A novel mouse model designed for studying cholangiocarcinoma tumorigenesis uncovers a role for a SOX17-Tensin 4 pathway

Short title: Tumor progression in cholangiocarcinoma

Mickaël Di-Luoffo[1], Sophie Pirenne[1,2,#], Thoueiba Saandi[1,#], Axelle Loriot[1,#], Claude Gérard[1], Nicolas Dauguet[3], Florence Lamoline[1], Katarzyna Konobrocka[1], Vitaline De Greef[1], Mina Komuta[2,4], Patrick Jacquemin[1], and Frédéric P. Lemaigre[1,§]

[1] de Duve Institute, Université catholique de Louvain, Brussels, Belgium.

[2] Department of Pathology, Cliniques Universitaires Saint-Luc, Université catholique de Louvain, Brussels, Belgium.

[3] CYTF platform, de Duve Institute, Université catholique de Louvain, Brussels, Belgium.

[4] Institute of Experimental and Clinical Research, Université catholique de Louvain, Brussels, Belgium.

[#] These authors contributed equally

Grant support: The work is supported by the National Fund for Scientific Research FRS-FNRS (Belgium, grants J003717 and J.0105.19 to FPL), Télévie (Belgium, grants 7.4544.18 to FPL and MK, and 7.8515.18 to PJ), the Foundation against Cancer (Belgium, grants 2014-125 and 2018-078 to FPL; 2016-089 and 2018-076 to PJ), the D.G. Higher Education and Scientific Research of the French Community of Belgium (grant ARC 15/20-065 to FPL), and grants from the Fonds Joseph Maisin (Belgium) to FPL and PJ. SP holds a PhD fellowship from Télévie (Belgium, grant 7.4544.18). ND is upported by the Belgian National Lottery. PJ is Senior Research Associate at the FRS-FNRS (Belgium).

Abbreviations: BilIN, biliary intraepithelial neoplasia; CK19, cytokeratin 19; DDC, 3,5-diethoxycarbonyl-1,4 dihydrocollidine; EGF, epidermal growth factor; EGFR,

epidermal growth factor receptor; EpCAM, epithelial cell adhesion molecule; ERK, extracellular-regulated kinase; ERBB2, v-erb-b2 erythroblastic leukemia viral oncogene homolog 2/human epidermal growth factor receptor 2; eYFP, enhanced yellow fluorescent protein; HES1, hairy and enhancer of split-1; H&E, hematoxylin and eosin; IL-33, interleukin-33; iCCA, intrahepatic cholangiocarcinoma; IPNB, intraductal papillary neoplasm of bile ducts; KRAS, v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog; MAPK, mitogen activated protein kinase PI3K, phosphoinositide-3-kinase ; SMAD2, small mothers against decapentaplegic 2; SOX17, SRY-related HMG box transcription factor 17; TCGA, The Cancer Genome Atlas; TGFβ, transforming growth factor β; TNS4, tensin 4.

§ Corresponding author: Frédéric P. Lemaigre, de Duve Institute, Université catholique de Louvain, Avenue Hippocrate 75-B1-7503, B-1200 Brussels, Belgium. Email: frederic.lemaigre@uclouvain.be. Phone:+32 2 764 7583. Fax +32 2 764 7573.

Conflict of interest statement. The authors disclose no conflicts.

Transcript profiling: GEO accession number: GSE141511.

Author contributions:

Study concept and design: Mickaël Di-Luoffo, Patrick Jacquemin and Frédéric P. Lemaigre.

Acquisition of data: Mickaël Di-Luoffo, Sophie Pirenne, Thoueiba Saandi, Nicolas Dauguet, Katarzyna Konobrocka, Vitaline De Greef, Florence Lamoline.

Analysis and interpretation of data: Mickaël Di-Luoffo, Sophie Pirenne, Thoueiba Saandi, Axelle Loriot, Claude Gérard, Mina Komuta, Patrick Jacquemin, and Frédéric P. Lemaigre

Drafting of the manuscript: Mickaël Di-Luoffo and Frédéric P. Lemaigre

Obtained funding: Mina Komuta, Patrick Jacquemin, and Frédéric P. Lemaigre

Abstract

Background & Aims. Although earlier diagnosis and treatment of intrahepatic cholangiocarcinoma (iCCA) is necessary to improve therapy, there is still limited information available about formation and evolution of iCCA precursor lesions. Therefore, there is need to identify mechanisms driving formation of precancerous lesions and their progression towards invasive tumor using experimental models that faithfully recapitulate human tumorigenesis.

