RESEARCH COMMUNICATION

Proteomics Analysis and Evaluation of Biomarkers for Detection of Cholangiocarcinoma

Pramote Sriwanitchrak[1], Vithoon Viyanant[1], Wanna Chaijaroenkul[1], Petcharin Srivatanakul[2], Hans Rudi Gram[1], Veerachai Eursiddhichai[1], Kesara Na-Bangchang[1]*

Abstract

Cholangiocarcinoma (CCA) is a rare but devastating neoplasm that accounts for about 3% of all gastrointestinal cancers and about 15% of all primary liver cancers worldwide. The lack of early detection and limited therapeutic options are major problems in controlling CCA. The current study attempted to identify novel serum markers which can substitute the carbohydrate antigen CA19-9, or can improve, when measured together, the diagnostic accuracy of CA19-9. Differentially expressed proteins in pooled and individual plasma samples obtained from patients with CCA and control subjects (10 each) were identified by using two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MALDI-TOF). Out of a total of 21 protein spots separated and identified, five spots were found to be up-regulated in plasma from CCA patients. The up-regulation of α1-antitrypsin (AP1) was observed in all of the ten samples from CCA patients with protein intensity significantly higher than control subjects. Based on results of binary logistic regression analysis of the three serum biomarkers (CA19-9, AP1 and α-fetoprotein: AFP), serum levels of at least CA19-9 together with AP1 were the minimum requirement to obtain prediction accuracy of greater than 80% in a battery test for diagnosis of CCA. However, in order to obtain high predictability of 100% or approaching, an addition of at least one of the three liver function enzymes (alkaline phosphatase: ALP; aspartase transaminase: AST; alanine trasaminase: ALT) is required. Serum biomarkers may be a useful diagnostic or prognostic monitoring tool for CCA. Further evaluation of larger number samples is needed to support their applicability in a clinical setting as diagnostic and prognostic tools. Determination of clinical utility of these marker models in early diagnosis of CCA requires study in animal models with disease progression.

Keywords: Cholangiocarcinoma - biomarker - *Opisthorchis viverrini* - diagnosis - cancer

Asian Pacific J Cancer Prev, **12**, 1573-1580

Introduction

Cholangiocarcinoma (CCA) is an adenocarcinoma arising from cholangiocyte, the epithelial cell lining the bile duct apparatus. Several epidemiologic studies have demonstrated an increase in the incidence of CCA in Southeast Asia (Jinawath et al., 2006). Infection with liver fluke Opisthorchis viverini (OV) has been identified as one significant factor associated with the risk of development of CCA (Srisomsap et al., 2004). OV is endemic in Southeast Asia, particularly in the northeastern part of Thailand, where the daily habit of eating raw and salt-fermented fresh water fish repeatedly exposed this local population to both OV and nitrosamine-contaminated food. The prevalence of OV infection is up to 70.8%, and the incidence of CCA in this region is approximately 317.6 per 100,000 person per year (Khuhaprema et al., 2007). A pathological diagnosis of CCA is often difficult due to its location, size and demoplastic characteristics (Gatto, 2010). Furthermore, tumor masses are not often even identifiable by computer tomography, ultrasound or magnetic resonance imaging (Gatto et al., 2010). Even those with operable tumor, the recurrence rate was extremely high, with a 5-year survival rate of less than 40% (Thongprasert, 2005).

At present, there is no effective tool or specific biomarkers that can detect the early stage or monitor status of CCA. Established serum tumor markers include carbohydrate antigen 19.9 (CA19-9) and carcinoembryonic antigen (CEA). However, these markers are not always helpful, with sensitivities of approximately 70 and 50%, respectively. CEA, which is mainly used for colorectal cancers, is of scarce utility, since it is unspecific and is increased only in approximately 30% of patients with CCA (Dbouk et al., 2007; Park et al., 2007; Sajid et al., 2007). CA19-9 is the most widely used serum marker for CCA

*[1]Thailand Center of Excellence on Drug Discovery Development, Thammasat University, Pathumthani, [2]National Cancer Institute, Ministry of Public Health of Thailand, Bangkok, Thailand, [3]Institute of Parasitology, McGill University, Sainte Anne de Bellevue, Quebec, Canada *For correspondence: kesaratmu@yahoo.com*

but is also elevated in pancreatic cancer, gastric cancer, primary cirrhosis, smokers and may also be transiently increased during cholangitis or cholestasis. A high value of sensitivity and specificity has been reported for CA19-9 in patients with CCA depending on the study population and the cut-off values, including CCA complicating primary sclerosing carcinoma (PSC). Furthermore, elevated CA 19-9 usually allows CCA diagnosis in advanced stages when radical treatments are not allowed. Other biomarkers such as trypsinogen-2 (Lempinen et al., 2007), platelet-lymphocyte ratio (PLR), mucin 5AC (Bamrungphon et al, 2007; Boonla et al, 2005), soluble fragment of cytokeratin-19 (CYFRA21-1) have been recently shown to assist in the diagnosis of CCA with, in some cases, a prognostic value (Uenishi et al., 2008).

