Abstract. The present study aimed to investigate the clinical significance and prospective molecular mechanism of cystatin (CST) genes in patients with hepatitis B virus (HBV)-related hepatocellular carcinoma (HCC). The role of CST genes in the molecular mechanism of HCC was revealed through bioinformatics analysis. The clinical significance of CST genes was investigated using GSE14520-derived data from patients with HBV-related HCC. Gene set enrichment analysis (GSEA) was used to identify pathways in which the CST genes were enriched, as well as the association between these pathways and HCC. The expression levels of CST1, CST2, CST5, CSTA and CSTB genes were higher in HCC tissue compared with in normal tissue; conversely, CST3 and CST7 were reduced in HCC tissue. Subsequent receiver operating characteristic analysis of the CST genes demonstrated that CST7 and CSTB genes may function as potential diagnostic markers for HCC. Furthermore, the expression levels of CST6 and CST7 were strongly associated with recurrence-free survival and overall survival of patients with HBV-related HCC. GSEA of the CST genes revealed that CST7 was significantly enriched in tumor evasion and tolerogenicity, cancer progenitors, liver cancer late recurrence, liver cancer progression and several liver cancer subclasses. In addition, CST genes demonstrated homology in terms of protein structure and were revealed to be strongly co-expressed. The present findings suggested that CST7 and CSTB genes may serve as potential prognostic and diagnostic biomarkers for HCC.

Introduction

Hepatocellular carcinoma (HCC) was reported to be the sixth most common cancer and the fourth most common cause of malignancy-associated mortality worldwide in 2018. Each year, ~841,000 new cases of HCC are diagnosed and 782,000 deaths occur due to HCC worldwide (1). Notably, ~50% of newly diagnosed HCC cases and HCC-related deaths are thought to occur in China, with ~466,100 newly diagnosed patients and ~422,100 deaths occurring in China in 2015 (2). Primary liver cancer includes several pathological types, of which HCC is the predominant form that accounts for 75-85% of all cases, with an incidence of 6.20 cases per 100,000 (1,3). Compared with other types of cancer, HCC in China carried the worst prognosis between 2003 and 2005, with a 1-year survival rate of <50% and a 5-year-survival rate of only 10.1% (4,5). Hepatitis B virus (HBV) infection is the primary cause of the high incidence of HCC in China (6). Given the poor prognosis of this disease, early HCC detection and treatment is of the utmost importance (7). Recent advances in genetic research have promoted a comprehensive understanding of the role of genetic mutations in HCC, allowing for the identification of diagnostic and prognostic HCC biomarkers (8).

Cysteine proteases are critical in promoting the progression of various types of tumor (9). There are eight subfamilies in the cystatin (CST) family group (Family 1, Family 2, Family 3, HRG, Fetuins, CRP, Spp24 and CRES) (10). Cysteine proteases are inhibited by CSTs (11) and are concentrated in the leading edge of tumor cells, where they dissolve extracellular matrix (ECM) proteins to promote invasion (12,13), thus enhancing tumor progression. Several types of CST have been discovered to possess significantly distinct expression profiles in HCC compared with their expression in healthy tissues. CST3, CSTA
and CSTB are significantly greatly expressed in HCC tissue
compared with adjacent healthy tissue, and the expression
levels of CSTA and CSTB are strongly associated with node
metastasis for HCC (14,15). Additionally, it has been reported
that CST3 and CSTB may function as serum markers for
HCC (15,16). Therefore, further investigations into the role
of CST genes in HCC are warranted. The present study aimed
to uncover the prognostic and diagnostic values of Family 1 CSTs
(CSTA and CSTB) and Family 2 CSTs (CST1, CST2, CST3,
CST4, CST5, CST6, CST7 and CST8) in patients with HCC using
freely available data derived from public genomic databases.

