


RESEARCH

Open Access



# Mass spectrometry-based analysis of formalin-fixed, paraffin-embedded distal cholangiocarcinoma identifies stromal thrombospondin-2 as a potential prognostic marker

Johannes Byrling<sup>1</sup>, Theresa Kristl<sup>2</sup>, Dingyuan Hu<sup>1</sup>, Indira Pla<sup>2</sup>, Aniel Sanchez<sup>2</sup>, Agata Sasor<sup>3</sup>, Roland Andersson<sup>1</sup>, György Marko-Varga<sup>2</sup> and Bodil Andersson<sup>1\*</sup> 

## Abstract

**Background:** Distal cholangiocarcinoma is an aggressive malignancy with a dismal prognosis. Diagnostic and prognostic biomarkers for distal cholangiocarcinoma are lacking. The aim of the present study was to identify differentially expressed proteins between distal cholangiocarcinoma and normal bile duct samples.

**Methods:** A workflow utilizing discovery mass spectrometry and verification by parallel reaction monitoring was used to analyze surgically resected formalin-fixed, paraffin-embedded samples from distal cholangiocarcinoma patients and normal bile duct samples. Bioinformatic analysis was used for functional annotation and pathway analysis. Immunohistochemistry was performed to validate the expression of thrombospondin-2 and investigate its association with survival.

**Results:** In the discovery study, a total of 3057 proteins were identified. Eighty-seven proteins were found to be differentially expressed ( $q < 0.05$  and fold change  $\geq 2$  or  $\leq 0.5$ ); 31 proteins were upregulated and 56 were downregulated in the distal cholangiocarcinoma samples compared to controls. Bioinformatic analysis revealed an abundance of differentially expressed proteins associated with the tumor reactive stroma. Parallel reaction monitoring verified 28 proteins as upregulated and 18 as downregulated in distal cholangiocarcinoma samples compared to controls. Immunohistochemical validation revealed thrombospondin-2 to be upregulated in distal cholangiocarcinoma epithelial and stromal compartments. In paired lymph node metastases samples, thrombospondin-2 expression was significantly lower; however, stromal thrombospondin-2 expression was still frequent (72%). Stromal thrombospondin-2 was an independent predictor of poor disease-free survival (HR 3.95, 95% CI 1.09–14.3;  $P = 0.037$ ).

**Conclusion:** Several proteins without prior association with distal cholangiocarcinoma biology were identified and verified as differentially expressed between distal cholangiocarcinoma and normal bile duct samples. These proteins can be further evaluated to elucidate their biomarker potential and role in distal cholangiocarcinoma carcinogenesis. Stromal thrombospondin-2 is a potential prognostic marker in distal cholangiocarcinoma.

\*Correspondence: bodil.andersson@med.lu.se

<sup>1</sup> Department of Clinical Sciences Lund, Surgery, Lund University, and Skåne University Hospital, Lund, Sweden  
Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Keywords:** Distal cholangiocarcinoma, Biliary tract cancer, Mass spectrometry, Parallel reaction monitoring, Biomarker, Stroma, Thrombospondin-2

## Introduction

Cholangiocarcinoma (CCA), an epithelial tumor arising along the biliary tree, is a low- incidence malignancy accounting for approximately 3% of all gastrointestinal cancers [1]. CCA is a notably aggressive malignancy with a current overall 5-year survival rate of less than 5% [2]. Surgical resection is the only treatment that offers curative potential. However, an invasive growth pattern and the absence of early symptoms contribute to presentation with metastatic disease in a majority of patients [3]. Based on its anatomical location along the biliary tree, CCA is classified as intrahepatic (iCCA) or extrahepatic (eCCA) which is further divided into the perihilar (pCCA) and distal (dCCA) subtypes [4]. Anatomically, dCCA is located between the insertion of the cystic duct and the ampulla of Vater [4, 5].

Mounting evidence points to a significant difference between the different subtypes with regards to not only clinical management but also tumor biology and molecular characteristics [3, 6, 7]. However, the majority of previous biomarker studies do not take this subclassification into account [3, 8]. Currently, only one biomarker, namely, plasma carbohydrate antigen 19-9 (CA 19-9) is in clinical use for the diagnosis of CCA. The sensitivity and specificity of CA 19-9 are not sufficient to allow screening of unselected patient populations [9].

A large effort has elucidated the spectra of genomic and transcriptomic alterations associated with malignancy along the biliary tree [6, 7]. However, changes at the mRNA level do not necessarily correlate with changes at the protein level [10, 11], and studies of the proteomic alterations that develop during carcinogenesis can provide complementary information and help identify new biomarkers [12, 13]. Modern mass spectrometry (MS) platforms utilizing online liquid-phase separation with tandem MS (LC-MS/MS) have been used for protein identification and quantification in human sample cohorts, enabling quantitative comparison of thousands of proteins [13–15]. Traditionally, MS-based proteomics has been performed using fresh frozen tissues. However, it has been demonstrated that the more widely available formalin-fixed, paraffin-embedded (FFPE) tissues can also be utilized successfully [16, 17].

A discovery LC-MS/MS experiment can identify a large number of differentially expressed proteins (DEPs) between the investigated biological conditions

providing information regarding proteomic alterations associated with malignancy as well as biomarker candidates. However, few identified potential biomarkers in cancer research have been able to progress to clinical utilization, in part due to the difficulty in validating the large number of biomarkers [18]. Parallel reaction monitoring (PRM) is an MS technique developed for targeted quantification [19, 20]. PRM has been shown to provide improved quantitative precision and to be able to quantify proteins at lower concentrations than a typical discovery LC-MS/MS experiment. Thus, PRM can be used for large scale verification of proteins identified from discovery experiments [21, 22].

The aim of the present study was to use a combined discovery LC-MS/MS and PRM-verification workflow to analyze resected dCCA and normal bile duct FFPE samples to identify DEPs. As a secondary aim, further validation of thrombospondin-2 (THBS2) expression was performed using immunohistochemistry (IHC).

## Materials and methods

### Study population

All consecutive patients who underwent surgery with curative intent for dCCA at the Department of Surgery, Skåne University Hospital, Lund and Malmö, between January 2000 and December 2015 were identified from hospital records. Only patients with no neoadjuvant treatment and no 30-day mortality were included. Sixty-four patients were identified. FFPE materials from 59 patients were available and subsequently retrieved from the local biobank at the Department of Pathology, Skåne University Hospital, Lund and Malmö. All samples underwent histopathological reevaluation by an experienced gastrointestinal pathologist (A.S) blinded to the original assessment. The cohort has been previously described [23, 24]. Staging was based on the American Joint Committee on Cancer (AJCC) 7th edition [5]. R1 resection was defined as cancer growth < 1 mm from the resection margin. Demographical and clinical data were retrospectively acquired through patients' medical records. Patients who received at least 3 cycles of postoperative chemotherapy were coded as having received adjuvant treatment. The clinicopathological data of the cohort are presented in Table 1. Patients were postoperatively followed for up to 5 years. Survival analysis was censored at 5 years after surgery for event-free patients. Survival status was recorded in September 2019.

**Table 1 Clinicopathological data in the distal cholangiocarcinoma cohort (N = 59)**

Variable	N	n (%), mean ± SD
Age	59	67 ± 8.1
Sex	59	
Female		20 (34%)
Adjuvant therapy		
Yes	58	30 (52%)
Tumor size (mm)	59	29 ± 9.6
Tumor differentiation	59	
High		0
Intermediate		16 (27%)
Low		43 (73%)
T stage	59	
I		2 (3.4%)
II		0
III		56 (95%)
IV		1 (1.7%)
Lymph node metastases	58	
Present		40 (69%)
AJCC stage	58	
IA		1 (1.7%)
IB		0
IIA		17 (29%)
IIB		39 (67%)
III		1 (1.7%)
R1 resection	59	
Present		34 (58%)
Lymph vessel invasion	57	
Present		42 (74%)
Nerve invasion	58	
Present		48 (83%)
Blood vessel invasion	58	
Present		37 (64%)
Adipose invasion	59	
Present		48 (81%)

AJCC American Joint Committee on cancer, R1 non-radical resection, SD standard deviation, T stage Tumor stage

Disease-free survival (DFS) was defined as the time until clinical diagnosis of dCCA recurrence or death from any cause. Overall survival (OS) was defined as the time until death of any cause.

The control tissues were identified from a prospective database of pancreaticobiliary surgery maintained at the Department of Surgery, Skåne University Hospital, Lund and Malmö. Patients who underwent pancreaticoduodenectomy due to a benign diagnosis without proximity to the distal bile duct were identified.

### MS tissue area selection

Samples were sectioned into 10- $\mu$ m sections. Hematoxylin and eosin (H&E)-stained slides were used to mark the areas of interest in each slide for macrodissection. Areas of interest in the dCCA samples were defined as areas containing only tumor epithelium and intratumoral stromal compartment without any adjacent tissues. The normal bile duct tissues were retrieved from the region between the cystic duct and the ampulla of Vater. The bile duct tissues were histopathologically evaluated, and any sample with dysplasia or inflammation of the bile duct was excluded. The macrodissected area was located from the epithelial lining and up to, but not including the adventitia or pancreatic parenchyma. Macrodissection was carried out using a scalpel with the marked H&E slides as a template for serial sections. An area corresponding to approximately 3 full slides was utilized when available. In some samples with scarce material the maximum available amount was retrieved although it did not correspond to 3 full sections. The material was stored at 4 °C.

### MS sample preparation

Twenty dCCA samples from the ten patients each with the worst and best OS were selected for the discovery study. The selection was performed to allow for additional comparison between prognostic groups that did not reveal significant results (data not shown).

All samples were processed in a manner blinded to patient identity and outcome. Preparation of all samples was performed in parallel. For deparaffinization and protein retrieval, macrodissected samples were incubated in 1 ml of EnVision FLEX Target Retrieval Solution (High pH) (K8004, Dako, Glostrup, Denmark) at a dilution of 1:50 at 97 °C for 10 min. Samples were then centrifuged at 14,000 rcf at 4 °C for 10 min. The supernatants were removed and the deparaffinization steps were repeated. Samples were resuspended in 150  $\mu$ l of 500 mM Tris-HCL (pH 8), followed by transfer to a new tube and incubation at 90 °C for 1 h. Then, 150  $\mu$ l of 6 M guanidine-HCL in 50 mM ammonium bicarbonate (AMBIC) was added, followed by probe sonication for 2  $\times$  4 min on ice. The samples were then centrifuged at 14,000 rcf for 10 min at 24 °C, and the supernatants were transferred to a new tube. Proteins were reduced by the addition of 6  $\mu$ l of 1 M dithiothreitol and incubated for 1 h at 56 °C. Alkylation was performed by the addition of 20  $\mu$ l of 1 M iodoacetamide for 30 min in dark. Precipitation was carried out overnight using pure ethanol at a ratio (v/v) (sample:ethanol) of 1:9, and the samples were stored at -20 °C. Following centrifugation at 14,000 rcf for 15 min at 4 °C the supernatants were carefully discarded, and the

pellets were resuspended in 100  $\mu$ l of 50 mM AMBIC. Protein determination was carried out using the Micro BSA Protein Assay Kit (Thermo Fisher Scientific, Rockford, IL, USA) in accordance with the manufacturer's instructions. Digestion was carried out using sequencing-grade trypsin (Promega, Madison, WI, USA) at a ratio (w/w) (trypsin:protein) of 1:50, followed by incubation at 37 °C overnight. The samples were dried using centrifugal evaporation and resuspended in 0.1% formic acid. Peptide determination was performed according to the manufacturer's instructions using the Pierce Quantitative Colorimetric Peptide Assay (Thermo Fisher Scientific, Rockford, IL, USA). The samples were then spiked with Pierce Peptide Retention Time Calibration Mixture (Thermo Fisher Scientific, Rockford, IL, USA) to evaluate chromatographic performance. A total of 25 fmol of Pierce Retention Time Calibration Mixture was added to every  $\mu$ g of peptide in the sample and the samples were diluted to a final concentration of 0.25  $\mu$ g/ $\mu$ l in 0.1% formic acid for injection.

