Targeted Proteins Reveal Cathepsin D as a Novel Biomarker in Differentiating Hepatocellular Carcinoma from Cirrhosis and Other Liver Cancers

Natthaya Chuaypen¹, Supachaya Sriphoosanaphan², Athaya Vorasittha³, Nutcha Pinjaroen⁴, Visith Thongboonkerd⁵,⁶, Pisit Tangkijvanich¹, Pongserath Sirichindakul¹³*  

Abstract

**Objective:** Hepatocellular carcinoma (HCC) represents a global health concern, particularly in Southeast Asia where hepatitis B virus (HBV) infection is common. In this study, we applied tissue-based proteomics to identify novel serological proteins for HCC and validated their performance in serum specimens. **Methods:** In a discovery set, liver tissue specimens of HBV-related HCC, intrahepatic cholangiocarcinoma (iCCA) and colorectal cancer with liver metastasis (CRLM) were analyzed using mass spectrometry (LTQ-Orbitrap-XL). A subset of proteins that showed highly expressed in HCC were then confirmed by Western blotting. Additionally, clinical significance of selected candidate proteins was tested in serum samples of 80 patients with HBV-related HCC, 50 patients with HBV-related liver cirrhosis and 30 healthy controls. **Results:** Based on LTQ-Orbitrap-XL mass spectrometer, various differentially expressed proteins (DEPs) between tumor and adjacent non-tumor tissues were identified. These included 77 DEPs for HCC, 77 DEPs for iCCA and 55 DEPs for CRLM. Among selected candidate proteins, annexin A2 and cathepsin D were confirmed to be overexpressed in HCC tissue by Western blot analysis. In a validate cohort, serum cathepsin D level, but not annexin A2, was significantly higher in HCC compared with the non-HCC groups. Serum cathepsin D level was also positively correlated with tumor size and tumor stage. Additionally, the combined assay of serum cathepsin D and alpha-fetoprotein had a high sensitivity in detecting early HCC (83%) and intermediate/advanced HCC (96%). Moreover, patients with low serum cathepsin D (<305 ng/mL) displayed significantly better overall survival than those whose serum levels were high (≥305 ng/mL). **Conclusions:** Proteomics and subsequent validation revealed cathepsin D as a novel biomarker for HCC. Apart from its diagnostic role, serum cathepsin D might also serve as a prognostic biomarker of HCC. Additional large-scale studies are needed to verify our findings.

**Keywords:** Cathepsin D- proteomics- biomarker- HCC- iCCA- CRLM

Introduction

Primary liver cancer (PLC) represents one of the most frequent and aggressive malignancies in global scale with its relatively low survival rates resulting from inadequate early diagnosis (Sia et al., 2017). Histologically, PLC is mainly comprised of two types of cancers namely hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (iCCA) (Sia et al., 2017). HCC, originating from the hepatocytes, has a high incidence found in Southeast Asia where hepatitis B virus (HBV) infection is common (Forner et al., 2018). The estimated incidence of HCC in Thailand is 38.6 and 17.2 per 100,000 person-years in males and females, respectively. For iCCA, this adenocarcinoma arises in the intrahepatic bile ducts and remains the second most common PLC after HCC (Sia et al., 2017). Despite its relative scarcity in global scale, iCCA is considerably prevalent in Southeast Asia, especially in the areas of Mekong River Basin countries with widespread liver fluke infestation (Mathema et al., 2015). For instance, Khon Kaen province in Thailand, is considered to be among...
the highest overall age-standardized incidence of iCCA worldwide. Thus, in the geographical areas where chronic viral hepatitis and liver flukes are common, it is essential to differentiate HCC from iCCA as both cancers might present with similar clinical features. Besides PLC, the liver is a common site for tumors spreading from other organs, accounting for 25% of all cases. Among these, colorectal cancer with liver metastasis (CRLM) is the most common metastatic tumors undergoing hepatic resection (Bray et al., 2018). As their prognosis and management differ significantly, distinguishing HCC from iCCA and CRLM remains challenging in clinical practice.

