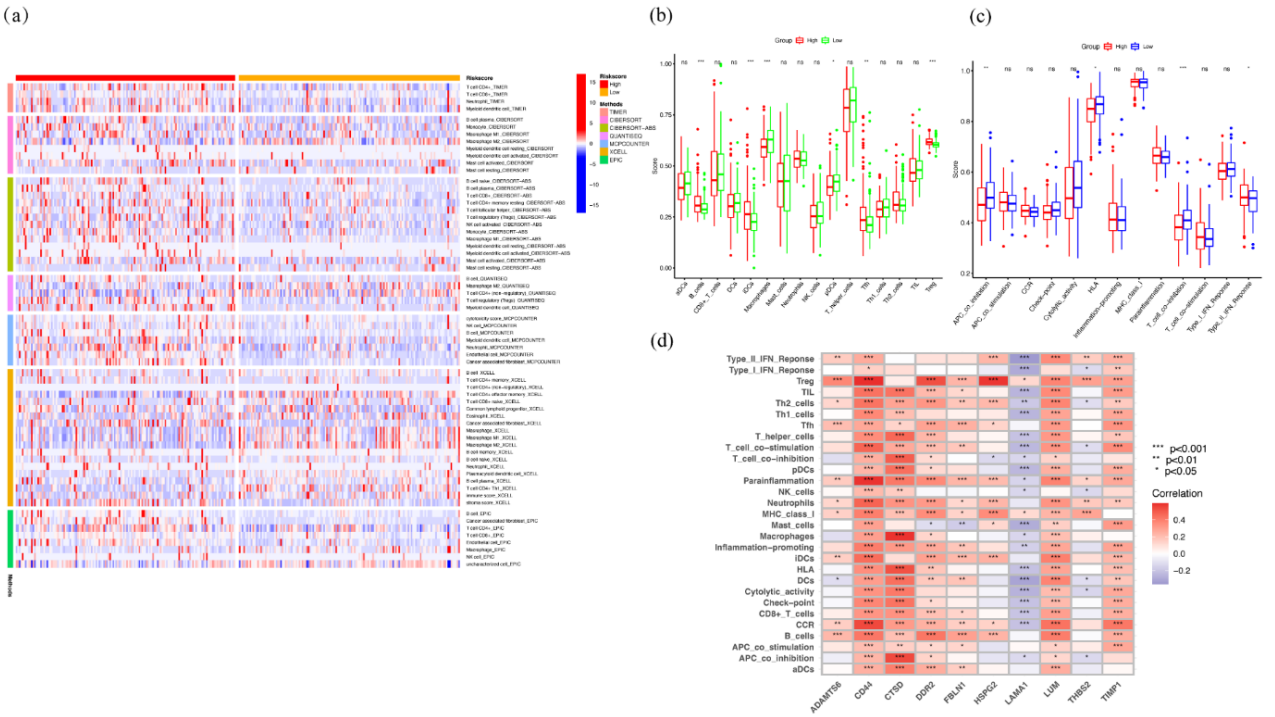

Figure 11. (a) Heat map of the difference in immunocyte infiltration between high-risk and low-risk groups. (b) Boxplot showing different infiltrating scores in 16 immune cell types between the high-risk and low-risk groups. (c) Boxplot showing the difference in infiltrating score for the activity of 13 immune-related pathways between the high-risk and low-risk groups. (d) Heat map of correlations of 10-signature genes with immune cells and immune pathways.

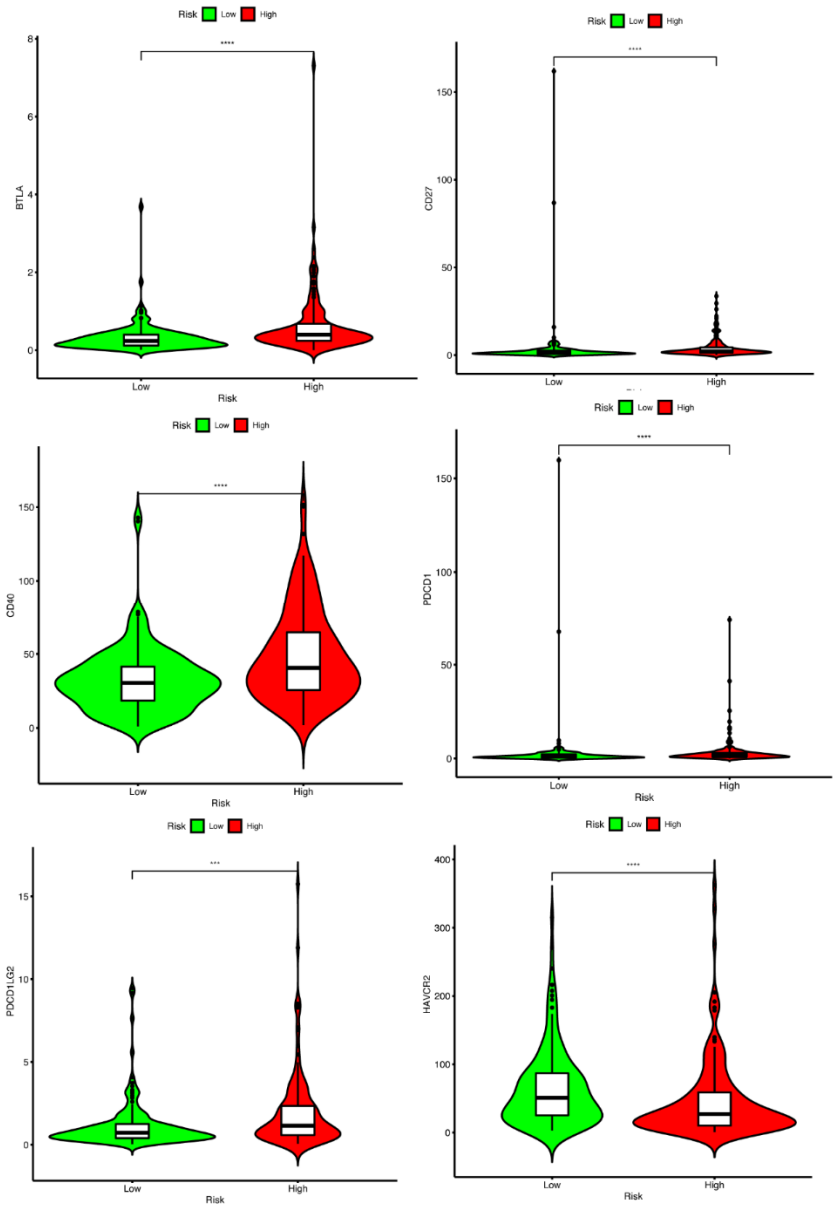

Figure 12. Violin plots visualize significantly different immune checkpoints between the high-risk and low-risk groups.

