BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Supplemental Information

Integrative analysis reveals early and distinct genetic and epigenetic changes in intraductal papillary and tubulopapillary cholangiocarcinogenesis

Benjamin Goeppert, Damian Stichel, Reka Toth, Sarah Fritzsche, Moritz A. Loeffler, Anna Melissa Schlitter, Olaf Neumann, Yassen Assenov, Monika Nadja Vogel, Arianeb Mehrabi, Katrin Hoffmann, Bruno Köhler, Christoph Springfeld, Dieter Weichenhan, Christoph Plass, Irene Esposito, Peter Schirmacher, Andreas von Deimling, Stephanie Roessler

Table of content

Supplemental Materials and Methods **2**

Supplemental Figures **5**

Figure S1: Representative images of immunohistochemical staining of IPNB and ITPN with the neuroendocrine markers synaptophysin and chromogranin A. 5

Figure S2: Kaplan-Meier survival analysis of invasive CCA cases with or without IPNB/ITPN precursor lesions. 6

Figure S3: Schematic diagram depicting the number of cases with intrahepatic IPNB or ITPN and corresponding large-duct or small-duct type iCCA. 7

Figure S4: Representative images of intrahepatic IPNB and ITPN with associated large-duct iCCA. 9

Figure S5: Representative images of intrahepatic IPNB and ITPN with associated small-duct iCCA. 11

Figure S6: Unsupervised clustering of DNA methylation profiles. 12

Supplemental Tables **13**

Table S1: Patient characteristics of ITPN of the pancreas (ITPN-P) cases (N=9). 13

Table S2: Subclassification of iCCA. 14

Table S3: Underlying hepatobiliary diseases of the CCA study cohort (N=54). 15

Table S4: Japan-Korea classification of 44 IPNB cases included in this study. 16

Table S5: Comparison of mutation frequencies of this study and published data sets*. 17

Table S6: Genetic alterations in IPNB (N=39) and ITPN (N=9) cases. 18

Table S7: Genetic alterations in IPNB/ITPN (N=26) and paired invasive CCA (N=26) cases. 18

Supplemental References **19**

Goeppert B, *et al. Gut* 2021;0:1–11. doi: 10.1136/gutjnl-2020-322983

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Supplemental Materials and Methods

Tissue microarray fabrication, diagnostic criteria and immunohistochemical analyses

The morphological classification into IPNB and ITPN and histomorphological subtyping was supported by immunohistochemical analyses of MUC1, MUC2, MUC5AC, MUC6 and CDX2. ITPN were diagnosed using stringent criteria [1, 2]. All ITPN showed a predominant tubulopapillary or trabecular architecture and were MUC5AC negative by immunohistochemical analysis. In addition, neuroendocrine neoplasias were ruled out by immunohistochemistry of synaptophysin and chromogranin A. Intraductal lesions were evaluated by immunohistochemical staining using a tissue microarray (TMA). The TMA was fabricated including all 54 biliary intraductal precursor lesions (Table 1). Tissue cores (duplicates, each 1 mm in diameter) of the marked regions were punched out from the donor blocks and embedded into a new paraffin array block using an automated tissue microarrayer (TMA Grand Master Fa. Sysmex, Germany). For immunohistochemistry (IHC), 3 µm sections of the TMA were cut, deparaffinized and rehydrated. For heat-induced epitope retrieval Ultra CC1 (Cell Conditioning Solution, Ventana Medical Systems, Tucson, AZ, USA) was used. After blocking of endogenous peroxidase, slides were incubated with primary antibodies anti-MUC1/EMA (M0613; Agilent Dako, Santa Clara, CA, USA), anti-MUC2 (760-4388; Ventana Medical Systems), anti-MUC5AC (sc-33667, clone CLH2; Santa Cruz Biotechnology, Heidelberg, Germany), anti-MUC6 (sc-33668, clone CLH5; Santa Cruz Biotechnology), anti-CDX2 (ready-to-use antibody, clone EPR2764Y; Ventana Medical Systems), anti-synaptophysin (ready-to-use antibody, clone MRQ-40; Ventana Medical Systems), and anti-chromogranin A (ready-to-use antibody, clone LK2H10; Ventana Medical Systems). Biotin-free OptiView DAB IHC Detection Kit (Ventana Medical Systems) including OptiView Universal Linker, OptiView HRP Multimer and DAB-Chromogen was used. Finally, the slides were counterstained with hematoxylin.

Genomic DNA isolation

Genomic DNA of the formalin-fixed paraffin-embedded (FFPE) samples was extracted using the AllPrep DNA/RNA FFPE Kit (Qiagen, Hilden, Germany) as recommended by the manufacturer with following modifications: After addition of xylene, samples were incubated at 56°C for 2 min followed by two ethanol washes. The first proteinase K digestion was done with 20 µl at 56°C for 30 min. The DNA was eluted twice with 30 µl H_2O.

Sequencing, Variant Calling and Annotation

To analyze the samples for genetic variations, a custom gene panel for massive parallel next generation sequencing (NGS) was used as published previously [3]. This panel consists of 285 amplicons covering 165 exons within 40 genes frequently mutated in biliary tract cancers.

The concentration of DNA was measured fluorimetrically using the QuBit 2.0 DNA high sensitivity kit (Thermo Fisher Scientific, Waltham, USA) and by a qPCR assay (RNAseP assay, Thermo Fisher Scientific) to quantify the amount of amplifiable DNA. Amplicon library preparation was performed with the Ion AmpliSeq Library Kit v2.0 (Thermo Fisher Scientific) and Ion Xpress Barcode Adapters, (Life Technologies, Carlsbad, USA) using 10 ng of DNA. Briefly, the DNA was mixed with the primer pool, containing all primers for generating the 284 amplicons and the AmpliSeq HiFi Master Mix and transferred to a PCR cycler (BioRad, Munich, Germany). After PCR amplification, primer end sequences of the libraries were partially digested using FuPa reagent, followed by the ligation of barcoded sequencing adapters (Ion Xpress Barcode

