Identification of PC4 as a diagnostic and prognostic biomarker in hepatocellular carcinoma

Liangliang Bai
Yan'an University

Guan Liu
Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University

Gang Dou
Xi'an Medical University

Xiaojun He
Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University

Chenyu Gong
Xi'an Medical University

Hongbin Zhang
Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University

Kai Tan
Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University

Xilin Du (duxl0705@fmmu.edu.cn)
Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University

Research Article

Keywords: hepatocellular carcinoma, diagnosis, Prognosis, Risk signature, bioinformatics

Posted Date: June 1st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2981470/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Background

Human positive cofactor 4 (PC4) is associated with the development and therapeutic resistance of several malignancies. Nonetheless, the role of PC4 in HCC remains obscure, warranting further investigation.

Methods

PC4 differential expression between HCC and normal tissues and its association with the clinicopathological characteristics of HCC patients were explored in the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) datasets. Subsequently, the prognostic and diagnostic significance of PC4 in HCC patients was analyzed. Functional enrichment analyses were conducted by Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis and Gene set enrichment analysis (GSEA) analysis to explore the biological functions of PC4. The CIBERSORT algorithm was used for immune infiltration analysis. The risk signature based on PC4-related genes was constructed by LASSO-COX regression and validated with the International Cancer Genome Consortium (ICGC) dataset. The qRT-PCR was used to verify the expression levels of all genes. Tumor Immune Dysfunction and Exclusion (TIDE) analysis evaluated the correlation between PC4 level and risk score and immunotherapy response in HCC patients. Finally, based on online databases, PC4-related ceRNA networks were constructed to explore the regulatory mechanism of PC4 in HCC.

Results

PC4 levels were significantly upregulated in HCC compared to normal tissues and positively correlated with the pathological grade and clinical stage. The PC4-high expression group showed a worse prognosis. In addition, PC4 could help distinguish tumor tissues from normal ones with an area under the curve (AUC) of 0.965. The PC4 level was associated with immune checkpoints and immune cell infiltration. In the training and validation sets, the eight-gene risk signature (Risk score = 0.074*CEP55 + 0.091*TRIP13 + 0.100*BRIX1 + 0.081*STIP1 + 0.191*PIGU + 0.038*CFL1 + 0.017*RBM17 + 0.178*OLA1) was strongly correlated with HCC patient prognosis and was considered as an independent prognostic factor. TIDE analysis showed that patients in both the PC4-low and low-risk groups were more likely to benefit from immunotherapy. Finally, the IncRNAs/miRNA-101-3p/PC4 network was constructed.

Conclusion

We confirmed PC4 as a diagnostic and prognostic marker in HCC patients. And, we developed and validated an 8-gene risk signature, which could help for clinical decision-making. The CeRNA network
Introduction

Primary liver cancer (PLC) is a common malignancy with high annual incidence and mortality. According to the latest cancer statistics, PLC accounted for 900,000 new cases and 830,000 deaths globally in 2020, posing a severe threat to human life and health (1). As the most common histopathological type of PLC, hepatocellular carcinoma (HCC) is highly invasive and insidious, accounting for the high number of patients with advanced-stage disease at diagnosis who miss the optimal timing for radical surgical treatment (2). In addition, patients who underwent radical surgical treatment had a recurrence of nearly 70% during 5 postoperative years (3). The past few years have witnessed unprecedented advances in research on therapeutic drugs, targeted therapy and immunotherapy, which have improved the prognosis of this patient population to some extent. However, significant heterogeneity surrounds drug efficacy and resistance, emphasizing the need for further research (4, 5). Accordingly, finding effective biomarkers and potential therapeutic targets is urgent to help diagnose HCC and improve the prognosis of these patients.

As a highly conserved nuclear protein during evolution, human positive cofactor 4 (PC4-human, SUB1-yeast) was extracted by purification from the human upstream stimulatory activity (USA) fraction in 1994 (6). Earlier studies have shown that PC4 participates in the initiation, elongation and termination of RNA polymerase II-mediated transcription (7). However, with extensive research on the PC4 gene, it is now recognized that PC4 is also closely associated with DNA damage repair, chromatin formation, and cell cycle regulation (8–11). Given that PC4 can interact with the P53 gene and form positive feedback (12), it has long been thought to be a tumor suppressor gene. In recent years, the tumor-promoting effect of PC4 has recently received much attention from scholars. Research has revealed a strong correlation between PC4 and the proliferation, metastasis and therapy resistance of malignant tumors like esophageal squamous cell carcinoma, non-small cell lung cancer (NSCLC), pancreatic ductal adenocarcinoma, breast cancer, prostate cancer (13–17). However, the role of PC4 in HCC is still under investigation.