Table 1. Comparison of immune cells in Low and High by TIMER.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
B cell	(0.10 ± 0.08)	(0.12 ± 0.09)	0.08
T cell CD4+	(0.11 ± 0.04)	(0.14 ± 0.06)	5.80×10^{-5}
T cell CD8+	(0.20 ± 0.19)	(0.21 ± 0.07)	0.02
Neutrophil	(0.11 ± 0.03)	(0.12 ± 0.04)	0.01
Macrophage	(0.08 ± 0.11)	(0.07 ± 0.09)	0.6
Myeloid dendritic cell	(0.45 ± 0.12)	(0.51 ± 0.14)	4.20×10^{-5}

Table 2. Comparison of immune cells in Low and High by CIBERSORT.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
B cell naive	$(8.8 \times 10^{-3} \pm 0.02)$	(0.02 ± 0.03)	0.07
B cell memory	$(6.2 \times 10^{-3} \pm 0.01)$	(0.01 ± 0.02)	0.28
B cell plasma	(0.03 ± 0.04)	(0.05 ± 0.06)	5.20×10^{-3}
T cell CD8+	(0.09 ± 0.07)	(0.08 ± 0.06)	0.78
T cell CD4+ naive	$(1.0 \times 10^{-4} \pm 1.2 \times 10^{-3})$	$(1.5 \times 10^{-3} \pm 0.01)$	0.31
T cell CD4+ memory resting	(0.17 ± 0.08)	(0.18 ± 0.08)	0.26
T cell CD4+ memory activated	$(7.7 \times 10^{-5} \pm 9.3 \times 10^{-4})$	$(3.4 \times 10^{-4} \pm 3.7 \times 10^{-3})$	0.31
T cell follicular helper	(0.01 ± 0.02)	(0.02 ± 0.02)	0.25
T cell regulatory (Tregs)	(0.02 ± 0.02)	(0.02 ± 0.02)	0.36
T cell gamma delta	$(6.2 \times 10^{-3} \pm 0.02)$	$(5.6 \times 10^{-3} \pm 0.02)$	0.24
NK cell resting	$(2.6 \times 10^{-3} \pm 9.6 \times 10^{-3})$	$(3.5 \times 10^{-3} \pm 0.01)$	0.69
NK cell activated	(0.05 ± 0.03)	(0.05 ± 0.03)	0.25
Monocyte	(0.03 ± 0.03)	(0.04 ± 0.04)	0.02
Macrophage M0	(0.04 ± 0.07)	(0.03 ± 0.05)	0.38
Macrophage M1	(0.01 ± 0.02)	(0.02 ± 0.02)	3.10×10^{-3}
Macrophage M2	(0.43 ± 0.15)	(0.36 ± 0.14)	2.90×10^{-5}
Myeloid dendritic cell resting	$(5.5 \times 10^{-4} \pm 2.3 \times 10^{-3})$	$(3.8 \times 10^{-3} \pm 0.02)$	0.05
Myeloid dendritic cell activated	$(1.4 \times 10^{-3} \pm 6.0 \times 10^{-3})$	$(3.9 \times 10^{-3} \pm 0.01)$	7.60×10^{-3}
Mast cell activated	(0.06 ± 0.08)	(0.08 ± 0.09)	0.02
Mast cell resting	(0.02 ± 0.04)	$(1.0 \times 10^{-2} \pm 0.03)$	0.04
Eosinophil	$(3.9 \times 10^{-5} \pm 4.6 \times 10^{-4})$	$(2.3 \times 10^{-4} \pm 1.8 \times 10^{-3})$	0.31
Neutrophil	$(8.1 \times 10^{-3} \pm 0.02)$	$(7.4 \times 10^{-3} \pm 0.02)$	0.59

Table 3. Comparison of immune cells in Low and High by CIBERSORT-ABS.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
B cell naive	$(2.4 \times 10^{-3} \pm 4.4 \times 10^{-3})$	$(4.9 \times 10^{-3} \pm 8.4 \times 10^{-3})$	0.03
B cell memory	$(1.8 \times 10^{-3} \pm 3.3 \times 10^{-3})$	$(5.4 \times 10^{-3} \pm 0.01)$	0.07
B cell plasma	$(8.4 \times 10^{-3} \pm 0.01)$	(0.02 ± 0.04)	2.40×10^{-4}
T cell CD8+	(0.03 ± 0.04)	(0.04 ± 0.04)	0.02
T cell CD4+ naive	$(2.0 \times 10^{-5} \pm 2.4 \times 10^{-4})$	$(2.8 \times 10^{-4} \pm 2.0 \times 10^{-3})$	0.31
T cell CD4+ memory resting	(0.05 ± 0.03)	(0.07 ± 0.05)	1.40×10^{-4}
T cell CD4+ memory activated	$(5.9 \times 10^{-5} \pm 7.1 \times 10^{-4})$	$(1.2 \times 10^{-4} \pm 1.3 \times 10^{-3})$	0.31
T cell follicular helper	$(4.5 \times 10^{-3} \pm 6.6 \times 10^{-3})$	$(7.2 \times 10^{-3} \pm 0.01)$	0.04
T cell regulatory (Tregs)	$(4.4 \times 10^{-3} \pm 4.9 \times 10^{-3})$	$(8.0 \times 10^{-3} \pm 9.7 \times 10^{-3})$	9.80×10^{-3}
T cell gamma delta	$(3.0 \times 10^{-3} \pm 0.01)$	$(2.6 \times 10^{-3} \pm 0.01)$	0.27
NK cell resting	$(5.5 \times 10^{-4} \pm 1.8 \times 10^{-3})$	$(1.1 \times 10^{-3} \pm 4.1 \times 10^{-3})$	0.67
NK cell activated	(0.01 ± 0.01)	(0.02 ± 0.01)	2.30×10^{-4}
Monocyte	$(8.6 \times 10^{-3} \pm 1.0 \times 10^{-2})$	(0.02 ± 0.02)	3.50×10^{-5}
Macrophage M0	(0.01 ± 0.02)	(0.01 ± 0.03)	0.74
Macrophage M1	$(3.7 \times 10^{-3} \pm 5.1 \times 10^{-3})$	$(8.7 \times 10^{-3} \pm 0.01)$	3.30×10^{-5}
Macrophage M2	(0.13 ± 0.09)	(0.14 ± 0.10)	0.33
Myeloid dendritic cell resting	$(2.5 \times 10^{-4} \pm 1.5 \times 10^{-3})$	$(1.9 \times 10^{-3} \pm 7.7 \times 10^{-3})$	0.04
Myeloid dendritic cell activated	$(2.7 \times 10^{-4} \pm 8.6 \times 10^{-4})$	$(1.1 \times 10^{-}3 \pm 2.9 \times 10^{-3})$	6.40×10^{-3}
Mast cell activated	(0.02 ± 0.02)	(0.03 ± 0.03)	1.80×10^{-4}
Mast cell resting	$(4.8 \times 10^{-3} \pm 0.01)$	$(3.6 \times 10^{-3} \pm 0.01)$	0.04
Eosinophil	$(2.2 \times 10^{-6} \pm 2.7 \times 10^{-5})$	$(5.7 \times 10^{-5} \pm 4.7 \times 10^{-4})$	0.31
Neutrophil	$(2.8 \times 10^{-3} \pm 7.4 \times 10^{-3})$	$(3.4 \times 10^{-3} \pm 8.0 \times 10^{-3})$	0.37

Table 4. Comparison of immune cells in Low and High by QUANTISEQ.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
B cell	$(2.9 \times 10^{-3} \pm 6.0 \times 10^{-3})$	$(5.9 \times 10^{-3} \pm 8.6 \times 10^{-3})$	2.80×10^{-8}
Macrophage M1	(0.02 ± 0.01)	(0.03 ± 0.02)	0.16
Macrophage M2	(0.05 ± 0.03)	(0.04 ± 0.03)	8.50×10^{-4}
Monocyte	(0.01 ± 0.02)	$(6.7 \times 10^{-3} \pm 0.02)$	0.07
Neutrophil	(0.08 ± 0.04)	(0.09 ± 0.04)	0.05
NK cell	$(0.01 \pm 8.0 \times 10^{-3})$	$(0.01 \pm 5.2 \times 10^{-3})$	0.52
T cell CD4+ (non-regulatory)	(0.04 ± 0.02)	(0.04 ± 0.02)	1.30×10^{-3}
T cell CD8+	$(8.9 \times 10^{-3} \pm 0.05)$	$(4.9 \times 10^{-3} \pm 0.01)$	0.06
T cell regulatory (Tregs)	$(2.9 \times 10^{-3} \pm 2.5 \times 10^{-3})$	$(6.0 \times 10^{-3} \pm 5.4 \times 10^{-3})$	8.80×10^{-11}
Myeloid dendritic cell	$(3.7 \times 10^{-3} \pm 6.9 \times 10^{-3})$	$(9.1 \times 10^{-3} \pm 0.01)$	1.50×10^{-6}
uncharacterized cell	(0.76 ± 0.07)	(0.76 ± 0.06)	4.20×10^{-1}