Methods. We generated a new mouse model which combines cholangiocyte-specific expression of *Kras*^{G12D} with 3,5-diethoxycarbonyl-1,4-dihydrocollidine diet-induced inflammation to mimic iCCA development in patients with cholangitis. Histological and transcriptomic analyses of the mouse precancerous lesions and iCCA were performed and compared with human analyses. The function of genes overexpressed during tumorigenesis was investigated in human cell lines.

Results. Mice expressing *Kras*^{G12D} in cholangiocytes and fed a DDC diet developed cholangitis, ductular proliferations, intraductal papillary neoplasms of bile ducts (IPNBs) and eventually iCCAs. The histology of mouse and human IPNBs were highly similar, and mouse iCCAs displayed the histological characteristics of human mucin-producing large duct-type iCCA. Signaling pathways activated in human iCCA were activated in mice. The identification of transition zones between IPNB and iCCA on tissue sections, combined with RNA-sequencing analyses of the lesions suggested that iCCAs derive from IPNBs. We provide evidence that a gene cascade which comprises *Kras*^{G12D}, SRY-related HMG box transcription factor 17 (SOX17) and Tensin 4 (TNS4), and which is activated by epidermal growth factor, promotes tumor progression.

Conclusions. We developed a novel mouse model that faithfully recapitulates human iCCA tumorigenesis and identified a gene cascade promoting tumor progression.

Keywords: intrahepatic cholangiocarcinoma; intraductal papillary neoplasm of bile ducts; KRAS; mouse model; SOX17; Tensin-4.

Introduction

Prognosis of intrahepatic cholangiocarcinoma (iCCA) remains dismal. Since this results in part from delayed diagnosis, there is a need to understand the mechanisms that drive development of precancerous lesions and their progression towards malignant tumor.[1] A range of perturbed signaling pathways and genomic alterations was found to be associated with late-stage cholangiocarcinoma, but there is still limited knowledge about how such pathways and mutations evolve throughout tumorigenesis.[2-8] Moreover, our understanding of this process is further complexified by the observation that iCCA can arise from several cell types, namely cholangiocytes, hepatocytes, progenitor cells and peribiliary glands.[9-16]

Histopathological studies identified two subtypes of iCCA in humans. Small duct subtype is typically located at the periphery of the liver and presents as tubular-type adenocarcinoma with cuboidal cells and ductular component without mucin production. The large duct subtype is predominantly found near the hilum and exhibits a tubular pattern with mucin-producing columnar cells. Histopathological studies classified biliary intraepithelial neoplasia (BilIN), intraductal tubulopapillary neoplasms and intraductal papillary neoplasms of bile ducts (IPNB) as non-invasive precursor lesions.[17-19] Whereas BilINs develop in large bile ducts and present as flat or micropapillary dysplasia of the bile duct epithelium, IPNBs are macroscopic lesions of the large ducts characterized by exophytic papillary structures with fibrovascular core and prominent intraductal growth.[20,21] Biliary micro-hamartomas, also called von Meyenburg complexes, are irregular dilations of intrahepatic bile ducts surrounded by dense stroma. These complexes are considered benign, but were reported to evolve towards iCCA in rare instances and potentially qualify as iCCA precursor lesion.[22]

Gain-of-function mutations of the small GTPase v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*) are detected in iCCA at a frequency of 15-25%, and are associated with a significant decrease of survival.[23-27] Importantly *KRAS* mutations were detected in BilINs and IPNBs in 25-50% of cases, strongly suggesting that such mutations contribute to initiate tumorigenesis. [28-30] Chronic inflammation is a well

known risk factor of iCCA. Patients with primary sclerosing cholangitis are at high risk for iCCA and molecular classification of late-stage cholangiocarcinomas underscored the importance of inflammatory gene expression in iCCA pathogeny.[31-33]

Animal models are invaluable for investigating intra- and extrahepatic CCA development, and several models were designed based on induction of cholestasis, administration of tumor-promoting agents, injection of tumor cells, and generation of genomic alterations.[34-36] However, limited attention was paid to tumor progression, and in several instances genetic alterations were induced at the prenatal stage or simultaneously in hepatocytes and cholangiocytes, unlike in humans. Here, we generated an original mouse model of iCCA tumorigenesis in which expression of the oncogenic mutant *Kras*^{G12D} is specifically induced in cholangiocytes and is combined with diet-induced inflammation that mimics chronic cholangitis. The mice developed IPNB similar to those in humans, and our data support that those lesions evolve toward malignant iCCA. We also identified a gene regulatory pathway that is activated during tumorigenesis and which promotes cell proliferation and migration.