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Adapters, Life Technologies). The final library was purified using AMPure XP magnetic beads (Beckman Coulter, Krefeld, Germany) and quantified using qPCR (Ion Library Quantitation Kit, Thermo Fisher Scientific) on a StepOne qPCR machine (Thermo Fisher Scientific). The individual libraries were diluted to a final concentration of 100 pM and 8-10 libraries were pooled and processed to library amplification on Ion Spheres using the Ion PGM™ Hi-Q OT2 200 Kit. Not enriched libraries were quality-controlled using Ion Sphere quality control measurement on a QuBit instrument. After library enrichment (Ion OneTouch ES, Thermo Fisher Scientific), the library was processed for sequencing using the Ion Torrent 200 bp sequencing v2 chemistry and the barcoded libraries were loaded onto a chip. Sequencing was conducted on an Ion PGM™ machine (Thermo Fisher Scientific). Our way of pooling 8 samples on a 318 chip resulted in a mean coverage of 2000-fold per amplicon. For the fusion analysis, 10 ng RNA was reverse transcribed with the SuperScript™ VILO™ cDNA Synthesis Kit according to the manufacturer's instruction (Thermo Fisher Scientific).

Data analysis was performed using the Ion Torrent Suite Software (version 5.0.2). After base calling, the reads were aligned against the human genome (hg19) using the TMAP algorithm of the Torrent Suite. Variant calling was performed with the variant caller plugin (version 5.0.2.1) of the Torrent Suite Software and the IonReporter package using a corresponding bed-file containing the coordinates of the amplified regions. Only variants with an allele frequency > 5% based on a minimum coverage > 100 reads were taken into account. Variant annotation was performed using ANNOVAR [4]. Annotations included information about nucleotide and amino acid changes of RefSeq annotated genes, COSMIC and dbSNP entries and detection of possible splice site mutations. For data interpretation and verification, the aligned reads were visualized using the IGV browser (Broad Institute, Cambridge, USA) [5].

Out of the 54 precursor lesions of the bile duct, 6 samples failed the NGS quality control and out of the 34 invasive CCA tumor tissues 8 samples failed due to poor DNA quality. Therefore, 48 precursor and 26 paired invasive CCA tumor tissues were included in the sequencing analysis (Figure 1).

DNA methylation analysis using Infinium MethylationEPIC array data processing

DNA methylation profiles were determined by the Genomics and Proteomics Core Facility (DKFZ Heidelberg) using the Infinium MethylationEPIC BeadChip assay (Illumina, San Diego, CA, USA). The assay determined DNA methylation levels and allowed for the quantitative measurement of copy number alterations. FFPE tissue-derived genomic DNA was treated with bisulfite using the EZ DNA Methylation kit (D5002, Zymo Research, Irvine, CA, USA). Infinium MethylationEPIC arrays were performed according to the manufacturer's instructions and scanned on an Illumina HiScan. The assay determined DNA methylation levels at >850,000 CpG sites and allowed for the quantitative measurement of copy number alterations.

Out of the 54 precursor lesions of the bile duct, 3 samples failed the quality control and out of the 34 invasive CCA tumor tissues 7 samples failed due to poor DNA quality. Therefore, 51 precursor and 27 corresponding invasive CCA tumor tissues were included in the DNA methylation analysis (Figure 1).

The DNA methylation array data were processed with the R/Bioconductor package minfi (version 1.22). The t-SNE plot was computed via the R package Rtsne using the 10,000 most variable CpG sites according to standard deviation, 3000 iterations and a perplexity value of 10. Copy-number variations were calculated from the IDAT files using the R/Bioconductor package

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

conumee including an additional baseline correction (https://github.com/dstichel/conumee and http://bioconductor.org/packages/release/bioc/html/conumee.html).

For the MeDeCom analysis, DNA methylation was processed using the RnBeads R package [6]. Normalization and background correction was performed was with “bmiq” and “methylumi.noob” as implemented in the package. Probes measuring CpG sites with overlapping SNPs (minor allele frequency > 1%) were removed, as were the ones with more than 50% missing measurements or bad quality. Reference free deconvolution was performed using MeDeCom [7] and DecompPipeline [8] packages. Confounding factors, such as age and gender were adjusted using Independent Component Analysis (ICA) as implemented in the DecompPipeline package with the following parameters: “alpha.fact”=1e-5, “ntry”=10, “nmax”=10. The most variable 20,000 sites were selected for the analysis. MeDeCom was run setting k=2:7 and LAMBDA_GRID=1e-2, 1e-6. After optimization, a model with 4 LMCs and lambda=0.01 was selected. The same 20,000 sites were used to deconvolute the cases with their corresponding invasive samples. In this case, after optimization 4 LMCs and lambda=0.1 were selected. The methylation heatmaps were created using the ComplexHeatmap Bioconductor package [9]. After removing additional SNPs in close (+/-3bps) vicinity of the respective CpG sites as implemented in RnBeads, the 10,000 most variable CpG sites were selected.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Supplemental Figures