Table 5. Comparison of immune cells in Low and High by MCPCOUNTER.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
T cell	(34.95 ± 31.59)	(37.44 ± 30.90)	0.29
T cell CD8+	(4.60 ± 24.41)	(2.33 ± 3.56)	0.37
cytotoxicity score	(1.99 ± 5.29)	(2.21 ± 2.58)	0.01
NK cell	(0.16 ± 0.15)	(0.23 ± 0.24)	1.60×10^{-3}
B cell	(1.08 ± 1.05)	(2.92 ± 5.37)	7.60×10^{-9}
Monocyte	(26.67 ± 14.06)	(27.97 ± 16.34)	0.6
Macrophage/Monocyte	(26.67 ± 14.06)	(27.97 ± 16.34)	0.6
Myeloid dendritic cell	(1.89 ± 1.85)	(2.84 ± 2.71)	9.70×10^{-4}
Neutrophil	(11.49 ± 4.63)	(15.70 ± 6.34)	9.80×10^{-10}
Endothelial cell	(5.17 ± 3.49)	(8.43 ± 7.34)	3.10×10^{-11}
Cancer-associated fibroblast	(39.81 ± 36.08)	(109.45 ± 115.56)	5.1×10^{-17}

Table 6. Comparison of immune cells in Low and High by XCELL.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
Myeloid dendritic cell activated	(0.19 ± 0.08)	(0.20 ± 0.12)	0.74
B cell	(0.02 ± 0.02)	(0.02 ± 0.05)	3.50×10^{-4}
T cell CD4+ memory	$(2.5 \times 10^{-3} \pm 6.1 \times 10^{-3})$	$(3.9 \times 10^{-3} \pm 6.9 \times 10^{-3})$	8.90×10^{-3}
T cell CD4+ naive	$(1.2 \times 10^{-3} \pm 0.01)$	$(2.5 \times 10^{-3} \pm 0.01)$	0.12
T cell CD4+ (non-regulatory)	$(2.6 \times 10^{-5} \pm 3.2 \times 10^{-4})$	$(1.1 \times 10^{-18} \pm 1.9 \times 10^{-18})$	8.40×10^{-3}
T cell CD4+ central memory	$(9.1 \times 10^{-3} \pm 0.01)$	$(7.5 \times 10^{-3} \pm 0.01)$	0.47
T cell CD4+ effector memory	(0.06 ± 0.03)	(0.04 ± 0.03)	5.50×10^{-7}
T cell CD8+ naive	$(2.1 \times 10^{-3} \pm 3.4 \times 10^{-3})$	$(1.9 \times 10^{-3} \pm 4.1 \times 10^{-3})$	0.05
T cell CD8+	(0.02 ± 0.05)	(0.01 ± 0.03)	0.66
T cell CD8+ central memory	(0.02 ± 0.06)	(0.02 ± 0.04)	0.16
T cell CD8+ effector memory	$(4.0 \times 10^{-3} \pm 0.02)$	$(3.1 \times 10^{-3} \pm 0.01)$	0.4
Class-switched memory B cell	$(4.2 \times 10^{-3} \pm 6.3 \times 10^{-3})$	$(8.7 \times 10^{-3} \pm 0.02)$	0.07
Common lymphoid progenitor	(0.01 ± 0.01)	$(0.01 \pm 9.9 \times 10^{-3})$	1.70×10^{-3}
Common myeloid progenitor	$(4.0 \times 10^{-3} \pm 7.4 \times 10^{-3})$	$(3.3 \times 10^{-3} \pm 6.2 \times 10^{-3})$	0.93
Myeloid dendritic cell	$(8.0 \times 10^{-3} \pm 7.9 \times 10^{-3})$	$(9.5 \times 10^{-3} \pm 0.01)$	0.59
Endothelial cell	(0.03 ± 0.03)	(0.04 ± 0.06)	0.96
Eosinophil	(0.01 ± 0.01)	$(7.5 \times 10^{-3} \pm 7.7 \times 10^{-3})$	0.02
Cancer-associated fibroblast	(0.03 ± 0.03)	(0.06 ± 0.07)	3.40×10^{-5}
Granulocyte-monocyte progenitor	$(1.1 \times 10^{-3} \pm 3.5 \times 10^{-3})$	$(4.0 \times 10^{-3} \pm 9.0 \times 10^{-3})$	0.11
Hematopoietic stem cell	(0.06 ± 0.08)	(0.11 ± 0.13)	0.07
Macrophage	(0.06 ± 0.05)	(0.04 ± 0.04)	3.00×10^{-7}
Macrophage M1	(0.06 ± 0.04)	(0.05 ± 0.04)	8.30×10^{-5}
Macrophage M2	(0.07 ± 0.05)	(0.03 ± 0.04)	1.60×10^{-13}
Mast cell	$(7.8 \times 10^{-3} \pm 7.1 \times 10^{-3})$	$(7.4 \times 10^{-3} \pm 6.6 \times 10^{-3})$	0.76
B cell memory	$(4.7 \times 10^{-3} \pm 8.9 \times 10^{-3})$	$(6.9 \times 10^{-3} \pm 0.02)$	2.00×10^{-3}
Monocyte	(0.03 ± 0.03)	(0.02 ± 0.03)	0.05
B cell naive	$(2.0 \times 10^{-3} \pm 2.8 \times 10^{-3})$	$(2.3 \times 10^{-3} \pm 6.4 \times 10^{-3})$	0.01
Neutrophil	$(3.1 \times 10^{-4} \pm 1.7 \times 10^{-3})$	$(5.9 \times 10^{-4} \pm 1.9 \times 10^{-3})$	4.40×10^{-3}
NK cell	$(4.1 \times 10^{-18} \pm 7.5 \times 10^{-18})$	$(2.6 \times 10^{-18} \pm 4.7 \times 10^{-18})$	0.13
T cell NK	(0.02 ± 0.02)	(0.02 ± 0.02)	0.83
Plasmacytoid dendritic cell	(0.01 ± 0.01)	$(6.7 \times 10^{-3} \pm 0.01)$	2.80×10^{-11}
B cell plasma	$(6.6 \times 10^{-3} \pm 5.4 \times 10^{-3})$	$(7.0 \times 10^{-3} \pm 0.01)$	5.90×10^{-3}
T cell gamma delta	$(4.8 \times 10^{-4} \pm 4.8 \times 10^{-3})$	$(6.7 \times 10^{-6} \pm 6.9 \times 10^{-5})$	0.29
T cell CD4+ Th1	(0.02 ± 0.02)	$(9.9 \times 10^{-3} \pm 0.01)$	3.50×10^{-7}
T cell CD4+ Th2	(0.05 ± 0.04)	(0.06 ± 0.07)	0.32
T cell regulatory (Tregs)	$(1.4 \times 10^{-3} \pm 4.3 \times 10^{-3})$	$(1.4 \times 10^{-3} \pm 4.1 \times 10^{-3})$	0.91
immune score	(0.10 ± 0.09)	(0.08 ± 0.09)	1.00×10^{-3}
stroma score	(0.03 ± 0.03)	(0.05 ± 0.05)	1.10×10^{-4}
microenvironment score	(0.13 ± 0.09)	(0.13 ± 0.11)	0.87

Table 7. Comparison of immune cells in Low and High by EPIC.