Currently, serum alpha-fetoprotein (AFP), a fetal-specific glycoprotein, is the most commonly used biomarker for HCC (Sauzay et al., 2016). Since approximately 90% of HCC occur within underlying liver cirrhosis (LC), it is recommended that cancer surveillance should be performed in high-risk groups using ultrasound with or without AFP (Heimbach et al., 2018). Nonetheless, AFP has some limitations including lack adequate sensitivity in an early-stage HCC and its concentration could also be elevated in non-malignant liver diseases. Moreover, high serum AFP levels have been reported in some cases with iCCA (Sauzay et al., 2016). Thus, identifying reliable biomarkers participating in molecular pathogenesis of HCC will play a key role in early diagnosis and predicting disease prognosis. Recent evidence in molecular biology have demonstrated that repetitive destruction and regeneration of the hepatocytes are accountable for HCC occurrence (Arzumanyan et al., 2013). Among essential biological mechanisms, aberrant protein expressions are shown to be involved in the multi-step processes of hepatic oncogenesis (Xing et al., 2016). These data indicate the potential use of high-throughput quantitative proteomic techniques to identify differentially expressed proteins as novel biomarkers for diagnosis and monitoring of liver cancers (Beretta 2009; Xing et al., 2016). Currently, less evidence is focusing on new diagnostic biomarkers to distinguish HCC from iCCA and CRLM.

In this study, we aimed at assessing proteomic profiling of HCC in differentiating with the other types of liver cancer. Furthermore, a subset of proteins that showed highly expressed in HCC were confirmed by Western blotting in liver specimens. Additionally, we validated clinical significance of candidate proteins in another set of proteins that showed highly expressed in HCC were confirmed by Western blotting in liver specimens. Additionally, we validated clinical significance of candidate proteins in another set of samples obtained from patients with PLC, LC and healthy controls.

Materials and Methods

Patients' tissues and blood collection

To investigate the protein profiling, pair tissues of tumor and adjacent non-tumor from patients with HBV-related HCC, iCCA and CRLM who had undergone surgical resection at the Department of Surgery, Chulalongkorn University from 2017 to 2020 were collected. The diagnosis of tumor type was confirmed by histopathological examination. Liver tissue specimens obtained from surgical resection were immediately frozen in liquid nitrogen until analysis.

For a validate cohort in serum samples of patients with HBV-related HCC, the diagnostic of this cancer was established by imaging studies based on dynamic computed tomography (CT) or magnetic resonance imaging (MRI) according to the American Association for the Study of Liver Diseases (AASLD) guideline (Heimbach et al., 2018). Baseline clinical and laboratory parameters of these patients were collected. The overall survival (OS) of these patients characterized by interval between initial recruitment and death or the last follow-up visit was also recorded. For the control groups, serum samples were taken from healthy controls and patients with HBV-related LC without evidence of HCC.

The diagnosis of LC was based on clinical and imaging evidence and confirmed by transient elastography (TE) (F4 >12.5 kPa)(Singh et al., 2017). The diagnosis of chronic HBV infection was documented by the positivity of serum hepatitis B s antigen (HBsAg) for at least 6 months. Patients who were infected with hepatitis C virus (HCV) and/or human immunodeficiency virus (HIV) were excluded. All of the collected serum samples were stored at -70°C until assay.

The study was performed according to the principles of the Declaration of Helsinki and approved by the Institute Ethics Committee of Faculty of Medicine, Chulalongkorn University. All participants provided written informed consent prior to the enrollment.

Protein extraction

Tissue samples from both tumor and adjacent non-tumor of each patient were extracted for proteins. Briefly, the tissue was dissected into small pieces and snap-freeze by immersing in liquid nitrogen then add ice-cold lysis buffer and homogenize using electric homogenizer. After these processes, the tissues were agitated for 2 hours at 4°C. The supernatant was collected following centrifugation at 16,000×g and 4°C for 20 minutes and total protein concentration was determined by Bio-Rad Protein Assay (Bio-Rad Laboratories; Hercules, CA).

In-solution tryptic digestion by filter-aided sample preparation (FASP) method and nanoLC-ESI-LTQ-Orbitrap MS/MS.

Equal amount of total protein derived from each sample was pooled and digested by trypsin according to FASP protocol as described previously (Thanomkitti et al., 2018). Finally, the peptides were resuspended in 0.1% formic acid before running the tandem mass spectrometry (MS/MS).

Each sample was run in triplicates. Digested peptides were separated by EASY-nLC II (Thermo Scientific; Waltham, MA). Briefly, cooled peptides (7°C) were loaded into an in-house 3-cm-long pre-column containing 5-µm C18 resin (Dr.Maisch GmbH; Ammerbuch, Germany) and then to 10-cm-long analytical column packed with by mobile phase B (ACN/0.1% formic acid). Peptide sequences were then examined by LTQ-Orbitrap-XL (Thermo Scientific) in positive mode with ESI nano-sprayer ion source. Data were obtained in a collision-induced dissociation (CID) top-12 mode under the control of the Xcalibur 2.1.0 and LTQ Tune Plus 2.5.5 software (Thermo Scientific) according to the method.
described previously (Thanomkitti et al., 2018).

MS/MS spectral interpretation and quantitative analysis.

