Identification of prognostic signature of cancer-associated fibroblasts associated with castration resistance prostate cancer based on Weighted Gene Co-expression Network Analysis

Wenjun Zhang  
Guizhou Medical University

Bangwei Che  
Guizhou Medical University

Miao Liu  
Guizhou Medical University

Wei Li  
Guizhou Medical University

Tao Huang  
Guizhou Medical University

Ying Yu  
Guizhou Medical University

Shenghan Xu  
Guizhou Medical University

Jun He  
Guizhou Medical University

Cheng Zha  
Guizhou Medical University

Zheng Peng  
Guizhou Medical University

Kunyuan Huang  
Guizhou Medical University

Kaifa Tang (✉ tangkaifa@gmc.edu.cn)  
Guizhou Medical University

Method Article

**Keywords:** Prostate cancer, Cancer associated fibroblasts, Tumor microenvironment, Weighted gene co-expression network analysis
Abstract

Prostate cancer (PCa) is the most common cancer in men and often progresses to castration resistant prostate cancer (CRPC) after treatment, with a poor prognosis. Cancer associated fibroblasts (CAF) are a major components of tumor microenvironment (TME), which participate in angiogenesis and immunosuppression, promote metastasis and treatment drug resistance. In order to identify the CAF prognostic genes associated with CRPC, the RNA sequencing data of 745 PCa patients from Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases were downloaded. The CAF-related hub genes were identified by weighted gene co-expression network analysis (WCGNA). The CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and signature were developed by machine learning methodology. Meanwhile, verified the CAF prognostic model could predict biochemical recurrence, metastasis and immunotherapy response. In addition, CAF infiltration was associated with immunosuppressive microenvironment, positively correlated with tumor mutation burden and “p53 downstream pathway”, “MET promotes cell motility pathway” and “TGF-β signal pathway”. subsequently, verified the CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) were specifically expressed in fibroblast cell lines, and the protein expression were located in stromal cells. In conclusion, these results indicated that CAF infiltration promoted the progression of PCa and associated with PCa recurrence and poor prognosis. The PCa prognostic signature has a potential clinical application value and the prognostic markers in CAF might be targets for inhibiting the progression of PCa.

Introduction

Prostate cancer (PCa) is the most frequent cancer in male. There are 1.4 million new cases worldwide and accounts for 370,000 deaths annually (Siegel et al. 2022). Radical prostatectomy, radiotherapy and androgen deprivation therapy (ADT) are the standard treatments for localized prostate cancer (Cornford et al. 2021). However, part of patients will experience biochemical recurrence and metastasis after treatment, inevitably progress to castration-resistant prostate cancer (CRPC) with poor prognosis (West et al. 2014; Cornford et al. 2021).

Cancer-associated fibroblasts (CAF) are the major components in tumor microenvironment (TME), which affect cancer phenotype by constructing and remodeling extracellular matrix (ECM), and make contribution to invasion, metastasis, immune escape, angiogenesis and drug resistance in cancer (Bonollo et al. 2020; Henke et al. 2019; Owen et al. 2022). In addition, CAF regulate the composition of immune cells in the TME, promote the construction immunosuppressive microenvironment in tumor, resist immunotherapy (Stultz and Fong 2021; Jenkins et al. 2022). CAF has been confirmed to be involved in the progression of drug resistance in PCa and could be a promising therapeutic target (Stultz and Fong 2021; He et al. 2022). Further elucidation of CAF related gene markers in prostate cancer progression is beneficial for guiding treatment and predicting the progression of prostate cancer.