Proteomics of serum as well as bile are under investigation but definitive findings are currently unavailable. Information of the pattern of proteins that are up- or down-regulated in CCA patients would be exploited for development of diagnostic or prognostic tool for this type of cancer. The current study attempted to identify novel serum markers which can substitute CA19-9, or can improve, when measured together, the diagnostic accuracy of CA19-9. Differentially expressed proteins in plasma obtained from patients with CCA were identified by using two- dimensional gel electrophoresis (2-DE) and mass spectrometry (MALDI-TOF), in comparison with plasma from control subjects. In addition, the diagnostic predictability of the conventional biomarkers of liver diseases --α-fetoprotein (AFP) and α1-antitrypsin (AP1), and the liver function enzymes --alkaline phosphatase (ALP), aspartase transaminase (AST) and alanine transaminase (ALT) when used together with serum level of CA19-9 as a battery test for diagnosis of CCA was also investigated.

Materials and Methods

Sample collection

Plasma samples (30 samples, each) were collected from age- and sex-matched (case vs control) subjects (living in Ubon Ratchathani Province) who had received the diagnosis and treatment at Ubon Rachathani Cancer Center. The case group included patients with late stage CCA and the control group was a matched-paired control for the case group for sex (21 vs 25 males for control vs case) and age [median (range) 51 (32-73) vs 56 (36-72) years for control vs case] and consisted of health subjects or patients with other diseases not related to liver diseases, cancer and diabetes (Table 1). The diagnosis of CCA was based on abdominal ultrasound and serological tests with confirmation along the course of treatment. The study protocol was approved by the Ethics Committee of the Ministry of Public Health of Thailand. Written informed consents were obtained from all subjects before sample collection.

Blood samples (7 mL each) were obtained pre-operatively on the day before surgery or transplantation. From conservatively treated patients, the samples were taken at the time of diagnosis. Plasma samples were prepared for proteomics analysis and serum samples were prepared for the analysis of the tumor markers CA19-9, α-fetoprotein (AFP) and α1-antitrypsin (AP1), as well as Opisthorchis viverrini antibody (OV-Ab), hepatitis C virus antibody (HCV-Ab), HBs antigen (HBs-Ag), serum alanine transaminase (ALT), aspartase transminase (AST), and alkaline phosphatase (ALP). All samples were stored at -80 ˚C until analysis.

Preparation of plasma samples

Protein patterns from the pooled plasma samples were initially investigated to provide an overall picture of the protein profiles of samples from both groups. Equal volume of plasma from each subject was pooled for each group (case and control). The pooled plasma samples were then depleted to remove high abundant protein by using a Proteome Lab IgY-12 Spin Column from Beckman Coulter (Mississauga, ON, Canada). In brief, 10 μL of the pooled plasma sample was diluted 50-fold in manufacturer's dilution buffer (10 mM Tris-HCl, pH 7.4, 150 mM NaCl). Diluted plasma samples were added to the beads, mixed with them and the columns were incubated 15 min on a serological shaker at room temperature. Columns were centrifuged for 30 seconds at 400xg (18 oC) and were washed with 500 μl of dilution buffer and the flow-through was combined with the washes to obtain the maximal yield of low-abundance plasma proteins. The depleted pooled plasma sample was then applied to a Microcon YM-3 (YM-3, MWCO 3 KD, Millipore MA) and centrifuged at 5,000 xg (4 ˚C). The supernatant was then filtered by using 2-DE Clean-Up. Total protein content of depleted plasma samples was determined using the DC Protein assay (Bio-Rad, USA) with bovine serum albumin (BSA) as a standard protein.

In the second step, proteomic profile of each individual plasma sample in both groups (10 each) was examined. The procedures followed that described above for the pooled plasma.

Two-dimensional gel electrophoresis (2-DE)

For 2-DE, each 100 μg of depleted plasma was exchanged with sample buffer (8 M urea, 2 M thiourea, 2% CHAPS, 65 mM DTT, 0.5% ampholyte). After dilution to 125 μL with rehydration buffer (8 M urea, 4% CHAPS, 65 mM DTT, 0.001% bromphenol blue, 1% Bio-Lyte ampholytes), the samples were loaded onto IPG strips (7 cm, covering the pH ranges 3-10 NL; Bio-Rad) for isoelectric focusing using the PROTEAN IEF cell. After rehydration overnight at room temperature, the voltage/time profile applied was as follow: 250 V for 15 min, 4,000 V for 1 hour, and then 4,000 V to total of 15,000 V-hr. After equilibration with a solution containing 6 M urea, 0.375 M Tris-HCl, pH 8.8, 2% SDS, 20% glycerol, 2% (w/v) DTT, IPG strips were equilibrated again with the same solution containing 2.5% (w/v) iodoacetamide instead of DTT. IPG strips were placed on top of 12% SDS-PAGE for second dimension separation followed by 200 V for 45 min with Mini-PROTEAN 3 Electrophoresis. Triplicate analysis was performed for each sample.

