



# Prognostic subclass of intrahepatic cholangiocarcinoma by integrative molecular–clinical analysis and potential targeted approach

Keun Soo Ahn<sup>1</sup> · Daniel O'Brien<sup>2</sup> · Yu Na Kang<sup>3</sup> · Taofic Mounajjed<sup>4</sup> · Yong Hoon Kim<sup>1</sup> · Tae-Seok Kim<sup>1</sup> · Jean-Pierre A. Kocher<sup>2</sup> · Loretta K. Allotey<sup>5,6</sup> · Mitesh J. Borad<sup>7</sup> · Lewis R. Roberts<sup>5</sup> · Koo Jeong Kang<sup>1</sup> 

Received: 6 March 2019 / Accepted: 18 May 2019 / Published online: 18 June 2019  
© Asian Pacific Association for the Study of the Liver 2019

## Abstract

**Background** Although molecular characterization of iCCA has been studied recently, integrative analysis of molecular and clinical characterization has not been fully established. If molecular features of iCCA can be predicted based on clinical findings, we can approach to distinguish targeted treatment. We analyzed RNA sequencing data annotated with clinicopathologic data to clarify molecular-specific clinical features and to evaluate potential therapies for molecular subtypes.

**Methods** We performed next-generation RNA sequencing of 30 surgically resected iCCA from Korean patients and the clinicopathologic features were analyzed. The RNA sequences from 32 iCCA resected from US patients were used for validation.

**Results** Patients were grouped into two subclasses on the basis of unsupervised clustering, which showed a difference in 5-year survival rates (48.5% vs 14.2%,  $p = 0.007$ ) and similar survival outcome in the US samples. In subclass B (poor prognosis), both data sets were similar in higher carcinoembryonic antigen and cancer antigen 19-9 levels, underlying cholangitis, and bile duct-type pathology; in subclass A (better prognosis), there was more frequent viral hepatitis and cholangiolar-type pathology. On pathway analysis, subclass A had enriched liver-related signatures. Subclass B had enriched inflammation-related and TP53 pathways, with more frequent *KRAS* mutations. CCA cell lines with similar gene expression patterns of subclass A were sensitive to gemcitabine.

**Conclusions** Two molecular subtypes of iCCA with distinct clinicopathological differences were identified. Knowledge of clinical and pathologic characteristics can predict molecular subtypes, and knowledge of different subtype signaling pathways may lead to more rational, targeted approaches to treatment.

**Keywords** Cholangiocarcinoma · RNA sequence · Gene expression · Pathway · Mutation · *KRAS* · Target therapy

---

Keun Soo Ahn and Daniel O'Brien contributed equally.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s12072-019-09954-3>) contains supplementary material, which is available to authorized users.

✉ Mitesh J. Borad  
Borad.Mitesh@mayo.edu

✉ Lewis R. Roberts  
roberts.lewis@mayo.edu

✉ Koo Jeong Kang  
kjkang@dsmc.or.kr; kjkang@kmu.ac.kr

Extended author information available on the last page of the article

## Introduction

Intrahepatic cholangiocarcinoma (iCCA) is a malignant disease of the biliary tree and the second most common primary liver cancer, accounting for 10–15% of hepatobiliary neoplasms [1], but its incidence and associated mortality rate have recently increased [2]. iCCA is hard to diagnose in its early stages, so by the time they are diagnosed, most patients have an advanced stage. Since chemotherapy or targeted therapy is limited, managing this cancer remains challenging.

Recent molecular analyses have revealed several markers for poor prognosis including *KRAS* mutation and activation of the oncogenic pathways (e.g., HER2 and EGFR signaling) [3, 4]. At the molecular level, iCCA has subgroups that are associated with different outcomes [3–6]. However, the relation between molecular and clinicopathologic features

has not been fully known yet. Molecular characterization might be helpful for screening patients, predicting prognosis, and developing targeted therapies. However, applying molecular signatures in clinical settings is still not feasible because of the high cost for evaluating molecular characterization. If the molecular subtype of iCCA has specific clinical and pathologic features, the molecular subtype could be predicted from clinical findings, and the approach may influence treatment selection.

In this study, we aimed to answer three key questions: (1) Do different molecular subtypes of iCCA have distinct clinicopathologic features and can the molecular subtype be predicted on the basis of clinical findings? (2) Do molecular signatures determined in previous studies agree with each other and can they be validated? (3) Is there an effective subtype-specific treatment? We hypothesized that iCCA can be classified according to molecular characterization, that each molecular subtype would have corresponding clinical and pathologic features that are associated with different outcomes, and that treatment options may be determined according to the subtype of iCCA. To test the hypotheses, we used next-generation sequencing (NSG) to analyze RNA sequencing data obtained from frozen tissues of iCCA, annotated with clinical and pathologic data. Additionally, we evaluated the therapeutic potential of gemcitabine for treating cholangiocarcinoma (CCA) cell lines on the basis of molecular subtypes.

## Materials and methods

### Patients, tissue samples and clinical analysis

From March 2008 to August 2013, 56 patients with iCCA underwent curative-intent surgical resection at Keimyung University Dongsan Medical Center (DSMC), Daegu, Korea. For 42 cases, tumor tissues were available as frozen tissues. We excluded seven patients whose tissues did not qualify for RNA analysis. Additionally, five patients who did not have adenocarcinoma (combined HCC/iCCA or cholangiosarcoma) were excluded. Finally, surgical specimens from 30 patients with iCCA were enrolled. Surrounding normal liver tissues were available for 26 patients and were used as paired control tissues. The 56 fresh samples (30 primary tumor and 26 matched control tissues for surrounding normal liver) were saved frozen ( $-80^{\circ}\text{C}$ ) and used for analysis. Two pathologists reviewed all hematoxylin and eosin-stained slides and classified pathologic type into cholangiolar and bile duct type (Supplementary Fig. 1) [7]. As a validation data set, 32 curative, resected fresh tumor samples were collected from 2001 through 2014 from patients admitted to Mayo Clinic, MN, USA, with the same inclusion criteria

as those used for the Korean data set. The patients' clinical, pathologic and survival data were analyzed retrospectively.

### Statistics, bioinformatics, and molecular analysis

The methods of histologic, statistics, bioinformatics, and molecular analysis are described in Supplemental Methods online.

RNA sequencing data with clinical information are available in Gene Expression Omnibus(GEO; <http://www.ncbi.nlm.nih.gov/geo>) with accession number of GSE107943.

## Results

### Demographic characteristics and clinical outcomes

The DSMC (Korean) patient population consisted of 24 males and 6 females, with a median (range) age of 65 (49–79) years (Table 1). All of the Korean patients were Asian. Ten patients received palliative chemotherapy post-operatively after an iCCA recurrence. During the median (range) follow-up of 30.5 (9.3–115.7) months, 21 patients (70.0%) had disease recurrence, and 17 (56.7%) of them died. The 3- and 5-year overall survival rates were 45.8% and 27.1%, respectively, and the 3- and 5-year disease-free survival rates were 34.2% and 20.5%, respectively.