Material and Methods

Animals

Osteopontin-iCreERT^2 (Opn-Cre^{ER}), loxP-stop-loxP-Kras$^{G12D/+}$ (LSL-Kras$^{G12D/+}$) and Rosa26R^{eYFP} mice were described[37-39] and bred in a CD1-enriched background. All mice were maintained at 23°C +/- 2°C with a 12 h day/night cycle. For Cre-LoxP recombination, 6 week-old mice were injected intraperitoneally three times at one day interval with tamoxifen (T5648; Sigma-Aldrich, Bornem, Belgium) dissolved in corn oil at a concentration of 30 mg/mL and at a dose of 150 mg/kg of body weight. Mice were fed a standard diet or a diet containing 3,5-diethoxycarbonyl-1,4 dihydrocollidine 0.1% (DDC; 137030 Sigma-Aldrich) and monitored for signs of illness as described in the Results section and in Supplementary Figure 1. Animals received humane care according to the Directive 2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific

purposes. Protocols were approved by the Animal Welfare Committee of the Université catholique de Louvain.

Plasmids, sub-cloning and mutagenesis

Kras^{G12D}, *Sox17*, and *Tns4* cDNAs were polymerase chain reaction (PCR)-amplified using the primers listed in Supplementary Table 1 and sub-cloned in pcDNA3.1+ (Thermo Fisher Scientific, Waltham, MA, USA) and pSBtet-GFP-Neomycin (pSBtet-GN; Addgene plasmid # 60501; a gift from R. Marschalek[40]). The SfiI site in mouse *Tns4* cDNA was mutated using the QuikChange II XL mutagenesis kit (Agilent, Santa Clara, CA, USA) and the primers listed in Supplementary Table 1. PCR amplification was performed using a CloneAmp™ HiFi PCR Premix (Takara, Mountain View, CA, USA) according to the manufacturer's protocol, and amplicons were digested 30 min to 1 h at 37°C with BamHI, KpnI, XbaI or XhoI or 50°C with SfiI (NEB Ipswich, MA, USA). The ligations were performed 30 min at room temperature using a T4 DNA ligase (Thermo Fisher Scientific). pCMV(CAT)T7-SB100 plasmid coding for Sleeping Beauty Tranposase (Addgene plasmid # 34879) was a gift from Z. Izsvak.[41]

Cell culture and transfections

Human EGI-1 cholangiocarcinoma cells (German Collection of Microorganisms and Cell Cultures, kind gift from L. Fouassier) and immortalized normal mouse[42] cholangiocytes (NMC; kind gift from T. Shimosegawa) were grown in Dulbecco's Modified Eagle's Medium (DMEM; Westburg, Leusden, Netherlands), 10% fetal bovine serum (FBS; Merck, Darmstadt, Germany), 2 mM L-Glutamine (Thermo Fisher Scientific), Penicillin-Streptomycin 50 U/mL and 50 µg/mL respectively (Gibco™, Waltham, MA, USA), 2.5 µg/mL Amphotericin B (Gibco™) and 10% MEM Non-Essential Amino Acids (NEAA; Gibco™) in a humidified 5% CO_2 incubator at 37°C. For transient transfections, cells were grown in 60 mm dishes and transfected with 3 µg plasmid, using jetPRIME® (Polyplus-Transfection, Illkirch-Graffenstaden, France) for 24 h, in at least three independent experiments. DNA was transfected 24 h after plating the cells on 60 mm dishes, the medium was changed every day. For stable transfections, NMC or EGI-1 cells were grown in 35 mm dishes. Cells in each well were transfected with 1.9 µg plasmid (pSBtet-GN, pSBtet-GN-SOX17 or pSBtet-GN-TNS4) and 100 ng of pCMV(CAT)T7-SB100 vector, using jetPRIME® (Polyplus-Transfection, Illkirch-Graffenstaden, France). Twenty-four hours after transfection,

cells were subjected to 1 mg/mL G418 for 7 days. GFP-positive cells were sorted using FACSAria III (BD, Franklin Lakes, USA) and distributed one cell per well for clonal selection. After growth for 5 days in DMEM, 10% FBS, 2 mM L-Glutamine, Peni-Strep (50 U/mL and 50 µg/mL), 2.5 µg/mL Amphotericin B and 10% NEAA, each clone was treated with 1 mg/mL G418 for 2 weeks and the stability of the transfected vectors was monitored by detecting GFP fluorescence. Cells were treated with 0.5 µg/mL doxycycline for 5 days to induce SOX17 and TNS4.

