Immune prognostic implications of PSMD14 and its associated genes signatures in hepatocellular carcinoma

Chuan Tian (teechuan2015@126.com)
Guizhou Provincial People's Hospital https://orcid.org/0000-0003-3255-6487

Mubalake Abudooryimu
nanjing university

Xinrong Lin
Nanjing University

Hao Zhou
nanjing medical university

Xiaoyuan Chu
nanjing university

Rui Wang
nanjing university

Research

Keywords: HCC, PSMD14, TCGA, ICGC, prognosis, immune infiltrates

Posted Date: April 6th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-20812/v1

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

Version of Record: A version of this preprint was published at BIOCELL on January 1st, 2021. See the published version at https://doi.org/10.32604/biocell.2021.016203.
Abstract

Background PSMD14 played a vital role in the initiation and progression of hepatocellular carcinoma (HCC). However, PSMD14 and its-related genes for the immune prognostic implications of HCC patients have rarely been analyzed. Therefore, we aimed to explore gene signatures and immune prognostic values of PSMD14 and its-related genes in HCC.

Method Analyzed the expression of PSMD14 in multiple databases, and clinicopathologic characteristics associated with PSMD14 overall survival using Wilcoxon signed-rank test, logistic and Cox regression, Kaplan-Meier method. An immune prognostic signature (including RBM45, PSMD1, OLA1, CCT6A, LCAT and IVD) was constructed and validated using the co-expression and cox regression analyses in TCGA, ICGC and TIMER datasets. Gene Set Enrichment Analysis (GSEA) was performed using TCGA data set.

Results Increased PSMD14 expression in HCC was significantly associated with poor prognosis and clinicopathologic characteristics (grade, histologic stage, surgical approach and T stage, all p-values < 0.05). A total of six PSMD14-related genes were detected, which markedly related to overall survival and immune infiltrating levels in HCC patients. Using cox regression analysis, the PSMD14 and its-related genes were found to be an independent prognostic factor for HCC survival. Calibration curves confirmed good consistency between clinical nomogram prediction and actual observation. Immune prognostic model suggests that patients in the high-risk group shown significantly poorer survival than patients in the low-risk group.

Conclusion We screened potential immune prognostic genes and constructed and verified a novel PSMD14-based prognostic model of HCC, which provides new potential prognostic biomarkers and therapeutic targets and lays a theoretical foundation for immunotherapy of HCC.

Introduction

Hepatocellular carcinoma (HCC) is one of the most common malignancy in the world and high rate of metastasis is a vital biological feature that leads to unfavorable prognosis [1–2]. Although surgery, radiofrequency ablation and chemoembolization have been widely applied for HCC therapy, the survival rate of HCC patients is still low, partly due to high heterogeneity of HCC [3]. Furthermore, due to the fact that biological processes involved in the occurrence and progression of HCC are very complicated, there have not been effective prognostic biomarkers until now [4]. Therefore, it is necessary to explore new HCC-related molecules for the diagnosis, prognosis and treatment of HCC.

The deubiquitinase (DUB) 26S proteasome non-ATPase regulatory subunit 14 (PSMD14, RPN11 or POH1), belongs to the JAMM domain metalloprotease family of DUBs and is an important part of 19S regulatory cap in 26S proteasome [5]. PSMD14 has been confirmed to play vital roles in gene ontology and related pathway, including protein stability [6], tumour formation [7], transcriptional regulation [8], double-strand DNA break repair [9], senescence [10], apoptosis and resistance [11], growth and metastasis [5], BMP6 signaling [6], TGF-β signaling [8]. Recently, PSMD14 has been overexpressed in
many human cancers, such as HCC, colorectal cancer, multiple myeloma, esophageal squamous cell carcinoma and breast cancer, which it This is related to the poor prognosis of the patients [5,6,11–13]. PSMD14 acts as an oncogene that promotes tumor progression by deubiquitinating of different protein substrates. For example, Deubiquitinase PSMD14 acts as a positive regulator for the initiation of the BMP6 signalling pathway, resulting in increased stability of the ALK2, either PSMD14 or ALK2 depletion significantly decreases colorectal cancer growth and chemoresistance [6]. Similarly, PSMD14 inhibits degradation of GRB2 by deubiquitinating this oncoprotein and stabilizing GRB2, which it enhances hepatocellular carcinoma growth and metastasis [5]. POH1 deubiquitinates the TGF-β receptors and CAV1, contributes to hyperactivation of TGF-β signaling and facilitates hepatocellular carcinoma metastasis, which it negatively regulates lysosome pathway-mediated turnover of TGF-β receptors [8]. However, other mechanisms of the contribution of PSMD14 in the progression of HCC remained to be further explored.

Here, we evaluated the PSMD14 mRNA and protein levels and the prognostic value in Multiple databases, increased PSMD14 expression associated with poor survival in HCC. Moreover, we screened prognostic genes signature closely related to PSMD14 through bioinformatics analysis and constructed the PSMD14-related prognostic model, and validated in The Cancer Genome Atlas (TCGA) and independent International Cancer Genome Consortium (ICGC) database. In addition, confirmed the correlation PSMD14 expression with clinicopathological characteristics, signature (RBM45, PSMD1, OLA1, CCT6A, LCAT and IVD) and immune infiltration of hepatocellular carcinoma patients. In conclusion, our study confirms the important role of PSMD14 and its associated genes in Liver cancer as well as correlated with prognosis and immune infiltrating levels in HCC patients, providing new insights relevant to individualized treatment.