In the Mayo (USA) cohort used for validation, there were 15 male and 17 females, with a median (range) age of 64 (39–83) years. Twenty-three of the patients were of Caucasian race/ethnicity, and 9 were American African, Hispanic, or other. During the median (range) follow-up of 40.7 (3.7–128.3) months, disease recurred in 13 patients, and 17 patients died. Data regarding recurrence were unavailable for six patients, including for three patients who died post-operatively. Seven patients underwent chemotherapy: one neoadjuvant; one adjuvant; and five palliative. The 3- and 5-year overall survival rates were 47.1% and 32.5%, and the 3- and 5-year disease-free survival rates were 29.5% and 22.7%, respectively.

In both cohorts, all patients had mass-forming (MF) type of CCA.

### Identification of two distinct iCCA subclasses by RNA expression profiling

Unsupervised hierarchical clustering by Pearson correlation in all 56 tissues (30, tumor; 26, surrounding normal liver) of the DSMC data set showed two distinct clusters between tumor and normal tissues (Supplementary Fig. 2). A total of 2379 differentially expressed genes (DEGs) between tumor and adjacent normal liver tissues were identified and used for further analysis for identifying

**Table 1** Clinical and pathologic features according to molecular subclass

Characteristic	Korean (DSMC) data			US (Mayo) data			Combined Korean–US data		
	Subclass A ( <i>n</i> = 12)	Subclass B ( <i>n</i> = 18)	<i>p</i> value	Subclass A ( <i>n</i> = 21)	Subclass B ( <i>n</i> = 11)	<i>p</i> value	Subclass A ( <i>n</i> = 33)	Subclass B ( <i>n</i> = 29)	<i>p</i> value
Sex (male/ female)	9/3	15/3	0.93	8/13	7/4	0.26	17/16	22/7	0.07
Age, mean (SD)	63.1 (10.1)	67.3 (7.6)	0.20	64.0 (12.3)	59.2 (12.7)	0.31	63.7 (11.4)	64.3 (10.5)	0.85
CEA, mean (SD), ng/mL	1.7 (0.7)	9.4 (15.6)	0.07	1.2 (0.9)	6.6 (5.0)	0.21	1.6 (0.7)	8.7 (13.8)	0.03
CEA, no. (%), ng/mL			0.04			0.55			0.01
<5	11 (100) (91.7)	9 (56.2) (50)		2 (100) (9.5)	2 (40.0) (18.2)		13 (100) (39.4)	11 (52.4) (37.9)	
≥5	0 (0)	7 (43.8) (38.9)		0 (0)	3 (60.0) (27.3)		0 (0)	10 (47.6) (34.5)	
CA 19.9, mean (SD), U/mL	36.4 (44.6)	668.6 (1547.8)	0.11	65.0 (141.6)	4288.9 (6306.4)	0.13	52.4 (109.1)	1724.5 (3855.8)	0.045
CA 19.9, no. (%), U/mL			0.34			0.05			0.04
<37	8 (72.7) (66.7)	8 (47.1) (44.4)		8 (72.7) (38.0)	1 (14.3) (9.1)		16 (72.7) (48.5)	9 (37.5) (31.0)	
≥37	3 (27.3) (25.0)	9 (52.9) (50.0)		3 (23.8) (14.3)	6 (85.7) (54.5)		6 (27.3) (18.2)	15 (62.5) (51.7)	
Chemotherapy, no. (%)	4 (33.3)	6 (33.3)	>0.99	5 (23.8)	2 (18.2)	>0.99	9 (27.3)	7 (24.1)	
Tumor size (SD), cm	6.3 (3.9)	6.0 (2.9)	0.838	8.1 (3.9)	7.3 (3.3)	0.540	7.4 (3.9)	6.5 (3.1)	0.296
Differentiation			>0.99			0.18			0.14
Moderate	9 (75.0)	13 (76.5)		4 (19.0)	2 (18.2)		13 (39.3)	15 (53.6)	
Poor	3 (25.0)	5 (23.5)		16 (76.2)	6 (54.5)		19 (57.6)	11 (35.7)	
Undifferenti- ated	0 (0)	0 (0)		1 (4.8)	3 (27.3)		1 (3.1)	3 (10.7)	
Pathologic type, no. (%)			0.009			<0.001			<0.001
Bile duct	2 (6.1)	13 (72.2)		0 (0)	5 (45.4)		2 (6.1)	18 (62.1)	
Cholangiolar	10 (83.3)	5 (27.8)		21 (100)	5 (45.4)		31 (93.9)	10 (34.5)	
Uncertain	0 (0)	0 (0)		0	1 (9.0)		0 (0)	1 (3.4)	
Lymph node metastasis, no. (%)	2 (16.7)	5 (27.8)	0.792	4 (19.0)	6 (54.5)	0.098	6 (18.2)	11 (37.9)	0.146
AJCC stage, no. (%)			0.62			0.09			0.19
I	7 (58.3)	8 (44.4)		7 (33.3)	2 (18.2)		14 (42.4)	10 (34.5)	
II	3 (25.0)	3 (16.7)		7 (33.3)	1 (9.1)		10 (30.3)	4 (14.8)	
III	0 (0)	1 (5.6)		3 (14.3)	4 (36.4)		3 (9.1)	5 (17.2)	
IV	2 (16.6)	6 (33.3)		4 (19.0)	4 (36.4)		6 (18.2)	10 (34.5)	
Vascular inva- sion, no. (%)	5 (41.7)	7 (38.9)	1.00	4 (19.0)	1 (9.1)	0.82	9 (27.3)	8 (27.6)	1.00
Hepatitis B or C, no. (%)	4 (33.3)	0 (0)	0.04	5 (23.8)	0 (0)	0.21	9 (27.3)	0 (0)	0.007
Cholangitis, no. (%)	0 (0)	7 (38.9)	0.04	2 (9.5)	1 (9.1)	1.00	2 (6.1)	8 (27.6)	0.05
KRAS muta- tion, no. (%)	0 (0)	5 (27.8)	0.13	0 (0)	4 (36.4)	0.02	0 (0)	9 (31.0)	0.002