Laser capture microdissection (LCM), RNA extraction and sequencing

The livers were collected, embedded in OCT, frozen in liquid nitrogen and stored at −80 °C until processing. Microscope support glass, membrane slides and all tools were pre-treated to remove RNAse (RNaseZAP, Sigma-Aldrich). 10 µm sections were cut using Cryostar NX70 (Thermo Fisher Scientific) using -25°C for knife/blade temperature and -13°C for specimen temperature. Liver sections were deposited on slides adapted for LCM (SuperFrost Plus (25x75x1mm); VWR, Leuven, Belgium) and stored at -80°C inside plastic containers with silica granules to prevent tissue hydration. Within 12h, the liver sections were microdissected using Leica LMD7000 microscope (GIGA, Liège, Belgium) and RNA was directly extracted with RNAqueous®-Micro total RNA isolation AM1931 Kit (Thermo Fisher Scientific). RNA was stored at -80°C. RNA sequencing was performed by the Next Generation Sequencing Service (NGS Genewiz, South Plainfield, NJ, USA). Quality of samples was controlled by NanoDrop (NanoDrop 2000 (Thermo Fisher Scientific) and Qubit fluorometry (Invitrogen, Carlsbad, CA)) for measuring concentration and calculation of RNA integrity number (RIN). RNA was amplified using SMARTer® Universal Low Input RNA Kit for sequencing (Takara, Mountain View, CA, USA). Ribosomal RNA was depleted using RIboGone™-Mammalian rRNA removal kit (Takara, Mountain View, CA, USA), and cDNA libraries were generated using the TruSeq total stranded RNA-seq kit (Illumina, San Diego, CA, USA). cDNA libraries were sequenced using Illumina HiSeq 2x150bp configuration with estimated data output ≥350M raw paired-end reads per lane. RNA sequencing data include the generation of FASTQ-format files containing reads sequenced.

Bioinformatics analysis workflow

Read quality control was performed using FastQC software v0.11.7 (Babraham Institute, Cambridge, UK). Low quality reads were trimmed and adapters were removed using Trimmomatic software v0.35 (RWTH Aachen University, Germany). Reads were aligned using HISAT2 software v2.1.0 (Johns Hopkins University School of Medicine, Center for Computational Biology, Baltimore, MD, USA) on GRCm38 mouse genome. Gene counts were generated using HTSeq-count (v0.5.4p3) software[43,44] and gencode.vM15.annotation.gtf annotation file. Differential expression analysis were done using DESeq2 v1.22.2,[45] on R version 3.5.1. Differential expression analyses were performed using adjusted p-value <0.05 and log fold-change <0.01. All differential expression data are relative data from normalized transcript per million (TPM). Pearson correlation analyses were performed using 'rcorr' function from the Hmisc package (RStudio).

Cell proliferation assays

For nuclear fluorescence detection experiments, 7000 EGI-1 cells were plated at day 1 and grown for 48 h in 48 well plateq. Cells were rinsed with PBS, and stained with Hoechst (1/5000) for 30 min. Hoechst fluorescence was monitored using victor x4 2030 Multilabel Reader (PerkinElmer, Waltham, MA, USA) with excitation wavelength at 350 nm. For crystal violet staining experiment, EGI-1 cells were rinsed with PBS. $4x10^4$ cells were plated at day 1 in 60 mm dishes and grown for minimum 3 days. At day 3, 5 and 7 cells were rinsed with PBS and fixed using 10% acetic acid/10% methanol in H_2O for 15 min. Cells were stained with 0.5% crystal violet (#C6158, Sigma-Aldrich) in H_2O for 15 min. The cells were rinsed gently with PBS and the dishes were inverted on absorbent paper. The dishes were photographed and cells were counted using 16 pictures taken using a grid placed on each dish for each condition (Axiovert 200 fluorescent microscope 10X magnificence). All pictures were transformed into white/black binary pictures and pixels were quantified using ImageJ v1.52p.