Results

Patients characteristics

As shown in Table 1, 418 primary tumors with both gene expression and clinical data were downloaded from TCGA (LIHC-CHOL) data in November 2019. The median age was 61 years old at diagnosis. Moreover, Histopathologic distribution of HCC included well-differentiated (14.8%), moderately differentiated (81.7%) and poorly differentiated (3.5%). The cancer status included 254 tumor-free (65.6%) and 133 with tumor (34.4%). Stage I disease was found in 194 patients (49.2%), stage II in 98 (24.9%), stage III in 90 (22.8%) and stage IV in 12 (3.1%). Grade I disease was found in 55 patients (14.8%), Grade II in 180 (48.4%), Grade III in 124 (33.3%) and Grade IV in 13 (3.5%). The Vascular invasion include 17 Macro (5.3%), 94 Micro (29.3%) and 210 NO (65.4%). Most tumors (87.8%, n = 367) were of Hepatocellular Carcinoma (HCC), 9.8% (n = 41) were Cholangiocarcinoma (CHOL), 1.7% (n = 7) were Hepatob膽管癌 (Mixed) and 0.7% (n = 3) were Fibrolamellar Carcinoma (FLC). Tumor size (≤ 5 cm, 74.9%, n = 311) and (> 5 cm, 25.1%, n = 104). Eight of 298 (2.7%) cases had lymph node metastases. Eight of 311 (2.6%) cases had distant metastases.
### Table 1

**TCGA hepatic carcinoma patient characteristics**

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>Total (418)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at diagnosis (y)</td>
<td>61(16–90)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>272/146</td>
<td>65.1/34.9</td>
</tr>
<tr>
<td>Histologic grade</td>
<td>55/304/13</td>
<td>14.8/81.7/3.5</td>
</tr>
<tr>
<td>Status</td>
<td>133/254</td>
<td>34.4/65.6</td>
</tr>
<tr>
<td>TNM Stage</td>
<td>194/98/90/12</td>
<td>49.2/24.9/22.8/3.1</td>
</tr>
<tr>
<td>Grade</td>
<td>55/180/124/13</td>
<td>14.8/48.4/33.3/3.5</td>
</tr>
<tr>
<td>Vascular invasion</td>
<td>17/94/210</td>
<td>5.3/29.3/65.4</td>
</tr>
<tr>
<td>Risk Factors</td>
<td>78/222</td>
<td>26.0/74.0</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatocellular Carcinoma</td>
<td>367</td>
<td>87.8</td>
</tr>
<tr>
<td>Cholangiocarcinoma</td>
<td>41</td>
<td>9.8</td>
</tr>
<tr>
<td>Hepatocholangiocarcinoma (Mixed)</td>
<td>7</td>
<td>1.7</td>
</tr>
<tr>
<td>Fibrolamellar Carcinoma</td>
<td>3</td>
<td>0.7</td>
</tr>
<tr>
<td>Tumor size</td>
<td>≤ 5 cm</td>
<td>311</td>
</tr>
<tr>
<td></td>
<td>&gt; 5 cm</td>
<td>104</td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>Negative/ Positive</td>
<td>290/8</td>
</tr>
<tr>
<td>Distant metastasis</td>
<td>Negative/ Positive</td>
<td>303/8</td>
</tr>
</tbody>
</table>

**PSMD14 mRNA expression**

Oncomine database analysis demonstrated that significantly upregulation in 2 Liver cancer data set compared to the normal liver tissues (Fig. 1A). Complementary, PSMD14 mRNA expression was higher in most tumors, such as bladder cancer, brain and CNS, breast cancer, cervical, colorectal cancer, head and neck cancer, kidney cancer, liver cancer, lung cancer, lymphoma, myeloma, pancreatic cancer and sarcoma. To further assess PSMD14 expression of liver cancer, we verified PSMD14 mRNA expression by TIMER data set (Fig. 1B). PSMD14 mRNA expression was significantly higher in Liver Hepatocellular Carcinoma (LIHC) compared with adjacent normal tissues.

**PSMD14 expression and its clinical significance in HCC**

To further assess the role of PSMD14 in HCC, we first analyzed four independent microarray datasets from Oncomine database. The median rank of PSMD14 in up-regulated genes of HCC was 1197.5 based on a meta analysis across the four datasets, including 4 analyses using the Oncomine algorithms (760
samples, $P = 2.96 \times 10^{-5}$, Fig. 2A) [14–16]. In the meantime, we examined the protein expression of PSMD14 using the Human Protein Atlas (HPA). Images revealed a markedly high expression of PSMD14 could be observed compared to the normal Liver tissues using the HPA002114 antibody in Liver cancer (Fig. 2B). Next, we evaluated PSMD14 mRNA expression compared with normal tissues in multiple HCC studies from TCGA and ICGC databases, and found that the mRNA level of PSMD14 was significantly higher in HCC patients than normal liver tissues (Fig. 2C-D). Besides, the results of TCGA database, which were mainly analyzed by the web-based tools, UALCAN, GEPIA and LinkedOmics database, showed that PSMD14 expression was significantly overexpressed in human HCC tissues compared with normal tissues (Fig. 2E-F), and elevated PSMD14 expression level corresponds to a shorter overall survival in HCC (Fig. 2G). These findings suggested that PSMD14 is highly expressed in HCC and elevated PSMD14 may predict a poor outcome for HCC patients.