Protein visualization and image analysis

Following second-dimensional SDS-PAGE, analytical

Table 1. Demographics and Ultrasound and Clinical Diagnosis of CCA in the 30 Patients

No.	Sex	Age	Ultrasound
1	M	47	Stage-III: defined mass of 2x2 cm (diameter) at hepatic hilar with dilated IHBDs
2	M	62	Stage-III: defined mass of 9.6x9 cm (diameter) of low echoic mass at right liver, peripheral type
3	F	55	Stage-III: defined mass of 2.5x2.5 cm (diameter) at porta hepatic region causing predominant dilated left intrahepatic bile ducts (about 1 cm in diameter), multiple celiac axis lymphadenopathy
4	M	50	Stage-III: defined mass of 1.5x1.5 cm (diameter) at right liver with mild dilated intrahepatic bile ducts
5	M	36	Stage-III: defined hyperechoic lesion of 5x6 cm (diameter) at right lobe liver with adjacent dilated IHD near the lesion considered to intrahepatic type
6	M	53	Stage-III: CCA at distal CBD causing obstruction (CBD 2x2 cm in diameter), minimal ascites and celiac axis lymphadenopathy
7	M	63	Stage-III: defined mass of 2.1x2.5 cm (diameter) hilar type, mild dilated left IHBDs, mild to moderate left hydronephrosis, minimal ascites
8	M	68	Stage-III: defined mass of 7.5x7.5 cm (diameter)at posterior superior segment of right liver, mass with central calcifications and irregular dilated IHBDs nearby, CBD (5x5 cm) not dilated at this age, few small celiac axis nodes, mild to moderate right hydronephrosis
9	F	72	Stage-III: large infiltrative mass at right liver with dilated IHBDs, CBD suggestive of CCA, as well as suspicious 1.7x1.7 cm distal CBD stone (not well characterized), gall bladder inflammation, large amount of ascites
10	M	56	Stage-III: mild prominence IHBD at left lobe liver as seen from previous study and hyperechoic of 2.5x2.6 cm (diameter) at left lobe, intrahepatic type CCA is compatible, correlated with CT scan
11	M	49	Stage-III: defined hyperechoic lesion at left lobe liver of 5x7.5 cm (diameter) with mild dilated left IHBD near the lesion, considered to be intrahepatic type with small amount of ascites seen
12	M	55	Stage-III: moderate dilated IHBD both lobes liver about 0.8x0.9 cm (dimeter) with III-defined isoechoic lesion of 4x4 cm (diameter) at confluence of right and left hepatic ducts with non-communication of right and left hepatic ducts compatible with central type CCA
13	M	52	Stage-III: defined mass of 3.5x3 cm (diameter) at right lobe liver with dilated IHBDs and multiple lymphadenopathies
14	M	72	Stage-III: d defined mass of 3x3 cm (diameter), iffused dilated IHBDs and defined echo at porta region, multiple peripancreatic nodes, bilateral renal stones
15	M	61	Stage-III: defined mass of 3x3 cm hyperechoic at right lobe liver with dilated right IHBDs, suspected peripancreatic nodes
16	F	67	Stage-III: mild dilated IHBDs with 2.2x2 cm (diameter) hyperechoic lesion at right lobe liver
17	F	56	Stage-III: defined mass of 10x10 cm (diameter) edechoic mass at right lobe liver with dilated IHBDs
18	M	38	Stage-III: defined mass of 8 x 10 cm (diameter) hyperechoic at right lobe liver with mild dilated IHD near the lesion compatible with intrahepatic type
19	F	68	Stage-III: mild hepatomegaly with diffused dilated intrahepatic bile ducts and lesion filled in CBD, CCA at proximal CBD
20	F	61	Stage-III: defined hyperechoic at right lobe liver of 3.5x5 cm (diameter) with mild prominence IHBD within and near the lesion compatible with intrahepatic type
21	F	57	Stage-III: defined mass of 5x5 cm irregular lesion at porta hepatic region causing diffuse dilated intrahepatic bile ducts and intraluminal echo filled in CBD
22	M	60	Stage-III: defined hyperechoic at right lobe liver of 5x5 cm (diameter) with mild prominence IHBD near the lesion considered to intrahepatic type
23	M	57	Stage-III: diffused irregular dilatation of intrahepatic bile ducts and common bile duct
24	F	62	Stage-III: defined low density lesion at posterior segment of right lobe liver of 3x4 cm (diameter) with calcification and poor enhancement compatible with intrahepatic type and multiple matted lymphadenopathy of 2x5 cm (diameter) at portal, celiac, superior mesenteric and paraaortic regions (CT whole abdomen)
25	M	52	Stage-III:-defined hyperechoic lesion at right lobe liver near hilar region of 5x6 cm (diameter) with dilated IHD within and near the lesion
26	M	55	Stage-III: defined mass at portal region of 6 cm size with mild dilated left IHBDs, 10x5 cm hyperechoic lesion at right lobe liver, possibly fatty infiltration or tumor.
27	M	46	Stage-III: Irregular dilated left IHBDs and CHD, lymphadenopathy at pancreatic head region
28	M	49	Stage-III: defined mass of 9x7 cm (diameter) with central necrosis at right lobe liver, posterosuperior segment and dilated surrounding IHBDs, few small celiac axis nodes of 1 cm in size
29	M	65	Stage-III: mild hepatomegaly with diffused irregular dilated intrahepatic bile ducts, CCA at hilar region
30	F	51	Stage-III: defined hyperechoic lesion of 4x5 cm (diameter) at hilar region compatible with hilar type and evidence of peripancreatic node of 1.5x1.6 cm (diameter)

IHBD; intra-hepatic bile duct; CBD, common bile duct

gels were stained with Bio-Safe Coomassie stain. The gels were washed three times for 5 min each in 200 ml of ddH2O per gel. Water was removed all water and 50 ml Bio-Safe Coomassie stain was added and gently shaken overnight. The stain was then discarded, and gel was washed thoroughly with water. Protein patterns in the gels were recorded as digitalized images using a high-resolution scanner (Pharox FX, Bio-Rad). Gel image matching was done with PDQuestTM software (Version 7, Bio-Rad, USA). Scanned gel images were processed

to remove backgrounds, staining on the gel borders and to automatically detect spots. For all spot intensity calculation, normalized values were used. Normalization of spot intensity was done with Loess Regression Method and normalized spot intensities were expressed in ppm. Representative gel (master gel) from each group of pooled samples was determined based on at least 80% similarity of the protein patterns.

