J Hepatobiliary Pancreat Sci (2015) ••:••–••
DOI: 10.1002/jhbp.262

ORIGINAL ARTICLE

Mass spectrometry-based proteomic analysis of formalin-fixed paraffin-embedded extrahepatic cholangiocarcinoma

Shimpei Maeda · Takanori Morikawa · Tatsuyuki Takadate · Takashi Suzuki · Takashi Minowa · Nobutaka Hanagata · Tohru Onogawa · Fuyuhiko Motoi · Toshihide Nishimura · Michiaki Unno

Abstract
Background Extrahepatic cholangiocarcinoma is very difficult to diagnose at an early stage, and has a poor prognosis. Novel markers for diagnosis and optimal treatment selection are needed. However, there has been very limited data on the proteome profile of extrahepatic cholangiocarcinoma. This study was designed to unravel the proteome profile of this disease and to identify overexpressed proteins using mass spectrometry-based proteomic approaches.
Methods We analyzed a discovery set of formalin-fixed paraffin-embedded tissues of 14 extrahepatic cholangiocarcinomas using shotgun mass spectrometry, and compared proteome profiles with those of seven controls. Then, selected candidates were verified by quantitative analysis using scheduled selected reaction monitoring-based mass spectrometry. Furthermore, immunohistochemical staining used a validation set of 165 cases.
Results In total, 1,992 proteins were identified and 136 proteins were overexpressed. Verification of 58 selected proteins by quantitative analysis revealed 11 overexpressed proteins. Immunohistochemical validation for 10 proteins showed positive rates of S100P (84%), CEAM5 (75%), MUC5A (62%), OLFM4 (60%), OAT (42%), CAD17 (41%), FABPL (38%), AOFA (30%), K1C20 (25%) and CPSM (22%) in extrahepatic cholangiocarcinomas, which were rarely positive in controls.
Conclusions We identified 10 proteins associated with extrahepatic cholangiocarcinoma using proteomic approaches. These proteins are potential targets for future diagnostic biomarkers and therapy.

Keywords Biomarker · Extrahepatic cholangiocarcinoma · Mass spectrometry · Proteomics · Scheduled selected reaction monitoring

S. Maeda · T. Morikawa · T. Takadate · T. Onogawa · F. Motoi · M. Unno (✉)
Department of Surgery, Tohoku University Graduate School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai, Miyagi 980-8574, Japan
e-mail: m_unno@surg1.med.tohoku.ac.jp

S. Maeda
Department of Surgery, South Miyagi Medical Center, Miyagi, Japan

T. Suzuki
Department of Pathology and Histotechnology, Tohoku University Graduate School of Medicine, Sendai, Japan

T. Minowa · N. Hanagata
Nanotechnology Innovation Station, National Institute for Materials Science, Tsukuba, Japan

T. Nishimura
Department of Surgery I, Tokyo Medical University, Tokyo, Japan

Introduction

Cholangiocarcinomas account for 3% of all gastrointestinal cancers [1] and are classified according to their anatomic location as intrahepatic and extrahepatic cholangiocarcinomas (EHCC). EHCC is very difficult to diagnose at an early stage and has a poor prognosis, which has improved only marginally over the past 30 years [2]. Although complete surgical resection is the only opportunity for cure, the 5-year survival rate after complete resection is 39.1% for perihilar EHCC, and 44.0% for distal EHCC [3]. One of the factors responsible for these poor outcomes of EHCC is limitations of diagnostic modalities. Novel biomarkers for early diagnosis and optimal treatment selection are needed; however, there have been very limited data on the proteome profile of EHCC. Therefore, we made an attempt to unravel the proteome profile of EHCC to identify proteins overexpressed in EHCC compared with noncancerous bile duct tissues.

Mass spectrometry (MS) is reportedly valuable in both preclinical and clinical research [4] as well as for biomarker

discovery [5]. Shotgun proteomics is a method of identifying proteins in complex mixtures using liquid chromatography (LC) and MS to provide global proteome profiles [6]. Meanwhile, targeted proteomics based on selected reaction monitoring (SRM) is an appropriate method for accurate identification and quantitation of proteins of interest [7].

Identification of scarce biomarkers in serum remains challenging because of the complexity and wide dynamic range characterizing such samples [8]. Cancer-specific proteins exist at high concentrations in tumor tissues compared with other samples such as blood and bile juice. Thus, cancerous tissue itself is an important source for biomarker discovery. Formalin-fixed paraffin-embedded (FFPE) tissues have been extensively collected and stored in hospitals for various periods of time. Those are readily available, pathologically well-defined, and include all stages of cancer, even for rare diseases. Newly developed technology has made it possible to efficiently extract proteins from FFPE tissues, thereby allowing proteomic analysis [9]. Although successful MS analysis of FFPE tissue has recently been reported [10, 11], there are no reports of proteomic analysis from archived FFPE samples of EHCC.

We conducted a large-scale proteomic study to identify novel proteins overexpressed in EHCC using both shotgun and targeted proteomics. We further validated the candidate proteins by immunohistochemical analysis.

Materials and methods

Tissues

We retrospectively retrieved EHCC samples from patients undergoing resection between 1998 and 2008 at Tohoku University Hospital. Those given neoadjuvant therapy were excluded. Intrahepatic cholangiocarcinoma, carcinoma of the gallbladder and carcinoma of the papilla of Vater cases were also excluded. Non-cancerous bile duct tissues were obtained from pancreatic cancer patients with pancreaticoduodenectomy. In total, 186 FFPE tissues, 165 EHCCs and 21 non-cancerous bile ducts, were examined. Before analysis, hematoxylin and eosin stained sections from each sample were evaluated by a pathologist. Clinicopathologic features are shown in Table 1. According to the Union for International Cancer Control (UICC) 7th edition, the numbers of stage I, II, III and IV EHCC tissues were 33, 64, 24 and 44, respectively. For MS analyses, a discovery set of 21 samples from early EHCC (stage I, $n = 7$), advanced EHCC (stage II, III and IV, $n = 7$), and non-cancerous bile duct tissues ($n = 7$) were used. The remaining 165 samples including 151 EHCCs and 14 non-cancerous bile ducts served as a validation set. Representative slide showing the largest diameter of each carcinoma was used for both MS and immunohistochemical analyses. Cells of invasive area were dissected and immunohistochemically evaluated in the case of invasive carcinoma. The study design and composition of the discovery and validation sets are shown in Figure S1.