We analyzed 377 HCC samples with PSMD14 expression and clinical characteristics from the TCGA database. As shown in Fig. 3 (A-D). Significant increase in PSMD14 expression correlated with tumor histological grade ($p = 0.046$), histological stage ($p < 0.001$), surgical approach ($p = 0.001$) and T stage ($p = 0.013$). Categorical dependent variable using logistic regression, univariate analysis showed that PSMD14 expression was significantly associated with poor prognostic clinicopathologic characteristics (Table 2). Increased PSMD14 expression was significantly associated with grade (OR = 1.9 for well vs. moderate), stage (OR = 2.5 for I vs. II, III), surgical approach (OR = 0.26 for HL vs. SSLR), tumor size (OR = 1.65 for $\leq 5$ cm vs. $> 5$ cm) in HCC (all $p$-values < 0.05). The results suggested that HCC with high PSMD14 expression was progressed to a more advanced stage than those with low PSMD14 expression.
Table 2
PSMD14 expression associated with clinical pathological characteristics from TCGA (logistic regression)

<table>
<thead>
<tr>
<th>Clinical characteristics</th>
<th>Total (N)</th>
<th>Odds ratio in PSMD14 expression</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>376</td>
<td>0.99 (0.98–1.01)</td>
<td>0.444</td>
</tr>
<tr>
<td>BMI</td>
<td>341</td>
<td>0.99 (0.96–1.01)</td>
<td>0.409</td>
</tr>
<tr>
<td>Gender (female vs. male)</td>
<td>377</td>
<td>1.02 (0.66–1.57)</td>
<td>0.941</td>
</tr>
<tr>
<td>Grade (well vs. moderate)</td>
<td>359</td>
<td>1.91 (1.06–3.51)</td>
<td>0.033</td>
</tr>
<tr>
<td>Histology (FLC vs. HCC)</td>
<td>370</td>
<td>0.48 (0.02–5.06)</td>
<td>0.552</td>
</tr>
<tr>
<td>Status (tumor free vs. with tumor)</td>
<td>349</td>
<td>0.90 (0.57–1.41)</td>
<td>0.644</td>
</tr>
<tr>
<td>Vascular invasion (positive vs. negative)</td>
<td>321</td>
<td>0.73 (0.45–1.16)</td>
<td>0.180</td>
</tr>
<tr>
<td>Hepatitis B (positive vs. negative)</td>
<td>300</td>
<td>0.84 (0.49–1.41)</td>
<td>0.504</td>
</tr>
<tr>
<td>Stage (I vs. II or III)</td>
<td>353</td>
<td>2.51 (1.25–4.32)</td>
<td>0.001</td>
</tr>
<tr>
<td>Surgical approach (HL vs. SSLR)</td>
<td>113</td>
<td>0.26 (0.09–0.67)</td>
<td>0.007</td>
</tr>
<tr>
<td>Tumor size (≤ 5 cm vs. &gt;5 cm)</td>
<td>374</td>
<td>1.65 (1.03–2.67)</td>
<td>0.039</td>
</tr>
<tr>
<td>Distant metastasis (positive vs. negative)</td>
<td>276</td>
<td>1.00 (0.12–8.43)</td>
<td>1.000</td>
</tr>
<tr>
<td>Lymph nodes (positive vs. negative)</td>
<td>261</td>
<td>3.05 (0.38–62.07)</td>
<td>0.337</td>
</tr>
</tbody>
</table>

FLC: Fibrolamellar Carcinoma; HCC: Hepatocellular Carcinoma; HL: Hepatic Lobectomy; SSLR: single segment liver resection

Survival and independent prognostic analysis

Kaplan-Meier survival analysis indicated that HCC with PSMD14-high had a poor prognosis than the PSMD14-low in TCGA and ICGC database in Fig. 3 (E-F) (p < 0.001). The univariate analysis revealed that PSMD14-high correlated significantly with a poor OS (hazard ratio [HR]: 2.17; 95% confidence interval [CI]: 1.32–3.57; p = 0.002). Other clinicopathologic variables related to poor survival include status, stage, T stage and distant metastasis. Multivariate analysis with HR of 1.9 shown that the PSMD14 remained independently associated with overall survival (CI: 1.08–3.33, p = 0.026), along with status in TCGA (Table 3). Therefore, we first revealed that PSMD14 was associated with poor prognosis, as an independent prognostic factor for HCC survival.
Table 3
Univariate and multivariate analyses of overall survival in hepatocellular carcinoma patients of TCGA

<table>
<thead>
<tr>
<th>Variables</th>
<th>Univariate analysis</th>
<th></th>
<th></th>
<th>Multivariate analysis</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P-value</td>
<td>HR (95% CI)</td>
<td>P-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSMD14</td>
<td>2.17(1.32–3.57)</td>
<td>0.002</td>
<td>1.90(1.08–3.33)</td>
<td>0.026</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.02(1.00–1.05)</td>
<td>0.107</td>
<td>1.02(0.99–1.04)</td>
<td>0.190</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.60(0.34–1.06)</td>
<td>0.078</td>
<td>0.97(0.51–1.87)</td>
<td>0.936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vascular invasion</td>
<td>1.27(0.70–2.28)</td>
<td>0.435</td>
<td>0.84(0.44–1.62)</td>
<td>0.606</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td>1.18(0.80–1.75)</td>
<td>0.400</td>
<td>1.10(0.70–1.74)</td>
<td>0.677</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td>2.65(1.49–4.72)</td>
<td>0.001</td>
<td>1.91(1.02–3.55)</td>
<td>0.042</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage</td>
<td>1.72(1.28–2.32)</td>
<td>0.000</td>
<td>0.78(0.15–3.96)</td>
<td>0.764</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T stage</td>
<td>1.65(1.24–2.20)</td>
<td>0.001</td>
<td>1.91(0.42–8.67)</td>
<td>0.399</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distant metastasis</td>
<td>4.15(1.00-17.24)</td>
<td>0.050</td>
<td>0.95(0.17–5.18)</td>
<td>0.954</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>3.42(0.47–24.96)</td>
<td>0.226</td>
<td>3.33(0.08-143.94)</td>
<td>0.531</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Construction and validation of signature