This study explored the PC4 gene level in HCC and the association between gene expression and diagnosis, prognosis and immune infiltration. Additionally, this study constructed and validated a risk signature for effective assessment of patient prognosis and immunotherapy response. Finally, establishing the ceRNA network facilitates understanding the regulatory mechanism of the PC4 in HCC.

Materials and Methods

Data Acquisition and Preparation

The expression data and clinical data of 374 HCC samples and 50 normal ones were acquired as the primary analysis dataset from The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/), and GSE25097, GSE36376, GSE39791 and GSE57957 datasets derived from Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/) were applied to identify differentially expressed genes between HCC
and normal tissue. Expression and clinical data of 260 HCC samples were obtained from the International Cancer Genome Consortium (ICGC) dataset (https://dcc.icgc.org/) to validate the risk signature. The Xena database (https://xenabrowser.net/datapages/) was used to obtain the HCC patient data on progression-free survival (PFS) and disease-specific survival (DSS).

**Differential Expression Analysis**

The Tumor Immune Estimation Resource (TIMER) database (http://timer.cistrome.org/) analyzed PC4 gene expression in 33 malignant tumors and normal tissues. The "limma" package was adopted to analyze the differential expression of the PC4 gene between HCC and normal tissues from TCGA and four GEO datasets (18). The association between gene expression and different clinicopathological features of HCC patients was also analyzed by the "limma" package.

**Diagnostic and Prognostic Analysis**

The "pROC" package was used to plot ROC curves for assessing the ability of the PC4 gene to distinguish Stage I-IV tumor tissues from normal tissues (19). Time-dependent ROC curves for predicting HCC patient overall survival (OS) were plotted by the "timeROC" package (20). Independent prognostic factors in HCC patients were explored using univariate and multivariate analyses. Differences in survival were shown by Kaplan-Meier (K-M) curves.

**Identification of Differentially Expressed Genes (DEGs)**

The DEGs of the PC4-high and PC4-low groups were identified with the "limma" package. The filter criteria were P-value < 0.05 and Log2FC > 1/-1. Heatmaps and volcano plots were used to visualize the results.

**Enrichment Analyses**

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of DEGs were conducted to explore biological functions and potential mechanisms (21). GSEA enrichment analysis was conducted to find PC4-associated signaling pathways (22).

**Immune Infiltration Analysis**

The "CIBERSORT" package was adopted for assessing the level of immune infiltration of 22 immune cells in every HCC sample (23). Boxplot displayed the differences in immune cell infiltration between the PC4-high and PC4-low groups. The differences in the expression of 24 HLA genes and 8 immune checkpoint genes between the two groups were visualized in a boxplot.

**Construction and Verification of Risk Signature**

Co-expressed genes of PC4 were screened using the following criteria: Pearson correlation coefficient > 0.6 and P-value < 0.001. Differential expression analysis and univariate COX analysis were conducted for screening for genes with high expression in HCC and associated with patient prognosis. The "glmnet" package was used to construct the risk signature through the LASSO-COX regression (24). The ICGC dataset was used to validate the risk signature.
Prediction of Immunotherapy Response

Based on the Tumor Immune Dysfunction and Exclusion (TIDE) database (http://tide.dfci.harvard.edu/), patient immunotherapy response was predicted using two common mechanisms of immune escape, T cell dysfunction, as well as exclusion. A larger TIDE score indicated a higher probability of immune escape and a worse response to immunotherapy in patients (25).

Construction of CeRNA Network

With the ENCORI database (https://starbase.sysu.edu.cn/), the potential regulatory associations of miRNA-mRNA and lncRNA-miRNA were predicted (26). The results were screened based on correlation coefficient <-0.3 and P-value < 0.001. The miRNA-mRNA and lncRNA-miRNA-mRNA networks were visualized by Cytoscape software.

Cell Lines and qRT-PCR

Five cell lines (L02, BEL7402, HCCLM3, HepG2, and Huh7) were purchased from Procell Life Science & Technology (Wuhan, China) and cultured according to the manufacturer's instructions. qRT-PCR was used to verify the expression levels of all genes. The primer sequence information was listed in Supplementary Table 1.