Data were analyzed using Proteome Discoverer v.1.4.1.14 software (Thermo Scientific). Data were searched against SwissProt database of mammalian using Mascot software version 2.4.0 (Matrix Science; London, UK) with standard Mascot parameters for CID (Enzyme = trypsin, maximal number of missed cleavages = 1, peptide tolerance = ±2 ppm, MS/MS tolerance = ±0.2 Da, fixed modification = carbamidomethyl (C), variable modification = oxidation (M), charge states = 2+ and 3+, and decoy database on FDR <1%). Quantitative data of three most abundant peptides identified from each protein with average value of areas under curve (AUC) was considered. Background was subtracted from all peak areas (Thanomkitti et al., 2018).

**Western blot analysis**

For Western blot analysis, total protein from liver tissues (adjacent non-tumor and tumor) were extracted and collected in Laemmli lysis buffer. Total protein level was measured by Bio-Rad protein assay (Bio-Rad Laboratories, Inc., USA). Then, protein concentration at 50 ug/lane from each sample was run on 10% SDS-PAGE and transferred to nitrocellulose blotting membranes (GE Healthcare Life science, Germany). Subsequently, the membranes were blocked in 5%BSA solution (0.05% Tween 20) for 1 h at room temperature and then incubated with mouse anti-Cathepsin D (1:1,000, ab6313; Abcam, Cambridge, UK) and mouse anti-Annexin II (1:1,000, sc-28385; Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA) for 16 hours at 4°C. Membranes were washed and incubated with Mouse IgGκ light chain binding protein (m-IgGκ BP) conjugated to horseradish peroxidase (HRP) (1:5,000, sc-516102, Santa Cruz Biotechnology Inc., Santa Cruz, CA, USA) for 2 hours. After that, the membranes were probed with Amersham ECL Prime Western Blotting Detection Reagent (Cytiva, UK). The immunoreactive protein bands were visualized by ChemiDoc MP Imaging System (Bio-Rad Laboratories, Inc., USA). Band intensity data was identified using Image analysis in JAVA (international institute of health, USA). GAPDH was used as the internal control.

**Enzyme-linked immunosorbent assay (ELISA)**

The quantitative of annexin A2 expression in serum samples was performed according to manufacturer’s protocol. Briefly, serum samples were diluted at 1:10 in PBS and added in duplicate into the plate (ELISA Assay Kit for Annexin A2, Product No. SEB944Hu, CLOUD-CLONE CORP, USA) and then incubated for 1 h at 37°C. After that, the Detection Reagent A and Detection Reagent B were added and incubated at 37°C for 30 minutes. After washing, TMB Substrate Solution was added and then the reaction was conducted using a spectrophotometer at 450 nm.

For circulating cathepsin D expression, serum samples were diluted at 1:100 in sample diluent buffer and added in duplicate into the plate (ab119586-cathepsin D human Elisa kit, Abcam, Cambridge, UK) following manufacturer’s instructions. Briefly, 1X Biotinylated anti-Human Cathepsin D antibody was added into each well and incubate at 37°C for 60 minutes. After each incubation, the plates were washed and added 1X Avidin-Biotin-Peroxidase Complex working solution. In the final step, TMB agent was added into each well and the reaction was stopped with TMB Stop Solution. The results were reported as O.D. using a spectrophotometer at 450 nm.

**Global protein network analysis**

Protein molecular and biological functions; Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway were analyzed using PANTHER classification system (Mi et al., 2016). Protein–protein interaction of differential upregulated proteins from patients were analyzed using Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) version 10.0 (http://string.embl.de).

**Statistical analysis**

Data were reported as percentage or mean ± standard deviation, as appropriate. Comparisons between groups were tested by unpaired Student’s-t test or Mann-Whitney U test. The receiver-operating characteristics (ROC) curves were generated to calculate the diagnostic performance of serum biomarkers in discriminating HCC from the non-HCC groups. The Pearson correlation coefficient was applied to analyze the correlation between levels of serum biomarkers and related parameters. The assessment of overall survival of patients with HCC was determined by the Kaplan-Meier method and difference between groups was assessed by the log-rank test. P values less than 0.05 were considered statistically significant. Statistical analyses were undertaken by the SPSS software for windows 22.0 (SPSS Inc., Chicago, IL).

**Results**

**Differential protein expression profiles in liver cancers**

To investigate the proteomic profile in the discovery set, 34 pairs of tumor and adjacent non-tumor tissues of patients with HCC (12 cases), iCCA (12 cases) and CRLM (10 cases) were used to identify the different protein expression profile. Detailed baseline demographic and characteristics of these patients were shown in Supplementary Table 1.