**Table 1** (continued)

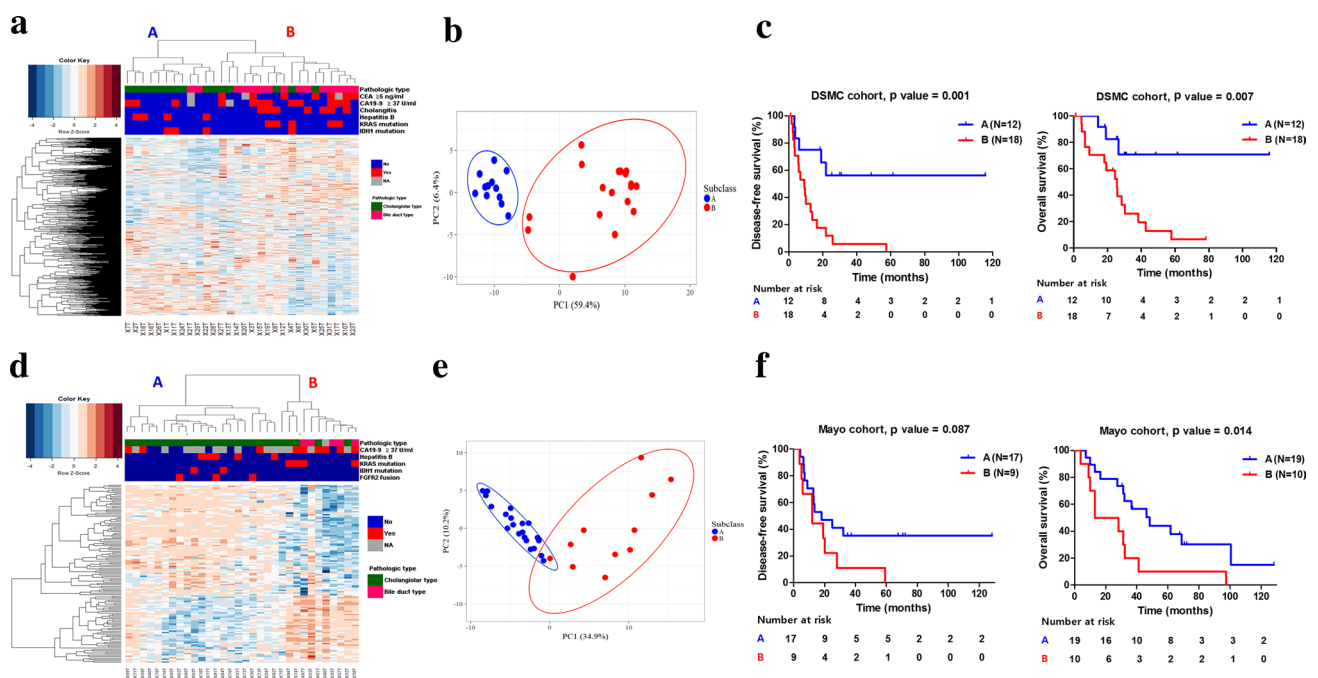
Characteristic	Korean (DSMC) data			US (Mayo) data			Combined Korean–US data		
	Subclass A (n = 12)	Subclass B (n = 18)	p value	Subclass A (n = 21)	Subclass B (n = 11)	p value	Subclass A (n = 33)	Subclass B (n = 29)	p value
<i>IDH1</i> mutation, no. (%)	3 (25.0)	1 (5.6)	0.32	2 (9.5)	0 (0)	0.77	5 (15.2)	1 (3.4)	0.26
<i>FGFR2</i> fusion, no. (%)	0 (0)	0 (0)	1.00	3 (14.2)	0 (0)	0.16	3 (9.1)	0 (0)	0.45

AJCC American Joint Committee on Cancer, CA cancer antigen, CEA carcinoembryonic antigen

tumor subtypes. In 30 tumor samples, unsupervised hierarchical cluster analysis separated the tumors into two distinct subclasses that were associated with survival (Fig. 1a, b). The 5-year disease-free and overall survival rates in subclass A (n = 12) were 56.2% and 70.7%, respectively, compared with 0% and 7.6% in subclass B (n = 18) (Fig. 1c).

A total of 148 DEGs between two subclasses of tumor and an FDR < 0.05 were identified (Supplementary Table 1). Of these 148 genes, 95 genes were up-regulated in the group with poor survival (subclass B) and 53 were

up-regulated in the group with better survival (subclass A). To validate tumor subtype, the 148 DEGs were applied to the validation group (Mayo cohort). By the same profiling with the Korean cohort, unsupervised cluster analysis revealed two distinct subgroups and successfully separated patients with poor survival (subclass B, n = 11) from those with better survival (subclass A, n = 21) (Fig. 1d–f). When the DSMC and Mayo cohorts were combined, there were significant differences in disease-free survival (p = 0.001) and overall survival (p = 0.001) according to subclass (Supplementary Fig. 3).



**Fig. 1** Tissue samples from both cohorts. Thirty samples from the Korean data set were separated into two subclasses with distinct clinical features by unsupervised hierarchical cluster (a) and principal component analysis (b). c Kaplan–Meier analyses showed worse disease-free and overall survival in subclass B than subclass A for the Korean patients. US tumor samples used for validation also revealed two distinct subclasses by unsupervised hierarchical cluster (d) and

principal component analysis (e). f Similarly, in the US cohort, the overall survival rates of subclass B were worse than that of subclass A. Data for disease recurrence were not available for six patients, three of whom died postoperatively; follow-up data for these patients were also unavailable. CA cancer antigen, CEA carcinoembryonic antigen, PC principal component

## Pathway analysis

Ingenuity pathway analysis (IPA) revealed that activation of cancer, cellular growth and inflammation-related pathways were predicted in subclass B, and metabolism-related pathways (lipid, carbohydrate and amino acid) were predicted by activation in subclass A (Supplementary Table 2).

According to gene set enrichment analysis (GSEA), several oncogenic pathways were enriched in subclass B, and over half of the significant gene sets (FDR < 0.05) or pathways were related to inflammatory and immune responses, including interferon alpha response, inflammatory response, IL-6, and the cytokine-mediated signaling pathway. KRAS-related signaling, TGF- $\beta$ , and TNF- $\alpha$  signaling were also enriched, and genes potentially repressed by p53 were up-regulated in subclass B. In contrast, subclass A was characterized by metabolic-related pathways, including bile acid and steroid metabolic process. Several liver-specific genes or liver cancer-related pathways were also more enriched in subclass A than in subclass B (Fig. 2, Supplementary Tables 3 and 4).

## mRNA variant analysis

Supplementary Table 5 shows variants that were finally selected. Mutations of *PMS2* (21%), *KRAS* (14.5%), *TP53*,

and *ERBB2* (9.7%) were frequently found in the DSMC and Mayo cohorts (Supplementary Fig. 4). *KRAS* mutations were seen in nine patients, all in subclass B ( $p=0.002$ ).