#### Liquid chromatography conditions

The liquid chromatography (LC) instrument used was an EASY-nLC (Thermo Fisher Scientific, Rockford, IL, USA). Previously prepared material from one patient was measured repeatedly at regular intervals as a quality control. One microgram of sample was injected and measured with a flow rate of 300 nL/minute, and a two-column setup consisting of an Acclaim PepMap RSLC column (75  $\mu$ m  $\times$  25 cm) as the analytical column and Acclaim PepMap 100 column (75  $\mu$ m  $\times$  2 cm) as the precolumn (both from Thermo Fisher Scientific, Rockford, IL, USA) was used for the separation. The LC gradient was created using solvent A (0.1% formic acid) and solvent B (0.1% formic acid in acetonitrile). Based on manual evaluation of the number of peptides eluted at different time points, a nonlinear gradient was developed. The nonlinear gradient started at 5% B and increased to 22% B at 95 min and 36% B at 150 min. All measurements in the discovery study were performed in technical duplicates. Samples were measured in a randomized order.

#### MS conditions

The samples were analyzed using a Q Exactive Plus mass spectrometer (Thermo Fisher Scientific, Rockford, IL, USA). The equipped ion source was an EASY-Spray (Thermo Fisher Scientific, Rockford, IL, USA). The system was operated in positive mode, data-dependent acquisition (DDA) was used. For peptide identification, a full MS survey scan was collected in the Orbitrap. Fifteen data-dependent higher energy collision dissociation MS/MS scans of the most intense precursors were performed.

The spray voltage was set to 1.75 kV, and the capillary temperature was 300 °C. Moreover, the S-lens radiofrequency level was fixed at 50. MS1 survey scans of the eluting peptides were executed with a resolution of 70,000, recording a window between m/z 350.0 and 1800. The automatic gain control (AGC) target was set to  $1 \times 10^6$  and the maximum injection time was 100 ms. The normalized collision energy (NCE) was set to 25.0% for all scans. The resolution of the data-dependent MS<sup>2</sup> scans was fixed to 35,000, and the values for the AGC target were set to  $1 \times 10^6$  with a maximum injection time of 120 ms.

#### Protein identification and quantification

The software Proteome Discoverer (PD) (version 1.4) (Thermo Fisher Scientific, Rockford, IL, USA) was used for protein identification. The selection of spectra employed the following settings: minimal and maximal precursor mass, 350 and 5000 Da, respectively; and signal-to-noise threshold of 1.5. Parameters for Sequest HT searches were set as follows: precursor mass tolerance, 10 ppm; fragment mass tolerance, 0.02 Da; trypsin; 1 missed cleavage site; UniProt human database; dynamic modification; oxidation (+15.995 Da; (M, P)); static modification; carbamidomethyl (+57.021 Da; (C)). A percolator was used for the processing node and the false discovery rate (FDR) was set to 1%. Proteins were identified using at least two peptides. A precursor ion area detector was used for quantification, and each protein was quantified from the average area of the 3 most abundant peptides identified for that particular protein. To increase the number of candidate proteins for PRM validation, we also performed protein identification and quantification using two alternative software programs, MaxQuant and OpenMS. The settings used for each software program are presented in Additional file 1.

#### PRM verification

To verify the DEPs identified from the discovery study, a targeted proteomic study was performed using PRM. Selection of suitable proteins for inclusion in the PRM study was performed through evaluation of the DEPs after quantification using PD, MaxQuant and OpenMS. In addition, proteins successfully quantified in only dCCA samples or controls respectively were considered. All proteins were submitted to a literature review, and proteins previously found to be dysregulated in CCA or with a known association to cancer biology were prioritized for inclusion. Peptide selection was based on the data from the PD evaluation of the discovery study. Only peptides with no missing cleavages were included. Peptide properties (charge state, precursor m/z, retention time) were extrapolated from PD.

Materials from the 20 dCCA samples used in the discovery study and 13 controls were prepared anew (additional material was available from 8 control samples analyzed in the discovery study, and 5 new samples were added). Sample preparation was performed in accordance with the same protocol utilized in the discovery study. The same LC instrumentation and settings and MS instrumentation used in the discovery study were used for the PRM study. Measurements were performed in technical triplicates when possible. For samples with a low amount of material, one or two technical replicates were deemed acceptable. For the measurements, 1  $\mu\text{g}$  of sample was injected into the instrument. PRM acquisition was performed without retention time scheduling over the complete chromatographic run. The MS<sup>2</sup> resolution was set at 70,000, with the AGC target set to  $5 \times 10^5$  and a maximum injection time of 70 ms. The chromatographic peak width was 30 s. The NCE was set to 26.0% and the isolation window was 2.0 m/z. MS1 ion chromatogram extraction and relative quantification were performed using Skyline [25].

## IHC

The entire dCCA cohort (N=59), including paired lymph node metastases in available samples (N=26) and control samples (N=10), was selected for IHC validation. FFPE samples were sectioned to 4  $\mu\text{m}$ , deparaffinized in xylene and rehydrated in graded ethanol solutions. Antigen retrieval was performed using EnVision FLEX Target Retrieval Solution (low pH) (K8005, Dako, Glostrup, Denmark) at 97 °C for 20 min using an automated PT Link (Dako, Glostrup, Denmark). Endogenous peroxidase activity was quenched with 0.3% hydrogen peroxide and 1% methanol in phosphate-buffered saline for 10 min at room temperature. After blocking with 5% goat serum and avidin/biotin blocking kit (SP-2001, Vector Labs, Burlingame, CA, USA) the sections were incubated with primary polyclonal rabbit anti-THBS2 antibody diluted 1:100 (Ab112543, Abcam, Cambridge, UK) overnight at 4 °C. A biotinylated goat anti-rabbit secondary antibody diluted 1:200 (BA-1000, Vector Labs, Burlingame, CA, USA) was added to the slides and incubated for 1 h at room temperature. The sections were subsequently incubated with avidin–biotin–peroxidase complex Vectastain Elite ABC (PK-6100, Vector Laboratories, Burlingame, CA, USA) for 30 min. For color development the slides were exposed to the chromogen 3,3'-diaminobenzidine (SK-4100, Vector Laboratories, Burlingame, CA, USA) for 8 min. Following counterstaining with Mayer's hematoxylin, the sections were dehydrated with graded ethanol solutions and mounted with Pertex. Omission of primary antibody was used as a negative control. Placental tissue sections were included as positive controls.

IHC staining was evaluated by an experienced gastrointestinal pathologist (A.S). The predominant staining pattern in the intratumoral area was evaluated.

Staining was reviewed under light microscopy at 200 $\times$  magnification. Given that both epithelial and stromal immunoreactivity was evident, separate scoring was performed for both compartments. Expression in >10% of cells was recorded as positive expression. The staining intensity was scored as 0 (negative), 1+ (low), 2+ (moderate) or 3+ (strong).

## Statistical analysis and bioinformatics

Statistical processing of the MS data from the discovery and PRM study was done using Perseus [26] and GraphPad Prism 8. In the discovery study, the mean value of technical replicates was used. The data was filtered, only proteins quantified in at least 50% of the dCCA and control samples were used for further analysis. Values were log<sub>2</sub> transformed and normalized by subtracting the median protein intensities in each sample. Missing values were replaced by imputing random numbers drawn from a normal distribution similar to the measured data (width=0.3, downshift=0). DEPs between the dCCA samples and controls were identified using a two-tailed t-test followed by multiple testing permutation-based FDR. Finally, proteins with a q value (adjusted p-value) < 0.05 (FDR 5%) and fold change (FC)  $\geq 2$  or  $\leq 0.5$  were considered significantly differentially expressed.

For statistical evaluation of the PRM data, the mean values of technical replicates were used. All data were log<sub>2</sub> transformed and normalized to the average value of peptides quantifying the housekeeping proteins glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (GALQNIIPASTGAAK, LISWYDNEFGYSNR) and tubulin beta chain (TUBB) (ISVYYNEATGGK) in each sample was performed. Group comparison of DEPs was performed identically to the discovery study. Gene ontology (GO) [27] classification was performed using the PANTHER online bioinformatics tool [28]. The PANTHER overrepresentation test of functional annotation and pathways was used. In addition the DAVID bioinformatics tool [29] was used to perform additional pathway analysis against the KEGG [30] and REACTOME [31] databases. All enrichment analysis were performed against the background of the total number of proteins identified in the study. FDR adjustment for multiple testing was employed, and a q < 0.05 (FDR 5%) was used to indicate significance in all enrichment tests. Protein–protein interactions were investigated using the STRING online bioinformatics tool [32]. An interaction score  $\geq 0.7$  (high confidence) was required for inclusion in the model.

Stata MP statistical package version 14.2 was used for analysis of IHC data. Comparison of continuous baseline



parameters was performed with the t-test or Mann–Whitney U test as appropriate. Categorical data were compared using Fisher's exact test. The exact McNemar's test was used to compare expression in primary tumors and paired lymph node metastases. Spearman correlation was used to compare expression in dCCA epithelial and stromal compartments. The survival rates were estimated using the Kaplan–Meier method and log-rank test. Multivariable Cox proportional hazards regression was performed to adjust for confounders. The multivariable model included age, sex, tumor differentiation, R1 resection, lymph node metastases and adjuvant treatment as covariables. The assumption of proportional hazards was evaluated using Schoenfeld residuals. All tests were two-tailed and a p-value of 0.05 was used to indicate significance.

## Results

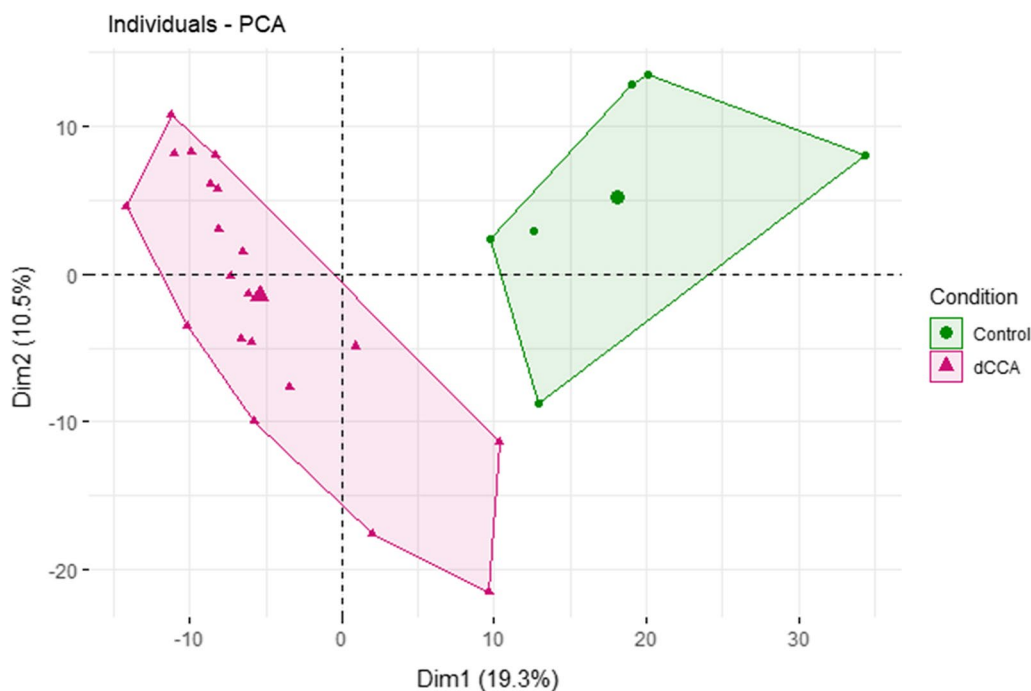
### Discovery study

A workflow including discovery LC–MS/MS followed by PRM verification using macrodissected archived FFPE samples was used to identify DEPs between dCCA and controls. For the discovery study materials from 20 dCCA samples and 10 controls were prepared for analysis. Four control samples could not be measured due to a low protein yield. The clinicopathological data of the patients whose samples were analyzed are presented in