We selected the genes signatures associated with PSMD14 in the TCGA-LIHC database. A total of 2478 PSMD14-associated genes (Pearson |R| > 0.4, p < 0.001) were chosen to generate prognosis gene signatures and the top 10 positively and negatively genes were selected for further analysis (Fig. 4A). All genes were analyzed by univariate Cox regression. A total of 17 genes were significantly related to the OS in TCGA-LIHC database (Fig. 4B). Then, the LASSO COX regression analysis and the regression coefficient was computed, the model achieved the best performance at 6 genes (Fig. 4C). Finally, constructed a risk signature for HCC using multivariate COX regression (Table 4). All patients were divided into high- and low-risk sets based on median risk score in the TCGA and ICGC database. Status, survival time and gene expression levels of patients were shown in TCGA (Fig. 4D) and ICGC (Fig. 4E).
Table 4  
The HR and p values of 6 Genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Gene description</th>
<th>HR</th>
<th>HR.95L</th>
<th>HR.95H</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBM45</td>
<td>RNA Binding Motif Protein 45</td>
<td>5.022</td>
<td>2.679</td>
<td>9.411</td>
<td>4.74E-07</td>
</tr>
<tr>
<td>PSMD1</td>
<td>26S Proteasome Regulatory Subunit S1</td>
<td>2.982</td>
<td>1.978</td>
<td>4.496</td>
<td>1.83E-07</td>
</tr>
<tr>
<td>OLA1</td>
<td>Obg Like ATPase 1</td>
<td>2.213</td>
<td>1.603</td>
<td>3.055</td>
<td>1.39E-06</td>
</tr>
<tr>
<td>CCT6A</td>
<td>Chaperonin Containing TCP1 Subunit6A</td>
<td>2.118</td>
<td>1.564</td>
<td>2.869</td>
<td>1.24E-06</td>
</tr>
<tr>
<td>IVD</td>
<td>Isovaleryl-CoA Dehydrogenase</td>
<td>0.636</td>
<td>0.495</td>
<td>0.816</td>
<td>0.000</td>
</tr>
<tr>
<td>LCAT</td>
<td>Lecithin-Cholesterol Acultransferase</td>
<td>0.775</td>
<td>0.684</td>
<td>0.877</td>
<td>5.63E-05</td>
</tr>
</tbody>
</table>

Note: HR and P values were generated from univariate Cox regression in TCGA dataset

The survival analysis indicated that the OS of the low-risk set was better than that of high-risk set in the TCGA database (P < 0.001) (Fig. 5A). The results were consistent in the ICGC database (P < 0.001) (Fig. 5B). The area under the ROC curve (AUC) for 1-, 3-, and 5-year OS were 0.723, 0.653, 0.645, 0.657, 0.705, 0.683 in the TCGA (Fig. 5C) and ICGC (Fig. 5D) cohorts, respectively. Together, Results revealed that the signature showed excellent performance for OS prediction.

Univariate and multivariate COX regression analysis

Univariate Cox regression indicated that stage, T stage, M stage and risk score in the TCGA database (stage: P < 0.001; T stage: P < 0.001; M stage: P = 0.026; risk score: P < 0.001; Fig. 6A), and gender, stage, risk score in the ICGC database (gender: P = 0.039; stage: P < 0.001; risk score: P < 0.001; Fig. 6C) were predictors for OS. Moreover, multivariate Cox regression analysis verified that age (HR = 1.019, 95% CI [1.003–1.034]; P = 0.017) and risk score (HR = 1.545; 95% CI [1.354–1.763]; P < 0.001; Fig. 6B) were significant independent risk factors in the TCGA database. Multivariate Cox regression further showed that gender (HR = 0.372, 95% CI [0.196–0.707]; P = 0.003); stage (HR = 2.320, 95% CI [1.599–3.367]; P < 0.001); priorMalignancy (HR = 2.500, 95% CI [1.065–5.868]; P = 0.035) and risk score (HR = 1.094; 95% CI [1.031–1.162]; P = 0.003; Fig. 6D) was significant independent risk factors in the ICGC database. These data demonstrated that this signature was an independent risk factor of HCC.

Nomogram construction

Based on the prognostic signature and clinical factors, such as age, gender, vascular invasion, tumor status and stage, a nomogram was constructed (Fig. 7A). The calibration curve was used to describe the prediction value of the nomogram and the 45-degree line indicates the actual survival outcomes. The results showed that the nomogram-predicted survival closely matched with the best prediction performance for predicting 1-, 3- and 5-year OS (Fig. 7B). The 1-year AUC was 0.765 for nomogram, and 0.481 for age, 0.490 for grade, 0.425 for status, 0.711 for stage. The 3-year AUC was 0.697 for
nomogram, and 0.508 for age, 0.525 for grade, 0.567 for status, 0.706 for stage. Moreover, the 5-year AUC was 0.715 for nomogram, and 0.594 for age, 0.508 for grade, 0.607 for status, 0.667 for stage (Fig. 7C-E). These showed that compared with a single clinical factor, the nomogram had great predictive accuracy combined the signature and clinical factors.