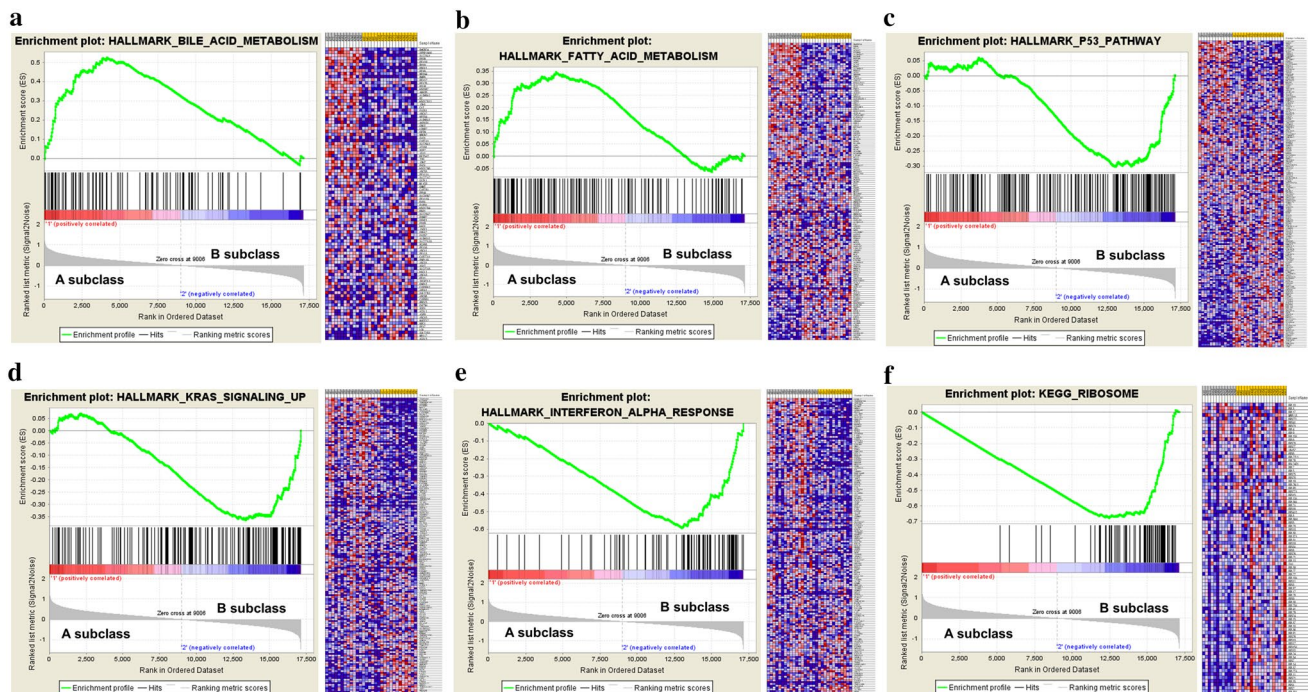
Because *KRAS* variant was significantly more frequent in subclass B, we validated the *KRAS* mutation by Sanger sequencing of two specific regions with available Korean samples. In the Korean data set, a *KRAS* mutation was found in five samples by RNA sequencing and all of them were matched with Sanger sequencing (Supplementary Fig. 5, Supplementary Table 6).

## RNA fusion

The RNA fusions in the DSMC and Mayo samples are listed in Supplementary Table 7. Three *FGFR2* fusions (*FGFR2-BICC1*, *FGFR2-CDYL*, *FGFR2-WAC*) were found in subclass A.

## Integrated molecular and clinical analysis

We then examined the association of each molecular subclass with clinical and pathologic features. In the DSMC cohort, hepatitis and cholangiolar-type pathologic findings were dominant in subclass A, and higher carcinoembryonic antigen (CEA) levels, more frequent bile duct pathologic type, and history of cholangitis were dominant in subclass B. In the Mayo cohort, patients with



**Fig. 2** Gene set enrichment analysis. **a** Subclass A was characterized by an enriched bile acid metabolism-related pathway and **b** a fatty acid metabolism pathway. **c** Subclass B was characterized by an

enriched P53-signaling pathway, **d** an enriched KRAS-signaling pathway, **e** inflammatory-response pathway and **f** ribosome-related pathway

a high cancer antigen (CA) 19-9, bile duct pathologic type, and advanced American Joint Committee on Cancer stage were more frequent in subclass B than subclass A (Table 1; Fig. 1).

In the combined DSMC–Mayo data, cholangiolar type was remarkably dominant in subclass A, and bile duct type was dominant in subclass B. Higher levels of CEA and CA 19-9 were observed in subclass B, and hepatitis B or C was more frequent in subclass A. Nine patients (31.0%) in subclass B had a *KRAS* mutation; no patients in subclass A had a *KRAS* mutation. Although an *IDH1* mutation was more common in subclass A than in subclass B (15.2% vs 3.4%), there was no statistical difference (Table 1).

On multivariate analysis in the combined DSMC–Mayo data set, subclass B was an independent risk factor ( $p < 0.05$ ) for disease-free survival, and advanced AJCC stage (III and IV) and subclass B were independent risk factors for overall survival (Table 2).

### Comparison of the present CCA classifier with other data sets

The recently published data on CCA gene expression was used to validate the molecular classification with gene

signatures by microarray analysis (GEO accession number GSE26566 [3], GSE33327 [5], GSE89749 [4]) and The Cancer Genome Atlas Project (TCGA) [6]. Their RNA expression data were downloaded from the TCGA portal (<https://portal.gdc.cancer.gov/>) and GEO, as well as clinical information. For comparison with clinicopathologic findings, we analyzed studies evaluating pathologic classification by Liao et al. [7] and Akita et al. [8]. For GSE26566 (cluster 1, good prognosis; cluster 2, poor prognosis) and GSE33327 (inflammation class, good prognosis; proliferation class, poor prognosis), gene lists for each subclass are available in supplementary tables. For TCGA (cluster 1 + 4 vs 2 + 3) and GSE89749 (cluster 1, poor prognosis vs 4, good prognosis), DEGs from the two subclasses according to sample-cluster annotations were identified.

In the GSEA, many DEGs of cluster 1 (GSE26566), cluster 4 (GSE89749), and cluster 1 + 4 (TCGA) which are good prognostic subclasses were overlapped with subclass A DEGs of the study (FDR < 0.001). Subclass B correlated with poor prognostic subclasses of cluster 2 (GSE26566), cluster 4 (GSE89749), and cluster 2 + 3 (TCGA) (FDR < 0.001) (Fig. 3a). Although inflammatory gene sets (good prognostic subclass) of GSE33327 were correlated with subclass A (FDR = 0.226), proliferative gene sets of GSE33327 did not correlate with either subclass A or B.