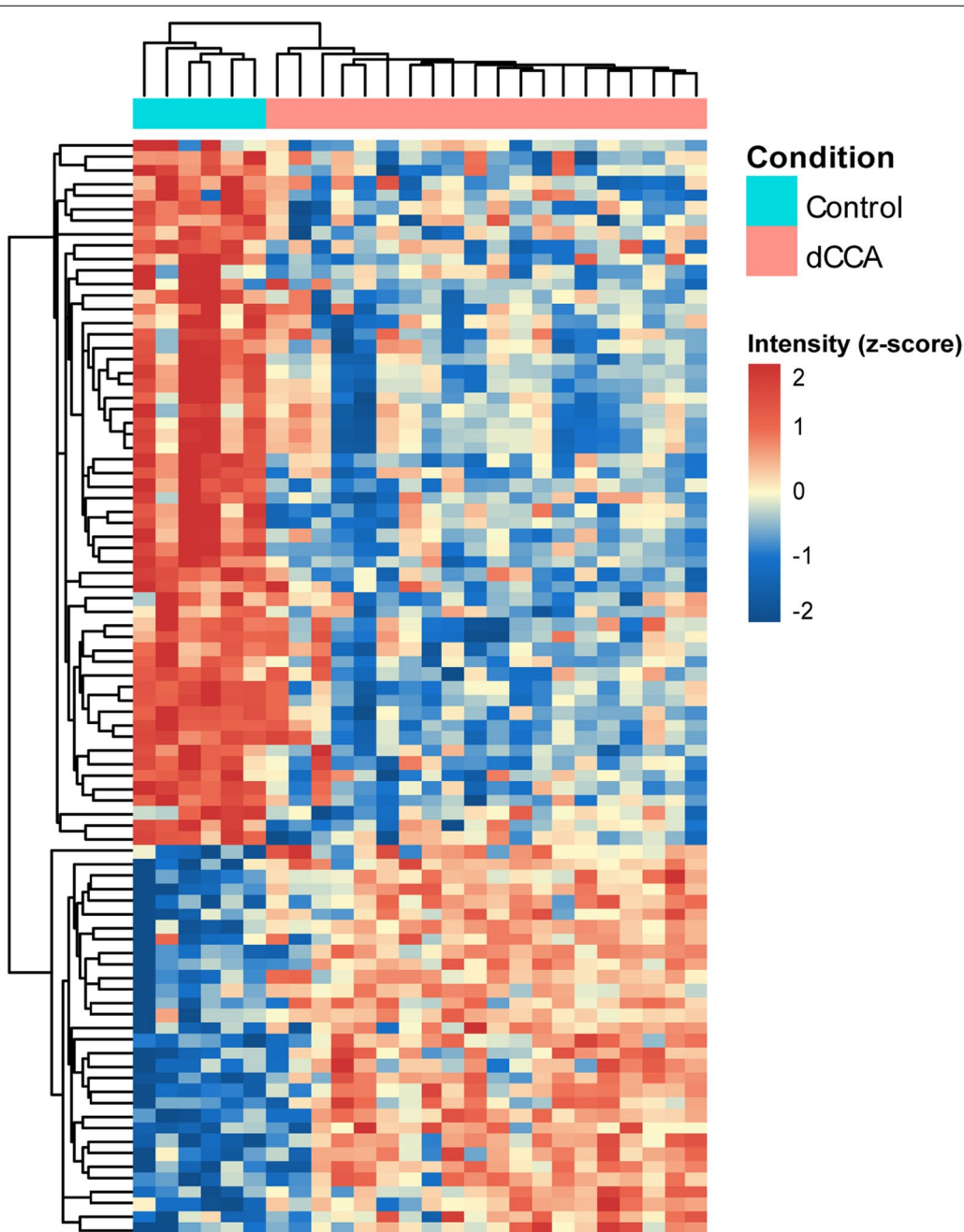
Additional file 2. In total, 3037 proteins were identified (Additional file 3). In the dCCA samples, 2967 proteins were identified, and in the controls, 1501 proteins were identified. After removal of all proteins with quantifiable values in fewer than 50% of the samples in each group, a total of 836 proteins remained. Principal component analysis (PCA) of the 836 proteins revealed the evident separation of the groups. Components 1 and 2 accounted for 19.1% and 9.9% of total variation, respectively (Fig. 1). A total of 87 DEPs ( $q < 0.05$ ,  $FC \geq 2$  or  $\leq 0.5$ ) were identified (Additional file 4). Of the 87 DEPs, 31 proteins were found to be upregulated in the dCCA samples relative to controls and 56 were downregulated. In addition to quantification using PD, quantification with MaxQuant and OpenMS generated 109 and 135 DEPs, respectively (Additional file 4).

### Bioinformatic analysis

Two-way unsupervised hierarchical clustering was applied to the 87 DEPs, and the results were visualized in a heat-map (Fig. 2). Two clusters of proteins that clearly separated dCCA and control samples were evident. GO analysis of the DEPs revealed the molecular functions in which the DEPs were most frequently involved to be binding, catalytic activity and molecular function regulation (Fig. 3a). The cellular components in which the DEPs were most frequently involved were cell, extracellular



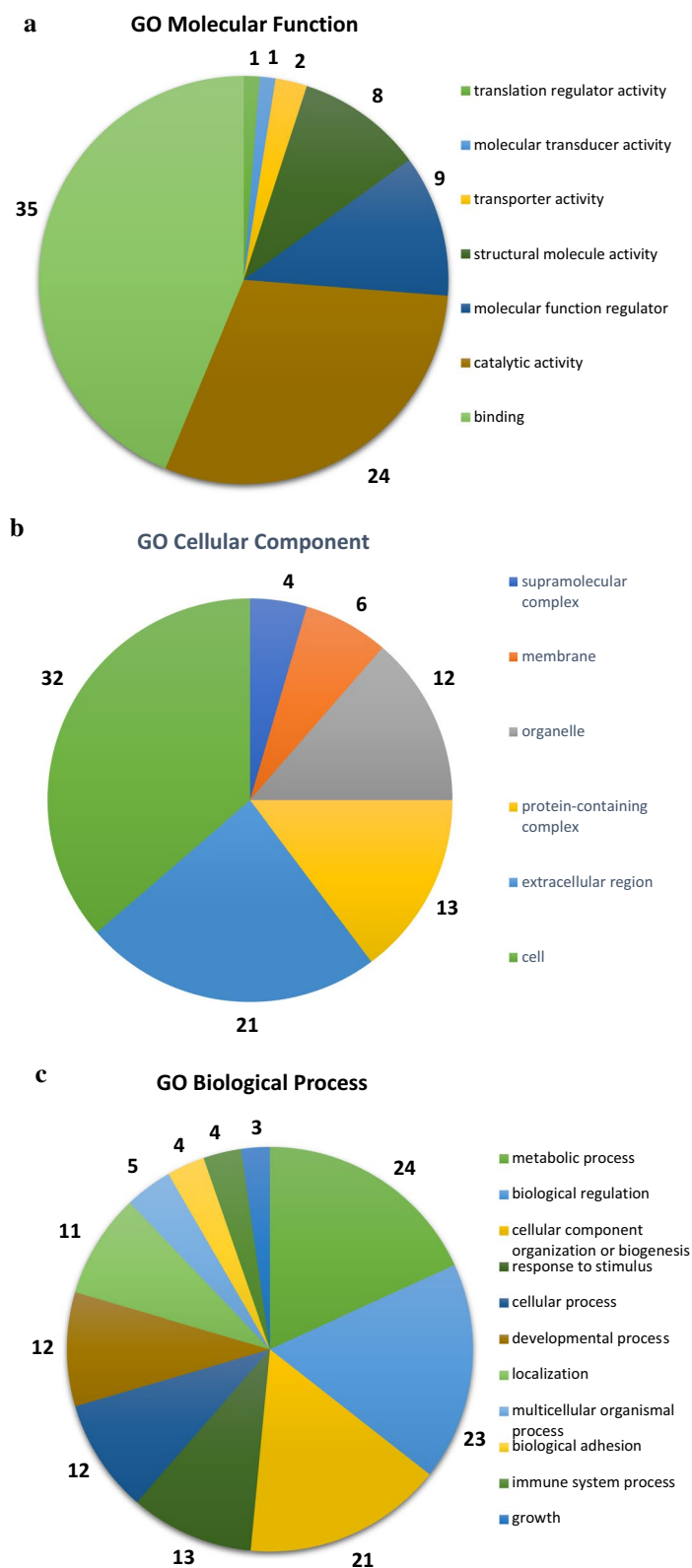
**Fig. 1** Principal component analysis of the 836 proteins that were quantified in at least 50% of distal cholangiocarcinoma samples (pink) and controls (green)



**Fig. 2** Heatmap showing the results of two-way unsupervised hierarchical clustering of the 87 proteins found to be differentially expressed ( $q < 0.05$ ,  $FC \geq 2$  or  $\leq 0.5$ ) between the distal cholangiocarcinoma samples (pink) and controls (teal)

region and protein-containing complex (Fig. 3b). The biological processes in which the DEPs were most frequently involved were metabolic process, biological regulation and biological adhesion (Fig. 3c). PANTHER-overrepresentation test was performed on the 87 DEPs. No statistically significant enrichment with regards to molecular function or biological process was seen. When cellular components were analyzed, collagen trimer, extracellular

space, extracellular matrix (ECM) and its lineage parent extracellular region and extracellular region part were significantly enriched in the DEPs. PANTHER pathway analysis revealed that the integrin signaling pathway was significantly enriched in the DEPs. KEGG pathway analysis revealed that protein digestion and absorption and ECM-receptor interaction were significantly enriched in the DEPs. When pathway enrichment was analyzed with



**Fig. 3** Gene ontology classification of the 87 differentially expressed proteins ( $q < 0.05$ ,  $FC \geq 2$  or  $\leq 0.5$ ) between the distal cholangiocarcinoma samples and controls using the PANTHER database, with the number of genes presented. **a** Molecular function. **b** Cellular component. **c** Biological process



REACTOME, ECM proteoglycans, integrin cell surface interactions, collagen degradation, assembly of collagen fibrils and other multimeric structures, collagen biosynthesis and modifying enzymes, neural cell adhesion molecule 1 interactions, ECM organization, signaling by platelet-derived growth factor and scavenging by class A receptors were found to be significantly enriched in the DEPs. Detailed results of the enrichment analysis are presented in Additional file 5. The STRING-database was employed to identify protein–protein interactions within the DEPs. In total, 46 protein interactions were found, and these interactions were significantly enriched based on the given nodes ( $P < 0.001$ ). A complex pattern of protein–protein interactions was found, and clusters of ECM proteins, blood components and cytoskeleton-associated proteins could be discerned (Fig. 4).

#### Targeted verification

For PRM verification, in total 33 patient samples were selected (20 dCCA samples and 13 controls). Minute number of samples, many with low protein yields allowed for 25 samples to comply with the PRM analysis criteria, comprising 16 dCCA and 9 control samples.

Clinicopathological data of the samples analyzed by PRM are presented in Additional file 2. The median coefficient of variation (CV)% of technical replicates for each protein ranged from 3 to 23%. A total of 170 peptides mapping to 91 proteins, including the housekeeping

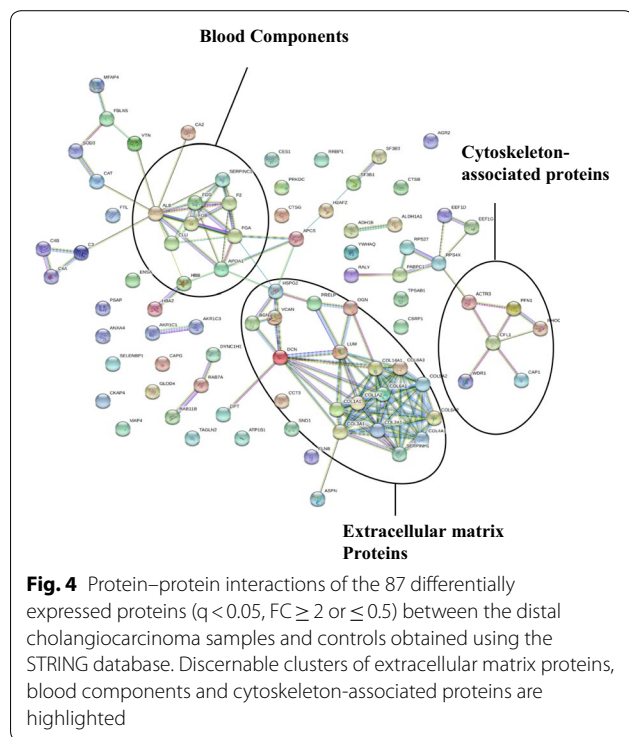
proteins GAPDH and TUBB, were included in the PRM spectral library (Additional file 6). A total of 122 peptides mapping to 79 proteins were successfully quantified (Additional file 7).