In-gel digestion

In-gel digestion of protein spots on Coomassie gels was carried out with 160 ng of Porcine Modified Trypsin (Sigma, USA) in 10% acetonitrile and 25 mM NH4HCO3 and performed essentially as described below. Briefly, after the completion of staining, the gel was washed twice with water for 15 min, and then twice with water/acetonitrile (1:1 v/v) for 15 min. The solvent volumes were about twice the gel volume. Liquid was removed, acetonitrile was added to the gel pieces and the mixture was left at room temperature for 5 min. Liquid was removed and the gel pieces were re-hydrated in 0.1 M NH4HCO3 for 5 min. Acetonitrile was added to give a 1:1 v/v mixture of 0.1 M NH4HCO3/acetonitrile and the mixture was incubated for 15 min. All liquid was removed with the digestion buffer containing 25 mM NH4HCO3 and 10 ng/μL of trypsin was added and all was incubated for 4 hours at 37°C. The supernatant was recovered and the extraction was carried out with 1% TFA/acetonitrile (1:1 v/v).

Protein identification by MALDI-TOF-MS

The protein spots from either the pooled plasma or individual plasma sample from both groups were used for identification of the proteins by peptide mass fingerprinting by MALDI-TOF-MS. The matrix solution, α-cyano-4-hydroxycinnamic acid (Bruker, USA) was prepared by dissolving to saturation in 70% acetronitrile/water with 0.1% trifluoroacetic acid. The peptide digest (1 μL) was mixed with 5-10 μL of matrix solution. The mixture was spotted onto a stainless steel target and left to dry at room temperature. The mixture was then irradiated with a 337 mm N2 laser, accelerated with 20 kV accelerating voltage in a two-stage gridless pulse-ion extraction source. The instrument was operated in the positive, reflectron mode. Typical mass accuracy was better than 200 ppm with external calibration. Known contamination peaks such as keratin, autoproteolysis peaks were removed prior to protein mass fingerprint database search (Mascot, Matrixscience, UK). The LC-MS/MS was performed using LCPackings nanoLC system (LCPackings, The Netherlands), equipped with FAMOS autosampler. Ultimate nanoLC and ESI Esquire300plus ion-trap mass spectrometers (Bruker, USA) was used for the LC-MS study, with a C-18 column (75 um ID x 15 cm). Eluents A and B were 0.1% formic acid and 0.1% formic acid in 90% acetronitrile, respectively. One μl of sample was injected onto the nanoLC system, and separation was performed using the following gradient: 0 min 98% A, 40 min60%A,50min10%A,60min10%A,62min98% A, and 90 min 98% A. The ion-trap instrument was set to auto MS/MS mode. The resulting (unprocessed) MS/MS data was submitted to MS/MS database search (Mascot, Matrixscience). Analysis of selected protein spots was performed using PDQuestTM software version 7 (BioRad, USA). Peptides identified was investigated with the MASCOT search engines (http://www. matrixscience. com).

Analysis of plasma tumor markers and serum liver function tests

Serum AP1 concentration was determined by the method of Nephelometry (NPL). Plasma was loaded into a tube of the commercial kit (BN Prospec Antisera to human A1-AT, Dade Behring, Behring Werke AG, Marburg, Germany). Immunological precipitation was measured by nephelometer according to the manufacturer's instructions. Serum levels of CA19-9, AFP, OV-Ab, HCV-Ab and HBsAg were determined by a solid phase enzyme-linked immunosorbent assay. The cut-off values for AP1, AFP and CA19-9 are 200 mg/dL, 9 ng/mL and 30 U/mL, respectively. The following laboratory tests were retrieved from clinical records: ALP, ALT, AST. The cut-off values for ALP, ALT and AST are 104, 40 and 40 unit/L, respectively.

Data analysis

Statistical analysis for comparing the significant difference of the quantitative data between the two groups was done by Mann Whitney U-test and Wilcoxon-Signed Rank test for unpaired and paired data, respectively. Difference in proportions was assessed by Chi-square test. Binary logistic regression analysis (stepwise forward and backward) was performed with the six covariates (serum levels of CA19-9, AFP, AP1, ALP, ALT, AST) from 54 subjects (22 and 23 for case and control group, respectively) in order to find the regression model(s) for these markers that would best predict the occurrence of CCA. In the next step, the applicability of the selected model(s) was then evaluated using data from 6 subjects (3 and 3 from case and control groups, respectively) based on the equation: $Y = \beta1X1 + \beta2X2 + \beta3X3..... + constant$. Y was the prediction of the occurrence of CCA; X1, X2, X3,.... were covariates included in the model; β1, β2, β3..... were slopes of the regression model when each covariate was included; and constant was the Y-intercept of the regression line. Statistical significant level for all tests was set at $\alpha = 0.05$.