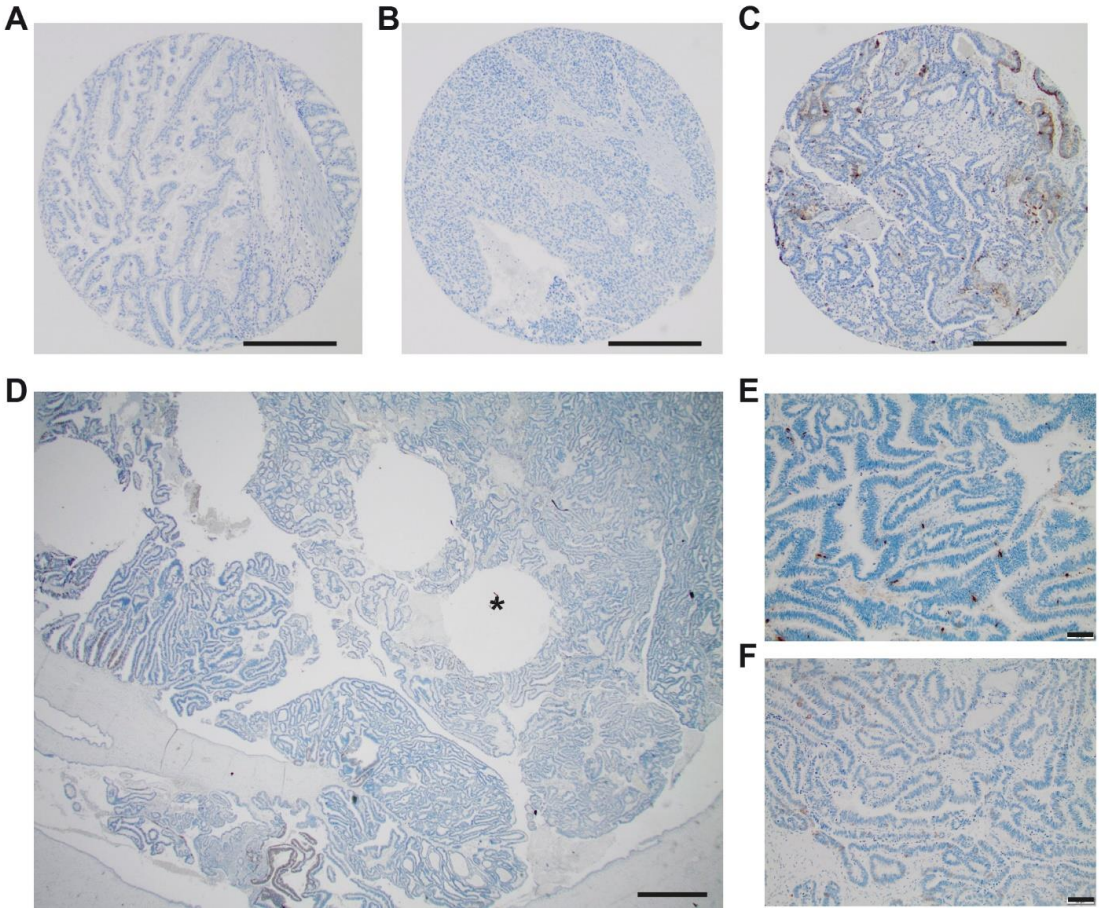


Figure S1: Representative images of immunohistochemical staining of IPNB and ITPN with the neuroendocrine markers synaptophysin and chromogranin A. Immunohistochemistry for the neuroendocrine markers, i.e. synaptophysin and chromogranin A, was performed for all 54 IPNB/ITPN using the tissue microarray technique (TMA). Forty-eight out of 54 samples (88.9%) were negative for both synaptophysin and chromogranin A. Six out or 54 IPNB/ITPN (11.1%) showed a focal, partly inconsistent and faint staining of synaptophysin and/or chromogranin A. These 6 cases were further evaluated by full-section staining which confirmed that less than 10% of cells were positive for synaptophysin or chromogranin A. **(A)** Representative image of chromogranin A immunohistochemistry of an IPNB and **(B)** an ITPN negative for both neuroendocrine markers (synaptophysin and chromogranin A), original magnification: 20x. **(C)** TMA dot of an IPNB case showing focal positivity for chromogranin A which was further analyzed by **(D)** full-section chromogranin A immunohistochemistry (* marks one of the punch-cavities after TMA construction. **(E and F)** Higher magnification of the same IPNB stained for chromogranin A and synaptophysin, respectively, original magnification: 100x. Scale bar in A-C: 200 µm, D: 500 µm, E-F: 50 µm.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