Cell type	Low (Mean ± std)	High (Mean ± std)	rank-sum test
B cell	$(2.2 \times 10^{-3} \pm 2.4 \times 10^{-3})$	$(5.2 \times 10^{-3} \pm 9.6 \times 10^{-3})$	2.00×10^{-7}
Cancer-associated fibroblast	$(4.2 \times 10^{-3} \pm 3.5 \times 10^{-3})$	(0.01 ± 0.02)	1.00×10^{-15}
T cell CD4+	(0.05 ± 0.02)	(0.05 ± 0.02)	6.50×10^{-4}
T cell CD8+	(0.02 ± 0.03)	$(0.02 \pm 8.0 \times 10^{-3})$	1.10×10^{-4}
Endothelial cell	(0.02 ± 0.02)	(0.03 ± 0.03)	1.30×10^{-9}
Macrophage	(0.03 ± 0.04)	(0.02 ± 0.03)	7.70×10^{-4}
NK cell	$(4.4 \times 10^{-5} \pm 2.0 \times 10^{-4})$	$(1.9 \times 10^{-4} \pm 9.1\times 10^{-4})$	2.00×10^{-2}
uncharacterized cell	(0.88 ± 0.06)	(0.86 ± 0.06)	3.00×10^{-5}

3.8. TIMER analysis

The TIMER database was used to investigate the link between immune cells and signatures. CD44, DDR2, FBLN1 and LUM were found to be favorably correlated with B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils and dendritic cells. LAMA1 was positively correlated with macrophages only, while it was negatively correlated with other cells. Meanwhile, ADAMTS6, HSPG2 and TIMP1 were only negatively correlated with macrophages. THBS2 was only negatively correlated with neutrophils. CTSD was negatively correlated with neutrophils and CD4+ T cells and positively correlated with other immune cells. Except for LUM, the remaining nine genes were negatively associated with purity (Supplement Figure S3).

3.9. Transcription levels of the signature and small molecule drugs identification

We explored the differences in methylation levels of 10 signature-related BMs between PRCC patients and normal controls using the UALCAN database. The results showed that ADAMTS6, CTSD, DDR2, HSPG2 and THBS2 had higher methylation levels in PRCC patients compared to normal kidney tissue; meanwhile, CD44, FBLN1 and LAMA1 gene methylation levels were lower in PRCC patients ($p < 0.05$) (Figure 13). Methylation levels of the TIMP1 gene were not significantly different between normal controls and PRCC patients and no information available for LUM was found. In addition, we explored the prognostic impact of high and low expression levels of these 10 genes. The results indicated that high expression of THBS2, LUM, HSPG2, FBLN1, DDR2 and ADAMTS6 genes in PRCC patients had a lower survival rate ($p < 0.05$) (Figure 14).

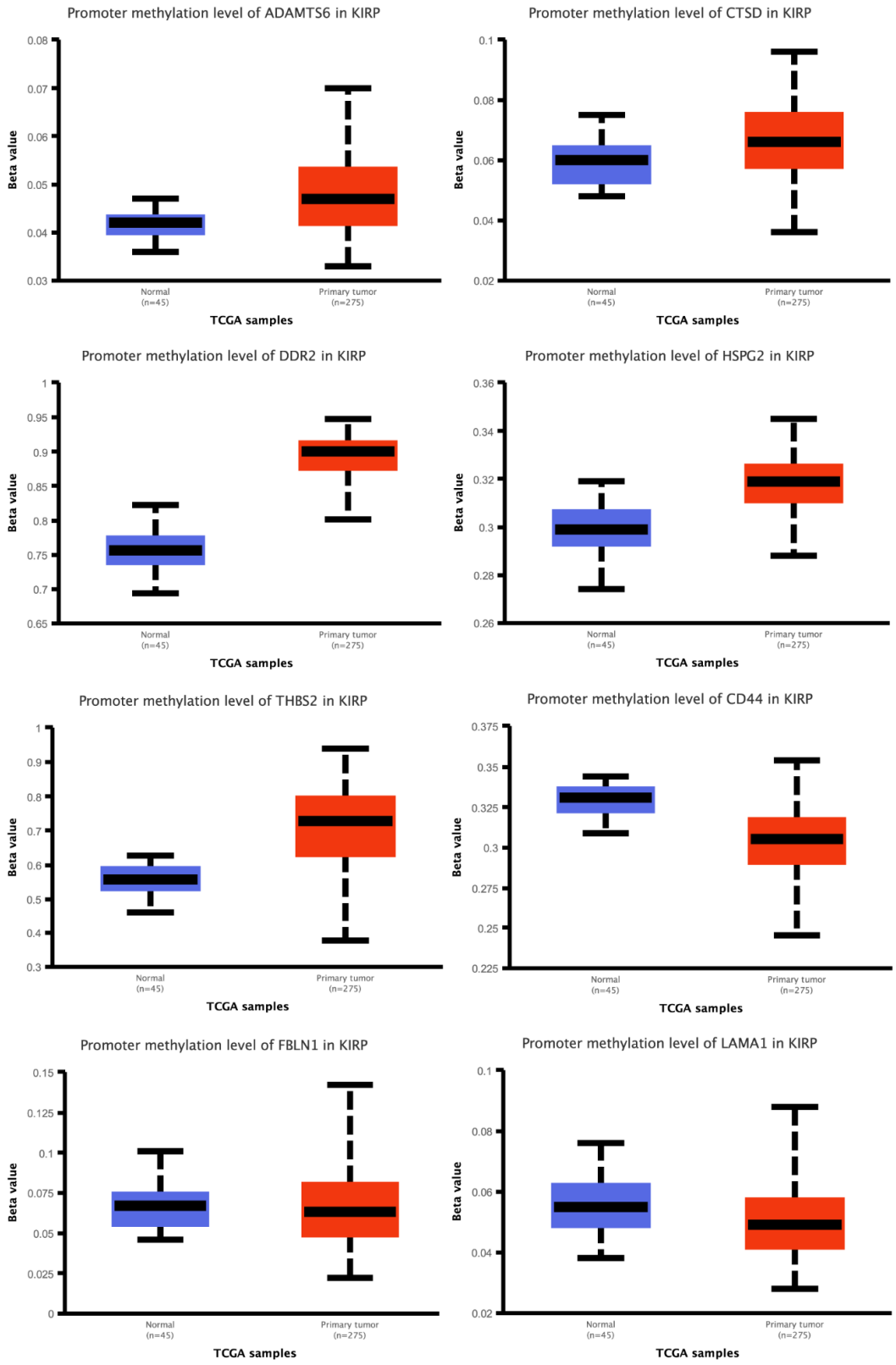

Figure 13. Boxplots showing the methylation levels of BM-related 8-signature genes between PRCC patients and healthy controls.

