Differential plasma proteomes of the patients with *Opisthorchiasis viverrini* and cholangiocarcinoma identify a polymeric immunoglobulin receptor as a potential biomarker

Sattrachai Prasopdee\(^a,b\), Yodying Yingchutrakul\(^c\), Sucheewin Krobthong\(^d,e\), Montinee Pholhelm\(^a,b\), Patompon Wongtrakoonpate\(^d,f\), Kritiya Butthongkomvong\(^g\), Jutharat Kulantiwong\(^h\), Teva Phanaksri\(^i\), Anthicha Kunjantarachot\(^b\), Thanakrit Sathavornmanee\(^j\), Smarn Tesana\(^k\), Veerachai Thitapakorn\(^a,b,\*\)

\(^a\) Thammasat Research Unit in Opisthorchiasis, Cholangiocarcinoma, and Neglected Parasitic Diseases, Thammasat University, Pathum Thani 12120, Thailand

\(^b\) Chulabhorn International College of Medicine, Thammasat University, Pathum Thani 12120, Thailand

\(^c\) Proteomics Research Team, National Omics Center, NSTDA, Pathum Thani 12120, Thailand

\(^d\) Center for Neuroscience, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

\(^e\) Interdisciplinary Graduate Program in Genetic Engineering, Kasetsart University, Bangkok 10900, Thailand

\(^f\) Department of Biochemistry, Faculty of Science, Mahidol University, 10400 Bangkok, Thailand

\(^g\) Medical Oncology Unit, Udonthani Cancer Hospital, Ministry of Public Health, Udonthani 41330, Thailand

\(^h\) Faculty of Science, Udon Thani Rajabhat University, Ministry of Public Health, Udon Thani 41000, Thailand

\(^i\) Chonburi Hospital, Ministry of Public Health, Chonburi 20000, Thailand

ARTICLE INFO

Keywords:
Plasma proteome
LC-MS/MS
Cholangiocarcinoma
*Opisthorchis viverrini*
S100A9
PIGR

ABSTRACT

In Southeast Asian countries, nitrosamine compounds and the liver fluke *Opisthorchis viverrini* have long been identified as carcinogens for cholangiocarcinoma (CHCA). In order to effectively treat *O. viverrini* infections and prevent the development of CHCA, methods for disease detection are needed. This study aims to identify biomarkers for *O. viverrini* infection and CHCA. In the discovery phase, technical triplicates of five pooled plasma pools (10 plasma each) of healthy control subjects (noOVCCA), *O. viverrini* subjects (OV), and cholangiocarcinoma subjects (CCA), underwent solution-based digestion, with the label-free method, using a Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer and UltiMate 300 LC systems. The noOVCCA, OV, and CCA groups demonstrated different profiles and were clustered, as illustrated by PCA and heat map analysis. The STRING and reactome analysis showed that both OV and CCA groups up-regulated proteins targeting immune system-related proteins. Differential proteomic profiles, S100A9, and polymeric immunoglobulin receptor (PIGR) were specifically expressed in the CCA group. During the validation phase, another 50 plasma samples were validated via the PIGR sandwich ELISA. Using PIGR > 1.559 ng/ml as a cut-off point, 78.00% sensitivity, 71.00% specificity, and AUC = 0.8216, were obtained. It is sufficient to differentially diagnose cholangiocarcinoma patients from healthy patients and those with *Opisthorchiasis viverrini*. Hence, in this study, PIGR was identified and validated as a potential biomarker for CHCA. Plasma PIGR is suggested for screening CHCA, especially in an endemic region of *O. viverrini* infection.

1. Significance of the study

By differentiating the proteomic profiles of healthy control subjects, *Opisthorchiasis viverrini* subjects, and cholangiocarcinoma subjects, the S100A9 and polymeric immunoglobulin receptor (PIGR) were identified as potential biomarkers for cholangiocarcinoma. The plasma PIGR concentration was further validated for its potential as a diagnostic biomarker. The results demonstrate that the plasma PIGR concentration was significantly higher in cholangiocarcinoma subjects than in patients only infected with *O. viverrini*, as well as control subjects (\(P < 0.05\)). The sensitivity, specificity, and accuracy were obtained. This suggests plasma PIGR is a potential biomarker for...
screening cholangiocarcinoma, especially in an endemic area of *O. viverrini*.

2. Introduction

Cholangiocarcinoma (CHCA), a bile duct cancer, is a nefarious disease due to its initial asymptomatic nature, resulting in a high mortality rate (Banales et al., 2020). Globally, CHCA is an uncommon disease; high prevalence and incidence rates have been reported in Southeast Asian countries, including Thailand, owing to the endemic liver fluke *Opisthorchis viverrini* (Steele et al., 2018). Infections due to O. viverrini and its sister taxa *O. felineus*, as well as *Clonorchis sinensis*, have been verified as carcinogens for CHCA (IARC, 1994). In non-endemic areas, however, primary sclerosing cholangitis (PSC), chronic inflammation, is a prema-

lignant condition that leads to CHCA (Banales et al., 2020). O. viverrini infections are present in raw or undercooked fishes containing the parasite’s metacercariae. Chronic infections with *O. viverrini* result in DNA damage due to reactive oxygen species, inflammation, accumulated mutations, and, ultimately, CHCA (Pinlaor et al., 2003, 2009; Prakobwong et al., 2011). To prevent CHCA, it is vital that O. viverrini infections be diagnosed early and treated with the antihelmintic drug praziquantel. Furthermore, the early detection of CHCA in its asymptomatic stages al-

lows for better chemotherapeutic response and overall survival. Detecting host-responsive proteins may help detect *O. viverrini* infections and/or early-stage CHCA. To identify specific proteins for use as biomarkers, the proteomics-based LC-MS/MS approach was applied. In this study, differential plasma proteome profiles of non-*O. viverrini* infection and non-cholangiocarcinoma subjects (nonOVCCA), *O. viverrini* subjects (OV), and cholangiocarcinoma subjects (CCA) were investigated. The Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrom-

eter was used to identify the potential biomarkers for *O. viverrini* infections and CHCA.