At present, bioinformatics analysis has become an important component of cancer research. Weighted gene co-expression network analysis (WGCNA) is a systematic analysis method, could maximize the
collection of co-expressed gene modules highly related to phenotype, and has been used to identify CAF markers in other cancers (Langfelder and Horvath 2008; Liu et al. 2021, 2020). In this study, we used WCGNA to identify hub gene modules related to CAF infiltration in PCa. Meanwhile, univariate Cox regression and the least absolute contraction and selection operator (LASSO) analysis were used to identify CAF prognostic signature and developed risk model. We tested the predicting potency of CAF prognostic model on biochemical recurrence, metastatic recurrence and immunotherapy response of PCa. Subsequently, explored the effects of high CAF infiltration on tumor microenvironment, including immune infiltration, pathway mechanism and tumor mutation load. In addition, we verified the gene and protein expression of CAF signature (FAP, SFRP2, COL1A1, VCAN) by cell line and immunohistochemistry. Our results showed that the prognostic markers in CAF might be potential targets for inhibiting the progression of PCa and CAF signature could provide a new idea for PCa diagnosis and treatment.

**Method**

**Data preparation and processing**

The data included RNA-seq, gene mutation and clinical characteristics of 498 patients with PCa were download from the cancer genome atlas (TCGA) database (Goldman et al. 2020), and transcripts per million (TPM) was log2(TPM + 1) transformed. The prognostic data were derived from TCGA prognostic supplementary documents published by LIU, Jianfang et al. (Liu et al. 2018), excluding competitive risk prognostic events. Form Gene Expression Omnibus (GEO) database (Barrett and Edgar 2006), obtained gene expression data, clinical information and prognostic data of 248 PCa patients who receiving ADT (GSE116918)(Jain et al. 2018).

**Analysis of immune infiltration**

CAF infiltration score was calculated using Estimate the Proportion of Immune and Cancer cells (EPIC) algorithm (Racle et al. 2017), xCell algorithm (Aran et al. 2017), microenvironment cell populations-counter (MCP-counter) algorithm (Becht et al. 2016) and Tumor Immune Dysfunction and Exclusion (TIDE) algorithms (Jiang et al. 2018). The proportions of 22 immune cell subtypes in PCa tissue was evaluated by CIBERSORT algorithm (Newman et al. 2015). Estimation of STromal and Immune cells in MAignant Tumours using Expression data (ESTIMATE) algorithm (Yoshihara et al. 2013) was used to calculated the stromal scores. Except for TIDE algorithm was completed by online database (http://tide.dfci.harvard.edu/), other algorithms were completed by invoking "IOBR" R package (Zeng et al. 2021).

In addition, the TIDE online database was used to estimate immunotherapy responses in TCGA and GEO cohorts. Subsequently, the ROC curve was used to evaluate the predicting potency of CAF prognostic model of immunotherapy response.

**Constructed CAF and Stromal Co-expression Network based on WCGNA**
The “WGCNA” R package (Langfelder and Horvath 2008) was used to construct co-expression networks and screened hub genes associated with CAF infiltrations as well as stromal scores. Inputted the top 10,000 genes of median absolute deviation (MAD) for network constructions in both TCGA and GSE116918 cohorts. At first, the Pearson’s correlation matrices and average linkage method were performed for all pair-wise genes. Then, a weighted adjacency matrix was constructed using a power function $A_{mn} = |C_{mn}|^\beta$, after determined the soft threshold $\beta$, the adjacency was transformed into a topological overlap matrix (TOM) and the corresponding dissimilarity (1-TOM) was calculated. Average linkage hierarchical clustering was conducted according to the TOM-based dissimilarity measure with a minimum size of 30 for the gene group, and then classify genes with similar expression profiles into gene modules. Modules with a distance of less than 0.25 are combined, and the first principal component expressed by each module was summarized as module eigengenes (ME). Evaluate the Pearson correlation between ME and EPIC-CAF score and Stromal score, then identified the hub module related to CAF infiltration and Stromal score. The gene significance (GS) for each gene’s traits and module membership (MM) in the hub module was measured with MM > 0.6 and GS > 0.6 as thresholds, the genes in the module were screened as potential CAF infiltration related hub genes.