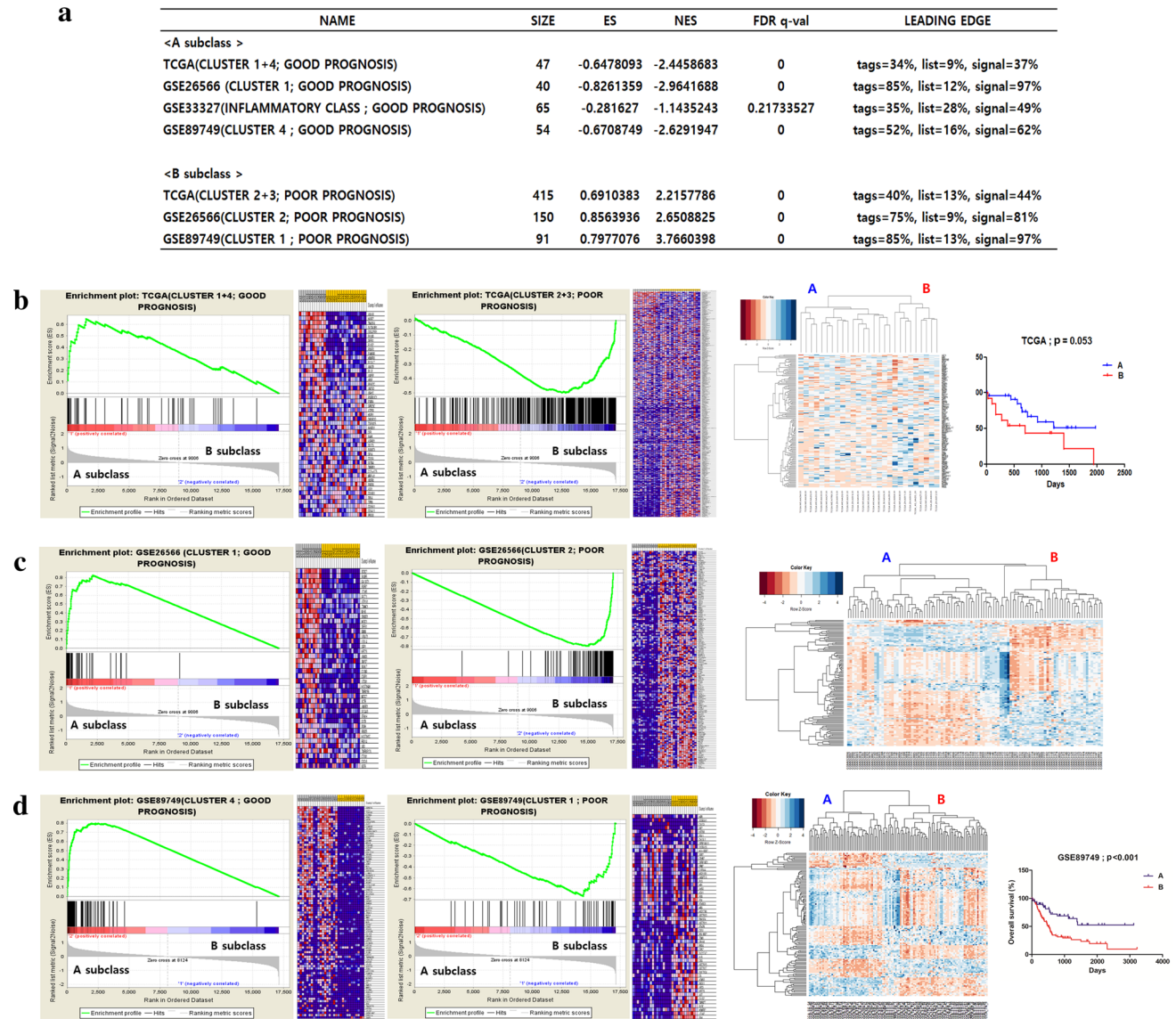
**Table 2** Risk factors for disease recurrence and overall survival in the combined Korean and US data sets

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	<i>p</i> value	HR (95% CI)	<i>p</i> value
Disease-free survival				
Sex, male	0.90 (0.46–1.69)	0.72		
High CEA, $\geq 5$ ng/mL	3.19 (1.28–7.97)	0.01		
High CA 19-9, $\geq 37$ U/mL	2.13 (1.00–4.52)	0.05		
AJCC stage III/IV	2.17 (1.13–4.20)	0.02		
Vascular invasion present	1.68 (0.83–3.40)	0.19		
Bile duct pathologic type	1.59 (0.83–3.03)	0.16		
Differentiation: poor or undifferentiated	1.12 (0.53–1.82)	0.79		
<i>KRAS</i> mutation	2.23 (0.97–5.17)	0.06		
Subclass B	2.88 (1.49–5.57)	0.002	8.43 (1.35–52.58)	0.02
Overall survival				
Sex, male	0.85 (0.45–1.60)	0.60		
High CEA, $\geq 5$ ng/mL	2.92 (1.20–7.11)	0.02		
High CA 19-9, $\geq 37$ U/mL	1.55 (0.74–3.24)	0.25		
Palliative chemotherapy	1.33 (0.68–2.63)	0.41		
AJCC stage III/IV	3.19 (1.67–6.12)	<0.001	9.76 (2.45–38.91)	0.001
Vascular invasion present	1.89 (0.92–3.90)	0.08		
Bile duct pathologic type	1.99 (1.05–3.79)	0.04		
Differentiation: poor or undifferentiated	0.82 (0.47–1.44)	0.51		
<i>KRAS</i> mutation	2.01 (0.91–4.42)	0.08		
Subclass B	2.95 (1.54–5.67)	0.001	7.45 (1.36–1.24)	0.02

AJCC American Joint Committee on Cancer, CA cancer antigen, CEA carcinoembryonic antigen, HR hazard ratio

Generally, the molecular features of subgroup according to the prognosis were similar among different data sets. Therefore, elevated levels of CEA and CA 19-9, bile duct-type (perihilar type) pathology, and a history of cholangitis can be specific features of molecular subgroup with a poor prognosis, whereas underlying hepatitis and cholangiolar-type (peripheral type) pathology are features of better prognostic molecular subgroup.

To validate tumor subtype of the present study, the 148 DEGs of the present study were applied to the three validation data sets (TCGA, GSE26566 and GSE89749). Unsupervised cluster analysis revealed two distinct subgroups in all data sets and successfully separated patients with poor survival (corresponding to subclass B of this study) from those with better survival (corresponding to subclass A) (Fig. 3b–d).