3. Materials and methods

3.1. Chemicals and reagents

Trypsin was obtained from Promega (Promega Co., Madison, WI, USA). UV-Vis spectroscopy was performed using the Synergy H1 micro-

plate reader (BioTek Corp., HT, USA). Sodium dodecyl sulfate (SDS), acetonitrile (analytical grade), acetic acid (glacial), acetone (AR), HEPES sodium salt hydrate, potassium hydroxide (KOH), ammonium bicarbon-

ate, Tris HCl, and formic acid (LC-MS grade) were all obtained from Sigma-Aldrich. An iodoacetamide (IAA) and dithiothreitol (DTT) were

purchased from J.T. Baker (Thermo Fisher Scientific, USA). An ionexchange column was purchased from GE Healthcare. RapidGest SF Surfactant was obtained from Waters. The acetonitrile (ultra-LC-MS) and LC-MS-grade water were purchased from J.T. Baker (Thermo Fisher Scientific, Loughborough, UK).

3.2. Subjects

The protocol of this study was approved by the Human Ethics Com-

mittee of Udonthani Cancer Hospital, Udon Thani, Ministry of Public Health, Thailand, protocol number UCH-CT 11/2563. The fecal exami-

nation for *O. viverrini* eggs, liver ultrasonography, and histopathology were used to classify subjects. During the discovery phase, 50 subjects of each group were analyzed using LC-MS/MS-based proteomics. During the validation phase, another 50 subjects in each group were analyzed using LC-MS/MS-based proteomics. The subjects were divided into three groups based on their health statuses. Subjects with a normal physical examina-

tion, abnormal liver ultrasonography, and CHCA confirmed on liver tissue

histopathology were assigned to the cholangiocarcinoma group (CCA). Subjects in the nonOVCCA and CCA groups were recruited from Udonthani Cancer Hospital, Udon Thani, Thailand. Subjects in the OV group were recruited during the annual health check-ups at health promotion hospitals in Nong Khai province, Thailand. All subjects in the OV group were asymptomatic with low intensity of infection or light infection (<1000 EPG) (Boondit et al., 2020; Maleewong et al., 1992). Recruitment criteria, group allocation, and a summary of the clinical statuses of subjects are summarized in Figure 1 and Table 1.

3.3. Blood collection and plasma preparation

Blood was collected from the median cubital vein using a 21 G needle. EDTA was used as an anti-coagulant. Plasma was collected by mixing by inverting the blood before spinning at 3000xg for 10 min at room tempera-

ture. The plasma was then collected and aliquoted to avoid freeze-

thaw cycles and kept at ~80 °C until used.

3.4. Sample preparation for gel-free-based proteomes

A triplicate of 5 pooled plasma pools (10 plasma each) of each group (technical triplicate of 5 biological replicates) was used in this study (Supplementary File 1). Protein concentrations of plasma were measured using Lowry’s method (DC Protein Assay, Bio-Rad, USA) with bovine serum albumin as the standard reference (2–10 μg/μL). The pooled plasma was prepared by pooling equal amounts of protein from each subject (100 μg each). For each group, the 5 pooled plasma samples were prepared by pooling 10 individual plasma samples each. Finally, 5 pooled plasma samples/group of all 3 groups were prepared (15 pooled plasma samples in total). The plasma proteins were prepared and analyzed as previously described (Krobthong et al., 2021) with minor modifications. Briefly, 50 μL of pooled plasma proteins were mixed with 50 μL of fresh lysis buffer solution (0.2% SDS, 20 mM DTT, in 10 mM HEPES-KOH, pH 8.0). The protein solution was precipitated by using 500 μL of cold acetone at ~20 °C for 24 h. The protein pellet was collected and reconstituted in 0.2% RapidGest SF (Waters Co., UK) in 10 mM ammonium bicarbonate. The protein concentration was re-measured, and 25 μg of protein was then subjected to trypsin digestion with the following pro-

tocol with minor modifications. Briefly, 5 mM DTT was used to reduce proteins by incubating at 60 °C for 20 min. Alkylation of the reduced proteins was conducted using 20 mM IAA for 30 min at room tempera-

ture in the dark. The alkylated proteins were cleaned up to remove the remaining DTT and IAA by a 7 kDa molecular weight cut-off (Zeba™ Spin Desalting Columns, Thermo Fisher Scientific, USA). Proteolytic digestion was performed using 500 ng of trypsin (Promega, Germany) and incubated at 37 °C for 3 h. The tryptic peptides were dried, resuspended in 0.1 % formic acid, and subjected to a TruView LC-MS vial (Waters).