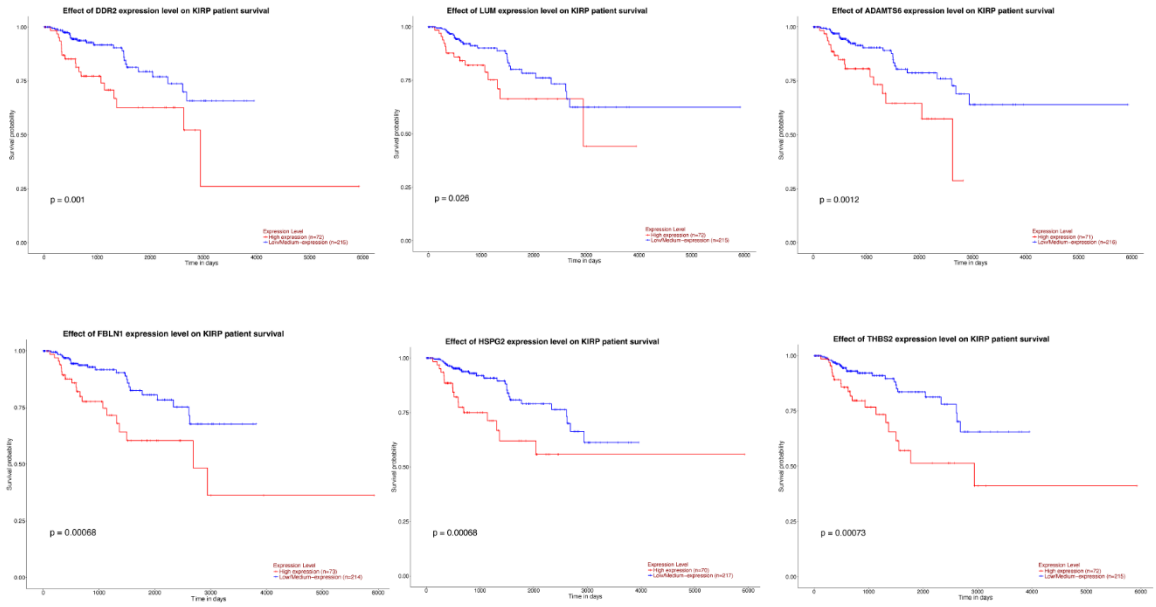

Figure 14. KM survival curve of prognostic differences in PRCC patients between high and low expression levels of signature genes in PRCC patients.

The 9 most potential small molecule drugs on the basis of BMs were obtained through the DSigDB database, including cytarabine, LAMININ, progesterone, Retinoic acid, genistein, Aziridine, alpha-Neu5Ac, T-2 TOXIN, carbamazepine, and all 9 drugs had p-values < 0.02 (Table 8).

Table 8. The 9 small molecule drugs of the DsigDB dataset analyses results.

Name of drugs	Overlap	Adjust P-value	Odds Ratio	Combined Score
cytarabine	5/594	0.0035	32.94	401.66
LAMININ	3/197	0.0186	43.73	399.69
progesterone	6/1915	0.0186	14.21	128.95
Retinoic acid	8/4258	0.0186	14.81	133.21
genistein	5/1231	0.0186	15.31	132.84
Aziridine	2/42	0.0186	124.69	1067.29
alpha-Neu5Ac	2/47	0.0186	110.81	923.40
T-2 TOXIN	2/47	0.0186	110.81	923.40
carbamazepine	2/47	0.0186	110.81	923.40

3.10. Verification of gene expression

The external data set of the GEO database was selected for verification. The comparative analysis found the expression of 5 genes (CD44, CTSD, FBLN1, LUM and TIMPI) was lower in the case than that in normal ($P < 0.05$). The expression levels of DDR2, HSPG2 and LAMA1 genes were higher in the case than those in normal ($P < 0.05$) (Figure 15).

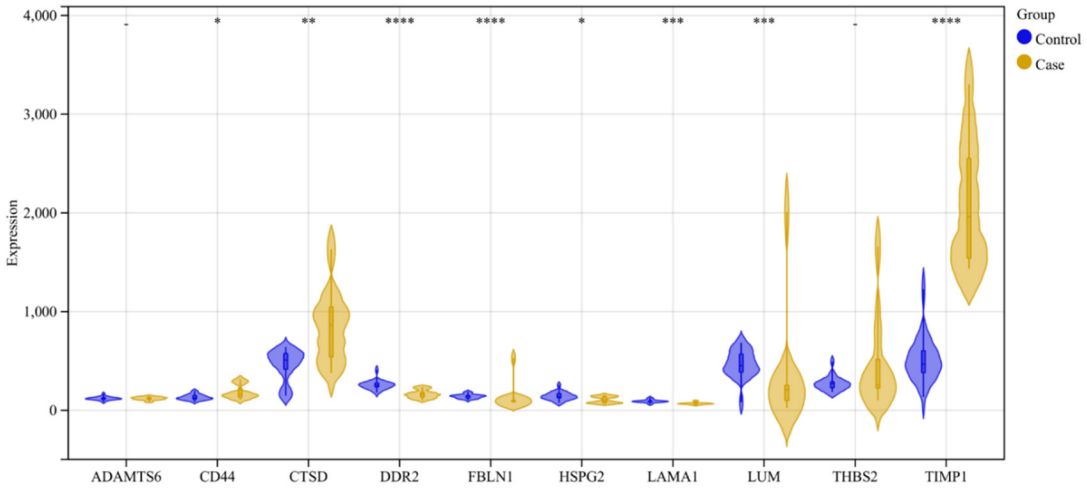

Figure 15. GSE15641 was used to validate the expression levels of BMs. Comparisons between two sets of data used non-parametric tests. P-values < 0.05 were considered statistically significant. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

4. Discussion

PRCC is the second most common subtype of RCC which originates in the renal tubules of the kidney. The BM is an important component of the proximal tubules of the kidney and has been shown to undergo structural and functional changes in most renal diseases. It also has been found to play an important role in the development of cancers [25]. Little is known so far about the molecular pathology of PRCC. Surgery and a number of alternative therapies, including immunotherapy, have been widely used in PRCC. The limited response rates and unsatisfactory outcomes have prompted further exploration of more appropriate approaches to improve treatment efficiency and achieve more personalized treatment from a predictive prognostic perspective [26]. However, no studies have comprehensively analyzed the prognostic significance of BM in PRCC.

In this study, we have comprehensively investigated the potential and prognostic value of BMs in PRCC using bioinformatic analysis. For the PRCC gene expression profiles acquired from TCGA database, we first looked for 98 BMs with significant differences in expression between PRCC tissue and healthy kidney tissue. The biological functions of the differentially expressed BMs in PRCC were then investigated by enrichment analysis. GO and KEGG pathway enrichment analyses were performed to explore the functions and pathways of BMs. The KEGG enrichment analysis revealed that these BMs were considerably enriched in the ECM-receptor interaction, focal adhesion and PI3K-Akt signaling pathways. ECM-receptor interaction and focal adhesion signaling pathways were primarily engaged in cancer metastasis and tumor invasion, whereas the PI3K-Akt signaling pathway regulates cell survival, cell cycle progression and cellular growth [27–29]. The results of the GO analysis indicated a considerable enrichment of collagen binding, BM and cell adhesion-related protein complexes. These enriched items were essentially ECM components that played a major role in the pathogenesis of cancer by promoting cell growth, invasion, migration and angiogenesis [30].

Using univariable and lasso-penalized Cox regression analysis, 10 BMs associated with the PRCC prognosis were discovered. We created and validated an outcome-related risk model based on these 10

BMs. ROC and survival evaluations were used to demonstrate the model's predictive power. Finally, the results of univariable and multivariable Cox analyses indicated that the signature based on 10 BMs was an independent prognostic indicator for PRCC patients. Additionally, we discovered that the signature was closely associated with immune cell infiltration and identified 3 chemotherapy agents and 9 small-molecule medicines for the treatment of PRCC patients.