**Fig. 3 a** In the gene set enrichment analysis, gene sets of cluster 1 (GSE26566), cluster 1+4 (TCGA) and cluster 4 (GSE89749) were significantly correlated with gene sets of subclass A in the study (FDR<0.001). Subclass B in the study corresponded to that of cluster 2 (GSE26566), and cluster 2+3 (TCGA) and cluster 1 (GSE89749) (FDR<0.001). Corresponding gene sets of the study with TCGA (b), GSE26566 (c) and GSE89749 (d). When 148 sub-

class-specific genes of the study were applied to those in the data set, unsupervised cluster analysis revealed two distinct subgroups in all data sets and successfully separated patients with poor survival (corresponding to subclass B of the present study) from those with better survival (corresponding to subclass A of the present study) (survival information was not available in the GSE26566 data set)

## Association between gemcitabine sensitivity and subclasses

Among gemcitabine-related genes [9], *SLC28A1* and *SLC29A1* were associated with a gemcitabine response and were up-regulated in subclass A. *CDA* and *RRM2*, which are related to gemcitabine resistance or toxicity, were up-regulated in subclass B. All patients who had higher expressions of *SLC28A1* and *SLC29A1* and lower expressions of *RRM2* and *CDA* were in subclass A, and patients who had opposite expression pattern were in subclass B (Fig. 4a).

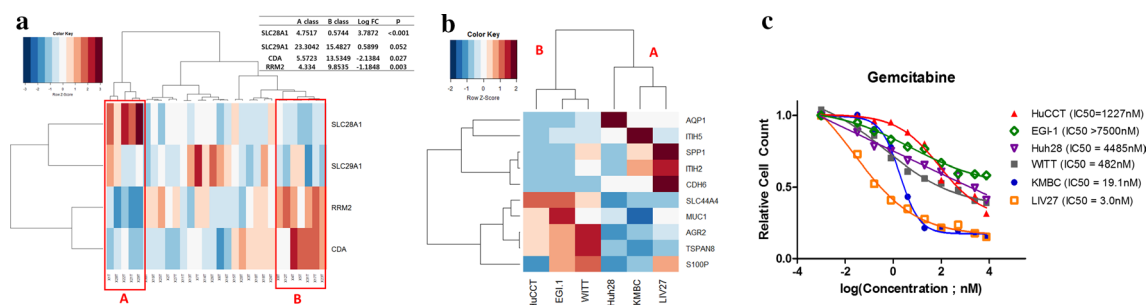
We selected ten genes that were up-regulated in each subclass on the basis of log-fold change and FDR (*SPP1*, *ITIH2*, *ITIH5*, *CDH6*, and *AQP1* for subclass A and *AGR2*, *S100P*, *TSPAN8*, *MUC1*, and *SLC44A4* for subclass B) and measured mRNA expression with quantitative polymerase chain reaction in 6 CCA cell lines. LIV27, KMBC, and Huh28 showed similar gene expression patterns as subclass A, and HuCCT, EGI-1, and WITT had similar features to subclass B (Fig. 4b).

We compared 6 CCA cell lines for sensitivity to gemcitabine. KMBC and LIV27 (gene expression patterns of subclass A) were sensitive to gemcitabine at concentrations between 3 and 7500 nM (Fig. 4c).

## Discussion

The two molecular subclasses in the study had specific clinical characterizations with distinct survival differences. In subclass A, patients with chronic viral hepatitis (B and C) were at a significant risk for CCA [10]. Most of these patients (93.9%) had cholangiolar-type (peripheral type) CCA and *CDH6*, which are commonly found in HCC [11, 12], and were up-regulated in subclass A. In the

aspect of mutation, five patients (15.2%) in subclass A had an *IDH1* mutation, which is usually found in patients with iCCA and not in those with extrahepatic CCA or gallbladder cancer [13, 14]. Moreover, all *FGFR2* fusions were in subclass A and they are typically found in iCCA, not extrahepatic or perihilar CCA [15]. In subclass B, eight patients (27.6%) had a history of cholangitis due to clonorchiasis, primary sclerosing cholangitis, or hepatolithiasis. In subclass B, bile duct type (perihilar type) of pathology was significantly more frequent than cholangiolar type. *S100P*, *AGR2*, and *MUC1* were up-regulated; these genes are related to bile duct-type CCA [7] and are frequently expressed in extrahepatic CCA [16, 17] or pancreatic cancer [18, 19]. Interestingly, serum levels of CEA and CA 19-9 were significantly higher in patients in subclass B than in those in subclass A. In subclass B, *CEACAM6*, one of the CEA family genes, was up-regulated. *CEACAM6* expression associated with serum CEA is common and a poor prognostic indicator [20]. CA19-9 is a cell surface antigen located on MUC genes-associated marker [21]. Several genes in the MUC family (*MUC1*, *MUC5B*, and *MUC13*), which have roles in early detection and prognostic prediction for pancreatic cancer [22], were up-regulated in the subclass B. Therefore, CEA and CA 19-9 may be useful for screening certain patients in subclass B. A *KRAS* mutation was found only in class B (31%). Usually, *KRAS* mutation is more frequent in extrahepatic CCA (~40%) and pancreatic cancer (~90%) than in iCCA (10–20%) [14, 23]. Therefore, the study revealed that MF type of iCCA has two molecular characterizations with distinct clinical features: one is peripheral type of iCCA with risk factor of hepatitis (subclass A) and another with pathological and clinical features with perihilar or extrahepatic CCA (subclass B).



**Fig. 4** Expression of gemcitabine-response genes (*SLC28A1* and *SLC29A1*) and resistance-related genes (*CDA* and *RRM2*) in the Korean Cohort. **a** Gemcitabine-response genes were up-regulated in subclass A and resistance-related genes in subclass B. **b** For ten genes that were up-regulated in subclass A (*SPP1*, *ITIH2*, *ITIH5*, *CDH6*, and *AQP1*) or subclass B (*AGR2*, *S100P*, *TSPAN8*, *MUC1*, and *SLC44A4*), mRNA expression was measured with quantitative poly-

merase chain reaction in six cholangiocarcinoma cell lines (HuCCT, EGI-1, WITT, LIV27, KMBC, and Huh28). LIV27, KMBC, and Huh28 had similar gene expression patterns as subclass A, and HuCCT, EGI-1, and WITT had features of subclass B. **c** KMBC and LIV27 were sensitive to gemcitabine at concentrations between 0.03 and 7500 nM