3.5. LC-MS/MS setting for protein identification and quantification

Proteomics data analysis was conducted in data-dependent mode using LC-MS/MS, Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer, and an UltiMate 3000 LC sys-

tem. The tryptic peptides were analyzed by LC-MS/MS with minor modification (Krobthong et al., 2022). Briefly, the tryptic peptides were desalted (C18 PepMap-100 trapping column) and resolved by C18 Pep-

Map™ capillary column (25 cm long) with a 70 min long gradient of acetonitrile and 0.1% formic acid at a flow rate of 0.3 μL/min. The MS spectra were acquired under the following conditions: m/z 400–1600, AGC target set at 3 × 106 ions, and a 60 k resolution. MS2 scan was initiated when the AGC target set reached 1 × 105 ions. An ion selection was performed within 12 s of the dynamic exclusion window. Proteome Discoverer software version 2.4 (Thermo Fisher Scientific), including SEQUEST, Percolator, and Minora algorithms, were used to analyze the raw spectrum files. The peptide was compared with the Human UniProtKB database (20,394 sequences on 12/02/2020). The following
parameters were used for protein identification and quantification: two trypsin missed cleavages as maximum, 10 ppm of precursor mass tolerance, 0.1 Da of fragment mass tolerance, carbamidomethylation (cysteine, static modifications), and oxidation (methionine, dynamic modifications). The false discovery rate (FDR) of 0.01 was selected and used for peptide and protein identification. Total peptide amount was used for normalizing relative protein abundance ratios. All proteome data were submitted to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaino et al., 2013). Dataset identifier: PXD028152; Project Name: Serum proteomics analysis of CCA; Project accession: PXD028152; Project DOI: 10.6019/PXD028152; Username: reviewer_pxd028152@ebi.ac.uk; and Password: IHy85xct.

3.6. Bioinformatics and identification of potential biomarkers

To investigate the heterogeneity of plasma variation in noOVCCA, OV, and CCA subjects, principal component analysis (PCA) was used to observe the differences by Proteome Discoverer. Three-dimensional PCA was conducted by Plotly (R Studio). The heatmap was constructed using the Proteome Discoverer software for hierarchical clustering. Jvenn was used to plot the Venn diagram (Bardou et al., 2014). The specific proteins of each proteome were then used for predicting protein–protein interactions by STRING (Szklarczyk et al., 2021). Reactomes were used to analyze biological networks based on the Reactome Pathway Database (http://www.reactome.org). The significant threshold for altered biological pathways was set at a \( P < 0.01 \) (Haw et al., 2011). The proteins found in the CCA group by the Venn diagram were found in all 15 replicates of the CCA group and were selected and identified as potential biomarkers.

3.7. Measurement of plasma PIGR by sandwich enzyme-linked immunosorbent assay

Validation of identified proteins for CHCA was performed on plasma PIGR using the sandwich ELISA method (Human PIGR ELISA Pair Set, Sino Biological Inc., Beijing, China) following the company-provided protocol. Plate washing of the following step was conducted using 300 μl of washing buffer (20 mM Tris, 150 mM NaCl, pH 7.4, and 0.05% Tween®20) (microplate washer, APW-100, Allsheng) 3 times each. Another 50 plasma samples in each subject group were used to verify the potential as a diagnostic biomarker for CHCA using plasma PIGR. For each sample, plasma diluted to 1:50 was used and performed in duplicates. The standard human PIGR of 0–312.5 ng/ml was prepared and used for a standard curve plot. The capture antibody was diluted by coating buffer pH 9.6 (0.05 M Na 2CO 3, 0.5M NaHCO 3, and filter sterile) to 2 μg/ml, and 100 μl of diluted capture antibody was then added to a 96-well ELISA plate (Maxisorp, NUNC) and incubated at 4 °C overnight. The plates were washed, blocked for non-specific binding using 300 μl of blocking buffer (2% bovine serum albumin in washing buffer), and incubated for 1 h at room temperature. The plates were washed, and each 100 μl of the diluted plasma samples and standards in sample dilution buffer (0.1% BSA in washing buffer) were added and incubated at room temperature for 2 h. The plates were washed, and 100 μl of 1 μg/ml detection antibody diluted in detection antibody dilution buffer (0.1% BSA in washing buffer) was then added and incubated for 1 h at room temperature. The plates were washed, 200 μl of TMB substrate solution (0.1 mg/ml 3,3,5,5-tetramethylbenzidine [Sigma] diluted in 10 ml of 0.05 M citrate
phosphate buffer pH 5.0, and 2 μl of 30% H2O2 was added, and they were incubated in the dark at room temperature for 20 min. To stop the reaction, 2 N H2SO4 was added, and the absorbance values were measured at 450 nm. The absorbance of each sample and standards were subtracted with zero standard absorbance. The log (PIGR standard concentration, ng/ml) values were plotted against subtracted OD450 and used as a standard curve for PIGR concentration calculation. PIGR concentrations (ng/ml) were then calculated using the linear equation shown in the graph below.

### 3.8 Statistical and diagnostic analyses

The IBM SPSS Statistics for Windows, version 26.0 (IBM Corp, Armonk, NY, USA) was used. Statistically significant differences between the PIGR concentration of each group were determined by one-way analysis of variance (ANOVA) followed by a subsequent Scheffe post-hoc test. GraphPad Prism 9 was used for analyzing the dot plot and the area under the receiver operating characteristic curve (AUC). Diagnostic parameters were calculated using the diagnostic test evaluation calculator (MedCalc Software, available at [https://www.medcalc.org/calc/diagnostic_test.php](https://www.medcalc.org/calc/diagnostic_test.php)), including the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy.