**Gene Ontology and the Kyoto Encyclopedia of Genes and Genomes Analyses**

Used "ClusterProfiler" R package (Yu et al. 2012), the hub genes were analyzed by Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses to determine the biological function and related pathways involved. (Threshold setting: $P < 0.05$ and adjusted $q < 0.05$ were considered to be statistically significant).

**Identify the CAF prognostic signature and construct prognostic model**

The GSE116918 cohort was selected to build the model cause all patients in it were explicitly treated with ADT, and TCGA was used as the verification cohort. Univariate Cox regression model was used to identify hub genes related to progression-free interval (PFI). Subsequently, Lasso regression analysis (Simon et al. 2011) was performed for gene with $P < 0.05$. After 7000 iterations to screen CAF prognostic signature and obtain the risk coefficient. Subsequently, constructed CAF prognostic model: CAF risk score = $(\beta_i \times \text{Exp}_i)$, where $\beta_i$ refers to the Lasso coefficient of $i$th gene, and $\text{Exp}_i$ represents the $i$th gene’s expression value. According to the median CAF risk score, the cohort divided into high and low risk groups, and the Kaplan-Meier curve was displayed the PFI differences between two groups in GSE116918 cohort. Similarly, the model was verified in the TCGA cohort.

**Prediction of the recurrence by CAF model**

We collated prognostic data of the biochemical recurrence and metastasis recurrent after ADT treatment in the GSE116918 cohort. Performed univariate and multivariate Cox regression analyses for clinic information (PSA, T stage, gleason score) and CAF prognostic model with "survival" R package. Using "timeROC" and "rmda" R package, the predicting potency of the CAF model on PCa progression was evaluated by time-dependent ROC curves. Subsequently, time-dependent AUC curves and decision curve...
analysis (DCA) compared the predicting potency of clinic information (PSA, T stage, gleason score) and CAF model.

**Relationship between CAF infiltration and tumor mutation burden**

According to the CAF model, the somatic mutation data of TCGA cohort was divided into high risk and low risk group. The tumor mutation burden (TMB) of each sample was calculated by "maftools" R package(Cibulskis et al. 2013). Displayed the top 20 mutation genes in the two group, and the TMB differences between two groups were compared. In addition, the spearman correlation between TMB and CAF model were analyzed.

**Gene set enrichment analysis**

Gene set enrichment analysis (GSEA) and single sample gene set enrichment analysis (ssGSEA) analysis uses the "GSVA" R package(Hanzelmann et al. 2013) to explored the pathway mechanism related to CAF infiltration, and the gene sets "c2.cp.kegg.v2022.1.Hs. symbols" and "c2.cp.v2022.1.Hs. symbols. gmt" were download from MSigDB database(Liberzon et al. 2015). In addition, the correlation between CAF infiltration and MET, TGF-β and p53 signal pathway enrichment scores was calculated by Spearman's correlations analysis.

**The correlation between CAF model and CAF infiltration.**

Spearman's correlations analysis was performed to evaluate the correlation between CAF model and CAF infiltration algorithms (MCP-counter, EPIC, xCell, ESTIMATE, TIDE). In addition, the CAF marker genes form published literature were collected(Gascard and Tlsty 2016; Han et al. 2020), and the expression correlation between CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and published CAF markers were verified in both TCGA and GES116918 cohort.

**Cell expression and localization validation**

The mRNA expression data of 37 human fibroblasts and 8 PCa cell lines were downloaded from the CCLE database (https://www.proteinatlas.org/)(Ghandi et al. 2019), and compared the expression of CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) in fibroblasts and PCa cell lines.

The human protein map (HPA) database (www.Portals.Broadinstitute.org)(Uhlen et al. 2015), was used to obtain immunohistochemical and immunofluorescence image of CAF prognostic markers. In order to observe the cellular and subcellular localization of the CAF prognostic markers.