Results

The 2-DE patterns of low abundance plasma proteins separated over the pH range between 4 and 7 from pooled plasma samples of patients with CCA (case group) and healthy subjects (control group) were compared following removal of high abundant proteins. There were a total of 21 protein spots separated from pooled plasma samples from both groups which could be accurately identified by MALDI-TOF (Table 2). Figure 1 (a, b) presents the analysis of separated spots from pooled plasma samples of CCA patients (2a) and healthy subjects (2b) using PDQuestTM software. Five proteins (ID 1375, 1427, 1512, 1324, 1368) were found to be up-regulated in pooled plasma samples from CCA patients, with the

Table 2. Protein Spots Separated by 2-DE and Identified by MALDI-TOF in Pooled Plasma Samples from CCA Patients in Comparison with Control Subjects

No.	Spot ID	Protein names	Accession numbers	Peptide matches	SQ	PS	MW	pI	Ratio (C/N)
1	1298	CLL-associated antigen KW-4 splice variant 2	19851928	12/60	26	44	92653	9.70	0.23
2	1306	Immunoglobulin heavy chain variable region	16075801	5/41	67	48	10233	9.25	0.30
3	1310	CLL-associated antigen KW-4 splice variant 2	19851928	9/44	20	41	92653	9.17	0.30
4	1312	RAB39, member RAS oncogene family	39930371	5/35	33	38	25390	7.57	0.31
5	1314	Transferrin	553788	12/39	33	90	55207	6.00	0.31
6	1321	Hemopexin precursor	386789	11/41	29	80	52254	6.57	0.31
7	1323	Chain A, structure of human serum albumin with S-naproxen and the Ga module	168988718	18/53	32	119	67773	5.63	0.16
8	1330	RAB39, member RAS oncogene family	39930371	8/40	52	68	25390	7.57	0.29
9	1340	Chain A, human serum albumin complexed with myristate and azapropazone	78101694	13/50	23	76	68398	5.57	0.22
10	1352	Kininogen 1 isoform 2	4504893	12/51	28	81	48936	6.29	0.20
11*	1368	Bruton's tyrosine kinase	7381350	6/38	48	51	17804	6.20	6.09
12*	1375	Alpha-1-antitrypsin	151302818	13/27	34	82	46978	5.51	4.39
13	1405	hCG1652647, isoform CRA_b	119569005	6/34	27	47	28863	10.13	0.27
14	1420	HIRIP3	3255985	7/55	32	50	27249	9.64	0.14
15*	1427	Unknown	-	-	-	-	-	-	2.50
16	1445	RAB39, member RAS oncogene family	39930371	7/47	46	52	25390	7.57	0.23
17	1449	p105MCM	1197636	10/40	13	44	94002	5.32	0.19
18	1494	Chain L, Mature Oxy-Cope antibody with hapten	2914180	5/29	29	45	23720	6.19	0.26
19	1499	DAZAP1/MEF2D fusion protein	56548953	7/35	19	48	49106	8.92	0.21
20*	1512	hp2-alpha	296653	9/15	24	88	42126	6.25	2.10
21*	1524	Chain A, A Covalent Dimer of Transthyretin Affecting The Amyloid Pathway	55669575	9/16	87	182	12836	5.33	4.17

SQ, sequence coverage (%); PS, protein score; C/N, ratio of spot density between case and control; *up-regulated protein spots

ratio of spot density between case and control (C/N) of > 2.0. Four spots were identified as AP1 (ID 1375), Bruton's tyrosine kinase (ID 1368), hp2-α (ID 1512) and chain A, a covalent dimer of ransthyretin that affects the amyloid pathways (ID 1524). Identification of one spot (ID 1427) was not possible due to poor resolution from the nearby spots. The ratio of protein density of Bruton's tyrosine kinase, AP1, unknown protein, hp2-α and chain A in pooled plasma collected from patients to that from control subjects were 6.09, 4.39, 2.50, 2.10, and 4.17, respectively. The up-regulation of AP1 was observed in all samples from patients with CCA, and in addition, the spot density was significantly higher ($p < 0.00001$; Mann-Whitney U test) in the patient group than the control group (75.5 vs 20.0).

Regarding tumor biomarkers (AP1, CA19-9, AFP) and liver function enzymes (ALT, AST, ALP) including results of serology tests (OV-Ab, HCV-Ab and HBs-Ag), the levels of all except AFP and liver enzymes were significantly higher in CCA compared with healthy group. In addition, the proportions of samples with significant increase in the levels (above the upper limit cut-off values) of the tumor markers AP1 and CA19-9 were markedly higher in CCA than the healthy control group. OV-Ab was detected in almost all samples (28/29) of the case group, but was also detected at high frequency in the healthy group (21/30), indicating previous or current infection with OV. No sample with HCV-Ab positive was found. HBs-Ag was found in 4 out of 30 of the case group. Table 5 et al, summarizes the best seven predicted model by stepwise forward and backward binary logistic regression analysis when the cut-off value (as described above) or actual value of each covariate was included in the models. In the later step, each model was evaluated for its accuracy