### 4. Results

#### 4.1 Analysis of plasma proteomes

The proteomics data was submitted to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD028152 (Vizcaíno et al., 2013). PCA and heatmap analyses of the triplicate of 5 pooled plasma pools of each group of noOVCCA, OV, and CCA showed differences in proteome profiles. The PCA results (Figure 2A, 2B, and 2C) were consistent with the heatmap (Figure 2D); within-group protein differences were smaller than between-group differences. By plotting PC1 against PC2 and PC1 against PC3, noOVCCA (orange circle) was more closely related to OV (green circle) than CCA (blue circle) (Figure 2A, 2B, and 2C), similar to the results found in the heatmap (Figure 2D). Differences in expressed plasma proteins among each group were clustered hierarchically, as presented in Figure 2D. The heatmap depicts their relative abundance with color coding (down- and up-regulated levels: green to red, respectively). The patterns of down- and up-regulation of protein profiles in CCA differed from the noOVCCA and OV groups. The green vertical line remarks the node of CCA down-regulated proteins, and red vertical line marks the node of CCA up-regulated proteins (Figure 2D).

#### 4.2 Differential proteome profiles

To investigate the pathogenesis-related proteins of *O. viverrini* infections and cholangio-carcinogenesis, differential plasma proteomes, including noOVCCA vs. OV, noOVCCA vs. CCA, and OV vs. CCA, were analyzed further by jvenn (Venn diagram), STRING, and reactome analysis. STRING and reactome analysis results were in agreement. The STRING results of CCA-specific proteins were targeted to the immune system, coagulation factors, and motor-related proteins (Figure 3), while OV-specific proteins were targeted to the motor and chromosome-related proteins and coagulation factors (Figure 3). Interestingly, the OV and CCA shared proteins were targeted to cell adhesion and activation of T cells (Figure 3). For reactome analysis, specific proteins of both OV and CCA groups

---

**Table 1. The demographic and clinical statuses of subjects.** The alcohol consumption, smoking, raw fish eating-habit, and history of *O. viverrini* infection were highest in CCA subjects when compared to noOVCCA and OV subjects.

| Discovery phase by LC/MS/MS | Validation phase by ELISA |
|-----------------------------|---------------------------|
| **noOVCCA** | **OV** | **CCA** | **noOVCCA** | **OV** | **CCA** |
| **Sample size** | 50 | 50 | 50 | 50 | 50 |
| **Sex (Male/Female)** | 16/34 | 28/22 | 38/12 | 21/29 | 29/21 |
| **Age (year)** | | | | | |
| **Min/Max** | 20/59 | 33/75 | 45/94 | 21/60 | 18/79 |
| **Mean ± SD** | 38.36 ± 9.45 | 55.56 ± 9.10 | 61.12 ± 7.70 | 38.44 ± 9.99 | 54.08 ± 13.64 |
| **Alcohol consumption** | | | | | |
| **No** | 33 (66%) | 26 (52%) | 10 (20%) | 20 (40%) | 14 (28%) |
| **Yes** | 17 (34%) | 24 (48%) | 40 (80%) | 30 (60%) | 36 (72%) |
| **Smoking** | | | | | |
| **No** | 42 (84%) | 32 (64 %) | 18 (36%) | 38 (76%) | 28 (56%) |
| **Yes** | 8 (16%) | 18 (36 %) | 32 (64%) | 12 (24%) | 22 (44%) |
| **Raw fish eating-habit (Source of *O. viverrini* infection)** | | | | | |
| **No** | 33 (66%) | 10 (20%) | 8 (16%) | 33 (66%) | 5 (10%) |
| **Yes** | 17 (34%) | 40 (80%) | 42 (84%) | 16 (32%) | 45 (90%) |
| **Uncertain** | 0 (0%) | 0 (0%) | 1 (2%) | 1 (2%) |
| **History of *O. viverrini* infection** | | | | | |
| **No** | 50 (100%) | 42 (84%) | 36 (68%) | 48 (96%) | 45 (90%) |
| **Yes** | 0 (0%) | 8 (16%) | 11 (22%) | 1 (2%) | 5 (10%) |
| **Uncertain** | 0 (0%) | 0 (0%) | 3 (6%) | 1 (2%) |
| **Fermented food eating-habit (Source of nitrosamine)** | | | | | |
| **No** | 2 (4%) | 0 (0%) | 0 (0%) | 3 (6%) | 0 (0%) |
| **Yes** | 48 (96%) | 50 (100%) | 50 (100%) | 47 (94%) | 50 (100%) |
| **Uncertain** | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) |
| **Cancer stage** | | | | | |
| **Stage 1 (%)** | 0 (0%) | 0 (0%) | 1 (2%) | 0 (0%) | 0 (0%) |
| **Stage 2 (%)** | 0 (0%) | 0 (0%) | 5 (10%) | 0 (0%) | 0 (0%) |
| **Stage 3 (%)** | 0 (0%) | 0 (0%) | 1 (2%) | 0 (0%) | 0 (0%) |
| **Stage 4 (%)** | 0 (0%) | 0 (0%) | 43 (86%) | 0 (0%) | 0 (0%) |
Figure 2. Principal component analysis (PCA) and heat map of noOVCCA, OV, and CCA groups. The orange, green, and blue colors represent noOVCCA, OV, and CCA groups, respectively. (A) PC1 to PC2 and (B) PC1 to PC3. (C) Three-dimensional plot of PCA. (D) Heat map of noOVCCA, OV, and CCA groups. The green to red color bar indicates the peak intensity from low to high, respectively. The R and T indicate biological replicate (R1-5) and technical replicate (T1-3), respectively. The green vertical line remarks the node of CCA down-regulated proteins, and the red vertical line remarks the node of CCA up-regulated proteins. (E) Identification of potential biomarkers by differential proteomes. The protein S100A9 and polymeric immunoglobulin receptor (PIGR) are identified as potential biomarkers for CCA. The green color box indicates the peptide peak is matched and found in each replicate and the white color box indicates peptide is not found. PCA and heat map clearly discriminated CCA from noOVCCA and OV.
belonged to the hierarchy of the immune system, homeostasis, signal transduction, and metabolism (Supplementary Files 2 and 3). For the comparison of OV to healthy controls, the immune system hierarchy (MHC class I, interferon, antigen processing, and presentation), metabolism (fat-soluble vitamins, especially vitamin A), and disease are shown (Supplementary File 2). Comparisons between CCA and healthy controls revealed specific pathways of the immune system (TLR, cytokine, antigen processing, and presentation), signal transduction (Rho GTPases and JAK/STAT), and disease (TLR signaling cascade of the immune system). The significant pathways are shown in Supplementary File 3. Overall, results suggested that OV- and CCA-specific proteins were targeted to the immune system and immune system-related signaling pathways.