**Statistical Analysis**

All statistical analyses were performed using the R software (v.4.2.2). The Wilcoxon test was applied for pairwise comparisons. The Kaplan–Meier analysis with the log-rank test was adopted for PFI comparisons. $P<0.05$ was considered statistically significant.
Results

High CAF infiltration was associated with PCa recurrence

The immune infiltration algorithm included EPIC, xCell, MCP-counter, and TIDE, were used to predict the sample CAF infiltration in GSE116918 cohort and TCGA cohort, and ESTIMATE algorithm was predicted the Stromal score. According to cutoff values of CAF infiltration score and Stromal score, the cohorts were divided into high and low risk groups to evaluate the difference of PFI via log-rank test. Kaplan–Meier curves showed that except the xCell in GSE116918 ($P = 0.744$) and TCGA ($P = 0.069$), the TIDE in the TCGA ($P = 0.064$) had no significant difference in the PFI, the other predicted CAF infiltration had shown the high CAF infiltration was associated with high risk of recurrence, as well as Stromal score (Fig. 1). This highlights the necessity of further exploration of CAF and stromal related genes for PCa. In order to obtain more comprehensive CAF infiltration related genes, we extracted The EPIC estimated CAF score and Stromal score as phenotype for subsequent WCGNA analysis, and other three CAF evaluated data as external validations.

Construction of co-expression gene network based on CAF and stromal score

GSE116918 cohort and TCGA cohort were analyzed by WCGNA used the top 10000 gene of variance to construct scale-free co-expression network. Soft threshold power ($\beta$) of GSE116918 is 2, and TCGA is 9 (Fig. 2A-B). The GSE116918 obtained 6 co-expression modules, and TCGA-PCAD obtained 15 co-expression modules. In GSE116918 cohort, the yellow module was highest correlation with EPIC-CAF score ($\text{cor} = 0.80, P < 0.001$) and the Stromal Score ($\text{cor} = 0.88, P < 0.001$) (Fig. 2C-D). The light green module in TCGA cohort was highest correlation with EPIC-CAF score ($\text{cor} = 0.81, P < 0.001$) and Stromal Score ($\text{cor} = 0.89, P < 0.001$) (Fig. 2E-F). Scatter map show the gene relationship (MM) and gene significance (GS) within the module of two cohort. From the yellow module (286 genes in total) and the light green module (195 genes in total), potential hub genes which were highly related to the EPIC-CAF score and stromal score were screened with threshold (GS > 0.6 and MM > 0.6).

Functional Analyses of CAF infiltration related hub genes

Extract 15 hub genes overlapped in TCGA cohort and GSE116918 cohort. GO analysis showed that extracted hub genes were mainly involved in biological process (BP) term “extracellular matrix organization”, ellular component (CC) term “collagen-containing extracellular matrix” and molecular function (MF) term “extracellular matrix structural constituent” (Fig. 3B). The KEGG analysis showed that hub genes were involved in “ECM-receptor interaction”, “Wnt signaling pathway” and “Focal adhesion” (Fig. 3C).

Identification of CAF prognostic signature and construction of prognostic model

In order to obtain CAF infiltration hub genes highly related to the progress in castration resistance of PC. We selected GSE116918 cohort, which was treated with ADT, as the training cohort, and TCGA cohort as the validation cohort. Univariate Cox regression analysis was performed on the 15 hub genes in the
training cohort (Fig. 3A), and 14 prognostic hub genes were identified ($P < 0.05$) (Fig. 3D). Subsequently, used Lasso regression to identify CAF prognostic signature (FAP, SFRP2, COL1A1, VCAN) and the minimum $\lambda$ value, and further constructed the calculation formula of CAF prognostic model (Fig. 3E-F). The samples in each cohort were divided into high and low risk groups according to risk CAF scores. Kaplan-Meier curves showed high CAF infiltration estimated by prognostic model had a higher risk of PCa recurrence in both training and verification cohorts (training cohorts, HR = 2.797, $P < 0.001$; validation cohorts, HR = 1.728, $P = 0.014$) (Fig. 3G-H).