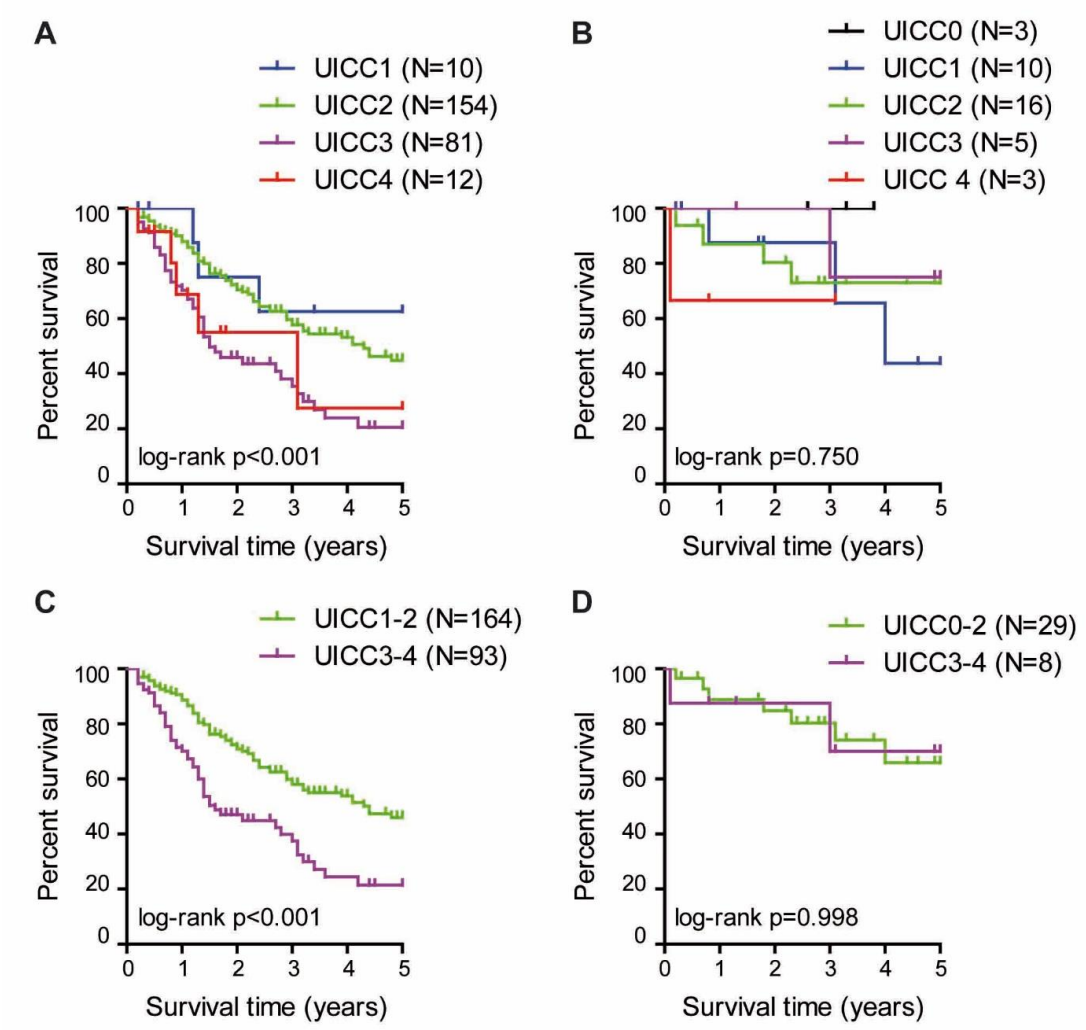


Figure S2: Kaplan-Meier survival analysis of invasive CCA cases with or without IPNB/ITPN precursor lesions. (A) Kaplan–Meier survival analysis of CCA cases without IPNB/ITPN precursor lesion separated by respective UICC stages. **(B)** Kaplan–Meier survival analysis of patients with IPNB/ITPN precursor lesion separated by respective UICC stages. **(C)** Kaplan–Meier survival analysis of CCA cases without IPNB/ITPN precursor lesion separated by UICC1-2 versus UICC3-4 stages. **(D)** Kaplan–Meier survival analysis of patients with IPNB/ITPN precursor lesion separated by UICC0-2 versus UICC3-4 stages.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

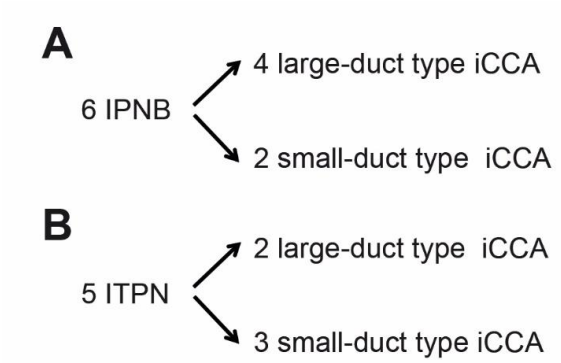
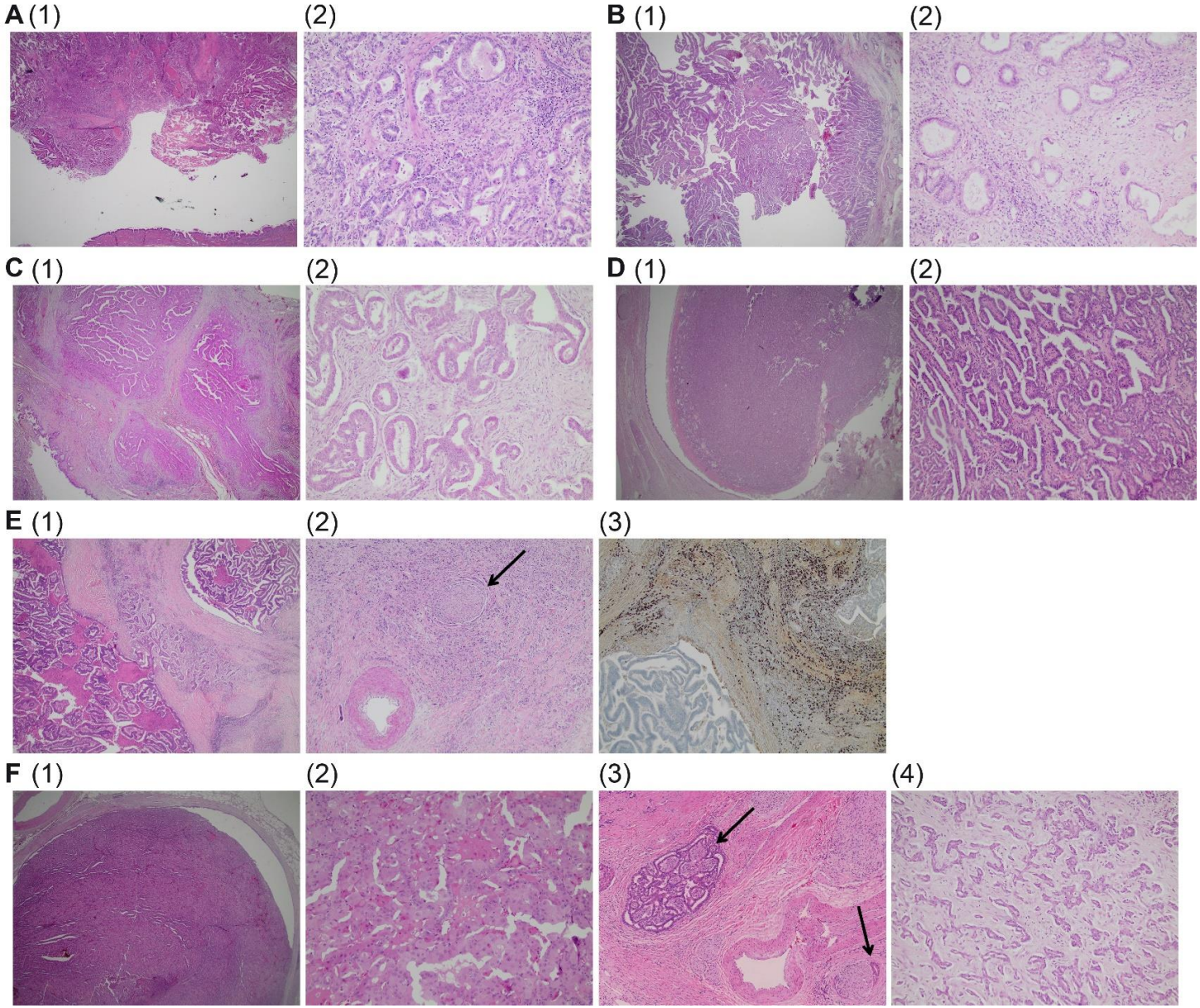


Figure S3: Schematic diagram depicting the number of cases with intrahepatic IPNB or ITPN and corresponding large-duct or small-duct type iCCA.