Cell migration assay

$7x10^3$ EGI-1 or $1.5x10^4$ NMC cells, wild-type or overexpressing doxycyclin-induced TNS4, were plated at day 1 in culture-insert 2 wells installed in μ-Dish^{35mm,high} (#81176, Ibidi Gmgh, Gräfelfing, Gemany) and grown for 48 h. Then, the culture-

insert 2 wells were removed leaving a wound of about 200 µm between the two wells. 6 adjacent pictures covering all the wound were taken in each condition at t=0h, t=1h, t=3h, t=6h and t=8h. Three wound distances were measured on each picture, at each time point, with 50 µm interval between each measurement (ZEN Blue v2.5 software, Carl Zeiss Vision, Oberkochen, Germany).

Statistical analyses

Single comparisons between two experimental groups were done using a paired Student's t test. To identify significant differences between multiple groups, statistical analyses were performed using a one-way ANOVA followed by Newman-Keuls test. For adjusted p-value, a Benjamini-Hochberg correction was applied. For all analyses, p-value<0.05 was considered significant. All statistical analyses were done using the Prism software (GraphPad softwares).

Western blotting, Reverse Transcription-quantitative PCR, cholangiocyte purification, tissue staining procedures

See supplementary information

Results

Cholangiocyte-specific expression of Kras^{G12D} combined with DDC-induced cholangitis causes sequential development of IPNB and iCCA.

To recapitulate the tumorigenic process of human iCCA we generated a mouse model which combines cholangiocyte-specific *Kras^{G12D}* expression with biliary sclerosis resulting from chronic cholangitis. Opn-Cre^{ER} mice, which express tamoxifen-inducible Cre^{ER} recombinase specifically in cholangiocytes, were mated with LSL-Kras$^{G12D/+}$ mice, which contain a mutated *Kras* allele coding for KRAS^{G12D} when a loxP-flanked cassette has been removed by CreER-mediated recombination.[37-39] The mice also contained a tamoxifen-inducible *Rosa26R^{eYFP}* allele. Chronic cholangitis and biliary sclerosis were induced by feeding the mice a

3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC) diet, as described by Fickert and coworkers.[46] *Kras*^{G12D} expression was induced by administration of tamoxifen to 6 week-old mice. Mice were then fed a normal diet for 4 weeks to enable wash out of tamoxifen. Thereafter, the mice received a DDC diet which was interrupted after 2 weeks because of weight loss. Following 1 week of weight recovery under normal diet, DDC diet was resumed until the duration of cholangitis extended to 8, 18 or 21 to 31 weeks (Figure 1A). During this period the mice accepted the DDC diet; the experiments were interrupted when animals showed signs of severe illness and had reached humane endpoints (Supplementary Figure 1A-B).

Histological analyses of control mice at 18 weeks of DDC treatment showed periductal fibrosis, porphyrin plugs, ductular proliferation throughout the liver and development of IPNB near the hilum (Figure 1B). Expression of *Kras*^{G12D} in cholangiocytes of tamoxifen-treated Opn-Cre^{ER}/LSL-Kras$^{G12D/+}$/Rosa26R^{eYFP} mice did not affect the histology of the biliary epithelium (Figures 1B and 2A; Supplementary Figure 1C). At that stage, livers of control mice treated with DDC did not differ significantly from *Kras*^{G12D}-expressing mice treated with DDC (Figure 1B). This indicated that until 18 weeks of DDC diet, the expression of *Kras*^{G12D} had limited impact on the effects of DDC, except for the fact that development of IPNB proceeded faster: at 8 weeks, ductal hyperplasia was observed in DDC-fed control mice, but IPNBs were already detected in DDC-fed mice that express *Kras*^{G12D} (Supplementary Figure 1C).