Materials and methods

Bioinformatics analysis of CST genes. Database for
Annotation, Visualization and Integrated Discovery (DAVID,
version 6.8; david.ncifcrf.gov/home.jsp) (17,18) was accessed
on December 17, 2018 for Kyoto Encyclopedia of Genes and
Genomes (KEGG) pathway annotation, Gene Ontology (GO)
fuctional annotation and enrichment analysis of CST genes.
An enrichment P<0.05 was considered to indicate a statistically
significant difference. Gene-gene interactions of CST genes
were constructed using GeneMANIA (www.genemania.org,
accessed December 17, 2018) (19,20), whereas protein-protein
interactions of CST genes were constructed using the Search
Tool for the Retrieval of Interacting Genes/Proteins (STRING;
string-db.org, accessed December 17, 2018) (21,22).

gov/geo/query/acc.cgi?acc=GSE14520, accessed December 17,
2018), which comprises clinical data of patients with
HBV-related HCC as well as their CST gene expression
profiles, was extracted from the Gene Expression Omnibus
database (23-25). Due to multiple probe sets in GSE14520,
the expression value of each gene was regarded as the average
value corresponding to the same gene and was normalized using
the limma package of the R platform (version 3.5.1.;
www.r-project.org).

Analysis of gene association and assessment of diagnostic
value. Correlations between the CST genes were analyzed
using Pearson’s correlation coefficient and were depicted using
the corrplot function of the R platform (version 3.5.1.;
www.r-project.org); P<0.05 as considered to indicate a statistically
significant difference. Differential expression of the CST
genes between healthy liver tissues and HCC tumor tissues
were statistically analyzed using Student’s t-test in SPSS
software (version 22.0; IBM Corp.); P<0.05 as considered to
indicate a statistically significant difference. Receiver oper-
ating characteristic (ROC) curve analysis was used to assess the
diagnostic value of CST genes in predicting HCC (26,27).

Survival analysis. Based on the median value of gene expression,
patients were grouped into either the low or high gene expres-
sion group. Each CST gene was analyzed for survival using
Kaplan-Meier analysis with log-rank test, and a Cox propor-
tional hazards regression model was conducted to analyze the associa-
tion of CST genes with clinical parameters that were strongly
associated with OS (P<0.05). The CST genes associated with
survival of patients with HCC (adjusted P<0.05) were analyzed
in combination to explore their joint effects on survival analysis
using Kaplan-Meier analysis and log-rank test, and Cox propor-
tional hazards regression model. Nomograms based on biological
and clinical variables were used to construct a statistical prog-
nostic model of overall survival (OS) for HCC in accordance
with survival analysis results and the Cox proportional hazards
regression model (28). Data processing and plot generation were
conducted in R platform (version 3.5.1.; www.r-project.org) with
rms package. A scale that was marked on both ends of the line
corresponding to each variable represented the value range of
the variable, and the length of the line segment reflected the
contribution of this factor to the outcome event.

Gene set enrichment analysis (GSEA). The biological pathways
targeted by CST genes were further explored with GSEA
(accessed December 17, 2018) (29) using data derived from the
Molecular Signatures Database of c2 (c2.all.v6.1 symbols) and
c5 (c5.all.v6.1 symbols) (30). GSEA-derived gene enrichment
sets that attained a false discovery rate (FDR) of <0.25 and
P<0.05 were determined to confer statistical significance.

Statistical analysis. Statistical data processing was conducted
using SPSS (version 22.0; IBM Corp.) and R (version 3.5.1.;
www.r-project.org). The relative risk of patients with HCC
based on CST gene expression was expressed in terms of 95% confidence intervals (CIs) and hazard ratios (HRs). Univariate
survival analysis of the CST genes and clinical parameters was
performed using Kaplan-Meier analysis with log-rank test.
CST genes and patient clinical parameters that were strongly
associated with OS (P<0.05) were further subjected to a multi-
variate Cox proportional hazards regression model. Pearson’s
correlation coefficient was used to assess the relationship
between co-expressed CST genes. P<0.05 was considered to
indicate a statistically significant difference. FDR control of
GSEA was achieved using the Benjamini-Hochberg procedure
and adjusted for multiple testing (31-33).

Results

Bioinformatics analysis of CST genes. Biological functions
(biological processes, cellular components and molecular
functions) of CST1, CST2, CST3, CST4, CST5, CST6, CST7,
CST8, CSTA and CSTB were subjected to a GO analysis using
DAVID. Each of these genes was markedly enriched in ‘extra-
cellular space’, ‘cysteine-type endopeptidase inhibitor activity’
and ‘protease binding’ (Fig. 1). KEGG pathway analysis using
DAVID suggested that CST1, CST2, CST3, CST4 and CST5
were involved in ‘salivary secretion’ (Fig. 1). CST1, CST2,
CST4, CSTA and CSTB genes and proteins had significant
co-expression relationships (Fig. 2B) and strong protein
homology (Fig. 2A) with each other.