**Prediction of CAF prognostic model for PCa recurrence**

To further evaluate the predicting potency of CAF prognostic model in PCa recurrence, we integrated the CAF risk score and clinical characteristics (age, PSA, T stage and gleason score, CAF) in GSE16918 cohort, which treated with ADT, for univariate and multivariate COX analysis of biochemical recurrence and tumor metastasis events. The results showed that CAF prognostic model could be used as a prognostic factor for PCa (biochemical recurrence, HR = 4.164, $P < 0.001$; metastasis recurrence, HR = 4.564, $P < 0.001$) (Fig. 4A, 4E), could effectively predict biochemical recurrence and tumor metastasis (Fig. 4B, 4F). Compared with PSA, T stage and gleason score, the CAF model had an advantage in identifying high-risk recurrence of PCa (Fig. 4C, 4G). In addition, CAF prognostic model combined with clinical characteristics could more accurately evaluate the risk of PCa recurrence (Fig. 4D, 4H).

**Effect of CAF infiltration on TMB and immune microenvironment**

We divided them into high-risk group and low-risk group according to the CAF prognostic model, evaluated the gene mutation data of two groups. The results showed high CAF infiltration was associated with higher frequency of TP53 mutations the frequency of TP53 mutation (high-risk vs low risk, 14% vs 5%) (Fig. 5A-B), and there was a significant difference in TMB between the two groups ($P < 0.049$) (Fig. 5C). In addition, the CAF risk score was positively correlated with TMB ($cor = 0.12, P = 0.0062$) (Fig. 5D).

Comparing the immune infiltration in the two groups, we found that the high CAF infiltration associated with the higher infiltration abundance of M1, M2 Macrophages, resting Dendritic cells, resting CD4 memory T cells, naive B cells, Monocytes and lower infiltration abundance of CD8 T cells, helper follicular T cells, Plasma cells (Fig. 5E-G).

**Prediction of immunotherapy response by CAF models**

The prediction of immunotherapy responses in TCGA and GSE116918 cohort through TIDE database. The results showed that there was a significant difference in the TIDE score between the high and low CAF infiltration ($P < 0.001$) (Fig. 6A, 6C), and the immunotherapy response rate was lower in patients with high CAF infiltration (Fig. 6B, 6D). In addition, ROC curve showed that CAF risk score could predict immunotherapy response (GSE116918, AUC = 0.633; TCGA, AUC = 0.760) (Fig. 6C, 6E).

**GSEA analysis of CAF infiltration**
We compared the difference of gene expression between high and low CAF infiltration in TCGA and GEO cohort, and GSEA enrichment analysis showed the major enriched signaling pathways associated with high CAF infiltration were “cell adhesion molecules cans”, “cytokine receptor interaction”, “ECM receptor interaction” (Fig. 7A-B). In addition, ssGSEA enrichment analysis showed high CAF infiltration was positively correlated with “p53 downstream pathway”, “MET promotes cell motility pathway” and “TGF-β signal pathway” (Fig. 7C-D).

**Correlation between CAF prognostic model and CAF infiltration**

In order to verify the correlation between CAF model and CAF infiltration in tumor microenvironment. We first analyzed the Spearman correlation between CAF model and CAF infiltration score (xCell, EPIC, MPC-counter, TIDE, ESTIMATE). The results showed that the CAF model was positively correlated with all CAF infiltration score (Fig. 8A, 8D). In addition, we validated the correlation between CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and the reported CAF markers used TCGA and GSE16918 expression data. The co-expression heat map showed the expression of CAF prognostic markers were positively correlated with the reported CAF markers, except for the negative correlation between CAV1 and OGN in GSE116918 (Fig. 8C-D, 8F-G).