4.3. Identification of potential biomarkers

In this study, the “present” and “absent” criteria were used to identify potential biomarkers for O. viverrini infection and cholangiocarcinoma. Unfortunately, no proteins specific to O. viverrini infection were identified in this study. However, polymeric immunoglobulin receptor (PIGR) and S100A9 were explicitly found in the CCA group proteome with potential as a cholangiocarcinoma biomarker (Figure 2E). S100A9 has been identified in Supplementary File 3. Overall, results suggested that OV- and CCA-specific proteins were targeted to the immune system and immune system-related signaling pathways.

Figure 3. Protein-protein interaction of noOVCCA, OV, and CCA-specific proteins by STRING. The motor and chromosome-related proteins and coagulation factors were linked to OV while the immune system-related proteins were additionally linked to the CCA protein network.
and validated as a potential biomarker for CHCA (Duangkumpha et al., 2019). Therefore, the PIGR was focused on in this study. From LC-MS/MS results, five peptides of PIGR were specifically identified in all CCA proteome replicates, including 123-YLCGAHSDQLQEGSPIQAWQLFVNEESTITPR-154, 193-LDIQGTGQLLSVAVLQAVENQLR-211, 232-NADLQVLKPEPELVYEDLRGSVTFCALGPEVANVAK-268, 297-ILLNPQDKDGSFSVITGLR-316, 526-LVSLTLNLVTR-536. These five peptides exclusively belonged to the CCA proteome; therefore, PIGR was identified as a potential biomarker for CHCA and further validated by sandwich ELISA.

4.4. Validation of plasma PIGR concentration as a potential biomarker

Plasma PIGR concentration based on sandwich ELISA was calculated from a linear equation ($R^2 = 0.9789$) of PIGR to a standard log (concentration) curve against an OD of 450 nm. The result showed that the concentrations of plasma PIGR were lowest in the OV group and highest in the CCA group (Figure 4). The average concentrations (ng/ml) ± SD of plasma PIGRs of noOVCCA, OV, and CCA were 1.4936 ± 0.6085, 1.3309 ± 0.5248, and 6.0942 ± 7.5840, respectively (Supplementary File 4). One-way ANOVA revealed that at least two groups had significantly different concentrations of PIGR ($P < 0.01$). The Scheffe test for multiple comparisons further revealed that the mean plasma PIGR concentration of the CCA group was significantly different from the OV group ($P < 0.01$; 95% CI: 2.585, 6.941) and the noOVCCA group ($P < 0.01$; 95% CI: 2.422, 6.778). However, there were no significant differences between the OV group and the noOVCCA group ($P = 0.98$; 95% CI: −2.340, 2.014). The AUC of noOVCCA and OV vs. CCA was 0.8216 (Figure 5). The highest AUC (0.8778) was obtained from OV vs. CCA (Figure 5). Using >1.559 ng/ml PIGR cut-off, 78.00% sensitivity, 71.00% specificity, 57.35% positive predictive value, 86.59% negative predictive value, and 73.33% accuracy were obtained (Table 2). However, using plasma PIGR >3.850 as a cut-off, the 100% specificity of the test was obtained with limited sensitivity (Table 2). Hence, the LC-MS/MS, Venn diagram, and ELISA results were consistent and indicated that PIGR was highly specific to CHCA.

5. Discussion

Proteome analysis by LC-MS/MS has been widely used for the characterization and quantification of proteins and peptides, especially in the evaluation of cancer biomarkers, including CHCA (Cevenini et al., 2020; Duangkumpha et al., 2019; Hoshino et al., 2020; Hristova and Chan, 2019; Huang et al., 2017; Janvilisri, 2015; Kim et al., 2009; Krug et al., 2020; Nunez-Naveira et al., 2019; Phanaksri et al., 2022; Prasopdee et al., 2022).