Data source. The present study derived its data only from the
Affymetrix HT Human Genome U133A Array of GSE14520
in order to avoid a batch effect. The majority of subjects in
this cohort had HBV-related HCC, whereas the remainder of
non-HBV-related cases and those with no survival data were
eliminated. This resulted in data from 204 adjacent healthy
liver tissues and 212 HBV-related HCC tumor tissues. Data
regarding patient prognosis were available for all patients.
Analysis of gene association and assessment of diagnostic value. CST gene co-expression in HCC neoplastic tissues was analyzed using the Pearson correlation coefficient. CST1, CST2 and CST4 were closely associated with each other in HCC neoplastic tissues. The expression levels of CST genes were compared between HCC and normal tissues in the GSE14520 dataset. The interaction and correlation analysis of CST genes were performed using STRING and GeneMANIA.

Figure 1. KEGG pathway and GO term analysis of cystatin genes. BP, biological process; CC, cellular component; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function.

Figure 2. Interaction and correlation analysis of CST genes, and the expression level of CST genes between HCC and normal tissues in the GSE14520 dataset. (A) STRING protein-protein association networks of the CST genes. (B) GeneMANIA gene-gene interaction networks of the CST genes. (C) Expression levels of CST genes between HCC and normal tissues in the GSE14520 dataset. (D) Matrix graphs of Pearson correlation analysis of the CST genes in the GSE14520 dataset. *P<0.05. CST, cystatin; HCC, hepatocellular carcinoma.
GSE14520 (Fig. 2D). Furthermore, CST1, CST2, CST5, CSTA and CSTB expression levels were markedly increased in HCC tumor tissue in the GSE14520 dataset, whereas CST3 and CST7 expression levels were markedly decreased in HCC tumor tissue (Fig. 2C). There was no significant difference in the expression of CST4, CST6 and CST8 between HCC tumor tissues and healthy liver tissues.

ROC analysis of CST genes revealed that the expression levels of CST7 and CSTB had significant diagnostic values in differentiating between healthy and malignant hepatic tissues. The area under the ROC curves of CST7 and CSTB were 0.702 (95% CI: 0.651‑0.753; Fig. 3G) and 0.919 (95% CI: 0.891‑0.948; Fig. 3J), respectively. The other CST genes did not exhibit significant diagnostic values.

Survival analysis. In GSE14520, patients with an advanced Barcelona Clinic Liver Cancer (BCLC) stage (34), larger tumor volume (diameter, >5 cm), higher serum α-fetoprotein (AFP; >300 ng/ml) and cirrhosis were at high risk of death due to HBV-related HCC (Table SI). Cirrhotic patients, males and those with advanced BCLC stages were also more at risk of recurrence of HBV-related HCC (Table SI). No other clinical parameters were revealed to impact recurrence-free survival (RFS) or OS.

The results of CST gene survival analysis indicated that the expression levels of CST6 and CST7 may be significantly associated with the recurrence and mortality of patients with HBV-related HCC. The combined impact of CST6 and CST7 on OS and RFS of patients with HBV-related HCC was then further analyzed. Patients were divided into four groups according to CST6 and CST7 expression: Group A, high CST6 and low CST7 expression; Group B, low CST6 and low CST7 expression; Group C, high CST6 and high CST7 expression; Group D, low CST6 and high CST7 expression. Patients who had low CST6 expression and high CST7 expression had a decreased risk of recurrence (adjusted P=0.003; adjusted HR=0.431; 95% CI: 0.264‑0.754; Table II; Fig. 6A) and mortality (adjusted P=0.001; adjusted HR=0.315; 95% CI: 0.115‑0.641; Table III; Fig. 6B) in HBV-related HCC. In addition, the nomogram indicated that both CST6 and CST7 may make a contribution to the prognosis of HCC (Fig. 6C).