ADAM Metallopeptidase with Thrombospondin Type 1 Motif 6 (ADAMTS6) is a member of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) protein family, which plays multiple functions in tissue morphogenesis and pathophysiological remodeling, inflammation and vascular biology. They have been connected to the development and progression of cancer, and both promoting and inhibiting effects have been identified. The majority of members of this class possess anti-angiogenic properties that prevent the progression of cancer. However, ADAMTS6 is elevated in certain cancers, including breast cancer, and some studies have hypothesized that this might be because ADAMTS6 produces a microenvironment that promotes tumor growth [31]. Notably, ADAMTS6 expression was similarly elevated in this study in PRCC and might be involved in the pathogenesis of PRCC by a mechanism similar to that described above. CD44 is a key ECM component and a co-receptor for numerous growth factors and cytokines [32]. Increasing evidence has suggested that CD44, particularly CD44v isoforms, were cancer stem cell (CSC) markers and key regulators of CSC features, including self-renewal, tumor initiation, metastasis and chemoradioresistance. A relevant meta-analysis has found that CD44 expression corresponded with the clinical characteristics of RCC and could be utilized as a marker to predict postoperative prognosis and tumor development [33]. Currently, target CD44 has shown significant promise for the treatment of life-threatening malignancies since they had the potential to eliminate the cancer stem cell (CSC) population [34]. Cathepsin D (CTSD) is a ubiquitously distributed lysosomal aspartic endoprotease. Numerous studies have demonstrated that the overexpression and hypersecretion of CTSD increased in numerous forms of cancer, such as breast cancer, lung cancer and prostate cancer. In contrast, Merseburger et al. reported an inverse association between CTSD expression and the advancement of RCC tumors, which was consistent with the findings of our investigation [35]. DDR2 is one of the receptors involved in tumor-ECM interactions [36]. Its alterations, including overexpression, amplification and mutations, have been observed in many cancer types and, in many instances, were known to promote an aggressive phenotype. Current research has indicated that targeting DDR2 could effectively boost the therapeutic efficacy of anti-PD-1 immunotherapy in tumor therapy [37]. There has been research on the application of anti-angiogenic tyrosine kinase inhibitors and immunotherapy combinations for the treatment of RCC with promising results in terms of feasibility and safety [38]. Fibulin 1 (FBLN1) is a multifunctional glycoprotein that functions as an ECM stabilizer by interacting with other ECM proteins. By modulating the interaction between malignant cells and the microenvironment, it played a crucial role in tumor development and metastasis [39]. Xiao et al. [40] postulated that dysregulation of FBLN1 was linked to the advancement of RRC, and they identified its tumor suppressor and angiogenesis inhibitor capabilities. Heparan Sulfate Proteoglycan 2 (HSPG2) is a large multi-domain ECM proteoglycan that has a role in the invasion, metastasis, and angiogenesis of solid tumors. Currently, high expression of HSPG2 has been indicated as a potential indication of prostate cancer grade, invasion potential and distant metastasis; however, there is no published research for renal papillary carcinoma [41]. Laminin Subunit Alpha 1 (LAMA1) codes for one of laminin's alpha 1 subunit. LAMA1 has been demonstrated to have prognostic significance in RCC in prior research [42]. Lumican (LUM) belongs to the small leucine-rich proteoglycan family and functions as

both an oncogene and a tumor suppressor gene [43]. Nonetheless, the impact of LUM on RCC has not been reported in detail. Thrombospondin 2 (THBS2) is a member of the matricellular calcium-binding glycoprotein family and interacts with growth factors, cell receptors and ECM. By interacting with matrix metalloproteinases (MMPs) and matrix serine proteases, THBS2 was known to behave as a strong inhibitor of tumor development and angiogenesis in a number of cancer types [44]. Down-regulation of THBS2 expression has been observed in numerous malignancies, including hepatocellular carcinoma and gastric cancer [44,45]. TIMP Metallopeptidase Inhibitor 1 (TIMP1) is a prominent member of the TIMP family that can inhibit the proteolytic activity of MMPs and regulate the equilibrium of matrix remodeling during the degradation of ECM in order to prevent tumor migration. Shou et al. believed that TIMP1 overexpression was an independent prognostic factor of RCC patients and that it could accelerate tumor growth through the epithelial-to-mesenchymal transition (EMT) signaling pathway [46].

The development, treatment and prognosis of PRCC were closely linked to immune cell infiltration. Increasing evidence has indicated that the BM has immunomodulatory effects, particularly its major component, laminins, which could influence the proliferation, migration and function of immune cells [47]. As a result, we examined the association between risk score and immune cell infiltration. The results revealed that the high-risk group had relatively greater levels of B cells, iDCs, Tfh cells and Treg cell infiltration, with the high levels of Treg cell infiltration suggesting that the high-risk group might have a suppressive anti-tumor T-cell response, which could be connected with their poor prognosis. Great immunological scores in the high-risk group indicated that immunotherapy might be more successful in this group. The substantial infiltration of macrophages in the low-risk group suggested that tumor aggressiveness might be quite high. However, the present research has only reported the role of M2 macrophages in cancer formation; therefore, additional confirmation of the type of infiltrating macrophages was required. ICI are currently a popular treatment option for renal clear cell carcinoma patients [48]. The expressions of more than 20 immune checkpoints, including PDCD1, BTLA, CD27, CD40, CD44, IDO1 and so on, were higher in the high-risk group of PRCC patients than in the low-risk group, suggesting that the unfavorable prognosis in the high-risk group might be due to the immunosuppressive microenvironment. Therefore, patients with PRCC in the high-risk group might benefit from checkpoint inhibitor immunotherapy [6].

Furthermore, we discovered via drug sensitivity analysis that high-risk patients benefited more from axitinib, pazopanib and erlotinib, which have some reference value for directing therapeutic medicine [49,50]. Pazopanib has not yet been documented for the treatment of PRCC, which might serve as a chemotherapeutic agent for PRCC in the future. Then, we forecasted BM-related small molecule medicines for prospective usage in the management of papillary renal cell cancer. Cytarabine was a pyrimidine antimetabolite that mainly acts on the S proliferative phase of cells and interferes with cell proliferation by inhibiting the synthesis of cellular DNA. It was mainly used for acute leukemia and has a certain effect on malignant lymphoma, lung cancer, digestive tract cancer, head and neck cancer and other cancers [51]. LAMININ was a non-collagenous sugar that constitutes the interstitial space of cells and was mainly synthesized by endothelial cells and fat-storing cells in the liver, constituting a component of the BM together with collagen. Its biological function was that cells adhere to mediators of the matrix and bind to a variety of BM components to regulate cell growth and differentiation. These small molecule drugs might have some help and clinical significance in the treatment of PRCC.

BM played an important role in kidney function as an important component of the kidney [52,53].

However, there were few studies between BM and PRCC in recent years. We were the first to explore the potential role of BM in PRCC by studying the expression, prognosis and immune infiltration of BMs in PRCC. The prognostic model and risk score constructed based on BMs could be used to predict the prognosis of PRCC patients. Meanwhile, our findings may provide ideas and theoretical basis for future studies on the molecular mechanisms and targeted therapies of BMs in PRCC [54].

Some limitations existed in our study. The mechanisms regarding the function of BMs and the regulation of the biological behavior of PRCC cells should be validated by in vivo and in vitro experiments. Due to the small number of PRCC samples in the relevant database and the incomplete clinical data, we did not conduct external validation of the model. Moreover, few related studies between BM and PRCC could give sufficient evidence to prove the role of BMs in PRCC, and thus, more research needs to focus on this in the future.