Table 1 Clinicopathologic features

Factor		No. patients
		EHCC ($n = 165$)
Median age (range)		67 (15–83) years
Sex	Male: female	112: 53
Location	Perihilar: distal	91: 74
T	1: 2: 3: 4	13: 68: 51: 33
N	0: 1	98: 67
M	0: 1	150: 15
Stage	I: II: III: IV	33: 64: 24: 44
Histological type	Papillary	10
	Well differentiated	28
	Moderately differentiated	108
	Poorly differentiated	18
	Adenosquamous	1

T2 includes T2a and T2b perihilar EHCC and T2 distal EHCC. Stage I includes stage I perihilar EHCC and stage IA and IB distal EHCC, the same applies to stage II, III and IV. The median age of 14 patients (4 males and 10 females) from whom non-cancerous bile ducts were obtained was 65 years (range 34–83). EHCC, extrahepatic cholangiocarcinoma.

Ethics statement

Informed consent was obtained from individual patients. This study was approved by the Tohoku University Ethics Committee, and conducted according to the principles expressed in the Declaration of Helsinki.

Laser Micro Dissection and protein extraction

Cancerous lesions and non-cancerous bile duct epithelium were identified on serial, hematoxylin and eosin-stained sections. For MS analysis, 10-μm sections were attached to DIRECTOR slides (Expression Pathology, Gaithersburg, MD, USA), de-paraffinized three times with xylene for 5 min, rehydrated with graded ethanol solutions and distilled water, and then stained with hematoxylin [12]. Stained, uncovered slides were air dried and about 30,000 cells (8 mm^2) were collected into the cap of a 0.2-mL polymerase chain reaction (PCR) tube using Leica LMD6000 (Leica Microsystems GmbH, Wetzler, Germany). 1-4 sections of carcinoma and 8-20 sections of non-cancerous bile duct were needed to obtain 30,000 cells. Peptides were extracted using a Liquid Tissue MS Protein Prep kit (Expression Pathology) [9] according to the manufacturer's instructions.

Exploratory shotgun analysis

LC-tandem mass spectrometry (MS/MS)

Peptide-mixture samples processed from FFPE tissues were used for LC-MS/MS using a Finnigan LXQ linear ion-trap mass spectrometer (Thermo Fisher, San Jose, CA, USA) [12]. A capillary reverse phase LC-MS/MS system (ZAPLOUS System; AMR, Tokyo, Japan), comprising a Paradigm MS4 (Michrom BioResources, Auburn, CA, USA), an HTC PAL autosampler (CTC Analytics, Zwingen, Switzerland) and Finnigan LXQ linear ion-trap mass spectrometer, was equipped with an ADVANCE nanospray ionization source (Michrom BioResources).

Data analysis and protein identification

Mascot software (version 2.2.03, Matrix Science, London, UK) was used for a database search against *Homo sapiens* entries in the Swiss-Prot 55.6 database (20,009 entries). Peptide and fragment mass tolerances were 2.0 Da and 1.0 Da, respectively, and trypsin specificity was applied with a maximum of two missed cleavages. Methionine oxidation and N-formylation including formyl (K), formyl (R) and formyl (N-terminus) were allowed as variable modifications. A *P*-value less than 0.05 was considered to be significant in protein identification. Reported results were obtained from triplicate LC-MS runs for each sample.

Semi-quantitative comparison using spectral counting

To compare protein expressions across all tissue samples, we used the spectral counting method. The number of peptide spectra with high confidence (Mascot ion score, $P < 0.05$) served as the spectral count value. Fold changes in the expressed proteins on a base-2 logarithmic scale were calculated using the protein ratio from spectral counting (Rsc) [13]. Relative abundances of identified proteins were also obtained by applying the normalized spectral abundance factor (NSAF) [14]. Candidate proteins differing between groups were chosen so that their Rsc would satisfy the ≥ 1 criterion, corresponding to a fold change ≥ 2, and with statistical significance at $P < 0.05$ by the *G*-test [15].

Targeted verification analysis by SRM mass spectrometry

Selected reaction monitoring-based MS analysis was conducted using a discovery set. Proteins for SRM verification were mainly selected from among candidates identified by shotgun analysis and spectral counting. Furthermore, several proteins that were not overexpressed in EHCC in our shotgun analysis but were previously reported to be potential biomarkers were added to the SRM list for verification [16–19]. All sequences were confirmed by the BLAST searches (National Center for Biotechnology Information) and compared with the Swiss-Prot human database.

An LC-MS/MS system was composed of a Paradigm MS4 (Michrom BioResources) connected to a 4000 QTRAP hybrid system (AB Sciex, Foster City, CA, USA) operating in positive ion mode [7]. For all SRM studies, quadruples were operated under conditions of unit/unit resolution, and the collision energy (CE) was determined using the equation: $CE = 0.044 \times m/z + 6$ for doubly-charged precursor ions. The scheduled SRM (sSRM) mode was used in this study, with the sSRM detection window set at 180 sec. The peptide AGFAGDDAPR (*m/z* 488.7) is a doubly-charged actin, beta (ACTB) peptide and its specific SRM transition to the singly charged fragment (*m/z* 630.3) served as the internal standard [20]. This internal standard is referred to as the in-sample internal standard (ISIS) since ACTB is a housekeeping protein [21]. Peak areas of each transition were normalized using the equation: Normalized peak area = peak area × (500,000/peak area of 488.7/630.3). The averaged values of early EHCC and advanced EHCC based on triplicate runs were each compared to those of non-cancerous bile duct tissues, and an expression difference of at least two-fold was defined as overexpression.