Statistical Analysis

In this study, all data analyses were conducted using R software (version 4.1.3). Differences between groups were compared using the Wilcoxon test. K-M analysis and log-rank tests were adopted for inter-group comparison of survival. Independent prognostic factors were screened by univariate and multivariate analyses. A P-value < 0.05 was statistically significant.

Results

High PC4 Expression in Tumor Tissues

Pan-cancer analysis of the TIMER database showed significant differences in PC4 expression between 13 malignant tumor tissues and adjacent normal ones. PC4 was highly expressed in nine malignancies, including cholangiocarcinoma, esophageal cancer, colon cancer, as well as HCC, while it was lowly expressed in kidney chromophobe, kidney renal clear cell carcinoma, kidney renal papillary cell carcinoma and thyroid cancer (Fig. 1A). According to differential expression analysis from the four GEO datasets, PC4 gene expression was significantly higher in HCC than in adjacent normal liver tissue (Fig. 1B). The same results were obtained for TCGA (Fig. 1C). In addition, qRT-PCR showed that the expression of PC4 in HCCLM3, HepG2 and Huh7 was significantly higher than that of L02 (Fig. 1D).

Association of PC4 Expression with the HCC Patients' Clinicopathological Features and Prognosis
Given the high PC4 gene expression in HCC, we further explored the association between PC4 expression and the clinicopathological features of HCC patients. PC4 expression was not associated with the age and gender of HCC patients but was positively correlated with pathological grade and clinical stage (Fig. 2). An association was also found between the expression of PC4 and the HCC patient prognosis. The PC4-high group had significantly lower OS, PFS and DSS than the PC4-low group (Fig. 3A). The cumulative K-M curves for the clinical stage of combined HCC patients showed the same results. Patients in the PC4-high group and late-stage had the shortest survival time, while those in the PC4-low group and early-stage had the longest survival time (Fig. 3B). Subsequently, we assessed the influence of PC4 on HCC patient prognosis in different subgroups of age, gender, pathological grade and clinical stages. The results implied a worse prognosis in the PC4-high group, irrespective of the subgroup (Fig. 3C).

**Diagnostic and Prognostic Value of PC4**

The area under the ROC curve (AUC) values were all larger than 0.9, suggesting that PC4 could distinguish Stage I-IV tumor tissues from normal tissues (Fig. 4A, B). In addition, PC4 exhibited good performance in predicting HCC patient prognosis. The 1-, 3-, and 5-year AUC were 0.69, 0.61, and 0.65, respectively (Fig. 4C). According to univariate and multivariate analyses, PC4 expression was an independent prognostic factor of HCC patients (Fig. 4D). Subsequently, we further constructed a nomogram to predict the 1-, 3-, and 5-year survival of HCC patients based on clinicopathological features and PC4 expression level (Fig. 4E). The calibration curve showed a good predictive power of the nomogram (Fig. 4F).

**Identification of DEGs between Groups and Enrichment Analysis**

We identified 2856 DEGs (2708 upregulated and 148 downregulated) in HCC (Supplementary Table 2). A heatmap was generated showing the top 50 most significantly upregulated and downregulated genes in HCC (Fig. 5A). A volcano plot was generated to visualize the distribution of all DEGs (Fig. 5B). To explore the potential mechanism of PC4 in HCC, GO, KEGG and GSEA enrichment analyses were further conducted. GO enrichment analysis revealed DEGs were mainly enriched in biological processes like "nuclear division", "chromosome segregation" as well as "organelle fission". KEGG enrichment analysis revealed that the DEGs were significantly enriched in the "cell cycle", "glycosphingolipid biosynthesis-lacto and neolacto series" and "cytokine-cytokine receptor interaction" pathways (Fig. 5C). According to GSEA enrichment analysis, the PC4-high group was primarily enriched in "cell cycle", "cytokine-cytokine receptor interaction" as well as "ECM receptor interaction" signaling pathways, while the PC4-low group in "glycine, serine and threonine metabolism", "fatty acid metabolism" as well as "primary bile acid biosynthesis" (Fig. 5D).

**Immune Infiltration Analysis between Groups**

The association between PC4 and tumor-infiltrating immune cells was further analyzed using the "CIBERSORT" algorithm. B Memory cells, Activated dendritic cells and M0 Macrophages were differentially expressed between PC4-high and PC4-low groups (Fig. 6A). Subsequently, correlation
analysis between PC4 expression and 24 HLA genes and 8 immune checkpoint genes showed that all genes were overexpressed in the PC4-high group (Fig. 6C, D). The difference in TIDE scores indicated a higher incidence of immune escape and a lower response rate to immunotherapy in the PC4-high group (Fig. 6B).