Importantly, starting at 21 weeks of DDC diet, iCCAs were observed, but only in Opn-Cre^{ER}/LSL-Kras$^{G12D/+}$/Rosa26R^{eYFP} mice (Figure 1C). iCCAs developed in one third of cases (12/36 males; 9/24 females), and were detected up to 31 weeks of DDC diet, after which animals were sacrified. The histology showed tubular structures with desmoplastic changes typical of human adenocarcinoma (Figure 1C). iCCAs were highly proliferative as illustrated by Ki67 staining, and expressed cytokeratin 19 (CK19), epithelial cell adhesion molecule (EpCAM) and eYFP, supporting their biliary origin (Figure 2A). The epithelium was columnar to cuboidal, and was mucin-positive as shown by alcian blue staining (Figure 2B). This is similar to the mucin-producing type iCCA. Some tumors appeared more glandular and less desmoplastic (Figure 2B). Interestingly, adjacent to large ducts, we often found areas of glandular

proliferation with expansive growth pattern and cytonuclear atypia, which we qualified as early stage of invasive cancer (Figure 2C). Importantly, focal transitions between IPNB and early invasive cancer were detected, suggesting that IPNB evolved to iCCA (Figure 2C).

We concluded that the mouse model develops IPNB and iCCA which faithfully recapitulate the histology of their human counterparts.

Signaling pathways known to drive iCCA development in humans are activated in the mouse model.

To investigate the status of signaling pathways often associated with human iCCA, we first evaluated mitogen activated protein kinase (MAPK) activity and looked at phosphorylated extracellular-regulated kinase (p-ERK).[47] p-ERK was detected in cholangiocytes of DDC-treated or *Kras*^{G12D}-expressing mice, and in iCCA, but not in control cholangiocytes (Figure 3A). Activation of p-ERK in tamoxifen-treated Opn-Cre^{ER}/LSL-Kras^{G12D} cholangiocytes in the absence of DDC provided further evidence for expression of *Kras*^{G12D} in cholangiocytes that otherwise keep normal morphology. Second, phosphoinositide-3-kinase (PI3K)[47] was found to be activated in DDC-treated mice and in iCCA of DDC-treated Opn-Cre^{ER}/LSL-Kras^{G12D} mice, as evidenced by staining with anti-phosphorylated AKT serine/threonine kinase (p-AKT) (Figure 3A). Third, we also detected accumulation of hairy and enhancer of split-1 (HES1)[48] in nuclei of iCCA, indicating that the Notch pathway was stimulated too (Figure 3A). Phospho-small mothers against decapentaplegic 2 (Smad2), a marker of Transforming Growth Factor β (TGFβ) activity, is detected at high levels only in IPNB, not in other conditions. Finally, v-erb-b2 erythroblastic leukemia viral oncogene homolog 2/human epidermal growth factor receptor 2 (ERBB2/HER2)[49] was expressed in control and *Kras*^{G12D}-expressing cholangiocytes, as well as in IPNB of DDC-treated mice and iCCA of DDC-treated *Kras*^{G12D}-expressing mice. Microdissection of IPNBs and iCCAs from tissues of Opn-Cre^{ER}/LSL-Kras^{G12D} mice treated for 25 weeks with DDC, followed by western blotting indicated that the amount of ERBB2 protein was higher in iCCA than in IPNB, and that ERBB2 was active, as evidenced by detection of phosphorylated ERBB2^{Y1248} (Figures 3A-B). Also, RT-qPCR analysis of total RNA from normal purified cholangiocytes and from

microdissected IPNB and iCCA showed higher expression of epidermal growth factor receptor (EGFR) in the lesions (Figure 3B).

We concluded that MAPK, AKT, Notch, TGFβ and ERBB signaling were all activated during tumorigenesis in our mouse model, like in human iCCA.

iCCA displays a gene expression profile similar to that of IPNB

To gain insight into the mechanisms of tumor progression we investigated the gene expression profile of ductular proliferations, IPNBs and iCCAs from Opn-Cre^{ER}/LSL-Kras^{G12D} mice. The three types of lesions co-existed in livers treated for 25 weeks with DDC and were isolated at that stage. Microdissected iCCAs displayed the histology of early stage of invasive cancer, as illustrated in Figure 2B. Microdissection was performed by laser capture microdissection (LCM) followed by RNA sequencing. The results were compared with those from normal cholangiocytes FACS-sorted from Opn-Cre^{ER}/Rosa26R^{eYFP} mice untreated with DDC. Principal component analysis (PCA) of the RNA sequencing data identified distinct groups. Ductular proliferations and normal cholangiocytes were well separated, while iCCAs and IPNBs clustered more closely. Combined with the location of iCCAs near large ducts containing IPNB and the existence of transitional zones between IPNB and iCCA (Figure 2B), the closely related gene expression pattern suggests that iCCA originate from IPNB. Nineteen genes were found to be differentially expressed between IPNB and iCCA (fold-change >2, adjusted p-value <0.05; Figure 4B). The nineteen genes differentially expressed between IPNB and iCCA showed intertumoral heterogeneity, as anticipated, but were also differentially expressed between iCCA and control cholangiocytes. To support the relevance with human iCCA, we looked at two human cohorts of iCCAs (The Cancer Genome Atlas cohort (TCGA); n=34; http://firebrowse.org) and Gene Expression Omnibus cohort GSE119336 (n=15)[50]) and found that most of these genes were differentially expressed between tumor and non tumor samples (Figure 4C) as in our mouse model.