Immunohistochemistry

A validation set of 165 samples was used. Sections (4-μm thick) after de-paraffinization with xylene were rehydrated with a graded ethanol series and distilled water. Protein S100-P (S100P) (HPA019502, Sigma, St. Louis, MO, USA), carcinoembryonic antigen-related cell adhesion molecule 5 (CEAM5) (HPA019758, Sigma), mucin-5AC (MUC5A) (OBT1746, AbD Serotec, Oxford, UK), olfactomedin-4 (OLFM4) (ab96280, Abcam, Cambridge, MA, USA), cadherin-17 (CAD17) (HPA023616, Sigma), keratin, type I cytoskeletal 20 (K1C20) (HPA027236, Sigma), and carbamoylphosphate synthase (ammonia) (CPSM) (ab54586, Abcam) immunostaining was achieved by heating slides in an autoclave at 120°C for 5 min in citrate acid buffer (10 mM citric acid, pH 6.0). Similarly, antigen retrieval for fatty acid-binding protein, liver (FABPL) (ab82157, Abcam) was performed in a microwave oven for 15 min in a citric acid buffer. No antigen retrieval was carried out for ornithine aminotransferase (OAT) (HPA040098, Sigma) or amine oxidase [flavin-containing] A (AOFA) (NBP1-19796, Novus Biologicals, Littleton, CO, USA). The dilutions of primary antibodies were as follows: S100P, 1: 3000; CEAM5, 1: 40; MUC5A, 1: 50; OLFM4, 1: 100; CAD17, 1: 1500; K1C20, 1: 100; CPSM, 1: 100; FABPL,

1: 300; OAT, 1: 100; AOFA, 1: 200. The sections were incubated overnight at 4°C with one of the primary antibodies. After blocking of endogenous peroxidase by methanol containing 0.3% hydrogen peroxidase, labeled antigens were detected with an EnVision$^+$ kit (Dako, Glostrup, Denmark) and visualized using 3,3'-diaminobenzidine tetrahydrochloride as a chromogen. Sections were counterstained with hematoxylin. Appropriate positive and negative tissue controls were used throughout, in part with reference to the Human Protein Atlas (URL: http://www.proteinatlas.org/).

After completely reviewing all slides of immunostained sections for each sample, three of the authors (S. M., T. Morikawa and T. S.) classified cases into two groups as described previously [11, 22–27]: those in which ≥ 10% of cells were positive for S100P, CEAM5, OLFM4, OAT, CAD17, AOFA, CPSM and K1C20 constituted the positive group, while the negative group was comprised of those in which < 10% of cells were positive. For MUC5A and FABPL, those with strong and moderate immunoreactivity were categorized as the positive group, while those with absent or weak staining constituted the negative group.

Fisher's exact test was used to assess the significance of differences among staining patterns. Differences with $P < 0.05$ were considered significant. Analyses were performed with JMP software version 9.0 (SAS Institute, Cary, NC, USA).

Results

Proteome profiles identified by shotgun proteomics and semi-quantitative comparison

A discovery set of 21 samples was used to identify proteins showing different expressions in cancerous and non-cancerous tissues. We identified 1,266 proteins in early EHCC, 1,143 in advanced EHCC, and 1,095 in non-cancerous bile ducts. In total, 1,992 proteins were identified. The identified proteins were compared semi-quantitatively using spectral counting. For proteins identified in non-cancerous bile ducts and EHCC, Figure 1 shows a plot of each Rsc value against the corresponding protein (X-axis) in increasing order from left to right. A positive value indicates greater expression in EHCC than non-cancerous bile ducts. Proteins with Rsc ≥ 1 and a significant difference by *G*-test were regarded as candidates for characterization of EHCC. A total of 136 of the 1,992 proteins identified had $P < 0.05$, indicating statistically significant overexpression. Meanwhile, ACTB, used as housekeeping protein, was commonly expressed in cancerous and non-cancerous cells with minimum variation.

Quantitative verification by SRM-based targeted proteomics

Selected reaction monitoring measurements were carried out for the discovery set to verify the spectral counting results. Preliminary analysis of the project control (mixtures of equal aliquots of all patient samples) was conducted to select well detectable SRM transitions and to confirm the retention time of each peptide. Finally, an SRM assay, comprising 56 proteins (102 peptides, 400 transitions), with sufficient sensitivity was developed. The 56 proteins included 48 proteins found to be overexpressed by our shotgun analysis, seven previously reported as potential biomarkers, and ACTB (ISIS).

Selected reaction monitoring quantitative analysis revealed that 11 proteins, S100P, CEAM5, MUC5A, OLFM4, OAT, CAD17, FABPL, AOFA, K1C20, CPSM and HMCS2, were