Construction and Validation of an Eight-Gene Risk Signature

To predict the HCC patient prognosis, we constructed a risk signature by co-expression analysis. Heatmap showed that 92 PC4-related genes were overexpressed in HCC (Fig. 7A). Univariate COX analysis further revealed that 77 genes were related to HCC patient prognosis (Fig. 7B). Subsequently, we constructed and validated an eight-gene risk signature (Risk score = 0.074*CEP55 + 0.091*TRIP13 + 0.100*BRIX1 + 0.081*STIP1 + 0.191*PIGU + 0.038*CFL1 + 0.017*RBM17 + 0.178*OLA1) (Supplementary Table 3). These 8 genes are highly expressed in HCC from TCGA (Fig. 8A). Meanwhile, the same results were shown in cell lines. The expression of all genes in HCCLM3, HepG2 and Huh7 was significantly higher than that of LO2 (Fig. 8B). The ROC curve in the training set confirmed the predictive power of this risk signature for HCC patient prognosis. The 1-, 2- and 3-year AUC values were 0.79, 0.70 and 0.68, respectively (Fig. 9B). The number of deaths in HCC patients increased correspondingly with an increase in the risk score (Fig. 9A). The same results were obtained during K-M survival curve analysis. The high-risk group presented a significantly worse prognosis than the low-risk one (Fig. 9C). According to univariate and multivariate analysis, the risk score was one independent prognostic factor for HCC patients (Fig. 9D, E). PCA and t-SNE analysis revealed that the risk signature could distinguish between high- and low-risk patients (Fig. 9F, G). The above results from the training set were confirmed in the validation set. The 1-, 2- and 3-year AUC values of the validation set were 0.75, 0.74 and 0.73, respectively (Fig. 10).

Additionally, the influence of the risk signature on HCC patient prognosis in different subgroups of age, gender, pathological grade and the clinical stage was further analyzed. The results revealed a worse prognosis in the high-risk group in any subgroup in the training set and validation set (Fig. 11). The clinical information of HCC patients from TCGA (training set) and ICGC dataset (validation set) is provided in Supplementary Table 4. Subsequently, we constructed a nomogram with the risk signature and clinicopathological features to predict the HCC patients' survival rates at 1, 2, and 3 years (Fig. 12A). The calibration curve revealed a good predictive power of the nomogram (Fig. 12B).

Finally, TIDE analysis assessed the correlation between the risk signature and immunotherapy response. Patients with high scores were unlikely to respond to immunotherapy, while the opposite was true for patients with low scores (Fig. 12C).

Construction of CeRNA Network based on PC4

To explore the potential upstream regulators of PC4, we constructed a PC4-related ceRNA network. First, 73 possible upstream miRNAs were obtained from the ENCORI online database, and the miRNA-mRNA network was constructed through Cytoscape software (Fig. 13A). Correlation analysis between PC4 and miRNAs further identified miRNA-101-3p as a potential regulatory miRNA for PC4 (Supplementary Fig. 1A). Subsequently, we obtained 49 possible IncRNAs to explore the upstream regulatory IncRNAs of
miRNA-101-3p. Correlation analysis of miRNA-101-3p with lncRNAs ultimately identified 5 potential lncRNAs. The expression of these lncRNAs was positively correlated with PC4 and negatively with miRNA-101-3p (Supplementary Fig. 1B-F2). Finally, the lncRNAs/miRNA-101-3p/PC4 network was constructed (Fig. 13B). Boxplots and K-M curves were generated to visualize the differences in expression and OS of miRNA-101-3p and 5 lncRNAs, respectively. miRNA-101-3p had low expression in HCC, and HCC patients with low expression showed a worse prognosis. In addition, all five lncRNAs were highly expressed in HCC and predicted a poorer prognosis (Fig. 13C, D).