(A) Of 6 intrahepatic IPNB with associated iCCA, 4 cases had associated large-duct type iCCA and 2 small-duct type iCCA. **(B)** Of 5 intrahepatic ITPN with associated iCCA, 2 cases had associated large-duct type iCCA and 3 small-duct type iCCA.

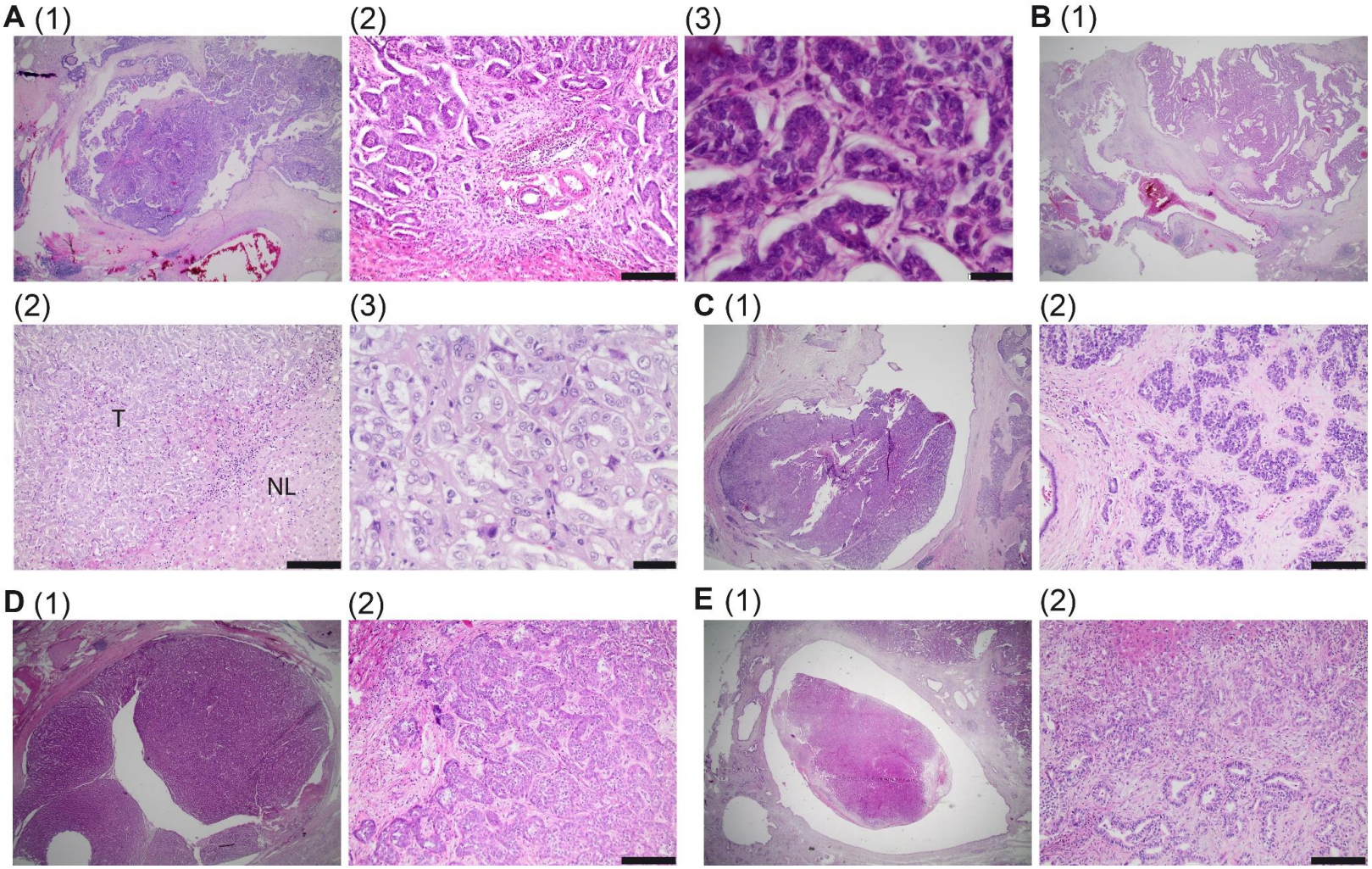
BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)



BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Figure S4: Representative images of intrahepatic IPNB and ITPN with associated large-duct iCCA. (A) (1): IPNB, HE, original magnification: 12,5x; (2): corresponding iCCA large-duct type, HE, original magnification: 100x. **(B)** (1): IPNB, HE, original magnification: 12,5x; (2): corresponding iCCA large-duct type, HE, original magnification: 100x. **(C)** (1): IPNB, HE, original magnification: 20x; (2): corresponding iCCA large-duct type, HE, original magnification: 100x. **(D)** (1): ITPN, HE, original magnification: 12,5x; (2): corresponding iCCA large-duct type, HE, original magnification: 100x. **(E)** (1): IPNB, HE, original magnification: 12,5x; (2): corresponding iCCA large-duct type with perineural tumor invasion (black arrow), HE, original magnification: 40x; center right: corresponding iCCA large-duct type, HE, original magnification: 100x; right: immunohistochemistry showing numerous IgG4-positive plasma cells, original magnification: 40x. **(F)** (1): ITPN, HE, original magnification: 40x; (2): higher magnification showing oncocytic histomorphology of precursor tumor cells, HE, original magnification: 100x; (3): corresponding iCCA large-duct type with perineural tumor invasions (black arrows), HE, original magnification: 60x; (4): corresponding iCCA showing also some areas with small-duct histology, HE, original magnification: 100x.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)