To identify significant differentially expressed proteins (DEPs) in HCC, iCCA and CRLM tissues, nanoLC-ESI-LTQ-Orbitrap tandem mass spectrometry (MS/MS) was performed with high-stringent criteria. We observed the differential expression patterns of proteins among the studied groups. In particular, significantly DEPs between tumor tissues and adjacent non-tumor tissues including 77 DEPs for HCC, 77 DEPs for iCCA and 55 DEPs for CRLM. Their identities, identification scores, related mass spectral parameters and quantitative data of HCC was summarized in Table 1, while the data regarding iCCA and CRLM were shown in Supplementary Table 2 and 3 respectively. From the discovery sets, 34 proteins were upregulated, and 15 proteins were downregulated in HCC. Additionally, 14 were upregulated and 21 were...
Natthaya Chuaypen et al
Asian Pacific Journal of Cancer Prevention, Vol 23
2020

downregulated in iCCA, and 25 were upregulated and 3 were downregulated in CRLM.

Gene ontology, KEGG pathway and protein network analysis

The upregulated proteins in HCC were subsequently analyzed to identify the function of respective DEPs. From Gene Ontology (GO) and KEGG pathway enrichment analysis, upregulated proteins were involved in biological process, cellular component and molecular function (Figure 1). These proteins were commonly located in cellular process (15, 23.8%), metabolic process (10, 15.9%), responsive to stimulus (11, 17.5%), cellular anatomical entity (21, 51.2%), intracellular functions (16, 39.0%), protein binding (11, 47.8%), catalytic activity (8, 34.8%) and structural molecular activity (3, 13.0%) (Figure 1A). In addition, the proteins were also enriched in the following pathways, including integrin signaling pathway, glycolysis, aminobutyrate degradation, methyl malonyl pathway, gamma-aminobutyric acid

Figure 1. Data Mining of the up-Regulated Proteins in HCC Tissues. (A) GO analysis; (B) KEGG pathway analysis; and (C) STRING analysis.

Figure 2. Protein Expression in Liver Tissues of Patients with HCC. (A) Cathepsin D level; (B) Annexin A2 level

synthesis, Alzheimer disease-prenilin pathway and pentose phosphate pathway (Figure 1B). Protein-protein interactions were identified among all the upregulated proteins (Figure 1C). These significantly enriched GO and KEGG pathway terms could facilitate a better understanding the role of the DEPs in the tumorigenesis of HBV-related HCC.

Assessment of candidate tissue proteins by Western blot analysis
To confirm the results of proteomics, the greatly up-regulated proteins including annexin A2, calnexin, cathepsin D, plectin and prohibitin were selected to verify the protein expression levels in tissue samples using Western blot analysis. Compared with adjacent non-tumor tissues, only annexin A2 and cathepsin D were significantly up regulated in HCC tissues (P<0.001 and P=0.023, respectively) as shown in Figure 2A and 2B. Moreover, there was no significant different between tumor and non-tumor tissues in tissue samples of iCCA and CRLM (data not shown).

Validation of protein expression in serum samples by ELISA
We subsequently validated the expression annexin A2 and cathepsin D in serum samples of 80 patients with HCC, 50 patients with LC and 30 healthy controls using commercial ELISA kits. Our results demonstrated that cathepsin D levels were significantly increased in patients with HCC (345.9±192.8 ng/mL) and LC (216.0±170.2 ng/mL) compared with healthy controls (93.3±41.9 ng/mL) (P<0.001). In addition, significant difference of cathepsin D was also observed between patients with HCC and LC (P<0.001; Figure 3A), which was comparable to the results of protein expression in liver tissues detected by Western blot analysis. Regarding serum annexin A2, there was no significant difference among HCC (31.7±21.0 ng/mL), LC (31.4±48.6 ng/mL) and healthy controls (28.5±9.4 ng/mL)(P > 0.05; Figure 3B).

Diagnostic performance of serum biomarkers
We further conducted the diagnostic performance of serum biomarkers in differentiating HCC from LC. As
### Table 1. Summary of Significantly Altered Proteins Levels in HCC Tissue vs. adjacent non-Tumor Tissue as Determined by Quantitative Proteomics