GSEA. A CST genome-wide RNA sequencing dataset was used for GSEA in order to uncover the potential biological mechanisms of CST6 and CST7 in HCC. The genome expression profile in GSE14520 was categorized based on CST6
and CST7 median gene expression values. GSEA results of the c2 reference gene set are shown in Table SII, in which increased CST7 expression was associated with tumor evasion and tolerogenicity, cancer progenitors, liver cancer late recurrence, liver cancer progression and several liver cancer subclasses (Fig. 7A-I). The enrichment results of c5 are shown in Table SIII; high CST7 expression was revealed to also be involved in positive regulation of the tumor necrosis factor subfamily cytokine production, positive regulation of NF-κB transcription factor and positive regulation of G1-S transition of mitotic cell cycle (Fig. 7J-L). Conversely, the GSEA results of CST6 did not exhibit a significant association between CST6 and biological pathways relevant to HCC.

Discussion

It has been reported that cysteine proteases are involved in the progression of several types of tumor (9). Destruction and remodeling of the ECM is an essential process in tumor progression (35), which can be promoted by cysteine proteases (36), particularly cathepsin B, a representative cysteine protease that serves a key role in tumor cell invasion (37-39). However, biological functions of cysteine proteases are inhibited by CSTs (11). Therefore, CSTs may also be associated with tumor progression. Notably, CSTs have been reported to be associated with the progression of various types of cancer, including bladder cancer (40), breast cancer (41,42), esophageal cancer (43), ovarian cancer (44) and prostate cancer (45). The results of GSEA in the present study indicated that CST7 was significantly enriched in liver cancer progression. Co-expression analysis of the CST genes in GSE14520 using Pearson’s correlation coefficient revealed that CST1, CST2 and CST4 are closely associated with each other, verifying the results of GeneMANIA and STRING.

CST7 expression was markedly decreased in HCC tissue samples, whereas CSTB was highly expressed in HCC tumor tissue. ROC analysis indicated that CST7 and CSTB exhibit significant diagnostic values and may serve as potential
diagnostic biomarkers. The diagnostic value of CST genes has been reported in previous studies. Having higher expression in HCC compared with in adjacent healthy tissues, the diagnostic value of CSTB for HCC was reported in previous research, and the present investigation provided validation for this (14,16,46). In addition to HCC, the diagnostic value of CSTB has been reported in other tumor types. For example, compared with in normal bladder tissue, CSTB immunohistochemical staining is more intense in bladder cancer tissue (40). CSTB has also been demonstrated to be a diagnostic biomarker of ovarian clear cell carcinoma, due to its high expression in tumor cells based on the results of immunohistochemical analysis, reverse transcription-PCR and western blot analysis (47).

The survival analysis of CST genes indicated that CST6 and CST7 may be strongly associated with the OS and RFS of patients with HBV-related HCC. The combined effects survival analysis indicated that the risks of mortality and recurrence in patients with HBV-related HCC were lowest in those with increased expression of CST7 and attenuated expression of CST6. Therefore, CST7 and CST6 may function as prognostic biomarkers for HCC. The prognostic value of CST genes has been reported across several malignancies. It has been reported that overexpression of CST6 promotes pancreatic cancer growth (48). The function of CST6 was revealed to be similar in the present study, where the survival analysis results indicated that high CST6 expression was associated with poor survival of patients with HCC. However, CST6 has also been reported to act as a human tumor suppressor gene in previous reports (45,49-54). Therefore, these findings indicated that CST6 may exert distinct effects in different types of cancer. The similar role of CST6 in the liver and pancreas may be a result of the liver and pancreas stemming from a common progenitor at the embryo stage (55). The fact that the molecular mechanism of CST6 serves different roles in various types of cancer still requires further exploration. Although no significant association was detected between CSTB and the prognosis of patients with HCC, differences have been reported in the expression of CSTB between tumor tissues and adjacent healthy tissues (16,46,56). In addition, the prognostic value of CSTB has been demonstrated in several

Figure 5. Kaplan-Meier survival curve analysis of overall survival for CST genes in hepatitis B virus-related hepatocellular carcinoma in the GSE14520 dataset. OS curves for (A) CST1, (B) CST2, (C) CST3, (D) CST4, (E) CST5, (F) CST6, (G) CST7, (H) CST8, (I) CSTA and (J) CSTB. CST, cystatin.
Table I. Prognostic values of CST gene expression in patients with hepatitis B virus-related hepatocellular carcinoma from the GSE14520 dataset.