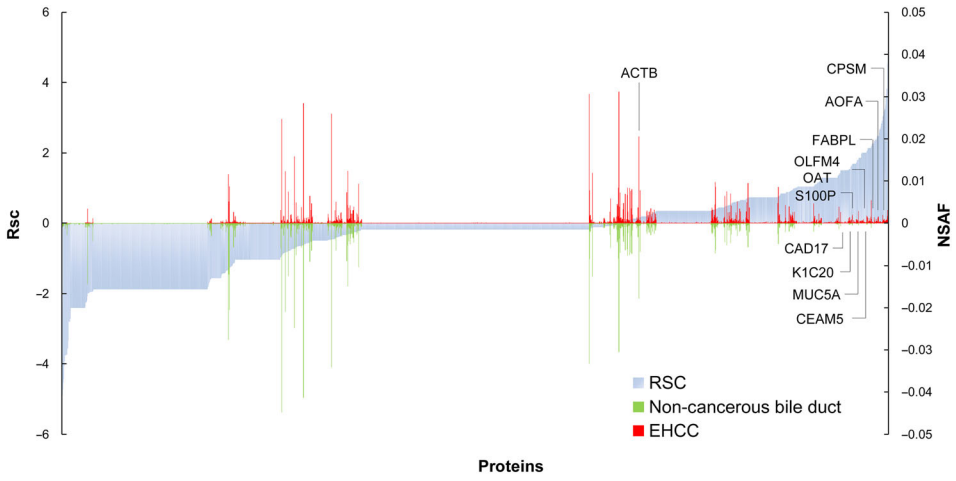

Fig. 1 Protein ratio from spectral counting (Rsc) and normalized spectral abundance factor (NSAF) values calculated for the proteins identified. Protein expressions are compared between extrahepatic cholangiocarcinoma (EHCC) and non-cancerous bile ducts. Proteins significantly overexpressed in EHCC are near the right side of the X-axis. ACTB is located near the center of the X-axis

overexpressed by at least two-fold and seemed to be potentially useful for detection of EHCC (Table S1). Figure 2 is a scatter plot for the normalized peak area. The seven potential biomarkers previously reported were not overexpressed in our SRM analysis. These results are consistent with those of spectral counting.

Validation by immunohistochemical analysis

Ten overexpressed proteins were further validated by immunohistochemical analysis using a validation set of 165 samples. HMCS2 was excluded because no appropriate antibody was commercially available. We confirmed the specific expressions of 10 proteins by immunohistochemistry (Fig. 3). Table 2 shows the positive rate for each protein in the non-cancerous bile ducts and EHCC. All non-cancerous bile duct samples were classified into the negative group in evaluation of each protein except for one sample with positive immunoreactivity for OAT. However, the positive rates of nine proteins, including S100P (84%), CEAM5 (75%), MUC5A (62%), OLFM4 (60%), OAT (42%), CAD17 (41%), FABPL (38%), AOFA (30%) and K1C20 (25%), differed significantly between non-cancerous bile ducts and EHCC ($P < 0.05$). Only in CPSM, the difference did not reach statistical significance ($P = 0.07$), although there was a tendency to have a higher positive rate (22%).

Discussion

We identified overexpressed proteins in EHCC using shotgun proteomics with spectral counting and targeted proteomics, and then validated our findings immunohistochemically. Shotgun and targeted proteomics revealed 11 proteins to be overexpressed in EHCC as compared with non-cancerous bile ducts. Ten of these 11 proteins were validated by immunohistochemistry. The significant overexpressions of nine proteins were also confirmed using a validation set, although the difference in CPSM did not reach statistical significance ($P = 0.07$).

The newly-identified markers of EHCC in our study are OLFM4, OAT, CAD17, FABPL, AOFA and CPSM. OLFM4, a member of the olfactomedin domain-containing protein family, is an anti-apoptotic factor promoting tumor growth. OLFM4 promotes pancreatic cancer cell proliferation by favoring transition from S to G2/M phase [28], and facilitates cell adhesion [29]. OLFM4 expression is related to differentiation and progression of gastric cancer [30]. OLFM4 was expressed in 60% of EHCC samples, while

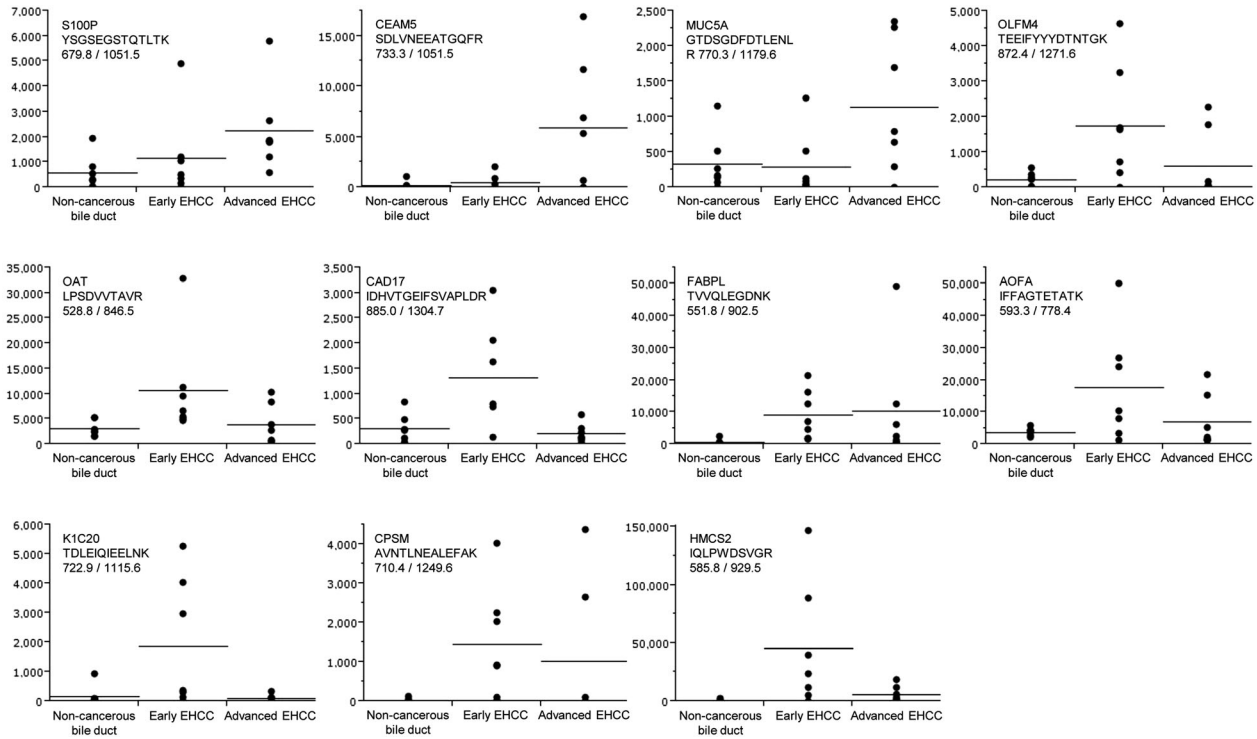

Fig. 2 Scatter plot of the normalized peak area by selected reaction monitoring (SRM)-based quantitative analysis. Proteins overexpressed by at least two-fold in early and/or advanced extrahepatic cholangiocarcinoma (EHCC) are shown