**Cell lines and Immunohistochemistry validation of CAF prognostic markers**

We obtained the RNA-seq data of 37 human fibroblasts and 8 PCa cell lines through CCLE database, and verified the high expression of CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) in fibroblasts (P < 0.001) (Fig. 9A-B). In addition, the immunohistochemical images from HPA database showed that the expression of FAP, COL1A1 and VCAN proteins were located in the tumor matrix. Unfortunately, SFRP2 immunohistochemical images are not included in the HPA database (Fig. 9C). Multidimensional verification showed that four central genes might be specific CAF markers of PCa.

**Discussion**

Biochemical recurrence and metastasis of PCa after ADT indicate the progression of CRPC, and this progression accompanied by the remodeling of the TME(Owen et al. 2022). CAF have been proved to be a major components of TME, via cell-cell communication participate in angiogenesis, invasion, metastasis and drug resistance of PCa(Owen et al. 2022; Bonollo et al. 2020). In this study, we observed that high CAF infiltration were associated with high-risk recurrence of PCa. Previous study confirmed that after ADT, CAF could regulate MAPK, STAT3 and PI3K/AKT signal pathways in PCa cells by secreting cytokines and interleukin, activate AR receptors in androgen-independent conditions, and enhance the resistance of PCa to enzalutamide and bicalutamide(Cioni et al. 2018; Ishii et al. 2018; Yang et al. 2003). Another study showed that CAF infiltration promotes PCa metastasis and neuroendocrine PCa by regulating WNT and mTOR signal pathways(Mishra et al. 2018). In our study the GO and KEGG analysis showed that the CAF hub genes were involved in “extracellular matrix remodeling”, “matrix receptor signal pathway” and “WNT pathways”. In addition, CAF infiltration was positively correlated with TMB, and the TP53 mutation
frequency was higher with high CAF infiltration. These evidences indicate that CAF could be a potential participant in the progression of PCa to CRPC (He et al. 2022).

The TGF-β signal pathway is a typical pathway of CAF activation, while the activated CAF participates in the construction of immunosuppressive microenvironment and promotes the progress of PCa (Owen et al. 2022; Stultz and Fong 2021). Previous studies showed CAF reduces the efficacy of immunotherapy by inhibiting CD8+ T cell infiltration, and CAF promotes the immunosuppressive microenvironment via ECM remodeling and TGF β signaling pathway, leads to resistant to immunotherapy (Stultz and Fong 2021; Jenkins et al. 2022; Henke et al. 2019). CAF recruit monocytes by secreting CXCL2 and promote M2-like macrophage phenotypic differentiation. meanwhile, M2-like macrophages further activate fibroblasts promote tumor angiogenesis and increase the escape of PCa cells cause tumor metastasis (Comito et al. 2014). In this study, GSEA analysis showed that high CAF infiltration was positively correlated with “p53 downstream signal”, “MET signal pathway” and “TGF-β signal pathway”. In addition, we found that high CAF infiltration was associated with higher mononuclear cells, M2 CAF M1 immune cell and dendritic cell infiltration and lower CD8+ T cell infiltration, which was consistent with previous studies. These indicated that high CAF infiltration was associated with immunosuppressive microenvironment and poor prognosis in PCa. Meanwhile, we further validation that CAF model was superior to commonly clinical characteristics (PSA, gleason score, T stage) in predicting the risk of PCa recurrence. CAF model combined with clinical characteristics, the recurrence of PCa could be predicted more effectively. Subsequent verification also showed that the CAF model could predict the response to immunotherapy. Although immunotherapy has limited efficacy in patients with advanced PCa, it can prolong the overall survival of partial patients and provide CRPC another treatment option (He et al. 2022). In conclusion, these results suggest that the CAF model could provide a new idea for the diagnosis and treatment evaluation of CRPC.

In order to verify the correlation between CAF model and CAF infiltration in PCa. we verified the positive correlation between CAF model and five algorithms (xCell, EPIC, MPC-counter, TIDE, ESTIMATE), as well as the co-expression relationship between CAF prognostic marker (FAP, SFRP2, COL1A1, VCAN) and reported CAF marker. Moreover, verified the CAF prognostic marker were specifically expressed in the fibroblast cell line, and the proteins expression were located in the cell matrix. These indicated that CAF model could reflect the degree of CAF abundance in PCa tissue.