<table>
<thead>
<tr>
<th>Gene expression</th>
<th>Patients (n=212)</th>
<th>No. of events</th>
<th>MRT (months)</th>
<th>Crude HR (95% CI)</th>
<th>Crude P-value</th>
<th>Adjusted HR (95% CI)</th>
<th>Adjusted P-value</th>
<th>RFS</th>
<th>OS</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>57</td>
<td>46</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>59</td>
<td>40</td>
<td>0.964 (0.670-1.388)</td>
<td>0.846</td>
<td>0.900 (0.624-1.299)</td>
<td>0.575</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>62</td>
<td>35</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>54</td>
<td>53</td>
<td>0.811 (0.563-1.169)</td>
<td>0.261</td>
<td>0.739 (0.511-1.068)</td>
<td>0.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>57</td>
<td>43</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>59</td>
<td>46</td>
<td>1.043 (0.725-1.501)</td>
<td>0.821</td>
<td>0.940 (0.647-1.365)</td>
<td>0.746</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>52</td>
<td>51</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>64</td>
<td>54</td>
<td>1.278 (0.886-1.843)</td>
<td>0.19</td>
<td>1.186 (0.821-1.713)</td>
<td>0.365</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>57</td>
<td>46</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>59</td>
<td>45</td>
<td>1.041 (0.723-1.498)</td>
<td>0.83</td>
<td>1.066 (0.738-1.539)</td>
<td>0.735</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>55</td>
<td>57</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>66</td>
<td>28</td>
<td>1.499 (1.037-2.166)</td>
<td>0.031</td>
<td>1.651 (1.136-2.398)</td>
<td>0.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>63</td>
<td>27</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>53</td>
<td>53</td>
<td>0.695 (0.482-1.003)</td>
<td>0.052</td>
<td>0.688 (0.475-0.966)</td>
<td>0.048</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CST8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>56</td>
<td>51</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>60</td>
<td>30</td>
<td>1.139 (0.791-1.641)</td>
<td>0.483</td>
<td>1.106 (0.764-1.602)</td>
<td>0.592</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>59</td>
<td>46</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>57</td>
<td>40</td>
<td>1.026 (0.713-1.477)</td>
<td>0.889</td>
<td>1.173 (0.811-1.696)</td>
<td>0.397</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSTB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>106</td>
<td>54</td>
<td>51</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>106</td>
<td>62</td>
<td>36</td>
<td>1.311 (0.910-1.889)</td>
<td>0.147</td>
<td>1.164 (0.803-1.698)</td>
<td>0.423</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RFS: recurrence-free survival; OS: overall survival; HR: hazard ratio; MRT: median recurrence time; MST: median survival time; CI: confidence interval; CST: cystatin; BCLC: Barcelona Clinic Liver Cancer; NA: not available.
other tumors. For example, high CSTB expression is associated with a more favorable prognosis in lung cancer (57). A similar scenario of CSTB functioning as a prognostic biomarker has been reported in gastric cancer, where it restrains tumor

Table II. Joint effects analysis of CST6 and CST7 expression in the recurrence-free survival of patients with hepatocellular carcinoma.

<table>
<thead>
<tr>
<th>Group</th>
<th>CST6</th>
<th>CST7</th>
<th>Patients</th>
<th>No. of events</th>
<th>MRT (months)</th>
<th>Crude HR (95% CI)</th>
<th>Crude P-value</th>
<th>Adjusted HR (95% CI)</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>High</td>
<td>Low</td>
<td>53</td>
<td>34</td>
<td>22</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>Low</td>
<td>Low</td>
<td>53</td>
<td>29</td>
<td>42</td>
<td>0.788 (0.480-1.294)</td>
<td>0.48</td>
<td>0.753 (0.458-1.237)</td>
<td>0.263</td>
</tr>
<tr>
<td>C</td>
<td>High</td>
<td>High</td>
<td>53</td>
<td>32</td>
<td>36</td>
<td>0.815 (0.503-1.322)</td>
<td>0.408</td>
<td>0.880 (0.541-1.4330)</td>
<td>0.608</td>
</tr>
<tr>
<td>D</td>
<td>Low</td>
<td>High</td>
<td>53</td>
<td>21</td>
<td>NA</td>
<td>0.452 (0.262-0.779)</td>
<td>0.004</td>
<td>0.431 (0.264-0.754)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

*Adjusted for sex, cirrhosis and Barcelona Clinic Liver Cancer stage and α-fetoprotein in the GSE14520 cohort using multivariate Cox proportional hazards regression model. CI, confidence interval; CST, cystatin; HR, hazard ratio; MRT, median recurrence time; NA, not available.