10

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Figure S5: Representative images of intrahepatic IPNB and ITPN with associated small-duct iCCA. (A) (1): IPNB with *ERBB2* mutation, HE, original magnification: 20x; (2): corresponding iCCA small-duct type with identical *ERBB2* mutation, HE, original magnification: 100x. (3): corresponding iCCA small-duct type with identical *ERBB2* mutation, HE, original magnification: 400x. **(B)** (1): IPNB, HE, original magnification: 12,5x; (2): corresponding iCCA small-duct type (T) and non-neoplastic liver (NL), HE, original magnification: 100x. (3): corresponding iCCA small-duct type, HE, original magnification: 400x. **(C)** (1): ITPN with a *FGFR2* mutation, HE, original magnification: 12,5x; (2): corresponding iCCA small-duct type with the identical *FGFR2* mutation, HE, original magnification: 100x. **(D)** (1): ITPN, HE, original magnification: 12,5x; (2): corresponding iCCA small-duct type, HE, original magnification: 100x. **(E)** (1): ITPN with *IDH1* mutation, HE, original magnification: 20x; (2): corresponding iCCA small-duct type with identical *IDH1* mutation, HE, original magnification: 100x. The black scale bar represents 100µm for panel (2) and 20µm for panel (3) of each case.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

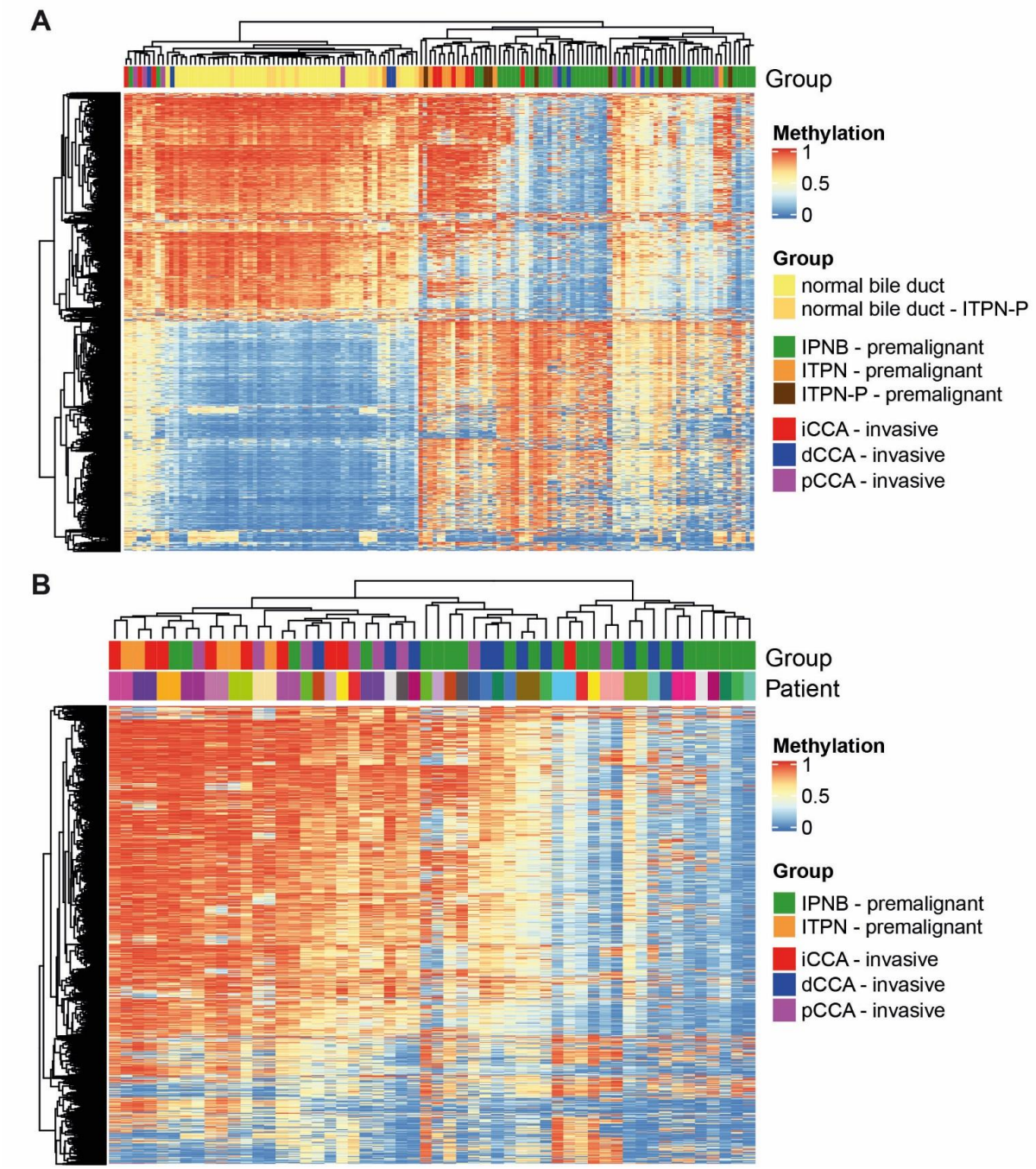


Figure S6: Unsupervised clustering of DNA methylation profiles. (A) Unsupervised clustering of all samples (N=137) included in this study: non-neoplastic normal samples, precursor and corresponding invasive CCA samples of a total of 51 patients with high-grade intraductal neoplasms of the bile duct as well as the non-neoplastic normal samples and precursor samples of 9 patients with ITPN-P. **(B)** Unsupervised clustering of the precursor and corresponding invasive CCA samples of a total of 27 CCA patients for whom paired samples were available.

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Supplemental Tables

Table S1: Patient characteristics of ITPN of the pancreas (ITPN-P) cases (N=9).

Number (percent)		**ITPN-P**
Parameter		9 (15.0)
Age	*median years (range)*	65 (53-77)
	mean years (range)	65.9 (53-77)
Sex	*male*	6 (66.7)
	female	3 (33.3)
Histology precursor	*pancreatobiliary*	8 (88.9)
	intestinal	1 (11.1)
UICC#	*UICC 0*	1 (11.1)
	UICC 1	1 (11.1)
	UICC 2	7 (77.8)
pT	*Tis*	1 (11.1)
	T1	1 (11.1)
	T2	2 (22.2)
	T3	5 (55.6)
pN	*N0*	7 (77.8)
	N1	2 (22.2)
M	*M0*	9 (100.0)
	M1	0 (0.0)
Invasive component	*Yes*	8 (88.9)
	No	1 (11.1)

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Table S2: Subclassification of iCCA.
Based on current WHO criteria, iCCA were subclassified into small- and large-duct type iCCA (Table 8.15 of [2]).