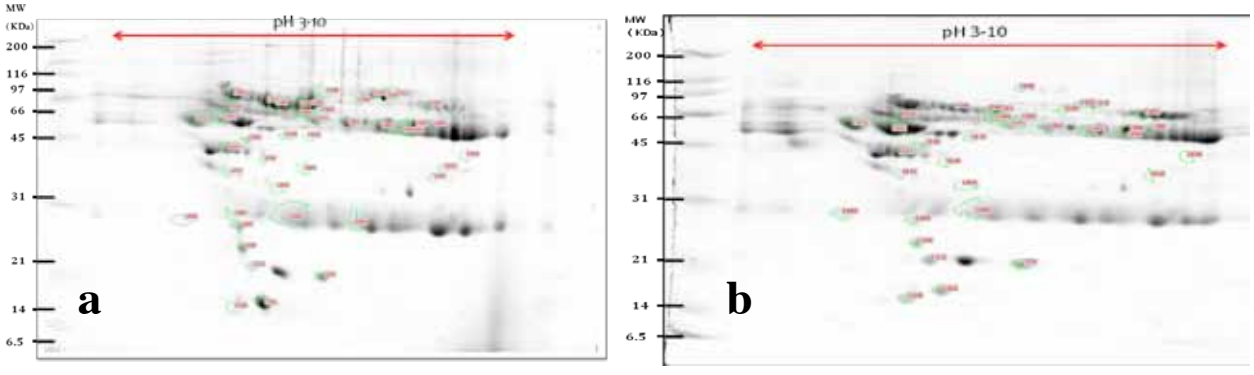

Figure 1. Protein Patterns Separated from Pooled Plasma Samples of (a) Healthy Subjects (Control Group) and (b) CCA Patients by 2-DE. Analysis of Separated Spots Was Done by PDQuestTM Software

of predictability in a total of 5 samples (2 for the case and 3 for the control group). CA19-9 and AP1 when used alone (cut-off value) provided about only up to 75% accuracy. The accuracy of predictability of other markers/enzymes when used alone was lower than 50%. CA19-9was found to be a key predictor of CCA in all models when used in combination with other markers/enzymes. AP1 was the second best predictor and was included in all except one model (model 7). Model 6 based on stepwise backward analysis of actual values (Y = 0.060 CA19-9 + 0.409 AP1 + 0.475 AFP +0.593 ALP + 0.086 AST – 0.393 ALT – 165.289) which included all the six covariates could predict the occurrence of CCA with 100% accuracy during either model building step or model evaluating step. The other three models (3, 4, 5) which included CA19-9 and two additional covariates could predict the occurrence of CCA with accuracy varying between 92.6 and 100%. The predictability of model 1 and 2 with only two covariates (CA19-9, AP1) was only 88.9% during model building step and 80% (4/5) during model evaluating step.

Discussion

Proteomic analysis and quantitative protein expression profiling is currently considered to be a considerable potential and powerful tool for identifying possible biomarkers for disease diagnosis. High-resolution two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), followed by protein identification using MS has been used to study a wide variety of biomarker discovery (Joo et al., 2003). Comparing to other biological specimens, serum tumor markers are attractive and have been the objects of extensive investigation to assist CCA diagnosis because of the ease of obtaining samples and their relatively low cost. Unfortunately, none of these markers has reached adequate specificity for CCA (Park et al., 2007; Sajid et al., 2007).

The limitations of the conventional tumor markers CA19-9 and CEA are well recognized for their insufficient specificity and sensitivity (Nehls et al., 2004; Khan et al., 2002). Several studies have evaluated the accuracy of serum CA19-9, the most useful marker for CCA, in predicting the development of CCA in patients with PSC. With a cut-off value of 100 U/mL, the reported sensitivity ranges between 60 and 89% and specificity between 80 and 97%, respectively (Ramage et al., 1995; Chalasani et al., 2000). The diagnostic value of CEA for CAA is inferior to that of CA19-9 16 but the combined use of CEA and CA19-9 has been shown to improve the diagnostic accuracy (Ramage et al., 1995).

The current effort aimed to identify potential novel serum biomarkers that can substitute or at least used in conjunction with CA19-9 as a battery test for diagnosis of CCA. Novel biomarkers for CCA was firstly identified by comparative analysis of the protein patterns of plasma obtained from CCA patients (case) and control subjects (control) by 2-DE and MALDI-TOF. Results demonstrated, in general, similar patterns of plasma from CCA patients with respect to the number of protein spots (21) separated and identified. Spot ID 1375 however seems prominent in its up-regulation in all plasma samples (10/10), but with significantly higher protein density in plasma samples from CCA patients compared with control subjects (75.49 vs 19.98). This spot was identified by MALDI-TOF as α1-antitrypsin (AP1). This proteomic finding corresponded well with the observation of significantly higher serum concentrations of AP1 in CCA patients compared with control subjects (median value of 245 vs 142 mg/L). AP1 is the most important and abundant plasma-circulating protease inhibitor belonging to serpin super family (Poley 1994). It is a 52 kDa glycoprotein consisting of 394 amino acids. The physiologic serum concentration for adults ranges from 1.5 to 3.0 g/l (20- 52 μmoL/L) (Chignard et al., 1994). Under normal conditions, AP1 is constitutively produced and, as an acute phase protein, is up-regulated during inflammation, infection, cancer, and pregnancy, administration of contraceptives, surgical trauma (Carrell, 1986; Chidwick et al., 1994). The presence of adequate local concentrations of AP1 is crucial for maintenance of protease: antiprotease homeostatis and prevention of proteolytic tissue damage. Immuno-histochemical analysis of liver specimens from CCA patients is underway to confirm the expression of AP1 in CCA cells.Marked increase in serum AP1 was previously reported in CCA (Changbumrung et al., 1988). Recently, increased expression of AP-1 mRNA in CCA tissue has been found to be a strong predictor of aggressive CCA (Tonouchi et al., 2006). Given the high plasma concentration of AP1, it is likely that albumin may interfere with its quantitative detection. To overcome problems with this low yield, removal of high abundant proteins (albumin, immunoglobulins, transferrin and fribrinogen) was performed as a preliminary step for characterization of the plasma proteome.