Discussion

As the most common histopathological type of PLC, HCC is highly invasive and insidious, accounting for the difficulty clinicians face in diagnosing and treating HCC. Alpha-fetoprotein (AFP) is the most commonly adopted diagnostic and screening biomarker for HCC. However, its sensitivity and specificity still do not meet the clinical needs (27). In recent years, targeted therapy and immunotherapy for HCC have achieved remarkable efficacy, but the heterogeneity in efficacy and drug resistance remains to be addressed (4, 5). Thus, searching for effective biomarkers and potential therapeutic targets is urgent to offer new strategies for diagnosing and treating HCC. PC4 is a highly conserved nuclear protein implicated in transcriptional regulation, DNA damage repair and chromatin formation (28). Moreover, overexpression of PC4 is significantly associated with proliferation, metastasis and treatment resistance in various malignancies. PC4 accelerates the development of pancreatic ductal adenocarcinoma by activating the mTOR/p70s6k signaling pathway (16). PC4 can also bind directly to the c-Myc promoter and regulate its transcription, which induces the Warburg effect to promote breast cancer progression (29). However, the role of PC4 in HCC remains largely understudied. This study comprehensively analyzed the diagnostic and prognostic value of PC4 in HCC and constructed a risk signature to effectively assess the HCC patients' prognosis. The construction of the ceRNA network facilitated the understanding of the regulatory mechanisms of PC4 in HCC.

PC4 is reportedly highly expressed in various malignancies, with the most significant differential expression in HCC. In addition, a positive correlation was found between the PC4 expression and the pathological grade and clinical stage of HCC patients, suggesting that PC4 potentially participates in the development of HCC. The value of PC4 in the diagnosis and prognosis of HCC was analyzed in the present study. The diagnostic ROC curve substantiated the ability of PC4 to distinguish tumor tissue from normal tissue. Survival analysis of OS, PFS and DSS showed that high PC4 expression correlated with a poor prognosis of HCC patients. Additionally, the prognostic ROC curve and nomogram indicated that PC4 could effectively predict the 1-, 3-, and 5-year survival rates of HCC patients. The above results implied that PC4 has huge prospects as a diagnostic and prognostic biomarker for HCC.

To understand the role of PC4 in HCC, we identified DEGs and conducted functional enrichment analyses. The PC4-high group was mainly associated with the "cell cycle", "cytokine-cytokine receptor interaction" and "ECM receptor interaction" signaling pathways. This result is consistent with the functions of PC4 documented in the literature. Subsequently, we further explored the association between the expression
level of PC4 and immune infiltration and immunotherapy response. The PC4 expression positively correlated with numerous HLA genes and immune checkpoint-related genes. Additionally, patients with high PC4 expression have a higher probability of immune escape, a poorer response to immunotherapy, and shorter survival. This result is caused by the tumor-suppressive microenvironment induced by the overexpression of immune checkpoint genes (30, 31).

We found that the ability of a single gene to assess HCC patient prognosis was limited. Therefore, to improve the predictive power, we constructed and validated an 8-gene risk signature (Risk score = 0.074*CEP55 + 0.091*TRIP13 + 0.100*BRIX1 + 0.081*STIP1 + 0.191 *PIGU + 0.038*CFL1 + 0.017*RBMI7 + 0.178*OLA1). Most genes in this risk signature have been strongly associated with the prognosis of HCC or other malignancies. CEP55 is located in the chromosomal region 10q23 and is important in cell division (32). It has been shown that CEP55 promotes HCC invasion and migration through the JAK2/STAT3/MMPs signaling pathway (33). As an AAA + ATPase that can promote the assembly or degradation of protein complexes, TRIP13 is significantly upregulated in various malignancies, including HCC (34). Moreover, TRIP13 can interact with ACTN4 to induce HCC invasion and metastasis through AKT/mTOR pathway (35). A study that constructed a prognostic risk signature for HCC based on RNA-binding proteins showed an association between BRIX1 and poor prognosis of HCC patients (36). STIP1, called heat shock protein (HSP)-organizing protein, is reportedly overexpressed in HCC and accelerates cancer cell growth and migration by interacting with Axin to activate β-catenin/TCF signaling (37). In addition, it has been reported that STIP1 in serum can be used as an indicator of microvascular invasion and can be adopted for predicting prognosis and response to TACE treatment in HCC patients (38). PIGU, also called cell division cycle 91-like 1 (CDC91L1), promotes HCC progression by activating the NF-κB pathway and promoting immune escape (39). CFL1 is an important actin depolymerization factor (ADF) family member, widely found in eukaryotes. It has been shown that hypoxia can induce CFL1 expression and thus activate the PLD1/AKT pathway to promote HCC progression (40). RBM17 is also known as splicing factor 45 (SPF45). Knockdown of RBM17 significantly inhibited cancer cell proliferation and arrested cells in the G2/M phase (41). Moreover, OLA1 is implicated in cellular processes like protein translation, signal transduction and cell proliferation (42). It has been reported that OLA1 is significantly expressed in HCC, and the knockdown of this gene suppressed the progression of cancer cells (43).