Figure 4. The plasma PIGR concentration. The circle, square, and triangle represent noOVCCA, OV, and CCA, respectively. The group is on the X-axis and the plasma PIGR concentration (ng/ml) is on the Y-axis. The PIGR level is significantly increased in the CCA group. The asterisk (*) indicates a significant difference $P < 0.01$. 

Table 2. The AUC of noOVCCA vs. OV and CCA.

| Group   | AUC       |
|---------|-----------|
| noOVCCA | 0.8216    |
| OV      | 0.8778    |
| CCA     | 0.8778    |

Using >1.559 ng/ml PIGR cut-off, 78.00% sensitivity, 71.00% specificity, 57.35% positive predictive value, 86.59% negative predictive value, and 73.33% accuracy were obtained (Table 2). However, using plasma PIGR >3.850 as a cut-off, the 100% specificity of the test was obtained with limited sensitivity (Table 2). Hence, the LC-MS/MS, Venn diagram, and ELISA results were consistent and indicated that PIGR was highly specific to CHCA.
An individual and pooled sample were used with different advantages and disadvantages. It has been estimated that thousands of individual samples are needed to reduce the biological variability and false biomarker numbers. Pooled samples can help reduce the biological variability of samples, resulting in data with a high degree of confidence (Orton and Doucette, 2013). The time and cost required to acquire individual samples were significantly higher than pooled samples. Based on the incidence of CHCA, sample variability, time, and cost, pooled plasma samples were used in this study to identify potential biomarkers for diagnostic CHCA. Two potential biomarkers for CHCA were identified—S100A9 and the polymeric immunoglobulin receptor (PIGR). S100A9 has been identified as a biomarker for several cancers (Gunaldi et al., 2015; Hermani et al., 2005; Huang et al., 2018; Kim et al., 2009; Lim et al., 2016; Liu et al., 2019; Lv et al., 2020; Meng et al., 2019; Nunez-Naveira et al., 2019; Zhou et al., 2019), including both non-\textit{O. viverrini}-related CHCA (Shi et al., 2013) and \textit{O. viverrini}-related CHCA (Duangkumpha et al., 2019). Meanwhile, PIGR is a biomarker for CHCA in a non-endemic area of \textit{O. viverrini} (Arbelaiz et al., 2017). S100A9 expression was also high in PSC (Reinhard et al., 2012).

Duangkumpha and colleagues identified serum S100A9 as a potential biomarker for CHCA against premalignant periductal fibrosis (PDF) and healthy subjects (Duangkumpha et al., 2019). Shi and colleagues demonstrated that S100A9 is a biomarker for CHCA by 2D-DIGE. The S100A9 immunohistochemical results demonstrated 92.6% sensitivity and 75% specificity in CHCA diagnosis (Shi et al., 2013). Puatkasichonpasutha and colleagues reported that up-regulation of S100A9 gene expression was correlated with longer survival in patients with non-\textit{O. viverrini}-related CHCA (Shi et al., 2013) and \textit{O. viverrini}-related CHCA (Duangkumpha et al., 2019). Meanwhile, PIGR is a biomarker for CHCA in a non-endemic area of \textit{O. viverrini} (Arbelaiz et al., 2017). S100A9 expression was also high in PSC (Reinhard et al., 2012).

Therefore, in this study, only PIGR was identified and further validated as a potential biomarker for diagnosing CHCA in the endemic area of \textit{O. viverrini}.

Figure 5. The AUC of the ROC curve of noOVCCA vs. OV (upper left), noOVCCA vs. CCA (upper right), CCA vs. OV (lower left), and noOVCCA and OV vs. CCA (lower right). The % sensitivity is plotted against 100% - % specificity on the X-axis and Y-axis, respectively. The AUC of ROC curves are calculated and indicated in each curve. The highest AUC was obtained in OV vs. CCA.

- **noOVCCA vs. OV**
  - ROC Curve (AUC = 0.5436)

- **CCA vs. OV**
  - ROC Curve (AUC = 0.8778)

- **noOVCCA and OV vs. CCA**
  - ROC Curve (AUC = 0.8216)
PIGR is a receptor for polymeric immunoglobulins such as IgA and IgM and plays a critical role in the transcytosis of IgA and IgM from the basolateral to apical surface in response to immunogens. Aberrations in PIGR have been reported to correlate with an increase and decrease in gene expression and gene polymorphism, as found in several malignancies (Arunugam et al., 2017; Chang et al., 2005; Gologan et al., 2008; Hirunmatit et al., 2003; Okaz et al., 2012; Xiao et al., 2005). Serum levels of PIGR were high in cancer patients before surgery and decreased in the same patients post-operatively (Sogawa et al., 2016). Nevertheless, higher levels of PIGR expression were correlated with longer survival, while the low or absent expressions of PIGR were usually found in patients with lymph node metastases (Ai et al., 2011; Arunugam et al., 2017; Bernsten et al., 2014; Fristedt et al., 2014b; Gologan et al., 2008; Niu et al., 2014; Qi et al., 2016; Wang et al., 2014). Conversely, high levels of PIGR expression resulted in chemoresistance and poor prognosis in patients with pancreatic cancer (Ohkuma et al., 2020). Yue and colleagues demonstrated that PIGR promoted cell transformation and induced oncogenic growth by Yes activation in the MEK/ERK cascade (Yue et al., 2017).