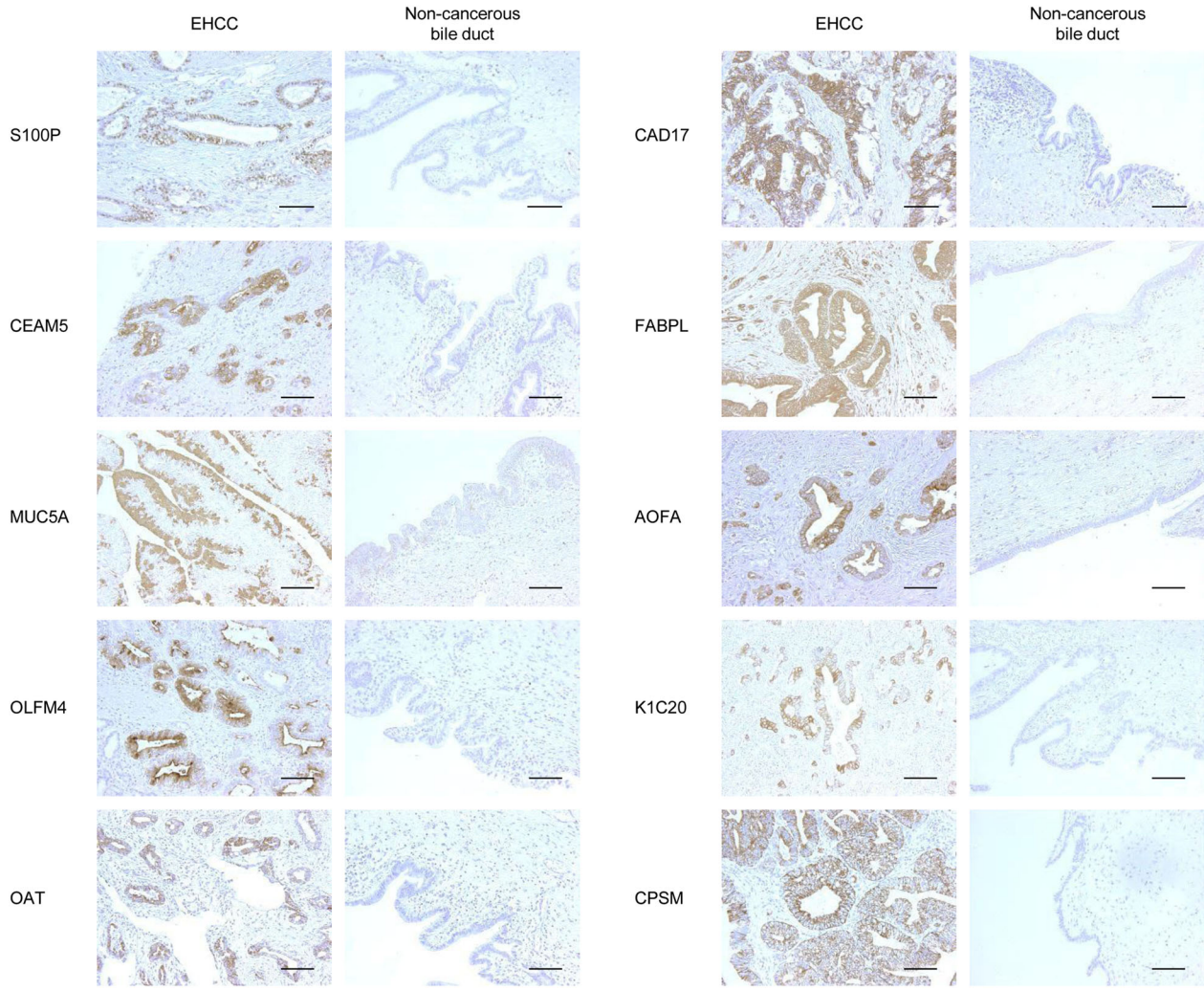

Fig. 3 Representative immunohistochemical staining results for the indicated proteins in extrahepatic cholangiocarcinoma (EHCC) and non-cancerous bile ducts. Positive staining is shown in *brown*. Scale bars represent 100 μm

Table 2 Immunohistochemical findings of the validation set

Antibody	Non-cancerous bile duct (n = 14)		EHCC (n = 151)		P-value	Stage I (n = 26)		Stage II (n = 62)		Stage III (n = 21)		Stage IV (n = 42)	
S100P	0	(0%)	127	(84%)	<0.01	21	(81%)	52	(84%)	16	(76%)	38	(90%)
CEAM5	0	(0%)	113	(75%)	<0.01	19	(73%)	47	(76%)	15	(71%)	32	(76%)
MUC5A	0	(0%)	94	(62%)	<0.01	10	(38%)	40	(65%)	13	(62%)	31	(74%)
OLFM4	0	(0%)	91	(60%)	<0.01	14	(54%)	34	(55%)	10	(48%)	33	(79%)
OAT	1	(7%)	63	(42%)	<0.01	11	(42%)	18	(29%)	11	(52%)	23	(55%)
CAD17	0	(0%)	62	(41%)	<0.01	15	(58%)	22	(35%)	7	(33%)	18	(43%)
FABPL	0	(0%)	57	(38%)	<0.01	9	(35%)	15	(24%)	10	(48%)	23	(55%)
AOFA	0	(0%)	45	(30%)	0.01	10	(38%)	15	(24%)	8	(38%)	12	(29%)
K1C20	0	(0%)	38	(25%)	0.04	6	(23%)	11	(18%)	7	(33%)	14	(33%)
CPSM	0	(0%)	33	(22%)	0.07	6	(23%)	13	(21%)	1	(5%)	13	(31%)

EHCC, extrahepatic cholangiocarcinoma.

non-cancerous bile ducts showed no immunoreactivity for this protein.