These studies overlap in their assertion that the risk score is correlated with the prognosis of HCC patients. Moreover, various analyses have shown that this risk score is reliable in predicting the HCC patients' prognosis. TIDE analysis showed that patients having high scores were more likely to be unresponsive to immunotherapy. This result corresponded to poor response to immunotherapy in the PC4-high group.

To better understand the role of PC4 in HCC, we constructed a lncRNAs/miRNA-101-3p/PC4 network. Most regulatory relationships in this network have been reported in malignancies. LncRNA SNHG6 improved E2F8 expression by associating with miR-101-3p, thus promoting proliferation and angiogenesis in cholangiocarcinoma (44). In HCC, SNHG6 upregulated ZEB1 expression through binding miR-101-3p and combined with downregulation of Smad7 to induce epithelial-mesenchymal transition (EMT) to speed up cancer cell metastasis (45). An SNX16 and PAPOLG-based study showed that the
GSEC/miR-101-3p/SNX16/PAPOLG axis was significantly associated with HCC patient prognosis (46). Reportedly, lncRNA SNHG1 interacts with miR-101-3p to promote tumor progression. The SNHG1/miR-101-3p/SOX9/Wnt/β-catenin axis has been reported to promote the progression of NSCLC (47). miR-101-3p was closely associated with programmed cell death, such as autophagy and apoptosis, playing an important role in HCC (48). It has been shown that miR-101-3p could bind to PYGB and inhibit its expression, while the overexpression of PYGB, in turn, inhibited the regulation of miR-101-3p on HCC cell invasion, proliferation as well as migration (49). However, the association between miR-101–3p and PC4 in HCC remains unknown. The construction of lncRNAs/miRNA-101-3p/PC4 network may bridge this knowledge gap.

In conclusion, this study demonstrated the potential of PC4 as a diagnostic and prognostic biomarker for HCC. Moreover, the risk signature based on PC4 could guide clinicians in evaluating prognosis and response to immunotherapy for HCC patients.

**Declarations**

**Acknowledgments**

We are grateful to the TCGA, GEO and ICGC databases for providing data for this study. We thank the "ENCORI" and "TIDE" databases and "R" and "Cytoscape" software for their help in data analysis.

**Authors’ Contributions**

LB, GL & GD conceived the study. XH, CG & HZ conducted data collection and collation. LB & GL was responsible for data analysis. LB wrote the manuscript, and KT & XD reviewed and revised it. All authors reviewed the manuscript. LB, GL & GD contributed equally to this work.

**Funding**

None.

**Data Availability**

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1 Yan'an University, Yan'an 716000, China.
2 Department of General Surgery, The Second Affiliated Hospital of Air Force Medical University, Xi'an 710038, China.
3 Xi'an Medical University, Xi'an 710068, China

References


32. Zhao WM, Seki A, Fang G. Cep55, a microtubule-bundling protein, associates with centralspindlin to control the midbody integrity and cell abscission during cytokinesis. Molecular biology of the cell. 2006;17(9):3881-96.doi:10.1091/mbc.e06-01-0015.


**Figures**
Figure 1

Expression status of PC4 in cancer. (A) Observation of PC4 expression level in pan-cancer from Timer database. (B) PC4 expression in HCC and normal tissues from 4 GEO datasets (GSE25097, GSE36376, GSE39791 and GSE57957). (C) PC4 expression in HCC (n=374) and adjacent normal liver tissues (n=50) from TCGA database. (D) PC4 expression in 5 cell lines (LO2, BEL7402, HCCLM3, HepG2 and Huh7).*p < 0.05, **p < 0.01, ***p < 0.001.
Figure 2

Correlation between PC4 expression level and age (A), gender (B), pathological grade (C) and clinical stage (D) of HCC patients.
Figure 3

Correlation between PC4 expression level and survival and prognosis of HCC patients. (A) Kaplan-Meier survival curves for overall survival, progression-free survival and disease-specific survival between the PC4-high and PC4-low groups. (B) Cumulative survival curves between the PC4-high and PC4-low groups under different clinical stages. (C) Kaplan-Meier survival curves for overall survival in different age, gender, grade and stage subgroups.
Figure 4