SRY-related HMG box transcription factor 17 (SOX17) and Tensin 4/C-terminal tensin-like protein (TNS4/CTEN) attracted our attention. Indeed, SOX17 cooperates with KRAS^{G12D} to promote pancreatic metaplasia and tumor progression, and the

focal adhesion protein TNS4 is upregulated by EGF-ERK signaling to promote tumor progression in hepatocellular carcinoma.[51,52] RNA sequencing and RT-qPCR showed overexpression of both genes in iCCA as compared to IPNB. Microdissection of IPNBs and iCCAs from tissues of mice expressing *Kras*^{G12D} and treated for 25 weeks with DDC, followed by western blotting confirmed the results for SOX17 and TNS4 at the protein level (Figure 5A). *Sox17* and *Tns4* were upregulated in iCCA as compared to control in mice as well as in humans (Figure 4C). *SOX17* and *TNS4* mRNA expression levels were correlated in human iCCA (Figure 4D, Supplementary Figure 2).

By immunostaining SOX17 was not detected in cholangiocytes from normal liver or in *Kras*^{G12D}-expressing cholangiocytes. In IPNBs from DDC-treated animals, a low number of cells expressed SOX17 at levels slightly higher than background. In contrast, patches of epithelial cells in iCCA expressed high levels of nuclear SOX17 (Figure 5B). SOX17 was also detected in endothelial CD31-positive cells, as expected.[53] Importantly, human liver displayed a similar profile with no detectable SOX17 in normal cholangiocytes, but with obvious expression in iCCA epithelial cells (Figure 5C). TNS4 was detected at the cell membrane in all tested conditions and was co-expressed with SOX17 in IPNB and iCCA (Figure 5B).

We concluded that iCCA and IPNB had similar RNA expression profiles and that SOX17 and TNS4 expression was induced in IPNB and increased in iCCA.

SOX17 and TNS4 are induced by EGF signaling and stimulate migration and proliferation

We next explored if ERBB2 activity, KRAS activation and expression of SOX17 and TNS4 were coordinated. To this end we used cholangiocarcinoma EGI-1 cells which are known to express mutant KRAS^{G12D}. Treating cultured cells with EGF stimulated expression of *KRAS*, *SOX17* and *TNS4* (Figure 6A). Similar results were found in EGF-treated NMC cells, an immortalized normal mouse cholangiocyte line that expresses high levels of cholangiocyte markers (Supplementary Figure 3A-B). Since KRAS is activated downstream of EGFR/ERBB2 and is known to stimulate SOX17 expression,[51] we envisaged the possibility that SOX17 activates TNS4 to constitute a EGFR/ERBB2-KRAS-SOX17-TNS4 cascade (Figure 6B). We generated EGI-1

clones that stably overexpress doxycyclin-inducible SOX17 or TNS4, and found that SOX17 stimulated TNS4 and to some extent *KRAS* expression, whereas TNS4 slightly induced *ERBB2* and *KRAS* (Figure 6C). Since EGI-1 cells constitutively express $KRAS^{G12D}$, we transfected NMC cells with $Kras^{G12D}$ to verify if the latter stimulates *Sox17* and *Tns4* in cholangiocytes; this was the case, further confirming the proposed gene cascade (Supplementary Figure 3C).

Since TNS4 is known to promote hepatocarcinoma cell migration,[52] we verified if this was the case for biliary cells. Doxycyclin-mediated stable increase of TNS4 in EGI-1 or NMC cells indeed accelerated cell migration in a wound healing assay (Figure 6D). The assay lasted for 8h, i.e. below the duration of doubling time. More important, doxycyclin-mediated increase of SOX17 or TNS4 stimulated proliferation of EGI-1 cells, as shown by crystal violet staining or Hoechst fluorescence intensity assays (Figure 6E).