iCCA is believed to originate from biliary epithelial cells in the intrahepatic bile duct or from hepatic progenitor cells. Hepatic bipotent progenitor cells along the small intrahepatic bile duct may be able to differentiate not only into hepatocytes but also into cholangiocytes, which can lead to iCCA [24, 25]. Therefore, two types of iCCA may develop: one originating from hepatic stem cell-derived lineages with stem-like molecular characteristics similar to those in HCC or combined HCC–CCA and the other originating from biliary tree progenitor stem cell-derived cholangiocytes found along the large intrahepatic bile duct with similar characteristics to those of perihilar or extrahepatic CCA [17, 26]. The clinical and molecular integration analysis showed that the results for the MF type of iCCA were divided, with different molecular pathways and clinical outcomes. Abundant liver-specific genes and HCC-related signaling, *IDH1* mutation, *FGFR2* fusion, underlying hepatitis, and cholangiolar types of pathology were characteristic of the group with a good prognosis (subclass A); in this group, iCCA may arise from the smaller intrahepatic bile duct or bipotential hepatic progenitor cells within portal areas. Meanwhile, the poor prognosis of the group (subclass B) characterized by *KRAS* mutations, underlying cholangitis, high serum CEA and CA 19-9, and bile duct-type pathology, likely results from malignant transformation of cholangiocytes that line the larger bile ducts (Supplementary Fig. 6) [7]. *KRAS* and P53-related pathways have strong impact on poor prognosis in many cancer including cholangiocarcinoma [3]; therefore, they might contribute to the worse prognosis in the B subclass than to the metabolic-related pathway in the A subclass. Most of recent cancer genomic studies focused on molecular findings and could not provide clear clinical relevance, especially for pathologic findings according to molecular subtypes. The molecular characteristics in the two subclasses could be validated to other public data sets and showed different survival outcomes as in this study (Fig. 3). Each of the two subclasses in the validation sets has similar molecular characteristics to those of this study. Although clinical information in validation sets is limited, subclass-specific clinical and pathological characteristics of this study can be applied also to validation sets. This study has the novelty of linking the gap between the clinicopathological [7] and molecular features [3, 5, 6] of iCCA, and from the clinicopathologic characteristics, the molecular subtypes could be predicted. Especially, the study revealed that the two different pathologic types are important features for predicting the molecular subtype.

CCA is highly resistant to chemotherapy. This study showed that patients with certain molecular characteristics may benefit from gemcitabine. On the basis of mRNA expression, some, but not all, patients in subclass A were expected to respond to gemcitabine because of a higher

expression of gemcitabine-response genes (*SLC28A1* and *SLC29A1*) and a lower expression of gemcitabine-resistant or toxicity-related genes (*CDA* and *RRM2*). LIV27 and KMBC cell lines, which had similar molecular features as subclass A and responded well to gemcitabine. Low initial serum levels of CEA or CA 19-9, a clinical feature of subclass A, more likely resulted from a good response to gemcitabine in patients with advanced CCA [27]. For subclass B, the presence of a *KRAS* mutation and high expression of *S100P* and *AGR2* are known to be important factors in resistance to gemcitabine [18, 28]. Therefore, gemcitabine-based chemotherapy can be considered for patients with advanced CCA that has similar molecular features to that of subclass A. Since this study could not show in vivo the clinical results of gemcitabine sensitivity, animal study or clinical study in non-resectable CCA or neoadjuvant chemotherapy is necessary for evaluation of the benefits of gemcitabine. Tyrosine kinase inhibitors (TKIs), especially erlotinib, are another potential therapeutic option for patients with advanced biliary cancer who have activated HER2 and EGFR signals, with overexpression of EGFR [29]. In the study, sorafenib was not sensitive in all CCA cell lines (data not shown). However, in the cell line studies of GSE26566, WITT, and HuCCT, which have similar molecular features to those of subclass B, were effective for lapatinib and trastuzumab [3]. Based on this study and the GSE26566 study [3], the effectiveness of combination therapy (gemcitabine for subclass A and TKIs for subclass B) should be evaluated in the future.

There are a few limitations in the present study. Among the three types of iCCA [MF type, periductal infiltrating (PI), and intraductal growing (IG) type], all patients were of MF type. PI and IG types are usually seen in extrahepatic CCA and have homogeneous histological features, bile duct type [7]. Therefore, most of the PI and IG types may correspond to subclass B. On the contrary, MF is the most frequent form of iCCA, accounting for more than 70% [30] and has heterogeneous histologic features [7]. It implies two different origins as seen in the study. Therefore, focusing on the MF type might be efficient for revealing the molecular–clinical relevance; however, further study for PI and IG is necessary. In addition, the study had limited case number. Although we used other public data for validation, there is lack of clinical and pathological information and further analysis is needed for validating pathological–molecular relevance.

In summary, we used clinicopathologic and molecular integration analysis to identify two molecular classifications of iCCA that have distinct clinical, pathologic, biologic, and prognostic differences. With this knowledge, molecular subtypes can be predicted. In addition, different subtype signaling pathways may lead to more rational, targeted approaches to treatment.

**Acknowledgements** The authors thank Li Li, Dehai Wu, Katsuyuki Miyabe, Tao Song, Ning Zhang, Jianbo Huang, Shaoqing Wang, Lin Yang, and Amy S Mauer (Mayo clinic, Rochester, MN) for their help in the experiment. The biospecimens and data used for this study were provided by the Biobank of Keimyung University Dongsan Medical Center, member of the Korea Biobank Network. This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea government (MSIP) (No. 2018R1C1B3004435). This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korea Government (MSIP) (No. 2014R1A5A2010008).

## Compliance with ethical standards

**Conflict of interest** Keun Soo Ahn, Daniel O'Brien, Yu Na Kang, Taofic Mounajjed<sup>4</sup>, Yong Hoon Kim, Tae-Seok Kim, Jean-Pierre A. Kocher, Loretta K. Allotey, Mitesh J. Borad, Lewis R. Roberts, and Koo Jeong Kang declare that they have no conflict of interest.

**Availability of data and materials** RNA sequencing data of this manuscript is registered at Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) with GEO accession number of GSE107943

**Ethics approval** This study was approved by institutional review boards at both institutions (Keimyung University Dongsan Medical Center: IRB No. 2014-12-066, Mayo clinic: IRB No. 16-007369) and included secondary use of human-derived materials.