<table>
<thead>
<tr>
<th>Accession no.</th>
<th>MS/MS identification score</th>
<th>%Cov</th>
<th>MW (kDa)</th>
<th>Abundance level (+10^6 arbitrary unit)</th>
<th>(Mean±SEM)</th>
<th>Ratio</th>
<th>P- value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td>Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins whose levels were significantly increased in tissues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P01834</td>
<td></td>
<td>1025.95</td>
<td>80.19</td>
<td>11.6</td>
<td>6.8±1.08</td>
<td>82.6±9.03</td>
<td>12.13</td>
</tr>
<tr>
<td>P08729</td>
<td></td>
<td>201.69</td>
<td>8.96</td>
<td>51.4</td>
<td>3.16±0.33</td>
<td>18.55±1.0</td>
<td>5.87</td>
</tr>
<tr>
<td>P11047</td>
<td></td>
<td>304.94</td>
<td>11.5</td>
<td>177.5</td>
<td>2.85±0.93</td>
<td>5.59±0.40</td>
<td>1.96</td>
</tr>
<tr>
<td>P22307</td>
<td></td>
<td>73.26</td>
<td>18.28</td>
<td>59</td>
<td>3.72±2.88</td>
<td>40.15±5.53</td>
<td>10.8</td>
</tr>
<tr>
<td>P15149</td>
<td></td>
<td>216.53</td>
<td>20.71</td>
<td>531.5</td>
<td>6.08±1.85</td>
<td>18.60±2.87</td>
<td>3.06</td>
</tr>
<tr>
<td>P35232</td>
<td></td>
<td>258.3</td>
<td>30.88</td>
<td>29.8</td>
<td>8.87±1.28</td>
<td>14.70±2.18</td>
<td>1.66</td>
</tr>
<tr>
<td>P05166</td>
<td></td>
<td>157.84</td>
<td>26.53</td>
<td>52.8</td>
<td>5.58±1.52</td>
<td>23.66±1.43</td>
<td>4.24</td>
</tr>
<tr>
<td>P07237</td>
<td></td>
<td>135.02</td>
<td>15.35</td>
<td>57.1</td>
<td>14.49±2.10</td>
<td>20.51±9.03</td>
<td>1.42</td>
</tr>
<tr>
<td>P13667</td>
<td></td>
<td>385.57</td>
<td>19.84</td>
<td>72.9</td>
<td>3.81±0.69</td>
<td>5.76±0.46</td>
<td>1.51</td>
</tr>
<tr>
<td>P02768</td>
<td></td>
<td>1384.23</td>
<td>45.65</td>
<td>69.3</td>
<td>51.86±4.24</td>
<td>151.07±0.93</td>
<td>2.91</td>
</tr>
<tr>
<td>Q9H299</td>
<td></td>
<td>36.55</td>
<td>29.03</td>
<td>10.4</td>
<td>2.11±1.0643</td>
<td>10.97±1.71</td>
<td>5.21</td>
</tr>
<tr>
<td>Q01082</td>
<td></td>
<td>39.67</td>
<td>11.25</td>
<td>274.4</td>
<td>3.04±0.93</td>
<td>18.74±1.03</td>
<td>6.17</td>
</tr>
<tr>
<td>P51649</td>
<td></td>
<td>94.95</td>
<td>17.76</td>
<td>57.2</td>
<td>4.65±0.33</td>
<td>18.74±1.03</td>
<td>6.17</td>
</tr>
</tbody>
</table>

Proteins whose levels were significantly decreased in tumor tissues

<table>
<thead>
<tr>
<th>Accession no.</th>
<th>MS/MS identification score</th>
<th>%Cov</th>
<th>MW (kDa)</th>
<th>Abundance level (+10^6 arbitrary unit)</th>
<th>(Mean±SEM)</th>
<th>Ratio</th>
<th>P- value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
<td>Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins whose levels were significantly decreased in tissues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P42765</td>
<td></td>
<td>333.45</td>
<td>31.49</td>
<td>41.9</td>
<td>18.83±3.72</td>
<td>9.04±2.48</td>
<td>0.48</td>
</tr>
<tr>
<td>P24752</td>
<td></td>
<td>94.33</td>
<td>39.34</td>
<td>45.2</td>
<td>4.73±0.62</td>
<td>1.74±0.42</td>
<td>0.37</td>
</tr>
<tr>
<td>Q8N556</td>
<td></td>
<td>30.14</td>
<td>10.14</td>
<td>80.7</td>
<td>31.09±6.10</td>
<td>8.74±4.37</td>
<td>0.28</td>
</tr>
<tr>
<td>P00505</td>
<td></td>
<td>141.22</td>
<td>3.49</td>
<td>47.5</td>
<td>22.92±2.25</td>
<td>9.83±0.91</td>
<td>0.43</td>
</tr>
</tbody>
</table>

**Note:** The table includes information on significantly altered proteins levels in HCC tissue vs. adjacent non-Tumor tissue as determined by Quantitative Proteomics. The data are presented for both increased and decreased proteins, with details on the MS/MS identification score, %Cov, MW (kDa), Abundance level (+10^6 arbitrary unit), (Mean±SEM), Ratio, and P-value.
shown in Figure 4, the area under the curve (AUROC) of HCC and LC was 0.728 [95% confidence interval (CI); 0.640-0.817, P<0.001] for AFP, 0.780 (95%CI; 0.691-0.869, P<0.001) for cathepsin D and 0.600 (95%CI; 0.488-0.711, P=0.069) for annexin A2. These data confirmed that cathepsin D, but not annexin A2, was a potential biomarker for HCC and its diagnostic performance was further evaluated.