In conclusion, proteomic analysis using solution-based tryptic digestion, a label-free method, Orbitrap analysis followed by high-energy collisional dissociation, and technical triplicates with five biological replicates was demonstrated to identify S100A9 and PIGR as potential biomarkers for CHCA. The plasma concentrations of PIGR showed substantial sensitivity, specificity, and accuracy as a diagnostic and screening tool for CHCA. Therefore, PIGR has great potential as a biomarker in differentiating patients with CHCA from patients with O. viverrini infections and healthy subjects. The application of plasma PIGR concentration holds promise in future CHCA screening and diagnostic protocols, especially in areas endemic to O. viverrini.

**Declarations**

**Author contribution statement**

Satkrachai Prasopdee; Yodying Yingchutrakul; Sucheevin Krothbong; Veerachai Thitapakorn, Ph.D: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Montinee Pholholm: Performed the experiments; Analyzed and interpreted the data.

Patomporn Wongtrakoongate; Kritiya Butthongkomvong: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Jutharat Kulsantiwong: Performed the experiments; Analyzed and interpreted the data.

Teva Phanakri; Anthicha Kunjantarachot; Thanakrit Sathavornmaenee: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Smarn Tesana: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

**Funding statement**

Veerachai Thitapakorn was supported by Thammasat Research Unit in Opisthorchiasis, Cholangiocarcinoma, and Neglected Parasitic Diseases, Thammasat University (TRU-OCN) and Thai Government Research Fund through Thammasat University (50/2559 and 64/2560). Satkrachai Prasopdee was supported by Thai Government Research Fund through Thammasat University (36/2562).

**Data availability statement**

Data associated with this study has been deposited at “Sequence proteomics analysis of CCA” under the accession number “PXD028152” [10.6019/PXD028152].
Declaration of interest’s statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e10965.

Acknowledgements

The authors acknowledge the Central Instrument Facility, Faculty of Science, Mahidol University for supporting Thermo Scientific™ Q Exactive™ HF hybrid quadrupole-Orbitrap mass spectrometer and UltiMate 300 LC system. Veerachai Thitapakorn was supported by Thammasat Research Fund through Thammasat University (50/2018) and (16/2019). Sattrachai Prasopdee was supported by Thai Government Research Fund through Thammasat University (50/2559 and 64/2018) and the National Research Council of Thailand (35/2018). The authors declare no conflict of interest.

References

Ai, J., Tang, Q., Wu, Y., Xu, Y., Feng, T., Zhou, R., Chen, Y., Gao, X., Zhu, Q., Yue, X., et al., 2011. The role of polymeric immunoglobulin receptor in inflammation-induced tumor metastasis of human hepatocellular carcinoma. J. Natl Cancer Inst. 103, 1696–1712.
Arbeláez, A., Azkargorta, M., Krawczyk, M., Santos-Laso, A., Lapiz, A., Perugorria, M.J., Errasti, O., Gonzalez, E., Jimenez-Agurro, P., Lacasta, A., et al., 2017. Serum extracellular vesicles contain protein biomarkers for primary sclerosing cholangitis and cholangiocarcinoma. Hepatology 66, 1125–1143.
Arumugam, P., Bhattacharya, S., Chin-Aleong, J., Capasso, M., Kocher, H.M., 2017. Calcium-binding proteins S100A8 and S100A9 as novel diagnostic molecular events related to multistep carcinogenesis. Int. J. Cancer 129, 88–96.
Balestra, C., Previtali, C., Alleva, E., 2013. Venn diagram viewer. BMC Bioinf. 15, 293.
Butthongkomvong, K., Tesana, S., Sathavornmanee, T., Thitapakorn, V., 2022. Colorectal polymeric immunoglobulin receptor expression is correlated with hepatitis metastasis and poor prognosis in colon carcinoma patients with hepatitis C. Hepatogastroenterology 69, 652–659.
Chang, Y., Lee, T.C., Li, J.C., Lai, T.L., Chua, H.H., Chen, C.L., Doong, S.L., Chou, C.K., Boondit, J., Suwannahitatorn, P., Siripattanapipong, S., Leelayoova, S., Mungthin, M., 2016. Proteomic work ow for biomarker identi cation of serum protein biomarker candidates for primary sclerosing cholangitis. Proteomics 17, 838–843.
Chenzuo, J., Sundgren, S., Bordini, P., Siripattanapong, S., Leelayoova, S., Mungthin, M., Tan-Ariya, P., Piyaraj, P., Naaglor, T., Ruang-Areerate, T., 2020. An epidemiological survey of Opisthorchis viverrini infection in a lightly infected community, eastern Thailand. Am. J. Trop. Med. Hyg. 102, 838–843.
Chen, Y., Li, X., Batten, G., Liu, F., Ye, P., Bi, T., Teng, H., Li, H., Wu, L., Wang, Y., et al., 2014. COLORECTAL polymeric immunoglobulin receptor expression is associated with hepatitis metastasis and poor prognosis in colon carcinoma patients with hepatitis C. Hepatogastroenterology 69, 652–659.
Cheung, Y., Li, X., Wu, Y., Feng, T., Chen, Y., Gao, X., Zhu, Q., Yue, X., et al., 2011. The role of polymeric immunoglobulin receptor in inflammation-induced tumor metastasis of human hepatocellular carcinoma. J. Natl Cancer Inst. 103, 1696–1712.
Chompoo, P., Bhattacharya, S., Chin-Aleong, J., Capasso, M., Kocher, H.M., 2017. Calcium-binding proteins S100A8 and S100A9 as novel diagnostic molecular events related to multistep carcinogenesis. Int. J. Cancer 129, 88–96.
Chompoo, P., Bhattacharya, S., Chin-Aleong, J., Capasso, M., Kocher, H.M., 2017. Calcium-binding proteins S100A8 and S100A9 as novel diagnostic molecular events related to multistep carcinogenesis. Int. J. Cancer 129, 88–96.
Prasopdee, S., Yingchutrakul, Y., Roityrakul, S., Pholhelm, M., Phanakri, T., Kunjastraachot, A., Kulatsitriwong, J., Buithongkoomvong, K., Teena, S., Sathavornmanee, T., Thinsaporn, V., 2022. Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit beta as a potential biomarker for Opisthorchis viverrini infection and cholangiocarcinoma. Parasitology 149, 171–180.