**Informed consent** Informed consent for the human-derived materials was obtained from all patients before surgery.

## References

- Blechacz BR, Gores GJ. Cholangiocarcinoma. *Clin Liver Dis* 2008;12:131–150, ix
- Patel T. Worldwide trends in mortality from biliary tract malignancies. *BMC Cancer* 2002;2:10
- Andersen JB, Spee B, Blechacz BR, Avital I, Komuta M, Barbour A, et al. Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. *Gastroenterology* 2012;142:1021–1031 e1015
- Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. *Cancer Discov* 2017;7:1116–1135
- Sia D, Hoshida Y, Villanueva A, Roayaie S, Ferrer J, Tabak B, et al. Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. *Gastroenterology* 2013;144:829–840
- Farshidfar F, Zheng S, Gingras MC, Newton Y, Shih J, Robertson AG, et al. integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-mutant molecular profiles. *Cell Rep* 2017;18:2780–2794
- Liau JY, Tsai JH, Yuan RH, Chang CN, Lee HJ, Jeng YM. Morphological subclassification of intrahepatic cholangiocarcinoma: etiological, clinicopathological, and molecular features. *Mod Pathol* 2014;27:1163–1173
- Akita M, Fujikura K, Ajiki T, Fukumoto T, Otani K, Azuma T, et al. Dichotomy in intrahepatic cholangiocarcinomas based on histologic similarities to hilar cholangiocarcinomas. *Mod Pathol* 2017;30:986–997
- Alvarellos ML, Lamba J, Sangkuhl K, Thorn CF, Wang L, Klein DJ, et al. PharmGKB summary: gemcitabine pathway. *Pharmacogenet Genom* 2014;24:564–574
- Kuper H, Ye W, Broome U, Romelsjo A, Mucci LA, Ekblom A, et al. The risk of liver and bile duct cancer in patients with chronic viral hepatitis, alcoholism, or cirrhosis. *Hepatology* 2001;34:714–718
- Zheng DL, Zhang L, Cheng N, Xu X, Deng Q, Teng XM, et al. Epigenetic modification induced by hepatitis B virus X protein via interaction with de novo DNA methyltransferase DNMT3A. *J Hepatol* 2009;50:377–387
- Sun HY, Li Y, Guo K, Kang XN, Sun C, Liu YK. Identification of metastasis-related osteopontin expression and glycosylation in hepatocellular carcinoma. *Zhonghua Gan Zang Bing Za Zhi* 2011;19:904–907
- Borger DR, Tanabe KK, Fan KC, Lopez HU, Fantin VR, Straley KS, et al. Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. *Oncologist* 2012;17:72–79
- Churi CR, Shroff R, Wang Y, Rashid A, Kang HC, Weatherly J, et al. Mutation profiling in cholangiocarcinoma: prognostic and therapeutic implications. *PLoS One* 2014;9:e115383
- Borad MJ, Gores GJ, Roberts LR. Fibroblast growth factor receptor 2 fusions as a target for treating cholangiocarcinoma. *Curr Opin Gastroenterol* 2015;31:264–268
- 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;22:683–691
- Komuta M, Govaere O, Vandecaveye V, Akiba J, Van Steenberghe W, Verslype C, et al. Histological diversity in cholangiocellular carcinoma reflects the different cholangiocyte phenotypes. *Hepatology* 2012;55:1876–1888
- Arumugam T, Ramachandran V, Logsdon CD. Effect of cromolyn on S100P interactions with RAGE and pancreatic cancer growth and invasion in mouse models. *J Natl Cancer Inst* 2006;98:1806–1818
- Dumartin L, Whiteman HJ, Weeks ME, Hariharan D, Dmitrovic B, Iacobuzio-Donahue CA, et al. AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. *Cancer Res* 2011;71:7091–7102
- Beauchemin N, Arabzadeh A. Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. *Cancer Metastasis Rev* 2013;32:643–671
- Akagi J, Takai E, Tamori Y, Nakagawa K, Ogawa M. CA19-9 epitope a possible marker for MUC-1/Y protein. *Int J Oncol* 2001;18:1085–1091
- Ringel J, Lohr M. The MUC gene family: their role in diagnosis and early detection of pancreatic cancer. *Mol Cancer* 2003;2:9
- Yang S, Che SP, Kurywchak P, Tavormina JL, Gansmo LB, Correa-de-Sampaio P, et al. Detection of mutant KRAS and TP53 DNA in circulating exosomes from healthy individuals and patients with pancreatic cancer. *Cancer Biol Ther* 2017;18:158–165
- Moeini A, Sia D, Bardeesy N, Mazzaferro V, Llovet JM. Molecular pathogenesis and targeted therapies for intrahepatic cholangiocarcinoma. *Clin Cancer Res* 2016;22:291–300
- Kordes C, Haussinger D. Hepatic stem cell niches. *J Clin Invest* 2013;123:1874–1880
- Cardinale V, Bragazzi MC, Carpino G, Torrice A, Fraveto A, Gentile R, et al. Cholangiocarcinoma: increasing burden of classifications. *Hepatobiliary Surg Nutr* 2013;2:272–280
- Lee BS, Lee SH, Son JH, Jang DK, Chung KH, Paik WH, et al. Prognostic value of CA 19-9 kinetics during gemcitabine-based chemotherapy in patients with advanced cholangiocarcinoma. *J Gastroenterol Hepatol* 2016;31:493–500

28. Ramachandran V, Arumugam T, Wang H, Logsdon CD. Anterior gradient 2 is expressed and secreted during the development of pancreatic cancer and promotes cancer cell survival. *Cancer Res* 2008;68:7811–7818
29. Borad MJ, Champion MD, Egan JB, Liang WS, Fonseca R, Bryce AH, et al. Integrated genomic characterization reveals novel, therapeutically relevant drug targets in FGFR and EGFR pathways in sporadic intrahepatic cholangiocarcinoma. *PLoS Genet* 2014;10:e1004135
30. Yamamoto J, Kosuge T, Takayama T, Shimada K, Makuuchi M, Yoshida J, et al. Surgical treatment of intrahepatic cholangiocarcinoma: four patients surviving more than five years. *Surgery* 1992;111:617–622

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Keun Soo Ahn<sup>1</sup> · Daniel O'Brien<sup>2</sup> · Yu Na Kang<sup>3</sup> · Taofic Mounajjed<sup>4</sup> · Yong Hoon Kim<sup>1</sup> · Tae-Seok Kim<sup>1</sup> · Jean-Pierre A. Kocher<sup>2</sup> · Loretta K. Allotey<sup>5,6</sup> · Mitesh J. Borad<sup>7</sup> · Lewis R. Roberts<sup>5</sup> · Koo Jeong Kang<sup>1</sup> 

Keun Soo Ahn  
ahnksmd@gmail.com

Daniel O'Brien  
obrien.daniel@mayo.edu

Yu Na Kang  
yunakang@dsmc.or.kr

Taofic Mounajjed  
Mounajjed.Taofic@mayo.edu

Yong Hoon Kim  
hbpsurgery@gmail.com

Tae-Seok Kim  
gskim80094@dsmc.or.kr

Jean-Pierre A. Kocher  
Kocher.JeanPierre@mayo.edu

Loretta K. Allotey  
Lorettallotey@gmail.com

<sup>1</sup> Department of Surgery, Dongsan Medical Center, School of Medicine, Keimyung University, 56 Dalseong-ro, Jung-gu, Daegu 700-712, Republic of Korea

<sup>2</sup> Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN 55905, USA

<sup>3</sup> Department of Pathology, Dongsan Medical Center, Keimyung University, Jung-gu, 42601, Daegu, Republic of Korea

<sup>4</sup> Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA

<sup>5</sup> Department of Gastroenterology and Hepatology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA

<sup>6</sup> University of Minnesota Medical School, Minneapolis, MN 55455, USA

<sup>7</sup> Division of Hematology and Medical Oncology, Mayo Clinic, 13400 E Shea Blvd, Scottsdale, AZ 85259, USA