OAT is a key enzyme in the pathway converting arginine and ornithine into the major excitatory and inhibitory neurotransmitters glutamate and gamma-aminobutyric acid. Although Miyasaka et al. demonstrated OAT overexpression in hepatocellular carcinoma using the suppression subtractive hybridization technique, little is known about the role of OAT in carcinogenesis [31].

CAD17, a cadherin superfamily member, is an important cell adhesion molecule and plays major roles in organ development and maintenance of tissue integrity. A low CAD17 level is associated with a poor prognosis in gastric cancer and intrahepatic cholangiocarcinoma, although the relationship between CAD17 expression and prognosis in gastric cancer is controversial [24, 32].

FABPL is a family of small, highly conserved, cytoplasmic proteins that bind free fatty acids, and is considered to play roles in fatty acid uptake, transport and metabolism. However, precise roles of FABPL have yet to be elucidated. Several reports suggest associations of FABPL with colon and pancreatic cancers [33, 34].

AOFA is a mitochondrial enzyme that degrades amine neurotransmitters, including dopamine, norepinephrine and serotonin. Little is known about AOFA functions in cancer. AOFA is reportedly overexpressed in poorly as compared to well differentiated prostate cancer, suggesting roles in progression and aggressiveness of tumors [35].

CPSM, expressed mainly in intestinal epithelial and liver cells, is a mitochondrial enzyme catalyzing the synthesis of carbamoyl phosphate from ammonia and bicarbonate, and is important for excess urea removal from cells. Previous reports have shown CPSM overexpression in gastric cancer [36], and under-expression in human hepatocellular carcinoma [37].

S100P, CEAM5, MUC5A and K1C20 are reportedly associated with EHCC [38–41]. The fact that we identified previously reported biomarkers in an unbiased fashion suggests our workflow to be extremely useful for biomarker research. CEAM5, which is identical to clinically-used CEA, was reported to be expressed in 79% of cholangiocarcinomas by immunohistochemistry, similar to our findings (75%) [39]. Recently, Hamada et al. demonstrated S100P to be a potentially novel biomarker of EHCC, and noted that detecting S100P expression levels in brushing cytology had diagnostic value [38].

Few reports have described MS-based proteomics analysis on EHCC using tissue [18], bile [42] and serum samples [19]. Compared with those studies, the outstanding feature of this investigation is that we analyzed EHCC tissue samples only, because the etiopathogenesis of EHCC and intrahepatic cholangiocarcinoma may differ [43]. Additionally, we obtained samples using Laser Micro Dissection (LMD) technology to isolate cancerous cells of interest from abundant stroma containing inflammatory cells and fibroblasts. Kawase et al. conducted MS-based proteomic analysis using six paired cancerous, including intrahepatic cholangiocarcinoma and EHCC, and non-cancerous bile duct cases [18]. The overexpressed proteins in their study, e.g. actinin-1 and actinin-4, were not confirmed by SRM analysis in our samples. Factors for this different result may include homogeneity of samples and LMD use, which presumably contribute to more precise analysis of protein expression.

Discoveries of novel biomarkers are based on identification of proteins with expressions that differ between disease and control samples. Global shotgun proteomics has advantages in terms of the number of proteins identified. With this approach, the proteomes of complex mixtures can be analyzed in a completely unbiased fashion with broad proteome coverage, thereby increasing chances to discover novel biomarkers. Then a rational approach to their prioritization before large scale validation is needed, because comparisons of global proteome profiles yield hundreds of candidate biomarkers. Verifications of hundreds of candidates by affinity-based methods, the broadly used Western blot or ELISA approaches, are impractical, because development of reagents of suitable specificity and affinity to support accurate detection and quantitation of target proteins remains expensive and time consuming. Furthermore, such methods are also hindered by marked limitations in abilities to detect multiple proteins in the same sample. Recently, quantification assays based on SRM MS have been extensively investigated for protein verification purposes. This technique provides two significant advantages, that is, biomarker candidates can be assessed simultaneously at high speed with good quantitative accuracy for verification. We have conducted MS-based verification using the leading-edge sSRM method [44]. Based on retention times, the sSRM method decreases the number of concurrent SRM transitions monitored at any one time-point, offering improved reproducibility and signal-to-noise ratio. Herein, we detected and quantified 56 proteins and 400 SRM transitions simultaneously in complex mixtures. The impact of MS technologies on biomarker discovery and clinical practice will be more important in near future.

These 10 proteins may improve diagnostic capability and have clinical utility. Because we conducted this proteomics study using tumor tissues in consideration of the wide dynamic range, this study lacks the validation of these proteins using serum and bile juice. However, detections of OLFM4, FABPL, S100P, CEAM5 and MUC5A in serum have been reported [40, 45–47]. In the near future, however, it might be possible to diagnose EHCC by measuring these proteins simultaneously using SRM-based MS.

In conclusion, we identified several unique proteins, newly discovered to be associated with EHCC, using MS-based proteomics approaches with archived FFPE tissues. These proteins are potential targets for future diagnostic biomarkers and anticipated to facilitate unraveling the molecular events that underlie this lethal disease. Many of these proteins are poorly-documented and further investigation about their roles in EHCC is needed.