Model gene verification

Consistent with our results, IVD and LCAT were found to be significantly downexpressed, while CCT6A and OLA1 were significantly overexpressed for liver cancer in the Oncomine (Fig. 8A), TIMER (Fig. 8B), TCGA (Fig. 8C) and ICGC (Fig. 8D) database. Though lack in the Oncomine database, the mRNA expression of PSMD1 and RBM45 were also found to be significantly overexpressed for liver cancer in the TIMER, TCGA and ICGC database. Taking together, we further verified aberrant expression of six prognostic genes and found that IVD and LCAT genes were significantly decreased, while CCT6A, OLA1, PSMD1 and RBM45 genes were increased in HCC tissues.

Gene set enrichment analyses

We carried on the Gene Set Enrichment Analysis (GSEA) between PSMD14 (high/low) expression and high/low risk model data sets in TCGA-LIHC. GSEA reveal significant difference (FDR < 0.05, NOM p-val < 0.05) in enrichment of MSigDB Collection (c2.cp.kegg.v7.0.symbols.gmts). We selected the most significantly enriched signaling pathways based on their normalized enrichment score (NES). A great majority of the enriched pathways were metabolism related, such as the purine metabolism, pyrimidine metabolism and ubiquitin mediated proteolysis are differentially enriched in PSMD14 high expression and high-risk group phenotype (Fig. 9A). Besides, the glycine serine and threonine metabolism, primary bile acid biosynthesis and retinol metabolism were enriched in PSMD14 low expression and low-risk group phenotype (Fig. 9B).

Seven genes expression is correlated with immune infiltration level

We analyzed the correlation between seven prognostic genes (PSMD14, RBM45, PSMD1, OLA1, CCT6A, LCAT and IVD) expression and 6 types of infiltrating immune cells (B cells, CD4 T cells, CD8 + T cells, neutrophils, macrophages and dendritic cells) in LIHC. The results showed that PSMD14 expression levels had a significantly positive correlation with infiltrating levels of B Cell (r = 0.245, P = 4.27e-06), CD8 + T cell (r = 0.205, P = 1.30e-04), CD4 + T cell (r = 0.266, P = 5.57e-07), Macrophage (r = 0.374, P = 9.16e-13), Neutrophils (r = 0.445, P = 3.30e-18) and Dendritic cells (r = 0.317, P = 2.33e-09). RBM45 expression levels had a significantly positive correlation with infiltrating levels of B Cell (r = 0.35, P = 2.39e-11), CD8 + T cell (r = 0.26, P = 1.07e-06), CD4 + T cell (r = 0.367, P = 2.12e-12), Macrophage (r = 0.461, P = 2.55e-19), Neutrophils (r = 0.471, P = 1.72e-20) and Dendritic cells (r = 0.419, P = 6.90e-16). PSMD1 expression levels had a significantly positive correlation with infiltrating levels of B Cell (r = 0.257, P = 1.31e-06), CD8 + T cell (r = 0.192, P = 3.50e-04), CD4 + T cell (r = 0.299, P = 1.55e-08), Macrophage (r = 0.393, P = 4.53e-14),
Neutrophils ($r = 0.447$, $P = 2.22e-18$) and Dendritic cells ($r = 0.355$, $P = 1.50e-11$). OLA1 expression levels had a significantly positive correlation with infiltrating levels of B Cell ($r = 0.359$, $P = 6.18e-12$), CD8 + T cell ($r = 0.27$, $P = 3.89e-07$), CD4 + T cell ($r = 0.328$, $P = 4.51e-10$), Macrophage ($r = 0.444$, $P = 6.18e-18$), Neutrophils ($r = 0.382$, $P = 1.98e-13$) and Dendritic cells ($r = 0.399$, $P = 1.93e-14$). CCT6A expression levels had a significantly positive correlation with infiltrating levels of B Cell ($r = 0.322$, $P = 9.24e-10$), CD8 + T cell ($r = 0.16$, $P = 3.06e-03$), CD4 + T cell ($r = 0.352$, $P = 1.72e-11$), Macrophage ($r = 0.462$, $P = 1.90e-19$), Neutrophils ($r = 0.367$, $P = 2.04e-12$) and Dendritic cells ($r = 0.313$, $P = 3.45e-09$). None of the above five genes had no significant correlations with tumor purity ($P > 0.05$) (Fig. 10A-E).