Currently, some CAF prognostic markers had been confirmed to be associated with PCa progression. A recent proteomic study on PCa paraffin-embedded (FFPE) tissue samples showed that protein VCAN is associated with biochemical recurrence of PCa, could be a prognostic marker (Lygirou et al. 2022). another study confirmed that FAP has potential as a diagnostic and therapeutic target, and PCa tissue could be better visualized in positron emission tomography/computed tomography (PET/CT) using a developer targeting FAP (Boinapally et al. 2022). In addition, previous in vitro studies confirmed that up-regulating SFRP in PCa cell lines could induced epithelial-mesenchymal transformation (EMT), transform PCa cells into osteoblast-like phenotypes that conducive to survival in bone metastases (Kato et al. 2019).
In summary, CAF infiltration promoted the progression of PCa and associated with PCa recurrence and poor prognosis. The PCa prognostic model based on CAF infiltration has potential clinical application value, could provide new ideas for the diagnosis and treatment of PCa. Although, we had explored the predictive potential of CAF model in the cancer progress via multi-dimensions, some limitations in our study should be noted. First, this study is a retrospective study based on open data sets, the accuracy of CAF model needs stringent clinical trials to verify. second, the mechanism and potential biological functions of CAF prognostic signature in PCa progression needs rigorous in vitro and vivo experiments to verify in the future.

### Conclusion

The CAF infiltration promoted the progression of PCa and associated with PCa recurrence and poor prognosis. The PCa prognostic signature has a potential clinical application value and the prognostic markers in CAF might be targets for inhibiting the progression of PCa.

### Declarations

#### Acknowledgements

We appreciate The Cancer Genome Atlas (TCGA) database, Gene Expression Omnibus (GEO) database, CCLE database, TIDE database and human protein map (HPA) database for using their data.

#### Founding

This study was supported by the Excellent Young Talents Plan of Guizhou Medical University (2022)107), and the Science and Technology Fund Project of Guizhou Health Commission (gzwkj2021-211).

#### Competing of interest

The authors declare no conflict of interest.

#### Author contributions

KaiFa Tang and WenJun Zhang conceived and designed the study; WenJun Zhang, BangWei Che, Miao Liu conducted data analysis, WenJun Zhang, BangWei, Miao Liu, SengHan Xu, Wei Li, Tao Huang, Ying Yu, Jun He, Cheng Zha, Zheng Peng, Kunyuan Huang collected data and visualized. WenJun Zhang wrote this manuscript. All authors read and approved the final manuscript.

### Data Availability

Not applicable.

### Ethics approval
Not applicable.

**Consent to participate**

Not applicable.

**Consent to publish**

Not applicable.

**References**


Figures
Figure 1

CAF infiltration is associated with PCa recurrence. (A-B) Kaplan-Meier curve showed high CAF infiltration was associated with high risk of PCa recurrence. A GSE116918 cohort. B TCGA cohort. CAF, Cancer-Associated Fibroblasts, PCa, Prostate cancer.
Figure 2

Construction of co-expression gene network based on CAF and stromal score. A Scale independence, mean connectivity and clustering dendrogram in GSE116918 cohort. B Scale independence, mean connectivity and clustering dendrogram in TCGA cohort. C Pearson correlation analysis of merging modules with CAF score and stromal score in GSE116918 cohort. D Scatter plots of MM and GS of each gene in the yellow module in GSE116918 cohort. E Pearson correlation analysis of merging modules with...
CAF score and stromal score in TCGA cohort. F Scatter plots of MM and GS of each gene in the lightgreen module in TCGA cohort. CAF, Cancer-Associated Fibroblasts, GS, Gene Significance, MM, Module Membership.