Binary logistic analysis was performed in order to predict the occurrence of CCA based on serum concentrations of three biomarkers and three liver function enzymes. Each individual marker/enzyme alone could not predict the occurrence of CCA with satisfactory accuracy. CA19-9 (either the actual and/or cut-off values) was found to be the key biomarker which was included in all of the seven regression models. AP1 was the second best predictor, which when used in conjunction with CA19- 9 in six of the seven models, provided the accuracy of predictability ranging from 88.9 (model-building step) or 80% (model-evaluating step) to 100% (both model-building and model-evaluating steps) depending on the addition of at least one of the three liver function enzymes (ALP, AST, ALT). AFP was a good predictor in the model only when used in combination with all of the other fives markers/enzymes (model 6) and the serum levels were similar between the CCA patients and healthy subjects.

To obtain high predictability of approaching 100%, the minimum of three covariates consisting of CA19-9, AP1 and one of the liver function enzymes is required. Unfortunately, protein spots of CA19-9 and AFP were not detected in the proteomic analysis. Detecting proteins that are present at lower levels in human plasma, for the identification of potential disease biomarkers, is complicated by a few highly abundant proteins. During

protein depletion step, these low abundant proteins may be removed together with the high abundance proteins. In addition, in our experiment, separation of proteins was performed using low gel tray due to limitation of plasma sample volume. In order to study the plasma proteome, the plasma sample has compositionally resembles serum but it is used more commonly for proteome studies because its constituents reflect patient pathological status more than do the serum constituents (Kim et al., 2007).

It is noted that, among the down-regulated protein spots, spot ID 1314 identified as transferring (Table 2), the iron binding protein, had markedly low expression in plasma of CCA patients. High rate of iron intake has been observed in most cancer cells (Shterman et al., 1991). In addition, the cells also express high cell surface concentration of transferrin receptors (Raaf et al, 1993) which allow binding of iron to the iron transporter holotransferrin. Iron then enters into the cells via a receptor-mediated endocytosis process (May et al., 1985). Therefore, the iron storage of tumor cells is generally greater in tumor cells than in normal cells (Shterman et al., 1991). Holo- transferrin and other iron sources have clearly been shown to increase the potency of artemisinin (antimalarial drug with anticancer activity) in killing cancer cells (Singh et al, 2001). In our recent study however, a controversial result was observed. Pretreatment with holo-transferrin did not alter the cytotoxicity of artemisinin and its derivatives against the CCA cell line-- CL6 (Chaijaroenkul et al., 2011). The possible explanation is that, unlike other cancers, expression of transferrin receptor in CCA may be lower than normal cell. In a previous study, the level of transferrin was shown to be significantly lower in intrahepatic CCA cell lines when compared with normal liver tissues which is specific for this protein (Wang et al., 2006). Further study to investigate the expression of transferrin receptor in CCA cells is required to support this supposition.

In conclusion, our preliminary results demonstrate significantly different protein patterns of plasma from patients with late stage CCA with respect to the up- or down-regulation, and more importantly, intensity of the up-regulated protein AP1. Serum biomarkers may be a useful diagnostic or prognostic tool for CCA. Measurement of serum levels of at least CA19-9 together with AP1 were the minimum requirement to obtain prediction accuracy of greater than 80% in a battery test for diagnosis of CCA. However, in order to obtain high predictability of 100% or approaching, additional of at least one of the three liver function enzymes (ALP, AST, ALT) is required. Further evaluation in larger number of samples from patients with CCA and other liver diseases is needed to support the applicability of these models in predicting CCA in a clinical setting. In addition, in order to evaluate the clinical utility of these models in early diagnosis of CCA, study in animal model with various disease progression is needed.

Acknowledgments

The study was supported by the Thailand Office of Higher Education Commission, Ministry of Education (NRU and RG projects) and Thammasat University.(NRU and RG projects) and Thammasat University.

References

Bamrungphon W, Prempracha N, Bunchu N, et al (2007). A new mucin antibody/enzyme-linked lectin-sandwich assay of serum MUC5AC mucin for the diagnosis of cholangiocarcinoma. *Cancer Lett*, **247**, 301-8.

Boonla C, Sripa B, Thuwajit P, et al (2005). MUC1 and MUC5AC mucin expression in liver fluke-associated intrahepatic cholangiocarcinoma. *World J Gastroenterol*, **11**, 4939-46.

Carrell RW (1986). Alpha 1-Antitrypsin: molecular pathology, leukocytes, and tissue damage. *J Clin Invest*, **78**, 1427-31.

Chaijaroenkul W, Viyanant V, Mahavorasirikul W, Na-Bangchang K (2011). Cytotoxic activities of artemisinin derivatives against cholangiocarcinoma cell line (CL6). *Asian Pac J Cancer Prevent*, **12**, 1-6??.

Chalasani N, Baluyut A, Ismail A, et al (2000). Cholangiocarcinoma in patients with primary sclerosing cholangitis: a multicenter case-control study. *Hepatology*, **31**, 7-11.