Table III. Joint effects analysis of CST6 and CST7 expression in the overall survival of patients with hepatocellular carcinoma.

<table>
<thead>
<tr>
<th>Group</th>
<th>CST6</th>
<th>CST7</th>
<th>Patients</th>
<th>No. of events</th>
<th>MST (months)</th>
<th>Crude HR (95% CI)</th>
<th>Crude P-value</th>
<th>Adjusted HR (95% CI)</th>
<th>Adjusted P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>High</td>
<td>Low</td>
<td>53</td>
<td>26</td>
<td>61</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>Low</td>
<td>Low</td>
<td>53</td>
<td>23</td>
<td>NA</td>
<td>0.849 (0.484-1.388)</td>
<td>0.567</td>
<td>0.785 (0.447-1.379)</td>
<td>0.399</td>
</tr>
<tr>
<td>C</td>
<td>High</td>
<td>High</td>
<td>53</td>
<td>21</td>
<td>NA</td>
<td>0.723 (0.407-1.286)</td>
<td>0.27</td>
<td>0.721 (0.396-1.310)</td>
<td>0.283</td>
</tr>
<tr>
<td>D</td>
<td>Low</td>
<td>High</td>
<td>53</td>
<td>12</td>
<td>NA</td>
<td>0.367 (0.260-0.779)</td>
<td>0.004</td>
<td>0.315 (0.115-0.641)</td>
<td>0.001</td>
</tr>
</tbody>
</table>

*Adjusted for sex, cirrhosis, Barcelona Clinic Liver Cancer stage and α-fetoprotein in the GSE14520 cohort using multivariate Cox proportional hazards regression model. CI, confidence interval; CST, cystatin; HR, hazard ratio; MST, median survival time; NA, not available.

Figure 6. Combined effect of CST6 and CST7 on the overall survival and recurrence-free survival of patients, and nomogram for predicting 1-, 2- and 3-year events. (A) Recurrence-free survival curves for the combined effect of CST6 and CST7; (B) overall survival curves for the combined effect of CST6 and CST7. Group A, high CST6 and low CST7 expression; Group B, low CST6 and low CST7 expression; Group C, high CST6 and high CST7 expression; Group D, low CST6 and high CST7 expression. (C) Nomogram for predicting 1-, 2- and 3-year events (death) that combine clinical data with CST6 and CST7 expression. AFP, α-fetoprotein; BCLC, Barcelona Center Liver Cancer; CST, cystatin.
Figure 7. GSEA results of CST7 in GSE14520. (A-H) GSEA results of c2-reference gene sets for groups with elevated CST7 expression.
development by suppressing proliferation and migration of neoplastic cells (58). The present study did not determine a prognostic value of CST3 in HCC; however, CST3 has been demonstrated to act as a tumor suppressor that restrains tumor cell invasion in previous studies (56,59). A recent study reported that the rate of glomerular filtration of creatinine and CST3 may serve as potential predictors of OS in HCC (60). By reviewing these studies, the different roles of CST genes in numerous types of cancer can be identified. The present results corresponded with the results of previous studies. Although in the same subfamily, the expression levels and biological functions of each CST gene are not the same as those of others, even in different types of cancer. CST genes may also function as oncogenes; however, further studies are needed to validate the present findings.

The GSEA conducted in the present study revealed that CST7 was enriched in tumor evasion and tolerogenicity, cancer progenitors, liver cancer late recurrence, liver cancer progression, several liver cancer subclasses, tumor necrosis factor subfamily cytokine production, regulation of NF-κB transcription factor and regulation of G1-S transition of the mitotic cell cycle. These results suggested that CST7 may be closely associated with liver cancer. However, the association between CST genes and HCC requires further validation in future studies. In addition, although GSEA of CST6 indicated that CST6 was not involved in any pathway or molecular mechanism associated with cancer, the effects of CST6 on different types of cancer have been confirmed by previous studies (45,49-54).