Based on the AUROC, a cut-off level of serum cathepsin D level reflected the highest accuracy for differentiating HCC from LC was 184 ng/mL. At this concentration, the overall sensitivity and specificity of this biomarker was 81.3% and 67.4%, respectively. Regarding AFP with the current recommendation cut-off point of 20 ng/mL, its sensitivity and specificity was 52.5% and 67.4%, respectively. Among early HCC (BCLC stages 0 and A), 13 of 36 (36.1%) patients had elevated levels of serum AFP ≥20 ng/mL, whereas 26 (72.2%) patients had serum cathepsin D ≥184 ng/mL. For intermediate/advanced HCC (BCLC stage B and C), the sensitivity of the corresponding biomarkers was 65.9% and 88.6%, respectively. When both serum AFP and cathepsin D were used together, the sensitivity of the combined assay was 83.3% for early HCC and 95.5% for intermediate/advanced HCC.

Correlation of serum cathepsin D with clinical characteristics and survival

Serum cathepsin D level was positively correlated with serum aspartate aminotransferase (AST) (r=0.407, P=0.001), alanine aminotransferase (ALT) (r=0.258, P=0.021) and AFP (r=0.323, P=0.003), but had negative correlation with albumin (r=-0.291, P=0.003). The biomarker also had a positive correlation with tumor size (r = 0.357, P=0.001). Among different tumor stages, serum cathepsin D in patients with early HCC (BCLC 0 and A) (276.3 ± 166.7 ng/mL) was significantly lower than those of patients with BCLC stages B (373.0 ± 194.6 ng/mL) and patients with BCLC stages C (455.2 ± 192.4 ng/mL) (P = 0.036 and P=0.001, respectively). Serum cathepsin D level in patients BCLC stages B was also lower that in patients with BCLC stage C, although did not reach the statistical significance (P = 0.184).

The potential prognostic value of serum cathepsin D was also analyzed. Among patients with HCC, the median value of serum cathepsin D (305 ng/mL) was used as a cut-off level the compare the overall survival. Our data showed that the overall survival of patients with low levels (<305 ng/mL) was significantly better than that of patients whose serum levels were ≥305 ng/mL (P=0.010 by log rank test) (Figure 5).

Discussion

HCC is regarded as a complex multistep malignant tumor, which arises from a variety of environmental, host genetic and viral factors (Arzumanyan et al., 2013; Forner et al., 2018). Among these etiological factors, chronic HBV infection is the most important risk of HCC development, accounting for at least 50-60% of all cases among Thai and Southeast Asian populations. Thus, the identification of potential diagnostic and prognostic biomarkers for HBV-related HCC is highly needed in clinical practice. In the present study, we applied proteomics strategy to detect and quantitate differentially expressed proteins in various types of liver cancer, including HCC, iCCA and CRLM. In particular, we revealed several HBV-related HCC associated proteins that might be used as novel biomarkers for the diagnosis and prognosis. In the discovery cohort, our data showed that annexin A2, cathepsin D, calnexin, plectin and prohibitin were significantly elevated in HCC tissues. We