**Figure 3**
Construction of CAF prognostic signature. **A** Hub gene shared by TCGA (blue) and GSE116918 (brown). **B** GO analyses of shared hub genes. **C** KEGG analysis of shared hub genes. **D** Univariate Cox analysis for shared hub genes in GSE116918. **E** Lasso regression analysis. **F** Formulation of the CAF prognostic risk model. **G** Kaplan–Meier analyses in different risk group in GSE116918 cohort. **H** Kaplan–Meier analyses in different risk group in TCGA cohort. CAF, Cancer-Associated Fibroblasts. Lasso, least absolute shrinkage and selection operator. GO, Gene Ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes.

**Figure 4**

Prediction of CAF prognostic model for PCa recurrence. **A** Univariate COX analysis of biochemical recurrence. **B** CAF model predicting biochemical recurrence at 3, 5, and 7 years. **C** Compared clinical characteristics and CAF model in predicting biochemical recurrence. **D** DCA curves showed combine CAF prognostic model with clinical characteristics had good risk identification efficacy of biochemical recurrence. **E** Univariate COX analysis of metastatic recurrence. **F** CAF model predicting metastatic recurrence at 3, 5, and 7 years. **G** Compared clinical characteristics and CAF model in predicting metastatic recurrence. **H** DCA curves showed combine CAF prognostic model with clinical characteristics had good risk identification efficacy of metastatic recurrence. CAF, Cancer associated fibroblasts.
Figure 5

Effect of CAF infiltration on TMB and immune microenvironment. A Top 20 gene mutation gene in high-risk group. B Top 20 gene mutation gene in low-risk group. C Compare TMB between high and low risk groups. D Spearman analysis between tumor mutational burden and CAF model. E Landscape of the proportions of 22 immune cell subtypes in TCGA cohort. F Landscape of the proportions of 22 immune cell subtypes in GSE116918 cohort. G TCGA and GSE116918 cohorts were combined to compare
immune infiltration between high and low risk group. CAF, Cancer associated fibroblasts. TMB, tumor mutational burden.

**Figure 6**

Prediction of immunotherapy response. **A** Compare TIDE scores between high and low risk groups in GSE116918 cohort. **B** Compare immunotherapy responses rate between high and low risk groups in GSE116918 cohort. **C** ROC curves showed CAF risk score in predicting immunotherapy responses in GSE116918 cohort. **D** Compare TIDE scores between high and low risk groups in TCGA cohort. **E** Compare immunotherapy responses rate between high and low risk groups in TCGA cohort. **F** ROC curves showed CAF risk score in predicting immunotherapy responses in TCGA cohort. TIDE, Tumor Immune Dysfunction and Exclusion.
Figure 7

Figure 8

Correlation between CAF prognostic model and CAF infiltration. **A** Spearman correlation between CAF model and CAF infiltration score (EPIC, MCP-counter, xCell, ESTIMATE and TIDE) in TCGA cohort. **B** Heatmaps showed expression of CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and reported CAF markers in TCGA cohort. **C** The co-expression heat map showed the relationship between CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and reported CAF markers in TCGA cohort. **D** Spearman correlation between CAF model and CAF infiltration score (EPIC, MCP-counter, xCell, ESTIMATE and TIDE) in GSE116918 cohort. **E** Heatmaps showed expression of CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and reported CAF markers in GSE116918 cohort. **F** The co-expression heat map showed the relationship between CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) and reported CAF markers in GSE116918 cohort. CAF, Cancer-Associated Fibroblasts.
Figure 9

Cell lines and Immunohistochemistry validation of CAF prognostic markers. **A** Comparison of CAF prognostic markers (FAP, SFRP2, COL1A1, VCAN) expression in prostate cancer cell lines and fibroblast cell lines. **B** Heat map showed prognostic markers expression in PCa and fibroblast cell lines. **C** Immunohistochemical images from the Human Protein Atlas database, showed the protein localization of prognostic markers.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table.S1.docx