Acknowledgments We thank Emiko Shibuya and Keiko Inabe (Tohoku University) for technical assistance. We thank Brent Bell (Tohoku University) for English editing.

Conflict of interest None declared.

Author contributions Study design: S.M., T. Morikawa, T.O. and M.U. Acquisition of data: S.M., T.T., T.S., T. Minowa, N.H. and F.M. Analysis and interpretation: S.M., T.O. and T.N. Manuscript drafted by: S.M. and T. Morikawa. Revision: S.M., T.T., T.S., T. Minowa, N.H., T.O., F.M., T.N. and M.U. Statistical advice: T. Minowa, N.H. and T.N.

References

1. Khan SA, Thomas HC, Davidson BR, Taylor-Robinson SD. Cholangiocarcinoma. Lancet. 2005;366:1303–14.
2. Shaib Y, El-Serag HB. The epidemiology of cholangiocarcinoma. Semin Liver Dis. 2004;24:115–25.
3. Ishihara S, Miyakawa S, Takada T, Takasaki K, Nimura Y, Tanaka M, et al. Status of surgical treatment of biliary tract cancer. Dig Surg. 2007;24:131–6.
4. Pandey A, Mann M. Proteomics to study genes and genomes. Nature. 2000;405:837–46.
5. Petricoin EF, Ardekani AM, Hitt BA, Levine PJ, Fusaro VA, Steinberg SM, et al. Use of proteomic patterns in serum to identify ovarian cancer. Lancet. 2002;359:572–7.
6. Washburn MP, Wolters D, Yates JR III. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001;19:242–7.
7. Anderson L, Hunter CL. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics. 2006;5:573–88.
8. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. Mol Cell Proteomics. 2002;1:845–67.
9. Prieto DA, Hood BL, Darfler MM, Guiel TG, Lucas DA, Conrads TP, et al. Liquid Tissue: proteomic profiling of formalin-fixed tissues. Biotechniques. 2005;38:32–5.
10. Hood BL, Darfler MM, Guiel TG, Furusato B, Lucas DA, Ringeisen BR, et al. Proteomic analysis of formalin-fixed prostate cancer tissue. Mol Cell Proteomics. 2005;4:1741–53.
11. Takadate T, Onogawa T, Fukuda T, Motoi F, Suzuki T, Fujii K, et al. Novel prognostic protein markers of resectable pancreatic cancer identified by coupled shotgun and targeted proteomics using formalin-fixed paraffin-embedded tissues. Int J Cancer. 2013;132:1368–82.
12. Kawamura T, Nomura M, Tojo H, Fujii K, Hamasaki H, Mikami S, et al. Proteomic analysis of laser-microdissected paraffin-embedded tissues: (1) Stage-related protein candidates upon non-metastatic lung adenocarcinoma. J Proteomics. 2010;73:1089–99.
13. Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, et al. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics. 2005;4:1487–502.
14. Zybailov B, Coleman MK, Florens L, Washburn MP. Correlation of relative abundance ratios derived from peptide ion chromatograms and spectrum counting for quantitative proteomic analysis using stable isotope labeling. Anal Chem. 2005;77:6218–24.
15. Sokal RR, Rohlf FJ. Biometry: the principles and practice of statistics in biological research, 3rd edn. New York: W. H. Freeman and Company; 1995.
16. Alvaro D. Serum and bile biomarkers for cholangiocarcinoma. Curr Opin Gastroenterol. 2009;25:279–84.
17. Bonney GK, Craven RA, Prasad R, Melcher AF, Selby PJ, Banks RE. Circulating markers of biliary malignancy: opportunities in proteomics? Lancet Oncol. 2008;9:149–58.
18. 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;8:4092–103.
19. Wang X, Dai S, Zhang Z, Liu L, Wang J, Xiao X, et al. Characterization of apolipoprotein A-I as a potential biomarker for cholangiocarcinoma. Eur J Cancer Care. 2009;18:625–35.
20. Nishimura T, Nomura M, Tojo H, Hamasaki H, Fukuda T, Fujii K, et al. Proteomic analysis of laser-microdissected paraffin-embedded tissues: (2) MRM assay for stage-related proteins upon non-metastatic lung adenocarcinoma. J Proteomics. 2010;73:1100–10.
21. Ferguson RE, Carroll HP, Harris A, Maher ER, Selby PJ, Banks RE. Housekeeping proteins: a preliminary study illustrating some limitations as useful references in protein expression studies. Proteomics. 2005;5:566–71.
22. Aishima S, Fujita N, Mano Y, Kubo Y, Tanaka Y, Taketomi A, et al. Different roles of S100P overexpression in intrahepatic cholangiocarcinoma: carcinogenesis of perihilar type and aggressive behavior of peripheral type. Am J Surg Pathol. 2011;35:590–8.
23. Nakanishi Y, Zen Y, Kawakami H, Kubota K, Itoh T, Hirano S, et al. Extrahepatic bile duct carcinoma with extensive intraepithelial spread: a clinicopathological study of 21 cases. Mod Pathol. 2008;21:807–16.
24. Takamura M, Yamagiwa S, Wakai T, Tamura Y, Kamimura H, Kato T, et al. Loss of liver-intestine cadherin in human intrahepatic cholangiocarcinoma promotes angiogenesis by up-regulating metal-responsive transcription factor-1 and placental growth factor. Int J Oncol. 2010;36:245–54.
25. Hodorova I, Rybarova S, Vecanova J, Solar P, Domorakova I, Adamkov M, et al. Comparison of expression pattern of monoamine oxidase A with histopathologic subtypes and tumour grade of renal cell carcinoma. Med Sci Monit. 2012;18:BR482–6.
26. Shimonishi T, Miyazaki K, Nakanuma Y. Cytokeratin profile relates to histological subtypes and intrahepatic location of intrahepatic cholangiocarcinoma and primary sites of metastatic adenocarcinoma of liver. Histopathology. 2000;37:55–63.
27. Hashimoto T, Kusakabe T, Watanabe K, Sugino T, Fukuda T, Nashimoto A, et al. Liver-type fatty acid-binding protein is highly expressed in intestinal metaplasia and in a subset of carcinomas of the stomach without association with the fatty acid synthase status in the carcinoma. Pathobiology. 2004;71:115–22.
28. Kobayashi D, Koshida S, Moriai R, Tsuji N, Watanabe N. Olfactomedin 4 promotes S-phase transition in proliferation of pancreatic cancer cells. Cancer Sci. 2007;98:334–40.
29. Liu W, Chen L, Zhu J, Rodgers GP. The glycoprotein hGC-1 binds to cadherin and lectins. Exp Cell Res. 2006;312:1785–97.
30. Liu W, Zhu J, Cao L, Rodgers GP. Expression of hGC-1 is correlated with differentiation of gastric carcinoma. Histopathology. 2007;51:157–65.