In total, 65 different peptides were found to be differentially expressed ( $q < 0.05$ ,  $FC \geq 2$  or  $\leq 0.5$ ). These peptides mapped to 46 proteins, of which 19 were identified as differentially expressed by 2 different peptides, and 27 proteins were identified as differentially expressed based on 1 peptide. Thirty-nine peptides mapping to 28 proteins were found to be upregulated, and 26 peptides mapping to 18 proteins were found to be downregulated when dCCA samples were compared to controls (Table 2). Previous associations between the identified proteins and CCA are presented in Table 2. Hierarchical clustering revealed a good separation between dCCA and controls based on the DEPs from the PRM verification (Fig. 5).

#### THBS2 expression in dCCA

The most upregulated protein verified by PRM was THBS2 (Fig. 6), which had not previously been studied in CCA; THBS2 was selected for further validation using IHC. THBS2 showed a membranous/cytoplasmic staining pattern in epithelial cells and stromal fibroblasts. In the normal bile ducts epithelial THBS2 immunoreactivity was focally positive but predominantly negative. Weak focal stromal expression was also detectable but predominantly negative. Occasional intratumoral lymphocytic infiltration with positive THBS2 expression was noted but not evaluated further. Among the dCCA samples, epithelial immunoreactivity was absent in 6 samples (10%) and scored as 1+ in 30 samples (51%), 2+ in 13 samples (32%) and 3+ in 4 samples (7%). In dCCA stromal compartment, THBS2 expression was absent in 5 samples (8%) and scored as 1+ in 18 samples (31%), 2+ in 25 samples (42%) and 3+ in 11 samples (19%). A weak correlation between epithelial and stromal THBS2 immunoreactivity was observed ( $r_s = 0.32$ ;  $P = 0.013$ ). The proportion of samples positive for epithelial THBS2 decreased from 25 (96%) to 14 (54%) when primary tumors and paired lymph node metastasis samples were compared ( $P = 0.001$ ). The proportion of samples positive for stromal THBS2 expression decreased from 25 (96%) to 18 (72%) when primary tumors and paired lymph node metastasis samples were compared ( $P = 0.031$ ). Representative examples of IHC-staining are presented in Fig. 7.

The distribution of clinicopathological variables between dCCA samples in the epithelial and stromal compartments with positive and negative THBS2 expression is presented in Additional file 8. Epithelial THBS2 positivity was associated with blood vessel invasion ( $P = 0.004$ ), and stromal THBS2 positivity was associated with R1 resection



**Table 2 Proteins found to be differentially expressed in the PRM analysis ( $q < 0.05$ ,  $FC \geq 2$  or  $\leq 0.5$ ) between the distal cholangiocarcinoma samples and controls**

#	Accession	Gene	Protein Name	Peptide Sequence	Values Cancer	Values Control	q-value	FC Ca/Co	Literature CCA
A									
1	P35442	TSP2	Thrombospondin-2	LVFNPDQEDLDGDGR	16	9	<0.0001	29.3	
2	P35442	TSP2	Thrombospondin-2	FDYIPPVNADDLSK	16	6	<0.0001	8.3	
3	Q6UX06	OLFM4	Olfactomedin-4	LLEYR	16	8	<0.0001	7.8	[71, 72]
4	P25815	S100P	Protein S100-P	YSGSEGSTQTLTK	16	9	<0.0001	7.3	[63, 64]
5	P40199	CEAM6	Carcinoembryonic antigen-related cell adhesion molecule 6	IGYSWYK	16	9	0.0007	7.2	[65]
6	P25815	S100P	Protein S100-P	ELPGFLQSGK	16	8	<0.0001	6.4	[63, 64]
7	Q04695	K1C17	Keratin, type I cytoskeletal 17	ASLEGNLAETENR	16	9	0.0044	5.5	[66, 67]
8	Q99439	CNN2	Calponin-2	GLQSGVDIGVK	16	8	<0.0001	4.8	
9	Q96CG8	CTHR1	Collagen triple helix repeat-containing protein 1	VLFSGSLR	16	9	<0.0001	4.5	
10	P08238	HS90B	Heat shock protein HSP 90-beta	NPDDITQEEYGEFYK	16	9	<0.0001	4.2	[70]
11	P19971	TYPH	Thymidine phosphorylase	MLAAQGVDPGLAR	16	9	<0.0001	4.1	[68, 69]
12	P31949	S10AB	Protein S100-A11	DGYNYTLTK	16	9	<0.0001	3.8	[96]
13	P19971	TYPH	Thymidine phosphorylase	VAAALDDGSALGR	16	9	<0.0001	3.6	[68, 69]
14	P19827	ITIH1	Inter-alpha-trypsin inhibitor heavy chain H1	AAISGENAGLVR	16	9	0.0030	3.5	
15	Q96HE7	ERO1A	ERO1-like protein alpha	LGAVDESLSEETQK	16	9	<0.0001	3.0	
16	Q01518	CAP1	Adenylyl cyclase-associated protein 1	VENQENVSNLVIETELK	16	9	<0.0001	2.9	
17	Q9UBR2	CATZ	Cathepsin Z	NVDGVNYASITR	16	9	0.0004	2.7	
18	P50454	SERPH	Serpin H1	AVLSAEQLR	16	9	<0.0001	2.7	[97]
19	P21291	CSRP1	Cysteine and glycine-rich protein 1	GYGYGQGAGTLSTDK	16	9	0.0018	2.7	
20	Q9UBR2	CATZ	Cathepsin Z	NSWGEPWGER	16	9	0.0007	2.6	
21	P02792	FRIL	Ferritin light chain	ALFQDIK	16	9	0.0002	2.6	
22	Q01518	CAP1	Adenylyl cyclase-associated protein 1	LSDLLAPISEIQIK	16	9	<0.0001	2.5	
23	P42224	STAT1	Signal transducer and activator of transcription 1-alpha/beta	TELISVSEVHPSR	14	4	0.0033	2.5	[98]
24	Q9NZM1	MYOF	Myoferlin	ANVTVLDTQIR	16	6	0.0020	2.4	
25	Q96CG8	CTHR1	Collagen triple helix repeat-containing protein 1	IIIEELPK	16	6	0.0020	2.4	
26	P00338	LDHA	L-lactate dehydrogenase A chain	SADTLWGIQK	16	9	<0.0001	2.4	[99, 100]
27	P00338	LDHA	L-lactate dehydrogenase A chain	VLTSEEER	16	9	<0.0001	2.3	[99, 100]
28	O75369	FLNB	Filamin-B	FNDEHIPESPYLVPVIAPSD-DAR	16	9	0.0005	2.3	
29	P43490	NAMPT	Nicotinamide phosphoribosyltransferase	STQAPLIIRPDGSGNPLDVLK	15	8	<0.0001	2.2	
30	Q9NZM1	MYOF	Myoferlin	GPVGTVSESQLAR	16	9	<0.0001	2.2	
31	P40121	CAPG	Macrophage-capping protein	ANAQAAALYK	16	9	0.0006	2.2	[87]
32	P38606	VATA	V-type proton ATPase catalytic subunit A	TVISQSLSK	16	9	<0.0001	2.2	

**Table 2 (continued)**

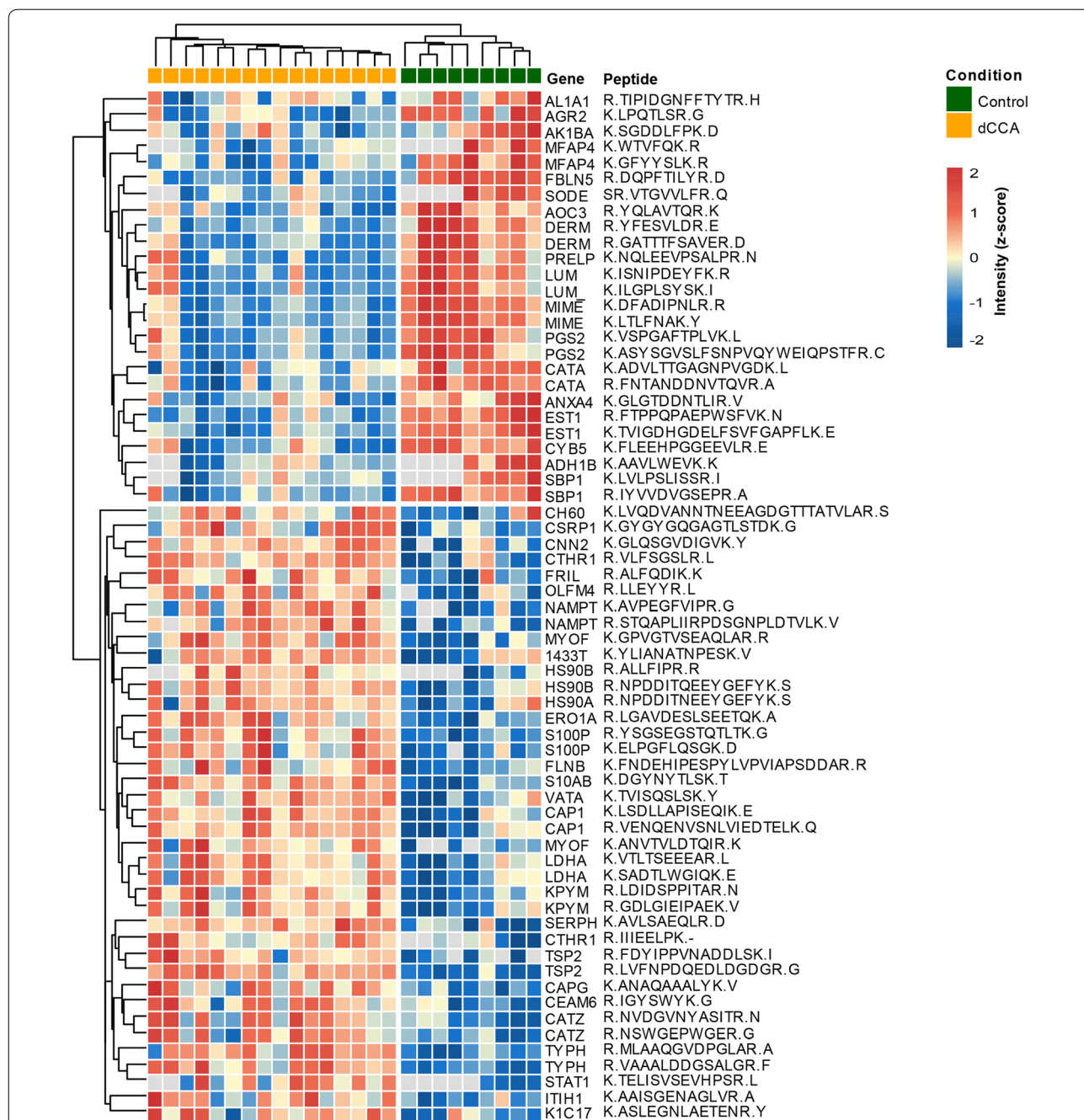
#	Accession	Gene	Protein Name	Peptide Sequence	Values Cancer	Values Control	q-value	FC Ca/Co	Literature CCA
33	P14618	KPYM	Pyruvate kinase PKM	LDIDSPPITAR	16	9	<0.0001	2.1	[101]
34	P08238	HS90B	Heat shock protein HSP 90-beta	ALLFIPR	14	5	0.0021	2.1	[70]
35	P27348	1433 T	14–3–3 protein theta	YLIANATNPESK	16	9	0.0078	2.1	[102, 103]
36	P07900	HS90A	Heat shock protein HSP 90-alpha	NPDDITNEEYGEFYK	16	9	0.0092	2.1	[70]
37	P43490	NAMPT	Nicotinamide phosphoribosyltransferase	AVPEGFVIPR	16	7	<0.0001	2.1	
38	P14618	KPYM	Pyruvate kinase PKM	GDLGIEIPAEK	16	9	<0.0001	2.1	[101]
39	P10809	CH60	60 kDa heat shock protein, mitochondrial	LVQDVANNTNEEAGDGTT TATVLAR	16	9	0.0408	2.0	[104]
<b>B</b>									
1	P00352	AL1A1	Retinal dehydrogenase 1	TIPIDGNFFTYTR	16	9	0.0457	0.46	[105]
2	P04040	CATA	Catalase	ADVLTTGAGNPVGDK	16	9	<0.0001	0.34	[106]
3	P04040	CATA	Catalase	FNTANDDNVTQVR	16	9	<0.0001	0.32	[106]
4	P09525	ANXA4	Annexin A4	GLGTDDNTLIR	16	9	0.0005	0.32	[83]
5	Q16853	AOC3	Membrane primary amine oxidase	YQLAVTQR	16	9	<0.0001	0.29	
6	O95994	AGR2	Anterior gradientprotein 2 homolog	LPQTLRSR	16	9	<0.0001	0.29	[107]
7	O60218	AK1BA	Aldo–keto reductase family 1-member B10	SGDDLFPK	16	9	0.0152	0.27	[83]
8	Q13228	SBP1	Methanethiol oxidase	IYVVDVGSEPR	16	9	<0.0001	0.26	[83]
9	P00167	CYB5	Cytochrome b5	FLEEHPGGEEVLR	16	9	<0.0001	0.25	[83]
10	Q13228	SBP1	Methanethiol oxidase	LVLPSLISSR	14	5	0.0004	0.24	[83]
11	P51884	LUM	Lumican	ISNIPDEYFK	16	9	<0.0001	0.18	
12	Q9UBX5	FBLN5	Fibulin-5	DQPFTILYR	16	9	<0.0001	0.18	[52]
13	Q07507	DERM	Dermatopontin	YFESVLDR	16	9	<0.0001	0.18	[55]
14	P51884	LUM	Lumican	ILGPLSYSK	16	9	<0.0001	0.16	
15	P55083	MFAP4	Microfibril-associated glycoprotein 4	GFYYSLK	16	9	<0.0001	0.16	[59]
16	P07585	PGS2	Decorin	VSPGAFTPLVK	16	9	<0.0001	0.16	[49]
17	P51888	PRELP	Prolargin	NQLEEVPSALPR	16	9	<0.0001	0.15	
18	Q07507	DERM	Dermatopontin	GATTTFSAVER	16	9	<0.0001	0.14	[55]
19	P55083	MFAP4	Microfibril-associated glycoprotein 4	WTVFQK	14	5	<0.0001	0.12	[59]
20	P00325	ADH1B	Alcohol dehydrogenase 1B	AAVLWEVK	14	5	<0.0001	0.12	
21	P07585	PGS2	Decorin	ASYSGVSLFSNPVQY-WEIQPSTFR	16	9	<0.0001	0.11	[49]
22	P08294	SODE	Extracellular superoxide dismutase	VTGVVLFRR	14	5	<0.0001	0.11	
23	P23141	EST1	Liver carboxylesterase 1	FPPQPAEPWSFVK	16	9	<0.0001	0.08	[52, 83]
24	P20774	MIME	Mimecan	LTLFNAK	16	9	<0.0001	0.07	[52]
25	P23141	EST1	Liver carboxylesterase 1	TVIGDHGDELFSVFGAPFLK	16	9	<0.0001	0.06	[52, 83]
26	P20774	MIME	Mimecan	DFADIPNLR	16	9	<0.0001	0.05	[52]