Data availability

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

References

1. G. Courthod, M. Tucci, M. Di Maio, G. V. Scagliotti, Papillary renal cell carcinoma: A review of the current therapeutic landscape, *Crit. Rev. Oncol./Hematol.*, **96** (2015), 100–112. https://doi.org/10.1016/j.critrevonc.2015.05.008
2. N. Mendhiratta, P. Muraki, A. E. Sisk Jr, B. Shuch, Papillary renal cell carcinoma: Review, *Urol. Oncol.: Semin. Orig. Invest.*, **39** (2021), 327–337. https://doi.org/10.1016/j.urolonc.2021.04.013
3. J. Cheng, Z. Han, R. Mehra, W. Shao, M. Cheng, Q. Feng, et al., Computational analysis of pathological images enables a better diagnosis of TFE3 Xp11.2 translocation renal cell carcinoma, *Nat. Commun.*, **11** (2020), 1778. https://doi.org/10.1038/s41467-020-15671-5
4. S. Steffens, M. Janssen, F. C. Roos, F. Becker, S. Schumacher, C. Seidel, et al., Incidence and long-term prognosis of papillary compared to clear cell renal cell carcinoma-a multicentre study, *Eur. J. Cancer*, **48** (2012), 2347–2352. https://doi.org/10.1016/j.ejca.2012.05.002
5. Q. Chen, L. Cheng, Q. Li, The molecular characterization and therapeutic strategies of papillary renal cell carcinoma, *Expert Rev. Anticancer Ther.*, **19** (2019), 169–175. https://doi.org/10.1080/14737140.2019.1548939
6. M. de Vries-Brilland, D. F. McDermott, C. Suárez, T. Powles, M. Gross-Goupil, A. Ravaud, et al., Checkpoint inhibitors in metastatic papillary renal cell carcinoma, *Cancer Treat. Rev.*, **99** (2021), 102228. https://doi.org/10.1016/j.ctrv.2021.102228
7. R. Reuten, S. Zendehroud, M. Nicolau, L. Fleischhauer, A. Laitala, S. Kiderlen, et al., Basement membrane stiffness determines metastases formation, *Nat. Mater.*, **20** (2021), 892–903. https://doi.org/10.1038/s41563-020-00894-0
8. S. E. Wilson, A. Torricelli, G. K. Marino, Corneal epithelial basement membrane: Structure, function and regeneration, *Exp. Eye Res.*, **194** (2020), 108002. https://doi.org/10.1016/j.exer.2020.108002

9. N. Khalilgharibi, Y. Mao, To form and function: on the role of basement membrane mechanics in tissue development, homeostasis and disease, *Open Biol.*, **11** (2021), 200360. https://doi.org/10.1098/rsob.200360
10. F. Kai, A. P. Drain, V. M. Weaver, The extracellular matrix modulates the metastatic journey, *Dev. Cell*, **49** (2019), 332–346. https://doi.org/10.1016/j.devcel.2019.03.026
11. R. W. Naylor, M. Morais, R. Lennon, Complexities of the glomerular basement membrane, *Nat. Rev. Nephrol.*, **17** (2021), 112–127. https://doi.org/10.1038/s41581-020-0329-y
12. R. Jayadev, M. Morais, J. M. Ellingford, S. Srinivasan, R. W. Naylor, C. Lawless, et al., A basement membrane discovery pipeline uncovers network complexity, regulators, and human disease associations, *Sci. Adv.*, **8** (2022), 2265. https://doi.org/10.1126/sciadv.abn2265
13. D. Szklarczyk, A. L. Gable, K. C. Nastou, D. Lyon, R. Kirsch, S. Pyysalo, et al., The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, *Nucleic Acids Res.*, **49** (2021), 605–612. https://doi.org/10.1093/nar/gkaa1074
14. D. Szklarczyk, A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, et al., STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, *Nucleic Acids Res.*, **47** (2019), 607–613. https://doi.org/10.1093/nar/gky1131
15. D. Warde-Farley, S. L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, et al., The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function, *Nucleic Acids Res.*, **38** (2010), 214–220. https://doi.org/10.1093/nar/gkq537
16. G. Zhou, O. Soufan, J. Ewald, R. Hancock, N. Basu, J. Xia, NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis, *Nucleic Acids Res.*, **47** (2019), 234–241. https://doi.org/10.1093/nar/gkz240
17. L. Danilova, W. J. Ho, Q. Zhu, T. Vithayathil, A. De Jesus-Acosta, N. S. Azad, et al., Programmed cell death Ligand-1 (PD-L1) and CD8 expression profiling identify an immunologic subtype of pancreatic ductal adenocarcinomas with favorable survival, *Cancer Immunol. Res.*, **7** (2019), 886–895. https://doi.org/10.1158/2326-6066.CIR-18-0822
18. T. Li, J. Fan, B. Wang, N. Traugh, Q. Chen, J. S. Liu, et al., TIMER: A web server for comprehensive analysis of tumor-infiltrating immune cells, *Cancer Res.*, **77** (2017), 108–110. https://doi.org/10.1158/0008-5472.CAN-17-0307
19. B. Li, E. Severson, J. C. Pignon, H. Zhao, T. Li, J. Novak, et al., Comprehensive analyses of tumor immunity: implications for cancer immunotherapy, *Genome Biol.*, **17** (2016), 174. https://doi.org/10.1186/s13059-016-1028-7
20. E. Y. Chen, C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. V. Meirelles, et al., Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool, *BMC Bioinformatics*, **14** (2013), 128. https://doi.org/10.1186/1471-2105-14-128
21. M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, et al., Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, *Nucleic Acids Res.*, **44** (2016), 90–97. https://doi.org/10.1093/nar/gkw377
22. Z. Xie, A. Bailey, M. V. Kuleshov, D. Clarke, J. E. Evangelista, S. L. Jenkins, et al., Gene set knowledge discovery with enrichr, *Curr. Protoc.*, **1** (2021), 90. https://doi.org/10.1002/cpz1.90

23. D. S. Chandrashekar, S. K. Karthikeyan, P. K. Korla, H. Patel, A. R. Shovon, M. Athar, et al., UALCAN: An update to the integrated cancer data analysis platform, *Neoplasia*, **25** (2022), 18–27. https://doi.org/10.1016/j.neo.2022.01.001
24. D. S. Chandrashekar, B. Bashel, S. Balasubramanya, C. J. Creighton, I. Ponce-Rodriguez, B. Chakravarthi, et al., UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses, *Neoplasia*, **19** (2017), 649–658. https://doi.org/10.1016/j.neo.2017.05.002
25. F. Chen, Y. Zhang, Y. Şenbabaoğlu, G. Ciriello, L. Yang, E. Reznik, et al., Multilevel genomics-based taxonomy of renal cell carcinoma, *Cell Rep.*, **14** (2016), 2476–2489. https://doi.org/10.1016/j.celrep.2016.02.024
26. T. Klatte, K. M. Gallagher, L. Afferi, A. Volpe, N. Kroeger, S. Ribback, et al., The VENUSS prognostic model to predict disease recurrence following surgery for non-metastatic papillary renal cell carcinoma: development and evaluation using the ASSURE prospective clinical trial cohort, *BMC Med.*, **17** (2019), 182. https://doi.org/10.1186/s12916-019-1419-1
27. Y. Bao, L. Wang, L. Shi, F. Yun, X. Liu, Y. Chen, et al., Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer, *Cell. Mol. Biol. Lett.*, **24** (2019), 38. https://doi.org/10.1186/s11658-019-0162-0
28. J. Shen, B. Cao, Y. Wang, C. Ma, Z. Zeng, L. Liu, et al., Hippo component YAP promotes focal adhesion and tumour aggressiveness via transcriptionally activating THBS1/FAK signalling in breast cancer, *J. Exp. Clin. Cancer. Res.*, **37** (2018), 175. https://doi.org/10.1186/s13046-018-0850-z
29. J. A. Fresno Vara, E. Casado, J. de Castro, P. Cejas, C. Belda-Iniesta, M. González-Barón, PI3K/Akt signalling pathway and cancer, *Cancer Treat. Rev.*, **30** (2004), 193–204. https://doi.org/10.1016/j.ctrv.2003.07.007
30. C. Xue, G. Li, J. Lu, L. Li, Crosstalk between circRNAs and the PI3K/AKT signaling pathway in cancer progression, *Signal Transduct. Target. Ther.*, **6** (2021), 400. https://doi.org/10.1038/s41392-021-00788-w
31. R. Kelwick, I. Desanlis, G. N. Wheeler, D. R. Edwards, The ADAMTS (A Disintegrin and Metalloproteinase with Thrombospondin motifs) family, *Genome Biol.*, **16** (2015), 113. https://doi.org/10.1186/s13059-015-0676-3
32. M. M. Gomari, M. Farsimadan, N. Rostami, Z. Mahmoudi, M. Fadaie, I. Farhani, et al., CD44 polymorphisms and its variants, as an inconsistent marker in cancer investigations, *Mutat. Res. Rev. Mutat. Res.*, **787** (2021), 108374. https://doi.org/10.1016/j.mrrev.2021.108374
33. X. Li, X. Ma, L. Chen, L. Gu, Y. Zhang, F. Zhang, et al., Prognostic value of CD44 expression in renal cell carcinoma: a systematic review and meta-analysis, *Sci. Rep.*, **5** (2015), 13157. https://doi.org/10.1038/srep13157
34. Y. Yan, X. Zuo, D. Wei, Concise review: Emerging role of CD44 in cancer stem cells: A promising biomarker and therapeutic target, *Stem Cells Transl. Med.*, **4** (2015), 1033–1043. https://doi.org/10.5966/sctm.2015-0048
35. A. S. Merseburger, J. Hennenlotter, P. Simon, P. A. Ohneseit, U. Kuehs, S. Kruck, et al., Cathepsin D expression in renal cell cancer-clinical implications, *Eur. Urol.*, **48** (2005), 519–526. https://doi.org/10.1016/j.eururo.2005.03.019
36. I. Romayor, M. L. García-Vaquero, J. Márquez, B. Arteta, R. Barceló, A. Benedicto, Discoidin domain receptor 2 expression as worse prognostic marker in invasive breast cancer, *Breast J.*, **2022** (2022), 5169405. https://doi.org/10.1155/2022/5169405