### Table 1. Continued

<table>
<thead>
<tr>
<th>Accession no.</th>
<th>MS/MS identification score</th>
<th>%Cov</th>
<th>MW (kDa)</th>
<th>Abundance level (&lt;10^6 arbitrary unit) Normal</th>
<th>(Mean±SEM)</th>
<th>Ratio (Cancer/Normal)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q93088</td>
<td>#Betaine-homocysteine S-methyltransferase 1</td>
<td>499.97</td>
<td>26.85</td>
<td>45</td>
<td>44.73±2.93</td>
<td>6.02±2.04</td>
<td>0.13</td>
</tr>
<tr>
<td>P04040</td>
<td>#Catalase</td>
<td>327.97</td>
<td>33.21</td>
<td>59.7</td>
<td>52.98±3.72</td>
<td>22.17±0.23</td>
<td>0.42</td>
</tr>
<tr>
<td>P02748</td>
<td>#Complement component C9</td>
<td>61.77</td>
<td>10.38</td>
<td>63.1</td>
<td>8.40±0.54</td>
<td>5.28±0.37</td>
<td>0.63</td>
</tr>
<tr>
<td>P33260</td>
<td>#Cytochrome P450 2C18</td>
<td>44.23</td>
<td>6.53</td>
<td>55.7</td>
<td>18.14±2.40</td>
<td>0.51±0.51</td>
<td>0.03</td>
</tr>
<tr>
<td>O75891</td>
<td>#Cytosolic 10-formyltetrahydrofolate dehydrogenase</td>
<td>229.11</td>
<td>19.18</td>
<td>98.8</td>
<td>10.26±1.45</td>
<td>5.92±0.51</td>
<td>0.58</td>
</tr>
<tr>
<td>O94905</td>
<td>#Erlin-2</td>
<td>163.33</td>
<td>34.51</td>
<td>37.8</td>
<td>11.99±1.89</td>
<td>2.44±0.21</td>
<td>0.2</td>
</tr>
<tr>
<td>P00367</td>
<td>#Glutamate dehydrogenase 1, mitochondrial</td>
<td>662.68</td>
<td>23.48</td>
<td>61.4</td>
<td>80.72±13.99</td>
<td>30.44±4.93</td>
<td>0.38</td>
</tr>
<tr>
<td>Q7Z4W1</td>
<td>#L-xylulose reductase</td>
<td>251.82</td>
<td>24.59</td>
<td>25.9</td>
<td>50.84±1.66</td>
<td>7.68±1.27</td>
<td>0.15</td>
</tr>
<tr>
<td>P50897</td>
<td>#Palmitoyl-protein thioesterase 1</td>
<td>71.99</td>
<td>15.03</td>
<td>34.2</td>
<td>4.82±0.08</td>
<td>2.29±0.10</td>
<td>0.48</td>
</tr>
<tr>
<td>P16083</td>
<td>#Ribosylhydronicotinamide dehydrogenase</td>
<td>74.47</td>
<td>25.11</td>
<td>25.9</td>
<td>12.80±1.75</td>
<td>0.62±0.62</td>
<td>0.05</td>
</tr>
<tr>
<td>Q16762</td>
<td>#Thiosulfate sulfurtransferase</td>
<td>99.98</td>
<td>26.26</td>
<td>33.4</td>
<td>44.36±6.06</td>
<td>7.87±0.32</td>
<td>0.18</td>
</tr>
</tbody>
</table>

%Cov, %Sequence coverage [(number of the matched residues/total number of residues in the entire sequence) x 100%]
fatty liver disease (NAFLD), circulating cathepsin D level was shown to be useful for the assessment of disease severity in some studies (Walenbergh et al., 2015; Walenbergh et al., 2016), although other recent data did not confirm this observation (Kamarajah et al., 2019). In patients with HCC, it was demonstrated that immunoreactivity scores of cathepsin D was significantly elevated in HCC tissue compared with surrounding non-cancerous liver specimens (Lin et al., 2016), which was in line with the findings of our report.

Our data also indicated that serum cathepsin D level was significantly higher in patients with HCC than patients with LC and healthy controls. Additionally, it was shown from the ROC curve analysis that serum cathepsin D could satisfactorily differentiate HCC from LC. In fact, serum cathepsin D might represent a better biomarker for detecting small HCC in comparison with AFP. Currently, serum AFP is regarded as the most commonly-used biomarker for HCC. However, this fetal glycoprotein has some restrictions in clinical use, particularly a relative low sensitivity in early-stage HCC. In our study, combined serum cathepsin D and AFP had the sensitivity of 80% for the diagnosis of early HCC and approached 96% for detecting intermediate and advanced stages. Thus, the combined usage of AFP and cathepsin D might be more beneficial than using AFP alone for identifying early HCC. However, further studies that include a larger number of patients are necessary to confirm this observation.

Our data also revealed that there was a positive correlation between increased serum cathepsin D and aggressive characteristics of HCC. In particular, enhanced serum cathepsin D level was detected more frequently in patients with larger tumor size and advanced BCLC stages. Indeed, the level of this biomarkers was steadily elevated in patients with more advanced tumor stages in comparison with early HCC. These findings might suggest that cathepsin D involves in the development and progression of HCC, similarly to the observations found in other cancers. Moreover, the Kaplan-Meier analysis indicated that overall survival of patients with high serum level was significantly inferior to that of the low expression group, supporting that serum cathepsin D could be unfavorable predictor of overall survival. To the best of our understanding, this study is the first to demonstrate the diagnostic and prognostic importance of serum cathepsin D in patients with HCC.