In addition, LCAT expression was significantly negatively related to tumor purity ($r = -0.124$, $P = 2.12e-02$) and had a significantly negative correlation with infiltrating levels of B Cell ($r = -0.197$, $P = 2.35e-04$), CD8 + T cell ($r = -0.103$, $P = 5.59e-02$), CD4 + T cell ($r = -0.174$, $P = 1.16e-03$), Macrophage ($r = -0.244$, $P = 5.21e-06$), Neutrophils ($r = -0.202$, $P = 1.53e-04$) and Dendritic cells ($r = -0.194$, $P = 3.12e-04$) (Fig. 10F). IVD expression was significantly positive related to tumor purity ($r = 0.121$, $P = 2.45e-02$) and had a significantly negative correlation with infiltrating levels of B Cell ($r = -0.173$, $P = 1.31e-03$), CD8 + T cell ($r = -0.097$, $P = 7.37e-02$), CD4 + T cell ($r = -0.135$, $P = 1.20e-02$), Macrophage ($r = -0.171$, $P = 1.49e-03$) and Dendritic cells ($r = -0.145$, $P = 7.53e-03$) and no significant correlations with Neutrophils ($r = -0.065$, $P = 2.271e-01$) (Fig. 10G).

**Discussion**

HCC remains a common malignant tumor in the world and the important reasons for the low survival rate. Thus, in view of the low survival rate for HCC, investigation of novel biomarkers and models is necessary. Recently, PSMD14-based signatures have revealed excellent potential in prognosis prediction of multiple cancers. The prognosis value of PSMD14 in cancers have been reported, including HCC, esophageal cancer, breast cancer [5,12–13]. Our results are in accordance with these studies, implying that PSMD14 could be a potential prognostic gene for HCC [5]. To our knowledge, the expression of PSMD14 and its potential immune prognostic impact on HCC has not yet been explored, the potential immune prognostic role of PSMD14 and its associated genes in HCC was the focus on the present study.

In the present study, we verified the high expression of PSMD14 in liver cancer through Multiple databases, including ONCOMINE, TIMER, HPA, TCGA, ICGC, UALCAN and GEPIA. Bioinformatic analysis from TCGA and ICGC database demonstrated that high expression of PSMD14 in HCC was associated with clinical pathologic characteristics (high grade, histologic stage, surgical approach, T stage), survival time and poor prognosis. Univariate and multivariate COX regression shows PSMD14 as an independent prognostic factor (Table 3). In addition, we identified the PSMD14 associated novel biomarkers base on TCGA data set, and confirmed PSMD14 associated genes model significantly correlated with prognosis using COX regression and LASSO analysis. Finally, six genes (RBM45, PSMD1, OLA1, CCT6A, LCAT and IVD) were selected to fabricate a prognostic signature for HCC, validated its efficiency in TCGA and ICGC data set. A robust nomogram consisted of the 6-genes signature, age, gender, vascular invasion, tumor status and the stage was constructed for prognostic prediction of HCC patients. At the same time, the
AUC value of the signature-based nomogram was nearly better than the AUC values of age, grade status and stage in 1-, 3, 5-years. Next, we further verified the mRNA expression of 6-genes (RBM45, PSMD1, OLA1, CCT6A, LCAT and IVD) for HCC patients in the Oncomine, TIMER, TCGA and ICGC database, which is consistent with the results of previous studies [17–20]. What's interesting, no reports of RBM45 and IVD genes in liver cancer. They may become new targets for the treatment of liver cancer, which need to be further studied.

A total of 7 key prognostic genes of HCC patients were selected in this paper. However, no review had been studied about intriguing mechanisms of these key genes in HCC except PSMD14 and CCT6A. POH1 was first found as a human pad1 homologue in 1997, a previously unidentified component of the human 26S proteasome, that degrades proteins targeted for destruction by the ubiquitin pathway [21]. Lv [5] reveal that PSMD14, as a novel posttranslational regulator of GRB2, inhibits degradation of GRB2 via deubiquitinating this oncoprotein in HCC cells. Overexpressing CCT6A Contributes To Cancer Cell Growth By Affecting The G1-To-S Phase Transition And Predicts A Negative Prognosis In Hepatocellular Carcinoma. CCT6A may contribute to HCC cell proliferation by accelerating the G1-to-S transition and predicts a negative prognosis in Hepatocellular Carcinoma [19]. Consistent with the previously published papers, the high expression of PSMD14 and CCT6A were found positively related to the poor prognostic of HCC patients in our research.

Enrichment analysis revealed many significantly enriched pathways for the signature, of which most were metabolism related. To further investigate the functions of PSMD14 and its associated genes in HCC, we performed GSEA using TCGA data, GSEA showed that purine metabolism, pyrimidine metabolism and ubiquitin mediated proteolysis are differentially enriched between PSMD14 gene high expression and PSMD14 associated genes high risk phenotype. Besides, glycine serine and threonine metabolism, primary bile acid biosynthesis and retinol metabolism are differentially enriched between PSMD14 gene low expression and PSMD14 associated genes low risk phenotype.

Another important aspect of this study, seven genes mRNA expression was correlated with diverse immune infiltration levels in LIHC. PSMD14, RBM45, PSMD1, OLA1 and CCT6A mRNA expression level was significantly positively correlated with infiltrating levels of B Cell, CD8 + T cell, CD4 + T cell, Macrophage, Neutrophils and Dendritic cells. In addition, LCAT mRNA expression had a significantly negative correlation with infiltrating levels of tumor purity, B Cell, CD8 + T cell, CD4 + T cell, Macrophage, Neutrophils and Dendritic cells. IVD mRNA expression was significantly positive related to tumor purity and had a significantly negative correlation with infiltrating levels of B Cell, CD8 + T cell, CD4 + T cell, Macrophage and Dendritic cells and no significant correlations with Neutrophils. Moreover, the correlation between seven immune genes expression and the signature genes of immune cells imply the role of its in regulating tumor immunology. Of note, among the 7 genes, only the PSMD14 and CCT6A genes, have been reported that correlated with diverse immune cells levels. For instance, POH1 deficiency in macrophages resulting in deubiquitination of pro-IL-1β that restrains inflammatory responses for the maintenance of immune homeostasis [22]. CCT6A research shown that CD8(+) T cells provide functional in cytokine secretion and lytic activity upon of their cognate antigens, it was used in personalised
adoptive T-cell therapy of melanoma [23]. These correlations could be indicative of a potential mechanism that 7 immune-related genes regulated immune cells in HCC. The above suggested that PSMD14 and signature genes could play a significant role in recruitment and regulation of immune infiltrating cells in HCC.