37. M. M. Tu, F. Lee, R. T. Jones, A. K. Kimball, E. Saravia, R. F. Graziano, et al., Targeting DDR2 enhances tumor response to anti-PD-1 immunotherapy, *Sci. Adv.*, **5** (2019), 2437. https://doi.org/10.1126/sciadv.aav2437
38. A. L. Laccetti, B. Garmezy, L. Xiao, M. Economides, A. Venkatesan, J. Gao, et al., Combination antiangiogenic tyrosine kinase inhibition and anti-PD1 immunotherapy in metastatic renal cell carcinoma: A retrospective analysis of safety, tolerance, and clinical outcomes, *Cancer Med.*, **10** (2021), 2341–2349. https://doi.org/10.1002/cam4.3812
39. M. M. Watany, N. M. Elmashad, R. Badawi, N. Hawash, Serum FBLN1 and STK31 as biomarkers of colorectal cancer and their ability to noninvasively differentiate colorectal cancer from benign polyps, *Clin. Chim. Acta.*, **483** (2018), 151–155. https://doi.org/10.1016/j.cca.2018.04.038
40. W. Xiao, J. Wang, H. Li, W. Guan, D. Xia, G. Yu, et al., Fibulin-1 is down-regulated through promoter hypermethylation and suppresses renal cell carcinoma progression, *J. Urol.*, **190** (2013), 291–301. https://doi.org/10.1016/j.juro.2013.01.098
41. X. Zhou, S. Liang, Q. Zhan, L. Yang, J. Chi, L. Wang, HSPG2 overexpression independently predicts poor survival in patients with acute myeloid leukemia, *Cell Death Dis.*, **11** (2020), 492. https://doi.org/10.1038/s41419-020-2694-7
42. J. W. Wragg, J. P. Finnity, J. A. Anderson, H. J. Ferguson, E. Porfiri, R. I. Bhatt, et al., MCAM and LAMA4 are highly enriched in tumor blood vessels of renal cell carcinoma and predict patient outcome, *Cancer Res.*, **76** (2016), 2314–2326. https://doi.org/10.1158/0008-5472.CAN-15-1364
43. X. Chen, X. Li, X. Hu, F. Jiang, Y. Shen, R. Xu, et al., LUM expression and its prognostic significance in gastric cancer, *Front. Oncol.*, **10** (2020), 605. https://doi.org/10.3389/fonc.2020.00605
44. K. Y. Ng, Q. T. Shea, T. L. Wong, S. T. Luk, M. Tong, C. M. Lo, et al., Chemotherapy-Enriched THBS2-Deficient cancer stem cells drive hepatocarcinogenesis through matrix softness induced histone H3 modifications, *Adv. Sci.*, **8** (2021), 2002483. https://doi.org/10.1002/advs.202002483
45. S. Zhang, H. Yang, X. Xiang, L. Liu, H. Huang, G. Tang, THBS2 is closely related to the poor prognosis and immune cell infiltration of gastric cancer, *Front. Genet.*, **13** (2022), 803460. https://doi.org/10.3389/fgene.2022.803460
46. Y. Shou, Y. Liu, J. Xu, J. Liu, T. Xu, J. Tong, et al., TIMP1 indicates poor prognosis of renal cell carcinoma and accelerates tumorigenesis via EMT signaling pathway, *Front. Genet.*, **13** (2022), 648134. https://doi.org/10.3389/fgene.2022.648134
47. T. Simon, J. S. Bromberg, Regulation of the immune system by laminins, *Trends Immunol.*, **38** (2017), 858–871. https://doi.org/10.1016/j.it.2017.06.002
48. Z. Gong, J. Xie, L. Chen, Q. Tang, Y. Hu, A. Xu, et al., Integrative analysis of TRPV family to prognosis and immune infiltration in renal clear cell carcinoma, *Channels*, **16** (2022), 84–96. https://doi.org/10.1080/19336950.2022.2058733
49. S. Negrier, N. Rioux-Leclercq, C. Ferlay, M. Gross-Goupil, G. Gravis, L. Geoffrois, et al., Axitinib in first-line for patients with metastatic papillary renal cell carcinoma: Results of the multicentre, open-label, single-arm, phase II AXIPAP trial, *Eur. J. Cancer*, **129** (2020), 107–116. https://doi.org/10.1016/j.ejca.2020.02.001
50. M. S. Gordon, M. Hussey, R. B. Nagle, P. N. Lara Jr, P. C. Mack, J. Dutcher, et al., Phase II study of erlotinib in patients with locally advanced or metastatic papillary histology renal cell cancer: SWOG S0317, *J. Clin. Oncol.*, **27** (2009), 5788–5793. https://doi.org/10.1200/JCO.2008.18.8821

51. J. E. Megías-Vericat, D. Martínez-Cuadrón, A. Solana-Altabella, J. L. Poveda, P. Montesinos, Systematic review of pharmacogenetics of ABC and SLC transporter genes in acute myeloid leukemia, *Pharmaceutics*, **14** (2022), 878. https://doi.org/10.3390/pharmaceutics14040878
52. M. Morais, P. Tian, C. Lawless, S. Murtuza-Baker, L. Hopkinson, S. Woods, et al., Kidney organoids recapitulate human basement membrane assembly in health and disease, *Elife*, **11** (2022), 73486. https://doi.org/10.7554/eLife.73486
53. S. Mikami, M. Oya, R. Mizuno, T. Kosaka, K. Katsube, Y. Okada, Invasion and metastasis of renal cell carcinoma, *Med. Mol. Morphol.*, **47** (2014), 63–67. https://doi.org/10.1007/s00795-013-0064-6
54. Y. Li, H. Liu, H. Yan, J. Xiong, Research advances on targeted-Treg therapies on immune-mediated kidney diseases, *Autoimmun. Rev.*, **22** (2023), 103257. https://doi.org/10.1016/j.autrev.2022.103257

Supplementary Materials

See Figures S1–3 in the Supplementary Material for comprehensive image analysis.