Proteins are ordered by descending FC. (A) upregulated proteins. (B) downregulated proteins

Ca Cancer, CCA cholangiocarcinoma, Co Control, FC Fold Change

( $P=0.011$ ). There was no association between epithelial THBS2 and survival. Kaplan–Meier analysis revealed a trend towards lower DFS ( $P=0.105$ ) and OS ( $P=0.079$ ) in samples positive for THBS2 in the stromal compartment

(Fig. 8). When confounding variables were adjusted with multivariable Cox proportional hazards regression (Additional file 9), positive stromal THBS2 expression was significantly associated with poor DFS (HR 3.95, 95% CI



**Fig. 5** Heatmap showing the results of two-way unsupervised hierarchical clustering of the 46 proteins found to be differentially expressed by PRM verification ( $q < 0.05$ ,  $FC \geq 2$  or  $\leq 0.5$ ) between the distal cholangiocarcinoma samples (yellow) and controls (green) respectively

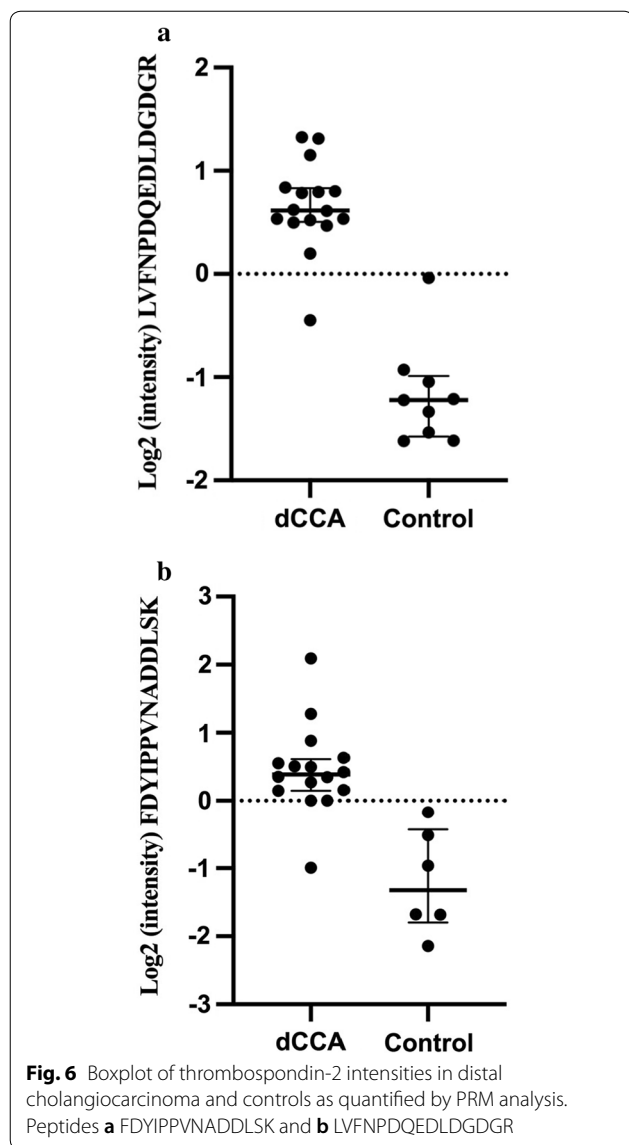
1.09–14.3;  $P = 0.037$ ) and tended to be associated with OS (HR 3.34, 95% CI 0.94–11.8;  $P = 0.062$ ).

**Discussion**

In this study, we identified DEPs between resected dCCA and normal bile duct samples using macro-dissected archived FFPE tissues. A workflow using

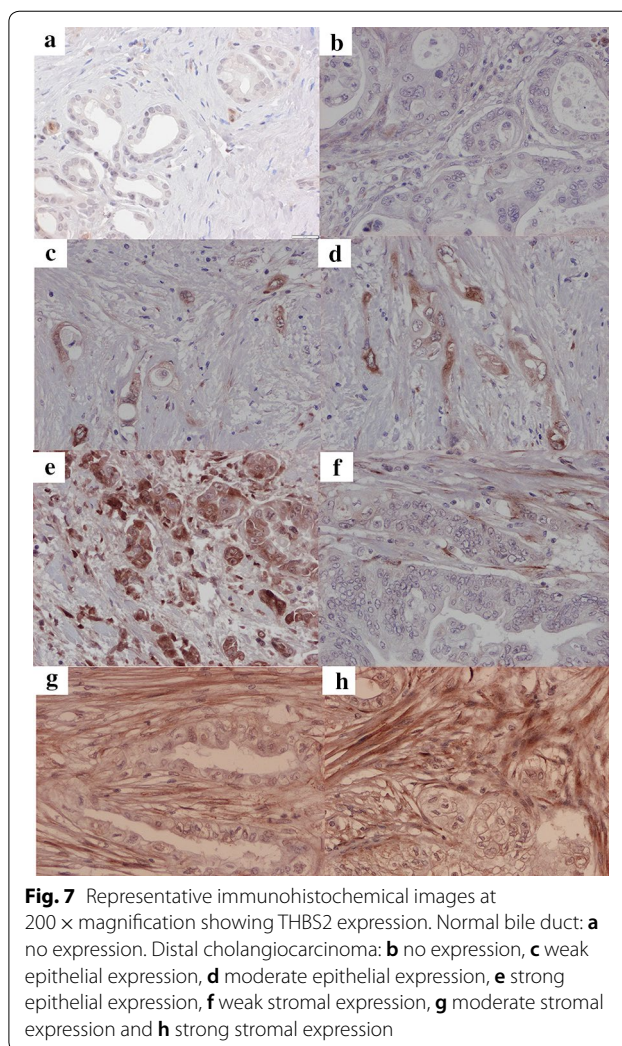
discovery LC–MS/MS followed by PRM verification of selected candidates was performed. In total, 46 proteins were successfully verified. Bioinformatic analysis highlighted alterations to the tumor reactive stroma (TRS) in dCCA. THBS2 was identified as upregulated in dCCA using MS and further validated to be upregulated in dCCA epithelium and stroma using IHC. There





was a significantly lower proportion of epithelial and stromal THBS2 in paired lymph node metastases compared to primary tumors. However, stromal THBS2 expression in lymph node metastases was frequent (72%). Stromal THBS2 was significantly associated with poor DFS (HR 3.95, 95% CI 1.09–14.3;  $P = 0.037$ ) when adjusted for confounding variables.

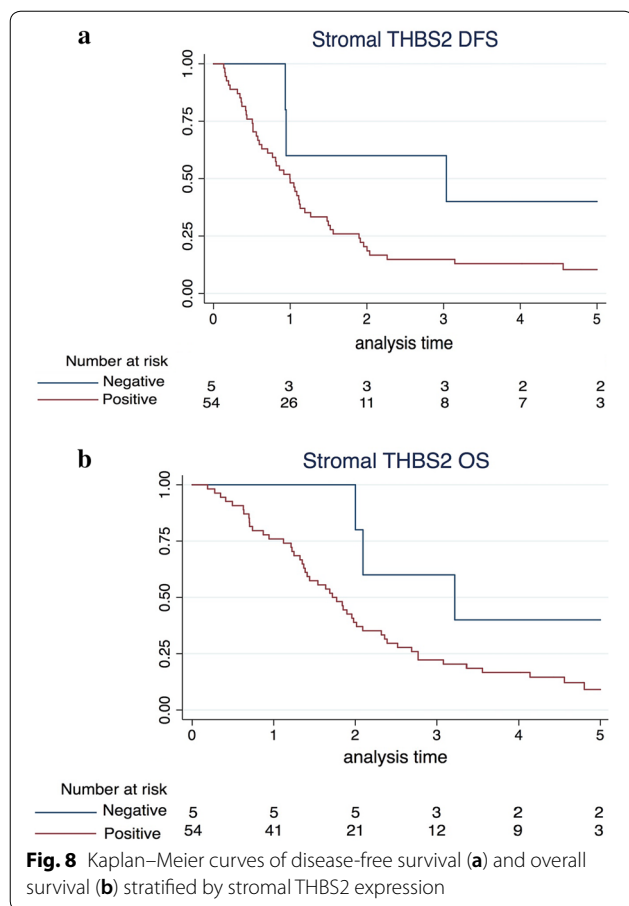
One of the characteristics of CCA tumor biology is the generation of a rich TRS [33]. The TRS consists of different cell types such as cancer-associated fibroblasts (CAFs) and tumor-associated macrophages; aberrant lymphatic vasculature; and a remodeled ECM, often with a dense desmoplastic reaction. Signaling interactions between cancer cells and the TRS play an important role in CCA carcinogenesis and treatment



resistance [34]. GO analysis of DEPs from the discovery study revealed their enrichment in several TRS components. In addition, the majority of enriched pathways have been implicated in cancer-TRS interactions. This finding highlights the importance of the proteomic alterations to the TRS that occurs in dCCA.