One limitation of the present study was that the sample size was insufficient, which could affect the validity of the results. Secondly, the clinical data obtained from the GSE14520 dataset are not complete, barring the opportunity to carry out a more comprehensive survival analysis using multivariate Cox proportional hazards regression model. In order to better evaluate the association between CST subfamily members and HCC prognosis, likely HCC risk factors, including the presence of a tumor capsule and vascular invasion, Child-Pugh score and alcohol intake, should be taken into consideration. Thirdly, the current investigation only explored the relationship between CST gene mRNA expression and HCC prognosis and did not explore the effects of CST protein levels on HCC.
prognosis. Finally, further studies are warranted to determine the effects of Family 3 and Family 4 CSTs on HCC.

Although there are several limitations to the present study, to the best of our knowledge, this is the first to discover the clinical significance of CST6 and CST7 in the prognosis of patients with HBV-related HCC. In addition, our result verified the findings of previous reports and suggested that CSTB may act as a diagnostic biomarker for HCC. Furthermore, CST7 was discovered to be enriched in several tumor-related signaling pathways and biological processes, including tumor evasion and tolerogenicity, cancer progenitors, liver cancer late recurrence, liver cancer progression, several liver cancer subclasses, tumor necrosis factor subfamily cytokine production, regulation of NF-κB transcription factor and regulation of G1-S transition of the mitotic cell cycle. The prospective molecular mechanisms underlying the effects of CST7 gene expression on patients with HBV-related HCC were determined using GSEA.

In conclusion, the gene expression levels of CST1, CST2, CST5, CSTA and CSTB were significantly increased in HCC tissue, whereas CST3 and CST7 were overexpressed in normal tissue compared with in HCC tissue. Notably, the present study revealed that CST7 and CSTB may serve as diagnostic markers for HCC, and survival analysis of CST genes indicated that CST6 and CST7 expression levels may be closely associated with the OS and RFS of patients with HCC. However, this investigation requires further validation using a sufficient sample size spread across multiple geographical regions.

Acknowledgements

Not applicable.

Funding

This study was supported in part by the National Natural Science Foundation of China (grant nos. 81560535, 81802874, 81072321, 30760243, 30460143 and 30560133), the Natural Science Foundation of Guangxi Province of China (grant no. 2017JBJ140189), the 2009 Program for New Century Excellent Talents in University (NCET), Guangxi Natural Sciences Foundation (grant no. GuiKeGong 1104003A-7) and the Guangxi Health Ministry Medicine Grant (Key Scientific Research Grant; grant no. Z201018). The present study was also partly supported by the Scientific Research Fund of the Health and Family Planning Commission of Guangxi Zhuang Autonomous Region (grant no. Z2016318), the Key Laboratory of High Incidence Tumor Prevention & Treatment (Guangxi Medical University), Ministry of Education (grant no. GKE2018-01), the Basic Ability Improvement Project for Middle-aged and Young Teachers in Colleges and Universities in Guangxi (grant no. 2018KY0110), the Innovation Project of Guangxi Graduate Education (grant no. JGY2018037) the 2018 Innovation Project of Guangxi Graduate Education (grant no. YCBZ2018036), and the Research Institute of Innovative Think-tank at Guangxi Medical University (the gene-environment interaction in hepatocarcinogenesis in Guangxi HCCs and its translational applications in the HCC prevention). The authors would also like to acknowledge the support provided by the National Key Clinical Specialty Programs (General Surgery & Oncology) and the Key Laboratory of Early Prevention & Treatment for Regional High-Incidence-Tumor (Guangxi Medical University), Ministry of Education, China. This work was also supported in part by the Natural Science Foundation of Guangxi Province of China (grant no. 2018GXNSFBA138013), the Key Laboratory of High-Incidence-Tumor Prevention & Treatment (Guangxi Medical University) and the Guangxi Key R & D Program (grant no. GKEAB1822109).

Availability of data and materials

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

Authors' contributions

XZ and TP conceived and designed the study. XW, KH, XL, CY, TY, JL, CH, GZ, HS, WQ, QH, ZL, JH, YG, XY and TP acquired the data and performed data analyses. XZ wrote the manuscript, and TP guided and supervised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patent consent for publication

Not applicable.

Competing interests

The authors declared that they have no competing interests.

References


Zhou et al: CLINICAL SIGNIFICANCE AND MOLECULAR MECHANISMS OF CST GENES IN HBV-HBcR- HCC


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.