	Small-duct type	**Large-duct type**
Main location and gross features	Peripheral hepatic parenchyma MF pattern	Proximal to hepatic hilar regions PI pattern, PI + MF pattern
Risk factors	Non-biliary cirrhosis, chronic viral hepatitis	Primary sclerosing cholangitis, hepatolithiasis, liver fluke infection
Precursors	Unknown	BilIN, IPNB
Histology	Small ductular components: tubular pattern with low columnar to cuboidal cells and desmoplastic reaction Ductular components: cuboidal epithelia showing ductular or cord-like pattern with slit-like lumen and desmoplastic reaction	Ductular or tubular pattern with columnar to cuboidal epithelium, with desmoplastic reaction
Mucin production	Non-mucin secreting glands	Mucin-secreting glands
Perineural/lymphatic invasion	-/+	++
Immune/molecular features	CD56, CRP, N-cadherin, *IDH1/2* mutation	MUC5AC, MUC6, S100, TFF1, *KRAS* mutation

MF, mass-forming; PI, periductal infiltrating

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Table S3: Underlying hepatobiliary diseases of the CCA study cohort (N=54). Some patients exhibited more than one hepatobiliary disease and these were included independently into the evaluation.

Number (percent)	**Total**	**IPNB**	**ITPN**	**p-value**
Parameter	54 (100.0)	44 (81.5)	10 (18.5)	54 (100.0)
Chronic cholecystitis (NOS*)	9 (16.7)	8 (18.2)	1 (10.0)	1.000
Chronic cholangitis (NOS)	6 (11.1)	6 (11.1)	0 (0.0)	0.580
Chronic pancreatitis (NOS)	5 (9.3)	5 (11.4)	0 (0.0)	0.571
Liver cirrhosis	2 (3.7)	1 (1.6)	1 (10.0)	0.339
Liver fibrosis	2 (3.7)	1 (1.6)	1 (10.0)	0.339
Hepatitis B virus	2 (3.7)	0 (0.0)	2 (20.0)	**0.031**
IgG4-associated cholangitis	1 (1.9)	1 (1.6)	0 (0.0)	1.000
Primary sclerosing cholangitis	1 (1.9)	0 (0.0)	1 (10.0)	0.185
Hemochromatosis	1 (1.9)	1 (1.6)	0 (0.0)	1.000
Caroli syndrome	1 (1.9)	1 (1.6)	0 (0.0)	1.000
None detected	26 (48.2)	22 (50.0)	6 (60.0)	0.262

*NOS: not otherwise specified

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Table S4: Japan-Korea classification of 44 IPNB cases included in this study.

	Number (percent)	Total	Class 1	Class 2	p-value
Parameter		44 (100.0)	10 (22.7)	34 (77.3)	
Age	median years (range)	68 (39-81)	68 (39-79)	69 (46-81)	
	mean years (range)	65.2 (39-81)	64.4 (39-79)	65.8 (46-81)	0.735 *
Sex	male	34 (77.3)	5 (50.0)	29 (85.3)	
	female	10 (22.7)	5 (50.0)	5 (14.7)	**0.033****
Location	intrahepatic	8 (18.2)	4 (40.0)	4 (11.8)	
	perihilar	10 (22.7)	2 (20.0)	8 (23.5)	
	distal	26 (59.1)	4 (40.0)	22 (64.7)	0.121***
Histology	pancreatobiliary	29 (65.9)	5 (50.0)	24 (70.6)	
	gastric	4 (9.1)	2 (20.0)	2 (5.9)	
	intestinal	11 (25.0)	3 (30.0)	8 (23.5)	
	oncocytic	0 (0.0)	0 (0.0)	0 (0.0)	0.313***
UICC#	UICC 0	3 (6.8)	1 (10.0)	2 (5.9)	
	UICC 1	12 (27.3)	1 (10.0)	11 (32.4)	
	UICC 2	19 (43.2)	5 (50.0)	14 (41.2)	
	UICC 3	3 (6.8)	1 (10.0)	2 (5.9)	
	UICC 4	3 (6.8)	1 (10.0)	2 (5.9)	
	NA	4 (9.1)	1 (10.0)	3 (8.8)	0.709***
pT	Tis	3 (6.8)	1 (10.0)	2 (5.9)	
	T1	14 (31.8)	1 (10.0)	13 (38.2)	
	T2	19 (43.2)	6 (60.0)	13 (38.2)	
	T3	7 (15.9)	2 (20.0)	5 (14.7)	
	T4	1 (2.3)	0 (0.0)	1 (2.9)	0.492***
pN	N0	30 (68.2)	7 (70.0)	23 (67.6)	
	N1	9 (20.5)	1 (10.0)	8 (23.5)	
	NA	5 (11.4)	2 (20.0)	3 (8.8)	0.653**
M	M0	41 (93.2)	9 (90.0)	32 (94.1)	
	M1	3 (6.8)	1 (10.0)	2 (5.9)	0.548**
G	G1	2 (4.5)	0 (0.0)	2 (5.9)	
	G2	31 (70.5)	9 (90.0)	22 (64.7)	
	G3	8 (18.2)	0 (0.0)	8 (23.5)	
	NA (Tis)	3 (6.8)	1 (10.0)	2 (5.9)	0.156***
Invasive	Yes	41 (93.2)	9 (90.0)	32 (94.1)	
	No	3 (6.8)	1 (10.0)	2 (5.9)	0.548**

* Mann Whitney U-test; ** Fisher's exact test; *** Chi-square test

Cases with pNx had no lymph nodes resected, therefore, UICC status could not be assessed

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Table S5: Comparison of mutation frequencies of this study and published data sets*.