In summary, our report showed that tissue-based proteomic profiles in HCC substantially differed from those observed in iCCA and CRLM. Among the selected candidate proteins, cathepsin D was shown to be highly expressed in HCC tissues compared with adjacent non-tumor tissues as assess by Western blotting. Moreover, circulating levels of cathepsin D were significantly elevated in patients with HCC compared to patients with LC and healthy controls. The combined assay of cathepsin D and AFP might also provide a promising tool to better differentiate early HCC from LC. Apart from its diagnostic role, serum cathepsin D might also serve as a prognostic biomarker of HBV-related HCC. Additional large-scale studies are needed to verify our findings and to explore whether this biomarker has a clinical utility in further confirmed their expression levels in liver tissues using Western blot analysis. However, only annexin A2 and cathepsin D were significantly up-regulated in HCC issues in comparison with adjacent non-tumor tissues. Moreover, these proteins were not found to be any difference in pair tumor and non-tumor tissues of iCCA and CRLM. Thus, measurement of annexin A2 and cathepsin D might represent specific tissue biomarkers for HCC in differentiating from iCCA and CRLM.

In the validated cohort using serum specimens, we observed that circulating annexin A2 level in patients with HCC did not differ from those found in patients with LC and healthy controls, indicating that this protein might not be a reliable biomarker of HCC. Annexin A2, a 36-kDa binding protein belonged to the annexin superfamily, is localized to the extracellular surface of endothelial cells, as well as in several types of tumor cells (Gerke et al., 2002). Indeed, annexin A2 was found to promote various processes related to HCC progression and their resistance to anticancer therapy (Qiu et al., 2020). Additionally, recent data showed that annexin A2 was involved in immune escape of HCC by inducing regulatory T cells and downregulating natural killer cells and dendritic cells (Qiu et al., 2020). Previous data also showed that annexin A2 was up-regulated in HCC tissues compared with non-cancerous tissues and its expression was correlated with aggressive tumor characteristics including tumor differentiation, portal vein thrombus and metastasis (Zhang et al., 2015). Moreover, annexin A2 was shown as an independent serological marker for detecting an early stage of HBV-related HCC (Sun et al., 2013). In contrast, some reports demonstrated that serum annexin A2 was not a good diagnostic and prognostic biomarker for HBV-related HCC, particularly among patients with underlying cirrhosis (Liu et al., 2013). Additionally, a recent report indicated that the prognostic value of annexin A2 was diminished in several clinical subgroups, including advanced HCC or severe impaired liver function (Huang et al., 2021). These inconsistent results in terms of diagnostic and prognostic value of annexin A2 are not known but might be related to the heterogeneity of studied populations, as well as the different in detection methods and distinct functions of intracellular and extracellular of annexin A2 proteins (Hitchcock et al., 2014).

Cathepsin D is a soluble lysosomal aspartic endopeptidase that has main functions involved in protein degradation and precursor activation in normal cells (Masson et al., 2010). It has been shown that cathepsin D is over-expressed and secreted by various human tumor cell types, including breast, ovarian, prostate and bladder cancers (Nicotra et al., 2010; Benes et al., 2008). Previous studies have demonstrated that cathepsin D plays a crucial role in promoting tumorigenesis including tumor growth, angiogenesis, apoptosis and metastatic properties (Masson et al., 2010). In breast cancer, for instance, a recent meta-analysis has revealed that high expression levels of cathepsin D are associated with poor prognosis and this biomarker could also be served as a potential target of anticancer therapy (Kang et al., 2020). Currently, the role of Cathepsin D in chronic liver disease has not been sufficiently explored. Among patients with nonalcoholic
other etiological causes of HCC.

Author Contribution Statement

PS and PT design initiated the study idea and design of the study, NC and VT performed experiment, interpretation, data analysis and visualization. SS, NP and AV collected and provided the clinical data. PS and AV collected and provided specimen. PT and NC drafted manuscript. PT and PS provided revision of scientific content to manuscript. NC re-write the manuscript. PS and PT provided funding. All authors read and approved this manuscript.

Acknowledgements

For Western blot and ELISA assays, we also would like to thank Wirin Bowonsumsari, Rattanaporn Kiatbumrung and Pornjira Somnark from Center of Excellence in Hepatitis and Liver Cancer, Faculty of Medicine, Chulalongkorn University.

Funding Statement

This study was supported by Ratchadapisek Somphot Endowment Fund (RA65/008), Faculty of Medicine, Chulalongkorn University. The study was supported by The Thailand Research Fund (TRF) Senior Research Scholar (RTA6280004) and Center of Excellence in Hepatitis and Liver Cancer, Faculty of Medicine, Chulalongkorn University. VT is supported by Mahidol University research grant.

Ethical Declaration

The Institute Ethics Committee of Faculty of Medicine, Chulalongkorn University approved the research.

Conflict of Interest

All authors have no conflict of interest.

References


This work is licensed under a Creative Commons Attribution-Non Commercial 4.0 International License.