Notably, using IHC, we found that the overexpression of THBS2 in dCCA identified by MS, could be due to both stromal and epithelial THBS2 overexpression. THBS2 is a primarily extracellular protein and has been shown to impact cellular functions such as angiogenesis [35], apoptosis [36], cytoskeletal organization and ECM remodeling [37]. THBS2 has been found to have both pro- and antitumorigenic functions and an association with both good and poor prognosis has been found in different cancers [38, 39]. In pancreatic cancer, THBS2 expression in cancer cells in vitro inhibited invasiveness through downregulation of matrix metalloproteinase-9 and urokinase-type plasminogen activator [40].





In contrast, stromal THBS2 expression in pancreatic stellate cells was found to promote cancer cell invasiveness in coculture experiments [41]. Discrepancies between the regulation and prognostic implications of epithelial and stromal THBS2 have been described [42, 43]. Recently, Kim et al. [44] demonstrated that plasma THBS2 is a promising diagnostic biomarker for pancreatic cancer. Given the biological similarities of dCCA and pancreatic cancer [45], further evaluation of THBS2 as a potential diagnostic biomarker of dCCA should be encouraged. We have not found any previous study in which the expression of THBS2 in paired lymph node metastases was investigated. Both epithelial THBS2 and stromal THBS2 were significantly less frequent in the lymph node metastases than in primary tumors. The loss of epithelial THBS2 during the dCCA metastatic process could be due to a tumor-suppressive function. Although significantly lower, stromal THBS2 expression was frequently retained in lymph node metastases (72%). The maintenance of THBS2 stromal cells through the metastatic process suggests a function of the THBS2 stroma in dCCA metastatic development. A large amount of

stroma in lymph node metastases has previously been associated with poor survival in various cancers [46].

We note that among the most substantially downregulated proteins according to PRM verification were several proteins with a known TRS association. Proteins belonging to the small leucine-rich proteoglycan family such as decorin, lumican, prolargin, dermatopontin and mimecan were all substantially downregulated. Small leucine-rich proteoglycans are known to be involved in cellular proliferation, differentiation, survival, adhesion, migration, the inflammatory response and cancer development [47]. Decorin has a tumor-suppressive effect and has been extensively studied as a therapeutic target in epithelial cancers [48]. Furthermore, decorin was previously shown to be downregulated and associated with negative prognosis in eCCA. In addition, decorin treatment could reverse CCA proliferation, invasion and migration in vitro [49]. Mimecan is downregulated in several solid cancers and has a tumor-suppressive function through its interaction with epidermal growth factor receptor signaling and tumoral immune infiltration [50, 51]. Mimecan was also found to be downregulated in CCA [52]. Dermopontin has been shown to be involved in transforming growth factor- $\beta$  signaling, collagen fibrillogenesis, cell adhesion, and cell proliferation [53]. Additionally, dermatopontin has a tumor-suppressive effect [54] and was found to be downregulated in gallbladder cancer [55]. Prolargin has not been extensively studied in a cancer context however, low prolargin expression has been associated with poor survival in pancreatic cancer [56]. Lumican is known to modulate collagen fibrillogenesis and integrin signaling and has been shown to have a tumor-suppressive effect in several cancers [57]. To the best of our knowledge, neither prolargin nor lumican has been associated with dCCA. Other substantially downregulated proteins include microfibril-associated glycoprotein 4, fibulin-5 and extracellular superoxide dismutase. Microfibril-associated glycoprotein 4 is known to be downregulated in the majority of human cancers with prognostic significance and was predicted to impact cell proliferation and elastic fiber formation [58]. Microfibril-associated glycoprotein 4 was found to be downregulated in CAFs isolated from iCCA [59]. Fibulin-5 is downregulated in several human cancers and has a tumor-suppressive function mediated through its interaction with matrix metalloproteinases [60]. Furthermore, fibulin-5 was found to be downregulated in CCA [52]. Extracellular superoxide dismutase, an important enzyme in the antioxidant defense system, is frequently downregulated in cancer and has a tumor-suppressive role [61]; however, it has not been described in CCA biology. In addition, catalase another essential member of the antioxidant defense system was found to be downregulated [62].

When looking at the most upregulated proteins identified in our study, several proteins, such as protein S100-P [63, 64], carcinoembryonic antigen-related cell adhesion molecule 6 [65], keratin, type I cytoskeletal 17 [66, 67], thymidine phosphorylase [68, 69], heat shock protein 90 alpha/beta [70] and olfactomedin-4 [71, 72], are well established as upregulated proteins and tentative biomarkers in CCA. The correct identification and quantification of known biomarkers gives us confidence in the biological relevance of the proteins identified with our study design. Other substantially upregulated proteins without previous association with CCA include calponin-2, collagen triple helix repeat-containing protein 1 and ero1-like protein alpha. Calponin-2 is an actin cytoskeleton-associated regulatory protein that can inhibit cell proliferation and migration. Calponin-2 has not been extensively studied in cancer, however, it was found to be a prognostic factor and to have tumor-suppressive effects in pancreatic and prostate cancers [73, 74]. Collagen triple helix repeat-containing protein 1 is involved in the regulation of cell motility through the regulation of collagen deposition. It has an oncogenic effect and has been shown to be upregulated and associated with poor prognosis in several cancers [75]. Ero1-like protein alpha, which is important for disulfide bond formation in secreted molecules, is upregulated in several cancer forms and has been shown to play an important role in tumor-mediated immunosuppression [76].

Several previous studies have employed various MS-based proteomics approaches to identify biomarkers of cholangiocarcinoma through analysis of resected human tissues [72, 77–88]. The majority of these previous studies have investigated iCCA or a mixture of different CCA subtypes. A study by Maeda et al. [72] was the first study to analyze an eCCA-only cohort. To the best of our knowledge, our study represents the first proteomic characterization of a dCCA-only cohort.

Some limitations to the present study are noted. A small number of samples were analyzed. All material was retrospectively acquired and stored for various time periods prior to analysis. However, archival storage time of FFPE tissues has not been found to negatively impact MS based proteomics [89–91]. For the MS analysis, tumor and control samples were unmatched with regards to sex due to the limited number of tissues available for use as controls. Although we cannot exclude some bias as a result of this imbalance, the effect is likely minor. We choose to use normal tissue controls as compared to the more widely used and widely available morphologically normal tumor adjacent tissues since comparative studies have found molecular aberrations in normal tumor adjacent tissues compared to normal tissues [92]. This is hypothesized to be caused by field cancerization or

microenvironment alterations influenced by the tumor [93]. Thus, normal controls can help identify additional biomarkers compared to normal tumor adjacent tissues [92]. In the PRM verification, a housekeeping normalization was used. Since there is currently no consensus on the most suitable housekeeping proteins in cholangiocarcinoma, we chose to use GAPDH and TUBB as housekeeping proteins since they are commonly used for this purpose and because suitable peptides were available in the discovery study data. However, expression of classic housekeeping genes was found to be less stable than assumed, especially in cancer tissues, which thus can be a source of bias [94]. IHC validation was also performed with a relatively small number of samples, particularly with regard to paired lymph node metastases. The survival analysis was underpowered for the detection of anything other than a large prognostic effect, especially as few cases were negative for THBS2. Notably, however, the effect size of the HR for stromal THBS2 was large.

## Conclusions

An MS-based workflow was used to identify and verify several proteins without a previous association with dCCA biology. The identified proteins can be further investigated to elucidate their function and potential as biomarkers in dCCA. THBS2 was validated as frequently expressed in the epithelium and stroma of dCCA. Stromal THBS2 is a potential prognostic marker; additionally, it was frequently retained in paired lymph node metastases. The use of stromal THBS2 as a prognostic marker in dCCA should be validated using separate larger cohorts, additionally the potential of THBS2 as a diagnostic biomarker in dCCA should be evaluated.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12967-020-02498-3>.

**Additional file 1.** Protein identification and quantification using MaxQuant and OpenMS.

**Additional file 2.** Clinicopathological data of analyzed samples.

**Additional file 3.** Total proteins identified by discovery LC-MS/MS.

**Additional file 4.** DEPs identified by discovery LC-MS/MS using PD, MaxQuant and OpenMS.

**Additional file 5.** Detailed enrichment analysis.

**Additional file 6.** Peptides included in the PRM spectral library.

**Additional file 7.** Peptides quantified by PRM.

**Additional file 8.** Distribution of clinicopathological variables in THBS2 positive and negative samples.

**Additional file 9.** Uni- and multivariable Cox proportional hazards regressions.

## Abbreviations

CCA: Cholangiocarcinoma; iCCA: Intrahepatic cholangiocarcinoma; eCCA: Extrahepatic cholangiocarcinoma; pCCA: Perihilar cholangiocarcinoma; dCCA: Distal cholangiocarcinoma; CA 19-9: Carbohydrate antigen 19-9; MS: Mass spectrometry; LC-MS/MS: Liquid phase separation on-line with tandem mass spectrometry; FFPE: Formalin-fixed paraffin-embedded; PRM: Parallel reaction monitoring; DEPs: Differentially expressed proteins; THBS2: Thrombospondin-2; IHC: Immunohistochemistry; AJCC: American Joint Committee on Cancer; R1-resection: Non-radical resection; DFS: Disease-free survival; OS: Overall survival; H&E: Hematoxylin and eosin; AMBIC: Ammonium bicarbonate; LC: Liquid chromatography; DDA: Data dependent acquisition; AGC: Automatic gain control; NCE: Normalized collision energy; PD: Proteome Discoverer; FDR: False discovery rate; GAPDH: Glyceraldehyde 3-phosphate dehydrogenase; TUBB: Tubulin beta chain; GO: Gene ontology; PCA: Principal component analysis; ECM: Extracellular matrix; CV: Coefficient of variation; TRS: Tumor reactive stroma; CAFs: Cancer-associated fibroblasts.

## Acknowledgements

The authors would like to thank Melinda Rezeli and Katarzyna Said Hilmersson for technical assistance.

## Authors' contributions

JB, RA, TK, GMV, BA conceived and designed the study. JB performed review of medical records. AgS performed clinicopathological evaluation, sample area selection and IHC-scoring. JB, TK performed sample preparation for MS. TK performed MS measurements and skyline evaluation of PRM data. JB, TK, IP, AnS performed statistical and bioinformatic analysis of MS data. JB performed statistical evaluation of IHC-data. JB wrote the manuscript. TK, DH, IP, AnS, AgS, RA, GMV, BA performed critical revision of the manuscript. All authors read and approved the final manuscript.

## Funding

This research work was supported by government grant for clinical research (<https://www.skane.se/fou/alf>), Erik and Angelica Sparres research foundation, and the Swedish Cancer Society (<https://www.cancerfonden.se/om-oss/about>). Open access funding provided by Lund University.

## Availability of data and materials

The mass spectrometry proteomics data has been deposited to the ProteomeXchange Consortium via the PRIDE [95] partner repository with the dataset identifier PXD017906.

## Ethics approval and consent to participate

The study was approved by the Regional Human Ethics Committee in Lund Sweden (2015/392).

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Department of Clinical Sciences Lund, Surgery, Lund University, and Skåne University Hospital, Lund, Sweden. <sup>2</sup> Department of Biomedical Engineering, Clinical Protein Science and Imaging, Lund University, Lund, Sweden. <sup>3</sup> Department of Clinical Sciences Lund, Pathology, Lund University, and Skåne University Hospital, Lund, Sweden.