In summary, here a systemic analysis was performed on PSMD14 expression and prognosis, which provided a further understanding of the biomolecular properties of HCC. Next, we identified and verified that immune-related genes (PSMD14, RBM45, PSMD1, OLA1, CCT6A, LCAT and IVD) as prognostic signatures for HCC. Therefore, it deepens our understanding of liver cancer immune-related genes and provides new potential biomarkers for prognosis and treatment.

Materials And Methods

Oncomine database

PSMD14 expression levels were identified in the Oncomine database (https://www.oncomine.org/resource/login.html) in various common of cancers. The data is defined as a P-value of 0.01, a fold change of 1.5, a top 10% gene ranking, and the data has to be from mRNA.

PSMD14 mRNA expression and survival in public databases

To investigate the expression level and prognostic role of PSMD14 mRNA in HCC, the Gene Expression Profiling Interactive Analysis (GEPIA) (http://gepia.cancer-pku.cn/index.html), UALCAN database (http://ualcan.path.uab.edu/analysis.html), and the Human Protein Atlas database (http://www.proteinatlas.org), LinkedOmics (http://www.linkedomics.org/admin.php), the Cancer Genome Atlas (TCGA) database (http://cancergenome.nih.gov/), the International Cancer Genome Consortium (ICGC) database (https://dcc.icgc.org/) were used to determine the prognostic significance.

TIMER database

We analyzed the correlation between signature gene expression and 6 types of immune infiltrating cells (B cells, CD4 + T cells, CD8 + T cells, neutrophils, macrophages and dendritic cells) in HCC patients via The Tumor IMMune Estimation Resource (TIMER) algorithm database (https://cistrome.shinyapps.io/timer/). Tumor purity is a vital factor that influences the analysis of tumor immune infiltration by genomic approaches.

Data collection

RNA-sequencing and clinical information for hepatic carcinoma were acquired from the Cancer Genome Atlas (TCGA) database and The International Cancer Genome Consortium (ICGC) database. The clinical features of patients with hepatic carcinoma in the TCGA cohorts are presented in Table 1.

Selection of immune-related prognostic genes
Genes significantly associated with PSMD14 (Pearson $|R| > 0.4$, $p < 0.001$) were filtered by Pearson correlation analysis in TCGA data sets. The top 10 genes that were positively and negatively correlated with PSMD14 were selected for further analysis.

**Identification and validation of the prognostic gene signature**

Using univariate and least absolute shrinkage and selection operator (LASSO) COX regression, we used to filter the independent risk prognostic genes. Moreover, multivariate COX regression was used to identify corresponding coefficients of HCC prognostic signature using the R package “glment”, “survminer” and “survival”. The risk score of every patient was calculated from the TCGA and ICGC database based on the signature. With the median score as cut-off value, all samples were randomly separated to high- and low-risk sets. In addition, survival analysis was evaluated for each set using the Kaplan-Meier curve and log-rank test. The receiver operating characteristic (ROC) curve and the area under the curve (AUC) were drawn using the R package “survivalROC”.

**Development of nomogram**

We construct a nomogram for age, gender, stage, T, N, M, and risk score using the survival and the rms package of R. Next, we used a calibration curve to assess the concordance between actual and predicted survival. In addition, The concordance index (C-index) was computed to evaluate the model performance for predicting prognosis, which was ranged from 0.5 to 1.0. Value of 0.5 and 1.0 represents a random chance and an excellent capacity for predicting survival with the model, respectively.

**Gene set enrichment analyses**

Gene set enrichment analysis (GSEA) was conducted to identify gene sets that differed significantly between the high- and low-risk patient groups, using GSEA 4.0.3 software (http://www.broadinstitute.org/gsea/index.jsp). In the analysis results, it was generally believed that the pathways were significantly enriched when $|NES| > 1$, NOM p-value $< 0.05$, and FDR q-value $< 0.25$.

**Statistical analysis**

All statistics were executed using the R software (v.3.4.3). Clinicopathologic characteristics associated with PSMD14 overall survival were analyzed using the Wilcoxon signed-rank test, logistic and Cox regression, Kaplan-Meier method. In addition, x2 test was used to check the association of risk scores with clinical characteristics and univariate and multivariate Cox proportional hazard regression analysis was performed to evaluate the association between risk score and OS. The Receiver Operating Characteristic (ROC) analysis was used to examine the sensitivity and specificity of survival prediction using the gene signature risk score. An area under the ROC curve (AUC) served as an indicator of prognostic accuracy. P-values $< 0.05$ were considered statistically significant.
Abbreviations

AUC: area under the curve; C-index: concordance index; GO: Gene Ontology; GSEA: Gene Set Enrichment Analysis; HR: hazard ratio; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: least absolute shrinkage and selection operator; HCC: hepatocellular carcinoma; CHOL: Cholangiocarcinoma; FLC: Fibrolamellar Carcinoma; HCC: Hepatocellular Carcinoma; HL: Hepatic Lobectomy; SSLR: single segment liver resection; OS: overall survival; ROC: Receiver Operating Characteristics; TCGA: The Cancer Genome Atlas; ICGC: International Cancer Genome Consortium.