	Goeppert IPNB (N=39)	Nakanuma IPNB (N=21)*	Yang IPNB (N=37)	Aoki IPNB (N=36)	Goeppert ITPN (N=9)	Schlitter ITPN (N=20)	Wardell dCCA (N=101)	Wardell iCCA (N=136)
TP53	53.8%	38.1%	24.3%	34.3%	11.1%	17%	44.6%	10.3%
SMAD4	25.6%	19.1%	n.d.	14.3%	0%	7%	14.9%	0.7%
FBXW7	10.3%	9.5%	n.d.	2.9%	0%	n.d.	11.9%	1.5%
BRAF	2.6%	4.8%	5.4%	5.7%	11.1%	0%	0.0%	5.1%
IDH1	0%	n.d.	n.d.	0.0%	11.1%	0%	0.0%	12.5%
KRAS	28.2%	9.5%	48.6%	31.4%	0%	6%	18.8%	11.8%
CTNNB1	17.9%	17.0%	10.8%	17.1%	0%	0%	2.2%	5.0%
ARID2	7.7%	n.d.	n.d.	5.7%	0%	n.d.	3.0%	3.7%
ARID1A	5.1%	n.d.	n.d.	5.7%	0%	n.d.	4.0%	8.1%
PIK3CA	5.1%	22.0%	n.d.	8.6%	0%	6%	5.9%	5.9%
GNAS	0%	9.5%	32.4%	11.4%	0%	0%	1.0%	1.5%

*Nakanuma et al. [10]; Yang et al. [11]; Aoki et al. [12]; Schlitter et al. [13]; Wardell et al. [14]; Goeppert et al. (present study); n.d. not determined

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Table S6: Genetic alterations in IPNB (N=39) and ITPN (N=9) cases.

	Total	IPNB	ITPN	p-value*
Number (percent)	**48 (100.0)**	**39 (81.3)**	**9 (18.8)**	
TP53	22 (45.8)	21 (53.8)	1 (11.1)	**0.028**
KRAS	11 (22.9)	11 (28.2)	0 (0)	0.095
SMAD4	10 (20.8)	10 (25.6)	0 (0)	0.172
CDKN2A	10 (20.8)	9 (20.5)	2 (22.2)	1.000
ERBB2	7 (14.6)	7 (17.9)	0 (0)	0.320
CTNNB1	7 (14.6)	7 (17.9)	0 (0)	0.320
FBXW7	4 (8.3)	4 (10.3)	0 (0)	1.000
BRAF	2 (4.2)	1 (2.6)	1 (11.1)	0.343
IDH1	1 (2.1)	0 (0)	1 (11.1)	0.188

*Fisher's exact test of IPNB versus ITPN

Table S7: Genetic alterations in IPNB/ITPN (N=26) and paired invasive CCA (N=26) cases.

	Precursor	Invasive	p-value*
Number (percent)	**26 (50.0)**	**26 (50.0)**	
TP53	11 (42.3)	10 (38.5)	1.000
KRAS	8 (30.8)	9 (34.6)	1.000
SMAD4	5 (19.2)	4 (15.4)	1.000
CDKN2A	6 (23.1)	3 (11.5)	0.465
FBXW7	2 (7.7)	5 (19.2)	0.419
ERBB2	3 (11.5)	3 (11.5)	1.000
CTNNB1	5 (19.2)	1 (3.8)	0.191
ROBO2	0 (0.0)	5 (19.2)	0.051
CDKN2B	1 (3.8)	3 (11.5)	0.610

*Fisher's exact test of precursor lesion versus corresponding invasive CCA

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

Supplemental References

1 Torbenson M, Zen Y, Yeh MM. AFIP Atlas of Tumor Pathology, Tumors of the Liver: ARP Press, 2018.

2 WHO Classification of Tumours Editorial Board. WHO Classification of Tumours: Digestive System Tumours, 5th Edition. World Health Organization 2019.

3 Goeppert B, Folseraas T, Roessler S, Kloor M, Volckmar AL, Endris V*, et al.* Genomic characterization of cholangiocarcinoma in primary sclerosing cholangitis reveals novel therapeutic opportunities. Hepatology 2020.

4 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 2010;**38**:e164.

5 Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G*, et al.* Integrative genomics viewer. Nat Biotechnol 2011;**29**:24-6.

6 Muller F, Scherer M, Assenov Y, Lutsik P, Walter J, Lengauer T*, et al.* RnBeads 2.0: comprehensive analysis of DNA methylation data. Genome Biol 2019;**20**:55.

7 Lutsik P, Slawski M, Gasparoni G, Vedeneev N, Hein M, Walter J. MeDeCom: discovery and quantification of latent components of heterogeneous methylomes. Genome Biol 2017;**18**:55.

8 Scherer M, Nazarov PV, Toth R, Sahay S, Kaoma T, Maurer V*, et al.* Reference-free deconvolution, visualization and interpretation of complex DNA methylation data using DecompPipeline, MeDeCom and FactorViz. Nat Protoc 2020;**15**:3240-63.

9 Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016;**32**:2847-9.

10 Nakanuma Y, Kakuda Y, Fukumura Y, Sugino T, Uesaka K, Serizawa M*, et al.* The Pathologic and Genetic Characteristics of the Intestinal Subtype of Intraductal Papillary Neoplasms of the Bile Duct. Am J Surg Pathol 2019;**43**:1212-20.

11 Yang CY, Huang WJ, Tsai JH, Cheng A, Chen CC, Hsu HP*, et al.* Targeted next-generation sequencing identifies distinct clinicopathologic and molecular entities of intraductal papillary neoplasms of the bile duct. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 2019;**32**:1637-45.

12 Aoki Y, Mizuma M, Hata T, Aoki T, Omori Y, Ono Y*, et al.* Intraductal papillary neoplasms of the bile duct consist of two distinct types specifically associated with clinicopathological features and molecular phenotypes. J Pathol 2020;**251**:38-48.

13 Schlitter AM, Jang KT, Kloppel G, Saka B, Hong SM, Choi H*, et al.* Intraductal tubulopapillary neoplasms of the bile ducts: clinicopathologic, immunohistochemical, and molecular analysis of 20 cases. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc 2016;**29**:93.

14 Wardell CP, Fujita M, Yamada T, Simbolo M, Fassan M, Karlic R*, et al.* Genomic characterization of biliary tract cancers identifies driver genes and predisposing mutations. Journal of hepatology 2018;**68**:959-69.