Received: 18 April 2020 Accepted: 21 August 2020

Published online: 04 September 2020

## References

- Bergquist A, von Seth E. Epidemiology of cholangiocarcinoma. *Best Pract Res Clin Gastroenterol*. 2015;29(2):221–32.
- Mosconi S, Beretta GD, Labianca R, Zampino MG, Gatta G, Heinemann V. Cholangiocarcinoma. *Crit Rev Oncol Hematol*. 2009;69(3):259–70.
- Banales JM, Cardinale V, Carpino G, Marzioni M, Andersen JB, Invernizzi P, et al. Expert consensus document: cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). *Nat Rev Gastroenterol Hepatol*. 2016;13(5):261–80.
- Razumilava N, Gores GJ. Classification, diagnosis, and management of cholangiocarcinoma. *Clin Gastroenterol Hepatol*. 2013;11(1):13–21.e1; quiz e3–4.
- Byrd DR, Carducci MA, Compton CC, et al. *AJCC cancer staging manual*. 7th ed. New York: Springer Verlag; 2010.
- Rizvi S, Khan SA, Hallemeier CL, Kelley RK, Gores GJ. Cholangiocarcinoma—evolving concepts and therapeutic strategies. *Nat Rev Clin Oncol*. 2018;15(2):95–111.
- Valle JW, Lamarca A, Goyal L, Barriuso J, Zhu AX. New horizons for precision medicine in biliary tract cancers. *Cancer Discov*. 2017;7(9):943–62.
- Byrling J, Andersson B, Marko-Varga G, Andersson R. Cholangiocarcinoma—current classification and challenges towards personalised medicine. *Scand J Gastroenterol*. 2016;51(6):641–3.
- Razumilava N, Gores GJ. Cholangiocarcinoma. *Lancet* (London, England). 2014;383(9935):2168–79.
- Schwanhauser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473(7347):337–42.
- Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol*. 2011;7:548.
- Ginestier C, Charafe-Jauffret E, Bertucci F, Eisinger F, Geneix J, Bechlian D, et al. Distinct and complementary information provided by use of tissue and DNA microarrays in the study of breast tumor markers. *Am J Pathol*. 2002;161(4):1223–33.
- Crutchfield CA, Thomas SN, Sokoll LJ, Chan DW. Advances in mass spectrometry-based clinical biomarker discovery. *Clin Proteomics*. 2016;13:1.
- Parker CE, Borchers CH. Mass spectrometry based biomarker discovery, verification, and validation—quality assurance and control of protein biomarker assays. *Mol Oncol*. 2014;8(4):840–58.
- Nilsson T, Mann M, Aebersold R, Yates JR, Bairoch A, Bergeron JJM. Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Meth*. 2010;7(9):681–5.
- Gustafsson OJ, Arentz G, Hoffmann P. Proteomic developments in the analysis of formalin-fixed tissue. *Biochem Biophys Acta*. 2015;1854(6):559–80.
- Steiner C, Ducret A, Tille J-C, Thomas M, McKee TA, Rubbia-Brandt LA, et al. Applications of mass spectrometry for quantitative protein analysis in formalin-fixed paraffin-embedded tissues. *Proteomics*. 2014;14(4–5):441–51.
- Polanski M, Anderson NL. A list of candidate cancer biomarkers for targeted proteomics. *Biomark Insights*. 2007;1:1–48.
- Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, Domon B. Targeted proteomic quantification on quadrupole-orbitrap mass spectrometer. *Mol Cell Proteomics*. 2012;11(12):1709–23.
- Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics*. 2012;11(11):1475–88.
- Bourmaud A, Gallien S, Domon B. Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: principle and applications. *Proteomics*. 2016;16(15–16):2146–59.
- Rauniyar N. Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry. *Int J Mol Sci*. 2015;16(12):28566–81.
- Byrling J, Sasor A, Nilsson J, Said Hilmersson K, Andersson R, Andersson B. Expression of peritumoral SPARC during distal cholangiocarcinoma progression and correlation with outcome. *Scand J Gastroenterol*. 2020;55:725–31.
- Byrling J, Sasor A, Nilsson J, Said Hilmersson K, Andersson R, Andersson B. Expression of fibroblast activation protein and the clinicopathological relevance in distal cholangiocarcinoma. *Scand J Gastroenterol*. 2020;55(1):82–9.
- Schilling B, Rardin MJ, MacLean BX, Zawadzka AM, Frewen BE, Cusack MP, et al. Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol Cell Proteomics*. 2012;11(5):202–14.

26. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016;13(9):731–40.
27. Gene Ontology C. Gene Ontology Consortium: going forward. *Nucleic Acids Res*. 2015;43(Database issue):D1049–D10561056.
28. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, et al. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. 2017;45(D1):D183–D189189.
29. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
30. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–D462462.
31. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The reactome pathway knowledgebase. *Nucleic Acids Res*. 2018;46(D1):D649–D655655.
32. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45(D1):D362–D368368.
33. Sirica AE, Gores GJ. Desmoplastic stroma and cholangiocarcinoma: clinical implications and therapeutic targeting. *Hepatology* (Baltimore, MD). 2014;59(6):2397–402.
34. Cadamuro M, Morton SD, Strazzabosco M, Fabris L. Unveiling the role of tumor reactive stroma in cholangiocarcinoma: an opportunity for new therapeutic strategies. *Transl Gastrointest Cancer*. 2013;2(3):130–44.
35. de Fraipont F, Nicholson AC, Feige JJ, Van Meir EG. Thrombospondins and tumor angiogenesis. *Trends Mol Med*. 2001;7(9):401–7.
36. Mirochnik Y, Kwiatek A, Volpert OV. Thrombospondin and apoptosis: molecular mechanisms and use for design of complementation treatments. *Curr Drug Targets*. 2008;9(10):851–62.
37. Calabro NE, Kristofik NJ, Kyriakides TR. Thrombospondin-2 and extracellular matrix assembly. *Biochem Biophys Acta*. 2014;1840(8):2396–402.
38. Kazerounian S, Yee KO, Lawler J. Thrombospondins in cancer. *Cell Mol Sci*. 2008;65(5):700–12.
39. Weng TY, Wang CY, Hung YH, Chen WC, Chen YL, Lai MD. Differential expression pattern of THBS1 and THBS2 in lung cancer: clinical outcome and a systematic-analysis of microarray databases. *PLoS ONE*. 2016;11(8):e0161007.
40. Nakamura M, Oida Y, Abe Y, Yamazaki H, Mukai M, Matsuyama M, et al. Thrombospondin-2 inhibits tumor cell invasion through the modulation of MMP-9 and uPA in pancreatic cancer cells. *Mol Med Rep*. 2008;1(3):423–7.
41. Farrow B, Berger DH, Rowley D. Tumor-derived pancreatic stellate cells promote pancreatic cancer cell invasion through release of thrombospondin-2. *J Surg Res*. 2009;156(1):155–60.
42. Matos AR, Coutinho-Camillo CM, Thuler LC, Fonseca FP, Soares FA, Silva EA, et al. Expression analysis of thrombospondin 2 in prostate cancer and benign prostatic hyperplasia. *Exp Mol Pathol*. 2013;94(3):438–44.
43. Chijiwa T, Abe Y, Inoue Y, Matsumoto H, Kawai K, Matsuyama M, et al. Cancerous, but not stromal, thrombospondin-2 contributes prognosis in pulmonary adenocarcinoma. *Oncol Rep*. 2009;22(2):279–83.
44. Kim J, Bamlet WR, Oberg AL, Chaffee KG, Donahue G, Cao X-J, et al. Detection of early pancreatic ductal adenocarcinoma with thrombospondin-2 and CA19-9 blood markers. *Sci Transl Med*. 2017;9(398):eaah5583.
45. Schmuck RB, de Carvalho-Fischer CV, Neumann C, Pratschke J, Bahra M. Distal bile duct carcinomas and pancreatic ductal adenocarcinomas: postulating a common tumor entity. *Cancer Med*. 2015;5(1):88–99.
46. Mesker WE, van Pelt GW, Tollenaar RAEM. Tumor stroma as contributing factor in the lymph node metastases process? *Oncotarget*. 2019;10(9):922–3.
47. Iozzo RV, Schaefer L. Proteoglycans in health and disease: novel regulatory signaling mechanisms evoked by the small leucine-rich proteoglycans. *FEBS J*. 2010;277(19):3864–75.
48. Sainio AO, Järveläinen HT. Decorin-mediated oncosuppression - a potential future adjuvant therapy for human epithelial cancers. *Br J Pharmacol*. 2019;176(1):5–15.
49. Yu X, Zou Y, Li Q, Mao Y, Zhu H, Huang G, et al. Decorin-mediated inhibition of cholangiocarcinoma cell growth and migration and promotion of apoptosis are associated with E-cadherin in vitro. *Tumor Biol*. 2014;35(4):3103–12.
50. Hu X, Li Y-Q, Li Q-G, Ma Y-L, Peng J-J, Cai S-J. Osteoglycin-induced VEGF inhibition enhances T lymphocytes infiltrating in colorectal cancer. *EBioMedicine*. 2018;34:35–45.
51. Hu X, Li Y-Q, Li Q-G, Ma Y-L, Peng J-J, Cai S-J. Osteoglycin (OGN) reverses epithelial to mesenchymal transition and invasiveness in colorectal cancer via EGFR/Akt pathway. *J Exp Clin Cancer Res*. 2018;37(1):41.
52. Miller G, Socci ND, Dhall D, D'Angelica M, DeMatteo RP, Allen PJ, et al. Genome wide analysis and clinical correlation of chromosomal and transcriptional mutations in cancers of the biliary tract. *J Exp Clin Cancer Res*. 2009;28(1):62.
53. Okamoto O, Fujiwara S. Dermato-pontin, a novel player in the biology of the extracellular matrix. *Connect Tissue Res*. 2006;47(4):177–89.
54. Yamatoji M, Kasamatsu A, Kouzu Y, Koike H, Sakamoto Y, Ogawara K, et al. Dermato-pontin: a potential predictor for metastasis of human oral cancer. *Int J Cancer*. 2012;130(12):2903–11.
55. Gu X, Li B, Jiang M, Fang M, Ji J, Wang A, et al. RNA sequencing reveals differentially expressed genes as potential diagnostic and prognostic indicators of gallbladder carcinoma. *Oncotarget*. 2015;6(24):20661–71.
56. Chen R, Dawson DW, Pan S, Ottenhof NA, de Wilde RF, Wolfgang CL, et al. Proteins associated with pancreatic cancer survival in patients with resectable pancreatic ductal adenocarcinoma. *Lab Invest J Tech Methods Pathol*. 2015;95(1):43–55.
57. Karamanou K, Franchi M, Piperigkou Z, Perreau C, Maquart FX, Vynios DH, et al. Lumican effectively regulates the estrogen receptors-associated functional properties of breast cancer cells, expression of matrix effectors and epithelial-to-mesenchymal transition. *Sci Rep*. 2017;7:45138.
58. Yang J, Song H, Chen L, Cao K, Zhang Y, Li Y, et al. Integrated analysis of microfibrillar-associated proteins reveals MFAP4 as a novel biomarker in human cancers. *Epigenomics*. 2019;11(1):1635–51.
59. Utispan K, Thuwajit P, Abiko Y, Charnkaew K, Paupairoj A, Chau-in S, et al. Gene expression profiling of cholangiocarcinoma-derived fibroblast reveals alterations related to tumor progression and indicates periostin as a poor prognostic marker. *Mol Cancer*. 2010;9:13.
60. Manders DB, Kishore HA, Gazdar AF, Keller PW, Tsunozumi J, Yanagisawa H, et al. Dysregulation of fibulin-5 and matrix metalloproteinases in epithelial ovarian cancer. *Oncotarget*. 2018;9(18):14251–67.
61. Griess B, Tom E, Domann F, Teoh-Fitzgerald M. Extracellular superoxide dismutase and its role in cancer. *Free Radical Biol Med*. 2017;112:464–79.
62. Glorieux C, Calderon PB. Catalase, a remarkable enzyme: targeting the oldest antioxidant enzyme to find a new cancer treatment approach. *Biol Chem*. 2017;398(10):1095–108.
63. Nakanuma Y, Uchida T, Sato Y, Uesaka A. An S100P-positive biliary epithelial field is a preinvasive intraepithelial neoplasm in nodular-sclerosing cholangiocarcinoma. *Hum Pathol*. 2017;60:46–57.
64. Sato Y, Harada K, Sasaki M, Nakanuma Y. Clinicopathological significance of S100 protein expression in cholangiocarcinoma. *J Gastroenterol Hepatol*. 2013;28(8):1422–9.
65. Rose JB, Correa-Gallego C, Li Y, Nelson J, Alseidi A, Helton WS, et al. The role of biliary carcinoembryonic antigen-related cellular adhesion molecule 6 (CEACAM6) as a biomarker in cholangiocarcinoma. *PLoS ONE*. 2016;11(3):e0150195-e.
66. Fernández Moro C, Fernández-Woodbridge A, Alistair D'souza M, Zhang Q, Bozoky B, Kandaswamy SV, et al. Immunohistochemical typing of adenocarcinomas of the pancreatobiliary system improves diagnosis and prognostic stratification. *PLoS ONE*. 2016;11(11):e0166067-e.
67. Stroescu C, Herlea V, Dragnea A, Popescu I. The diagnostic value of cytokeratins and carcinoembryonic antigen immunostaining in differentiating hepatocellular carcinomas from intrahepatic cholangiocarcinomas. *J Gastrointest Liver Dis*. 2006;15(1):9–14.
68. Won HS, Lee MA, Chung E-S, Kim D-G, You YK, Hong TH, et al. Comparison of thymidine phosphorylase expression and prognostic factors in gallbladder and bile duct cancer. *BMC Cancer*. 2010;10:564.
69. Thanasai J, Limpai-boon T, Jearanaikoon P, Sripa B, Pairojkul C, Tantimavanich S, et al. Effects of thymidine phosphorylase on tumor aggressiveness and