Puetkasichonpasutha, J., Namwat, N., Sa-Ngiamwibool, P., Titapun, A., Suthiphongchai, T., 2020. Evaluation of p53 and its target gene expression as potential biomarkers of cholangiocarcinoma in Thai patients. Asian Pac. J. Cancer Prev. APJCP 21, 791–798.

Qi, X., Li, X., Sun, X., 2016. Reduced expression of polymeric immunoglobulin receptor (pIgR) in nasopharyngeal carcinoma and its correlation with prognosis. Tumour Biol. 37, 11099–11104.

Reinhard, L., Rupp, C., Riedel, H.D., Ruppert, T., Flechtenmacher, C., Weiss, K.H., Kloeters-Plachky, P., Stremmel, W., Schirmacher, P., et al., 2012. S100A9 is a biliary protein marker of disease activity in primary sclerosing cholangitis. PLoS One 7, e29821.

Rizvi, S., Borad, M.J., Patel, T., Gores, G.J., 2014. Cholangiocarcinoma: molecular pathways and therapeutic opportunities. Semin. Liver Dis. 34, 456–464.

Shi, Y., Deng, X., Zhan, Q., Shen, B., Jin, X., Zhu, Z., Chen, H., Li, H., Peng, C., 2013. A prospective proteomic-based study for identifying potential biomarkers for the diagnosis of cholangiocarcinoma. J. Gastrointest. Surg. 17, 1584–1591.

Sogawa, K., Takano, S., Iida, F., Satoh, M., Tsuchida, S., Kawashima, Y., Yoshitomi, H., Sanda, A., Kodera, Y., Takizawa, H., et al., 2016. Identification of a novel serum biomarker for pancreatic cancer, C4b-binding protein alpha-chain (C4BPA) by quantitative proteomic analysis using tandem mass tags. Br. J. Cancer 115, 949–956.

Steele, J.A., Richter, C.H., Echaubard, P., Saenna, P., Stout, V., Sithithaworn, P., Wilcox, B.A., 2018. Thinking beyond Opisthorchis viverrini for risk of cholangiocarcinoma in the lower Mekong region: a systematic review and meta-analysis. Infect. Dis. Poverty 7, 44.

Szklarczyk, D., Gable, A.L., Nastou, X.C., Lyson, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legay, M., Fang, T., Bork, P., et al., 2021. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49, D605–D612.

Tiemin, P., Fanzheng, M., Peng, X., Jihua, H., Ruipeng, S., Yaliang, L., Junlin, X., Qingfu, L., Zhefeng, H., et al., 2020. MUC13 promotes intraperitoneal cholangiocarcinoma progression via EGFR/PI3K/AKT pathways. J. Hepatol. 72, 761–773.

Vizzarino, J.A., Cote, R.G., Cuadrado, A., Dianes, J.A., Fabregat, A., Foster, J.M., Griss, J., Alpi, E., Birim, M., Contell, J., et al., 2013. The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res. 41, D1063–D1069.

Wang, X., Li, X., Sun, X., 2014. Polymeric immunoglobulin receptor expression is correlated with poor prognosis in patients with osteosarcoma. Mol. Med. Rep. 9, 2105–2110.

Wu, Q., Fan, H., Lang, R., Li, X., Zhang, X., Lv, S., He, Q., 2020. Overexpression of 14-3-3sigma modulates cholangiocarcinoma cell survival by PI3K/akt signaling. BioMed Res. Int. 2020, 3740418.

Xiao, T., Ying, W., Li, L., Hu, Z., Ma, Y., Jiao, L., Ma, J., Cai, Y., Lin, D., Guo, S., et al., 2005. An approach to studying lung cancer-related proteins in human blood. Mol. Cell. Proteomics 4, 1480–1486.

Yue, X., Ai, J., Xu, Y., Chen, Y., Huang, M., Yang, X., Hu, B., Zhang, H., He, C., Yang, X., et al., 2017. Polymeric immunoglobulin receptor promotes tumor growth in hepatocellular carcinoma. Hepatology 65, 1948–1962.

Zhou, M., Li, M., Liang, X., Zhang, Y., Huang, H., Feng, Y., Wang, G., Liu, T., Chen, Z., Pei, H., Chen, Y., 2019. The significance of serum S100A9 and TNC levels as biomarkers in colorectal cancer. J. Cancer 10, 5315–5323.