Declarations

Acknowledgements

Not applicable.

Authors’ contributions

CT and MA wrote the paper; XRL and HZ assisted in the preparation of the manuscript and editing. XYC and RW coordinated all the research activities, from the design of the experiment to the final correction of the manuscript. All authors have read and approved the final manuscript.

Funding

NO.

Competing interests

The authors declare that they have no competing interests.

Funding

NO.

Availability of data and materials

The TCGA, ICGC and Other databases material is public available.

References


Figures
Figure 1

PSMD14 expression levels in human cancers. (A) PSMD14 in data sets of different cancers in the Oncomine database. (B) PSMD14 expression levels in different tumor types were determined by TIMER (*P < 0.05, **P < 0.01, ***P < 0.001).
PSMD14 expression is upregulated in HCC. (A) A meta-analysis of PSMD14 gene expression from four Oncomine databases where colored squares indicate the median rank for PSMD14 (vs. Normal tissue) across 4 analyses. Chen Liver (1), Roessler Liver (2-3), Wurmbach Liver (4). The P value is given for the median rank analysis. (B) The representative protein expression of the PSMD14 in HCC and normal liver tissue. Data were from the Human Protein Atlas (http://www.proteinatlas.org) database. (C) PSMD14 expression in TCGA LIHC Database. (D) PSMD14 expression in ICGC LIHC Database. (E) Expression of PSMD14 in LIHC. (F) PSMD14 expression by CoxPH analysis. (G) Survival analysis of patients with HCC based on PSMD14 expression.
significantly increased in human 374 LIHC tissues to 50 normal tissues using TCGA Database. (D) PSMD14 significantly increased in human 243 LIHC tissues to 202 normal tissues using ICGC Database. (E) Relative expression of PSMD14 in 50 normal and 371 LIHC samples using UALCAN. (F) Boxplot showing relative expression of PSMD14 in normal and tumor samples using GEPIA. (G) Impact of PSMD14 expression on overall survival in HCC patients in LinkedOmics (*P < 0.05, **P <0.01, ***P < 0.001).

Figure 5

Figure 7

Establishment and validation of the six genes prognostic signature. (A) Circle chart showing PSMD14 related genes (top 10 positive and negative related genes). (B-C) The procedure of the establishment of the prognostic signature. (D) Correlation between the prognostic signature and the overall survival of patients in the TCGA cohort (D) and ICGC (E) cohorts. The distribution of risk scores (upper), survival time (middle) and genes expression levels (below). The black dotted lines represent the median risk score cut-off dividing patients into low- and high-risk groups. The red dots and lines represent the patients in high-risk groups. The green dots and lines represent the patients in low-risk groups.
Figure 9

Kaplan-Meier survival and ROC curves of the six genes prognostic signature. (A, B) Kaplan-Meier survival curves of overall survival among risk stratification groups in the TCGA (A) and ICGC (B) set. (C, D) ROC curves with calculated AUCs for risk prediction in 1-, 3-, 5-years in the TCGA (C) and ICGC(D) sets.
Figure 11

Univariate and multivariate Cox regression analyses of clinical factors associated with overall survival. (A-B) Univariate Cox regression analyses of clinical factors associated with overall survival in the TCGA (A) and ICGC (B) set. (C-D) Multivariate Cox regression analyses of clinical factors associated with overall survival in the TCGA (C) and ICGC (D) sets.
Figure 13

Construction of a nomogram for overall survival prediction in HCC. (A) The nomogram consists of age, gender, grade, vascular invasion, tumor status, stage, and the risk score based on the six genes signature. (B) Calibration curves of the nomogram for the estimation of survival rates at 1-, 3-, 5-year. (C-E) The Kaplan-Meier curves of the risk subgroups stratified by the tertiles of total points derived from the nomogram.
Figure 15

Expression of the six predictive genes in HCC. (A) The expression profiles of the six genes in the Oncomine database. Data of PSMD1 and RBM45 in liver cancer were not found in the database. (B) The expression of the six predictive genes in TIMER-LIHC database. (C-D) The expression of six predictive genes in TCGA-LIHC and ICGC-LIHC database.
Figure 17

The significantly enriched KEGG pathways in TCGA-LIHC database by GSEA. (A) Six representative KEGG pathways in PSMD14 expression. (B) Six representative KEGG pathways in risk model.
Figure 19

Correlation of seven genes expression with immune infiltration level in LIHC. (A-E) RBM45, PSMD1, OLA1 and CCT6A mRNA expression level was significantly positively correlated with infiltrating levels of B Cell, CD8+ T cell, CD4+ T cell, Macrophage, Neutrophils and Dendritic cells. (F) LCAT expression was significantly negatively related to tumor purity and had a significantly negative correlation with infiltrating levels of B Cell, CD8+ T cell, CD4+ T cell, Macrophage, Neutrophils and Dendritic cells. (G) IVD expression was significantly positive related to tumor purity and had a significantly negative correlation with
infiltrating levels of B Cell, CD8+ T cell, CD4+ T cell, Macrophage and Dendritic cells and no significant correlations with Neutrophils.