Diagnostic and prognostic value of PC4 in HCC patients. (A) ROC curve for diagnosis to distinguish HCC from adjacent normal liver tissue. (B) ROC curve for diagnosis to distinguish Stage - HCC from adjacent normal liver tissue. (C) Time-dependent ROC curve to predict 1-, 3-, and 5-year survival rates of HCC patients. (D) Univariate and multivariate Cox analyses to identify independent prognostic factors. (E)
Nomogram for predicting 1-, 3- and 5-year survival rates of HCC patients by clinicopathological features and PC4 expression level. (F) Calibration curve for evaluating the nomogram.

Figure 5

Identification of differential gene and enrichment analysis. (A) Heatmap showing 100 differentially expressed genes between the PC4-high and PC4-low groups. (B) Volcano plot displaying the distribution...
of all differentially expressed genes based on LogFC>1/-1 and P<0.05. (C) GO and KEGG enrichment analysis of all differentially expressed genes. (D) GSEA enrichment analysis for pathway enrichment between PC4-high and PC-4 low groups.

Figure 6

Correlation of PC4 expression level with immune cell infiltration in HCC. (A) Differences in tumor-infiltrating immune cells between the PC4-high and PC4-low groups. (B) Boxplot showing the difference in TIDE scores between the two groups. (C, D) Expression differences of 24 HLA genes (E) and 8 immune checkpoint genes (F) between the PC4-high and PC4-low groups.*p < 0.05, **p < 0.01, ***p < 0.001.
Figure 7

Heatmap and forest plot of PC4 co-expressed genes. (A) Heatmap showing expression of PC4 co-expressed genes in HCC and adjacent liver tissues. (B) Forest plot displaying the results of univariate COX analysis of PC4 co-expressed genes.
Figure 8

Expression levels of 8 genes in the risk signature. (A) Expression levels of 8 genes in HCC (n=374) and adjacent normal liver tissues (n=50) from TCGA database. (B) Expression levels of 8 genes in 5 cell lines (LO2, BEL7402, HCCLM3, HepG2 and Huh7). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.
Figure 9

Survival analysis of the eight-gene risk signature in the training cohort (TCGA cohort). (A) Risk score distribution, survival status and gene expression in high-risk and low-risk groups. (B) Time-dependent ROC curve to predict 1-, 2-, and 3-year survival rates of HCC patients. (C) Overall survival Kaplan-Meier curve between the high-risk and low-risk groups. (D, E) Univariate (D) and multivariate (E) COX analysis to identify independent prognostic factors. (F, G) PCA (F) and t-SNE (G) analysis for the two groups.
Figure 10

Survival analysis of the eight-gene risk signature in the test cohort (ICGC cohort). (A) Risk score distribution, survival status and gene expression in high-risk and low-risk groups. (B) Time-dependent ROC curve to predict 1-, 2-, and 3-year survival rates of HCC patients. (C) Overall survival Kaplan-Meier curve between the high-risk and low-risk groups. (D, E) Univariate (D) and multivariate (E) COX analysis to identify independent prognostic factors. (F, G) PCA (F) and t-SNE (G) analysis for the two groups.
Figure 11

Overall survival Kaplan–Meier curves for the high-risk and low-risk groups of the training cohort (A) and test cohort (B).
Figure 12

Construction of prognostic nomogram and prediction of immunotherapy response based on the training cohort (TCGA cohort). (A) Nomogram for predicting the 1-year, 2-years, and 3-years overall survival probability of HCC patients. (B) Calibration curve for evaluating the nomogram. (C) Boxplot showing the correlation between risk score and immunotherapy response of HCC patients.
Figure 13

Construction of PC4-associated ceRNA network. (A) An miRNA-mRNA network composed of PC4 and 73 miRNAs. (B) A lncRNA-miRNA-mRNA network by PC4, miRNA-101 and 5 lncRNAs. (C) Expression differences of miRNA-101 and 5 lncRNAs in HCC and adjacent normal liver tissues based on TCGA cohort. (D) Overall survival Kaplan-Meier curves of miRNA-101 and 5 lncRNAs based on TCGA cohort.
Supplementary Files

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

- Supplementaryfigure1.tif
- Supplementaryfigure2.tif
- Supplementarytable1.xlsx
- Supplementarytable2.xlsx
- Supplementarytable3.docx
- Supplementarytable4.docx