- 5-fluorouracil sensitivity in cholangiocarcinoma. *World J Gastroenterol*. 2010;16(13):1631–8.
70. Shirota T, Ojima H, Hiraoka N, Shimada K, Rokutan H, Arai Y, et al. Heat shock protein 90 is a potential therapeutic target in cholangiocarcinoma. *Mol Cancer Ther*. 2015;14(9):1985–93.
  71. Lukic N, Visentin R, Delhaye M, Frossard JL, Lescuyer P, Dumonceau JM, et al. An integrated approach for comparative proteomic analysis of human bile reveals overexpressed cancer-associated proteins in malignant biliary stenosis. *Biochem Biophys Acta*. 2014;1844(5):1026–33.
  72. Maeda S, Morikawa T, Takadate T, Suzuki T, Minowa T, Hanagata N, et al. Mass spectrometry-based proteomic analysis of formalin-fixed paraffin-embedded extrahepatic cholangiocarcinoma. *J Hepatobiliary Pancreat Sci*. 2015;22:683–91.
  73. Qiu Z, Chu Y, Xu B, Wang Q, Jiang M, Li X, et al. Increased expression of calponin 2 is a positive prognostic factor in pancreatic ductal adenocarcinoma. *Oncotarget*. 2017;8(34):56428–42.
  74. Moazzem Hossain M, Wang X, Bergan RC, Jin JP. Diminished expression of h2-calponin in prostate cancer cells promotes cell proliferation, migration and the dependence of cell adhesion on substrate stiffness. *FEBS Open Biol*. 2014;4:627–36.
  75. Wu Q, Yang Q, Sun H. Role of collagen triple helix repeat containing-1 in tumor and inflammatory diseases. *J Cancer Res Ther*. 2017;13(4):621–4.
  76. Tanaka T, Kutomi G, Kajiwara T, Kukita K, Kochin V, Kanaseki T, et al. Cancer-associated oxidoreductase ERO1- $\alpha$  promotes immune escape through up-regulation of PD-L1 in human breast cancer. *Oncotarget*. 2017;8(15):24706–18.
  77. Darby IA, Vuillier-Devillers K, Pinault E, Sarraza V, Lepreux S, Balabaud C, et al. Proteomic analysis of differentially expressed proteins in peripheral cholangiocarcinoma. *Cancer Microenviron*. 2011;4(1):73–91.
  78. Kawase H, Fujii K, Miyamoto M, Kubota KC, Hirano S, Kondo S, et al. Differential LC-MS-based proteomics of surgical human cholangiocarcinoma tissues. *J Proteome Res*. 2009;8(8):4092–103.
  79. Padden J, Megger DA, Bracht T, Reis H, Ahrens M, Kohl M, et al. Identification of novel biomarker candidates for the immunohistochemical diagnosis of cholangiocellular carcinoma. *Mol Cell Proteomics*. 2014;13(10):2661–722.
  80. Le Faouder J, Laouirem S, Alexandrov T, Ben-Harzellah S, Leger T, Albuquerque M, et al. Tumoral heterogeneity of hepatic cholangiocarcinomas revealed by MALDI imaging mass spectrometry. *Proteomics*. 2014;14(7–8):965–72.
  81. Shi Y, Deng X, Zhan Q, Shen B, Jin X, Zhu Z, et al. A prospective proteomic-based study for identifying potential biomarkers for the diagnosis of cholangiocarcinoma. *J Gastrointest Surg*. 2013;17(9):1584–91.
  82. Matsuda A, Kuno A, Matsuzaki H, Kawamoto T, Shikanai T, Nakanuma Y, et al. Glycoproteomics-based cancer marker discovery adopting dual enrichment with Wisteria floribunda agglutinin for high specific glyco-diagnosis of cholangiocarcinoma. *J Proteomics*. 2013;85:1–11.
  83. Dos Santos A, Court M, Thiers V, Sar S, Guettier C, Samuel D, et al. Identification of cellular targets in human intrahepatic cholangiocarcinoma using laser microdissection and accurate mass and time tag proteomics. *Mol Cell Proteomics*. 2010;9(9):1991–2004.
  84. Kristiansen TZ, Harsha HC, Gronborg M, Maitra A, Pandey A. Differential membrane proteomics using 18O-labeling to identify biomarkers for cholangiocarcinoma. *J Proteome Res*. 2008;7(11):4670–7.
  85. Thanan R, Oikawa S, Yongvanit P, Hiraku Y, Ma N, Pinlaor S, et al. Inflammation-induced protein carbonylation contributes to poor prognosis for cholangiocarcinoma. *Free Radical Biol Med*. 2012;52(8):1465–72.
  86. Padden J, Ahrens M, Kalsch J, Bertram S, Megger DA, Bracht T, et al. Immunohistochemical markers distinguishing cholangiocellular carcinoma (CCC) from pancreatic ductal adenocarcinoma (PDAC) discovered by proteomic analysis of microdissected cells. *Mol Cell Proteomics*. 2016;15(3):1072–82.
  87. Morofuji N, Ojima H, Onaya H, Okusaka T, Shimada K, Sakamoto Y, et al. Macrophage-capping protein as a tissue biomarker for prediction of response to gemcitabine treatment and prognosis in cholangiocarcinoma. *J Proteomics*. 2012;75(5):1577–89.
  88. Onsurathum S, Haonon O, Pinlaor P, Pairojkul C, Khuntikeo N, Thanan R, et al. Proteomics detection of S100A6 in tumor tissue interstitial fluid and evaluation of its potential as a biomarker of cholangiocarcinoma. *Tumour Biol*. 2018;40(4):101042831876195.
  89. Craven RA, Cairns DA, Zougman A, Harnden P, Selby PJ, Banks RE. Proteomic analysis of formalin-fixed paraffin-embedded renal tissue samples by label-free MS: assessment of overall technical variability and the impact of block age. *Proteomics Clin Appl*. 2013;7(3–4):273–82.
  90. Balgley BM, Guo T, Zhao K, Fang X, Tavassoli FA, Lee CS. Evaluation of archival time on shotgun proteomics of formalin-fixed and paraffin-embedded tissues. *J Proteome Res*. 2009;8(2):917–25.
  91. Piehowski PD, Petyuk VA, Sontag RL, Gritsenko MA, Weitz KK, Fillmore TL, et al. Residual tissue repositories as a resource for population-based cancer proteomic studies. *Clin Proteomics*. 2018;15:26.
  92. Aran D, Camarda R, Odegaard J, Paik H, Oskotsky B, Krings G, et al. Comprehensive analysis of normal adjacent to tumor transcriptomes. *Nat Commun*. 2017;8(1):1077.
  93. Huang X, Stern DF, Zhao H. Transcriptional profiles from paired normal samples offer complementary information on cancer patient survival—evidence from TCGA pan-cancer data. *Sci Rep*. 2016;6(1):20567.
  94. Hu X, Du S, Yu J, Yang X, Yang C, Zhou D, et al. Common housekeeping proteins are upregulated in colorectal adenocarcinoma and hepatocellular carcinoma, making the total protein a better "housekeeper". *Oncotarget*. 2016;7(41):66679–88.
  95. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 2019;47(D1):D442–D45050.
  96. Zhang MX, Gan W, Jing CY, Zheng SS, Yi Y, Zhang J, et al. S100A11 promotes cell proliferation via P38/MAPK signaling pathway in intrahepatic cholangiocarcinoma. *Mol Carcinog*. 2019;58(1):19–30.
  97. Bertram S, Padden J, Kalsch J, Ahrens M, Pott L, Canbay A, et al. Novel immunohistochemical markers differentiate intrahepatic cholangiocarcinoma from benign bile duct lesions. *J Clin Pathol*. 2016;69(7):619–26.
  98. Meissl K, Macho-Maschler S, Müller M, Strobl B. The good and the bad faces of STAT1 in solid tumours. *Cytokine*. 2017;89:12–20.
  99. Thonsri U, Seubwai W, Waraasawapati S, Sawanyawisuth K, Vaeteewootacharn K, Boonmars T, et al. Overexpression of lactate dehydrogenase A in cholangiocarcinoma is correlated with poor prognosis. *Histol Histopathol*. 2017;32(5):503–10.
  100. Yu Y, Liao M, Liu R, Chen J, Feng H, Fu Z. Overexpression of lactate dehydrogenase-A in human intrahepatic cholangiocarcinoma: its implication for treatment. *World J Surg Oncol*. 2014;12:78.
  101. Huang QX, Cui JY, Ma H, Jia XM, Huang FL, Jiang LX. Screening of potential biomarkers for cholangiocarcinoma by integrated analysis of microarray data sets. *Cancer Gene Ther*. 2015;23:48.
  102. Singrang N, Kittisenachai S, Roytrakul S, Svasti J, Kangsamaksin T. NOTCH1 regulates the viability of cholangiocarcinoma cells via 14-3-3 theta. *J Cell Commun Signal*. 2019;13(2):245–54.
  103. Darby IA, Vuillier-Devillers K, Pinault E, Sarraza V, Lepreux S, Balabaud C, et al. Proteomic analysis of differentially expressed proteins in peripheral cholangiocarcinoma. *Cancer Microenviron*. 2010;4(1):73–91.
  104. Shen J, Wang W, Wu J, Feng B, Chen W, Wang M, et al. Comparative proteomic profiling of human bile reveals SSP411 as a novel biomarker of cholangiocarcinoma. *PLoS ONE*. 2012;7(10):e47476.
  105. Shuang ZY, Wu WC, Xu J, Lin G, Liu YC, Lao XM, et al. Transforming growth factor-beta1-induced epithelial-mesenchymal transition generates ALDH-positive cells with stem cell properties in cholangiocarcinoma. *Cancer Lett*. 2014;354(2):320–8.
  106. Loilome W, Kadsanit S, Muisook K, Yongvanit P, Namwat N, Techasen A, et al. Imbalanced adaptive responses associated with microsatellite instability in cholangiocarcinoma. *Oncol Lett*. 2017;13(2):639–46.
  107. Chevet E, Fessart D, Delom F, Mulot A, Vojtesek B, Hrstka R, et al. Emerging roles for the pro-oncogenic anterior gradient-2 in cancer development. *Oncogene*. 2013;32(20):2499–509.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.