31. Miyasaka Y, Enomoto N, Nagayama K, Izumi N, Marumo F, Watanabe M, et al. Analysis of differentially expressed genes in human hepatocellular carcinoma using suppression subtractive hybridization. Br J Cancer. 2001;85:228–34.
32. Lee HJ, Nam KT, Park HS, Kim MA, Lafleur BJ, Aburatani H, et al. Gene expression profiling of metaplastic lineages identifies CDH17 as a prognostic marker in early stage gastric cancer. Gastroenterology. 2010;139:213–25. e3.
33. Lawrie LC, Dundas SR, Curran S, Murray GI. Liver fatty acid binding protein expression in colorectal neoplasia. Br J Cancer. 2004;90:1955–60.
34. Sharaf RN, Butte AJ, Montgomery KD, Pai R, Dudley JT, Pasricha PJ. Computational prediction and experimental validation associating FABP-1 and pancreatic adenocarcinoma with diabetes. BMC Gastroenterol. 2011;11:5.
35. True L, Coleman I, Hawley S, Huang CY, Gifford D, Coleman R, et al. A molecular correlate to the Gleason grading system for prostate adenocarcinoma. Proc Natl Acad Sci U S A. 2006;103:10991–6.
36. Liu TH, Li DC, Gu CF, Ye SF. Carbamyl phosphate synthetase I. A novel marker for gastric carcinoma. Chin Med J (Engl). 1989;102:630–8.
37. Kinoshita M, Miyata M. Underexpression of mRNA in human hepatocellular carcinoma focusing on eight loci. Hepatology. 2002;36:433–8.
38. Hamada S, Satoh K, Hirota M, Kanno A, Ishida K, Umino J, et al. Calcium-binding protein S100P is a novel diagnostic marker of cholangiocarcinoma. Cancer Sci. 2011;102:150–6.
39. Minato H, Nakanuma Y, Terada T. Expression of blood group-related antigens in cholangiocarcinoma in relation to non-neoplastic bile ducts. Histopathology. 1996;28:411–9.
40. Boonla C, Wongkham S, Sheehan JK, Wongkham C, Bhudhisawasdi V, Tepsiri N, et al. Prognostic value of serum MUC5AC mucin in patients with cholangiocarcinoma. Cancer. 2003;98:1438–43.
41. Rullier A, Le Bail B, Fawaz R, Blanc JF, Saric J, Bioulac-Sage P. Cytokeratin 7 and 20 expression in cholangiocarcinomas varies along the biliary tract but still differs from that in colorectal carcinoma metastasis. Am J Surg Pathol. 2000;24:870–6.
42. Lankisch TO, Metzger J, Negm AA, Vobetakuhl K, Schiffer E, Siwy J, et al. Bile proteomic profiles differentiate cholangiocarcinoma from primary sclerosing cholangitis and choledocholithiasis. Hepatology. 2011;53:875–84.
43. Welzel TM, Graubard BI, El-Serag HB, Shaib YH, Hsing AW, Davila JA, et al. Risk factors for intrahepatic and extrahepatic cholangiocarcinoma in the United States: a population-based case-control study. Clin Gastroenterol Hepatol. 2007;5:1221–8.
44. Fillatre Y, Rondeau D, Jadas-Hecart A, Communal PY. Advantages of the scheduled selected reaction monitoring algorithm in liquid chromatography/electrospray ionization tandem mass spectrometry multi-residue analysis of 242 pesticides: a comparative approach with classical selected reaction monitoring mode. Rapid Commun Mass Spectrom. 2010;24:2453–61.
45. Oue N, Sentani K, Noguchi T, Ohara S, Sakamoto N, Hayashi T, et al. Serum olfactomedin 4 (GW112, hGC-1) in combination with Reg IV is a highly sensitive biomarker for gastric cancer patients. Int J Cancer. 2009;125:2383–92.
46. Lachmann RA, Werchan S, Schachtrup C, Haitsma JJ, Spener F, Lachmann B. Liver-type fatty acid binding protein in serum and broncho-alveolar lavage in a model of acute respiratory failure because of surfactant depletion--a possible marker for lung damage? Clin Physiol Funct Imaging. 2006;26:371–5.
47. Wang Q, Zhang YN, Lin GL, Qiu HZ, Wu B, Wu HY, et al. S100P, a potential novel prognostic marker in colorectal cancer. Oncol Rep. 2012;28:303–10.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1 Proteins to which the scheduled SRM MS analysis was applied.

Fig. S1 Study design for identification and validation. Shotgun proteomics with semi-quantitative spectral counting was conducted to identify proteins with expression profiles differing between EHCC and non-cancerous bile ducts. Selected candidates were verified by quantitative analysis using scheduled SRM-based targeted proteomics. The resulting proteins were then validated by immunohistochemical analysis. EHCC, extrahepatic cholangiocarcinoma; LC-MS/MS, liquid chromatography-tandem mass spectrometry; SRM, selected reaction monitoring.