Changbumrung S, Migasena P, Supawan V et al (1988). Serum protease inhibitors in opisthorchiasis, hepatoma, cholangiocarcinoma, and other liver diseases. *Southeast Asian J Trop Med Pub Hlth*, **19**, 299-305.

Chidwick K, Whichelow CE, Zhang Z, et al (1994). Relationship between alpha 1-antitrypsin inactivation and tumor necrosis factor alpha concentration in the synovial fluid of patients with rheumatoid arthritis. *Arthritis Rheum*, **37**, 1723-6.

Chignard M, Hazouard E, Renesto P, et al (1994). Plasma antiproteinase screen and neutrophil-mediated platelet activation. A major role played by alpha 1 antitrypsin. Biochem Biophys Acta, 1224, 433-40

Dbouk HA, Tawil A, Nasr F, Kandakarjian L, Abou-Merhi R (2007). Significance of CEA and VEGF as diagnostic markers of colorectal cancer in Lebanese patients. *Open Clin Cancer J*, **8**, 1-5.

Jinawath N, Chamgramol Y, Furukawa Y, et al (2006). Comparison of gene expression profiles between *Opisthorchis viverrini* and non-*Opisthorchis viverrini* associated human intrahepatic cholangiocarcinoma. *Hepatology*, **44**, 1025-38.

Gatto M, Bragazzi MC, Semeraro R, et al (2010). Cholangiocarcinoma: update and future perspectives. Dig Liver Dis, 42, 253-60.

Joo WA, Kang MJ, Son WK, et al (2003). Monitoring protein expression by proteomics: human plasma exposed to benzene. *Proteomics*, **3**, 2402-11.

Khan SA, Davidson BR, Goldin R, et al (2002). British Society of Gastroenterology. Guidelines for the diagnosis and treatment of cholangiocarcinoma: consensus document. *Gut*, **51**, VI1-9.

Khuhaprema T, Srivatanakul P: Liver and Bile Duct. In: Khuhaprema T, Srivatanakul P, H S, Wiangnon S, Sumitsawan Y, Attasara P (eds). Cancer in Thailand, 2007, Vol. IV, 1998–2000. Bangkok Medical Publisher, Bangkok pp. 36-8.

Kim MR, Kim CW (2007). Human blood plasma preparation for two-dimensional gel electrophoresis. *J Chromatogr B Analyt Technol Biomed Life Sci*, **849**, 203-10.

Lempinen M, Isoniemi H, Mäkisalo H, (2007). Enhanced detection of cholangiocarcinoma with serum trypsinogen-2 in patients with severe bile duct strictures. *J Hepatol*, **47**, 677-83.

May WS Jr, Cuatrecasas P (1985). Transferrin receptor: its biological significance. *J Membr Biol*, **88**, 205-15

Nehls O, Gregor M, Klump B (2004). Serum and bile markers for cholangiocarcinoma. *Semin Liver Dis*, **24**, 139-54.

Park JY, Lee KH (2007). Carcinoembryonic antigen and patterns of recurrence after curative resection of the colorectal cancer. *Hepato-gastroenterol*, **54**, 1966-9

Poley JR (1994). Malignant liver disease in alpha 1-antitrypsin deficiency. *Acta Paediatr*, **393**, 27-32

Raaf HN, Jacobsen DW, Savon S, Green R (1993). Serum transferrin receptor level is not altered in invasive adenocarcinoma of the breast. *Am J Clin Pathol*, 99, 232-7

Ramage JK, Donaghy A, Farrant JM, Iorns R, Williams R (1995). Serum tumor markers for the diagnosis of cholangiocarcinoma in primary sclerosing cholangitis. *Gastroenterol*, **108**, 865-9.

Sajid KM, Parveen R, Durr-e-Sabih, et al. (2007). Carcinoembryonic antigen (CEA) levels in hookah smokers, cigarette smokers and non-smokers. *J Pak Med Assoc*, **57**, 595-9.

Shterman N, Kupfer B, Moroz C (1991). Comparison of transferrin receptors, iron content and isoferritin profile in normal and malignant human breast cell lines. *Pathobiol*, **59**, 19-25

Singh NP, Lai H (2001). Selective toxicity of dihydroartemisinin and holotransferrin toward human breast cancer cells. *Life Sci*, **70**, 49-56

Srisomsap C, Sawangareetrakul P, Subhasitanont P, et al (2004). Proteomic analysis of cholangiocarcinoma cell line. *Proteomics*, **4**, 1135-44.

Thongprasert S (2005). The role of chemotherapy in cholangiocarcinoma. *Ann Oncol*, **16**, 93-6

Tonouchi A, Ohtsuka M, Ito H et al (2006). Relationship between pancreatic secretory trypsin inhibitor and early recurrence of intrahepatic cholangiocarcinoma following surgical resection. *Am J Gastroenterol*, **101**, 1601-10.

Uenishi T, Yamazaki O, Tanaka H, et al (2008). Serum cytokeratin 19 fragment (CYFRA21-1) as a prognostic factor in intrahepatic cholangiocarcinoma. *Ann Surg Oncol*, **5**, 583-9.

Wang AG, Yoon SY, Oh JH et al (2006). Identification of intrahepatic cholangiocarcinoma related genes by comparison with normal liver tissues using expressed sequence tags. *Biochem Biophys Res Commun*, **345**, 1022-32.