We concluded that a EGFR/ERBB2-KRAS-SOX17-TNS4 gene cascade is activated during iCCA tumorigenesis and contributes to cell migration and proliferation.

Discussion

Several genetic mouse models of cholangiocarcinoma have been generated by Cre recombinase-mediated gene alterations.[34-36] These models often induce genetic modifications at prenatal stages, or simultaneously in multiple liver cell types. For these reasons, they are not optimally suited for analysis of tumor progression starting from a well-identified adult liver cell type. Here, we selected OPN-CreER for its high efficiency and cholangiocyte-specificity[37,38] to induce expression of $Kras^{G12D}$, a gene frequently mutated in human iCCA, including in precursor lesions. $Kras^{G12D}$ is expressed from its endogenous locus, meaning that the level of expression is within physiological limits. This was combined with chronic cholangitis in order to mimic tumor development in a context of chronic inflammation which is a major risk factor for iCCA. The cell type of origin is clearly identified, and iCCA in our mouse model

develops from cholangiocytes as assessed by eYFP expression. Tumorigenesis is associated with activation of the ERK, AKT, NOTCH and EGF pathways as in many human iCCAs. Also, the mouse iCCAs display a tubular pattern with mucin-producing columnar cells, which are criteria typical for human large duct type iCCA; intense desmoplasia was also detected in a subset of malignant tumors. Therefore, our animal model displays many histological and biochemical features of human iCCA. Moreover, the transcriptomic profile of IPNB and early stage of invasive iCCA, together with the histological identification of junctions between the two lesions strongly suggest that iCCA derive from IPNB. Interestingly, in humans, mutant *KRAS* is detected already in human low-grade IPNB, and tumor progression is associated with decreased TGFβ signaling,[30] as in our model which shows strong nuclear staining of P-SMAD2 in IPNB, but no detectable P-SMAD2 in iCCA. We conclude that our mouse model is well suited for the study of iCCA tumorigenesis.

Fickert and coworkers described the inflammation occuring in DDC-treated livers.[46] iCCA is often associated with activation of EGF signaling,[32,54] a feature observed here in our model. Moreover, we monitored an increase in interleukin-33 (IL-33) expression throughout the tumorigenic process (Supplementary Figure 4). IL-33 is a biliary mitogen which induces iCCA when combined with expression of constitutively active AKT and Yes-associated protein, and with obstructive cholestasis[55,56]. IL-33 is therefore likely to contribute to tumor growth in our model.

During the tumorigenic process we found that SOX17 expression in the intrahepatic cholangiocyte lineage becomes detectable at near background level in IPNB. In iCCA, SOX17 was heterogeneously distributed, with large epithelial cell patches displaying high expression of SOX17, while other tumor cells lacked detectable SOX17. The higher level of SOX17 in microdissected iCCA as compared to IPNB reflects the higher number of cells expressing the gene in iCCA, and likely the fact that SOX17-positive epithelial cells in iCCA express higher levels of SOX17 than in IPNB cells, as evidenced by more intense immunostaining in iCCA than in IPNB.

SOX17 protein is not expressed in normal human or mouse intrahepatic cholangiocytes, but regulates development of the extrahepatic biliary tract and is detected in adult gallbladder epithelium.[57-60] SOX17 is also strongly expressed in

endothelial cells, which means that care is needed to interprete transcriptomic and epigenomic analyses of SOX17 expression from sampled tumor tissue composed of multiple cell types. Our data suggest that higher expression of SOX17 in iCCA results, at least in part, from EGFR/ERBB2-KRAS-mediated stimulation. SOX17 is regarded as a tumor suppressor in iCCA and downregulation of its expression by promoter hypermethylation negatively affects the patients' prognosis.[61,62] However, our data showing that SOX17 promotes proliferation point toward a pro-tumorigenic role. Similar to our findings, and in line with the similarity of tumorigenic processes in liver and pancreas, SOX17 was shown earlier to be induced in precursor lesions of pancreatic ductal adenocarcinoma, and to cooperate with oncogenic $KRAS^{G12D}$ to enhance transformation of normal pancreas.[51] Also, SOX17 is overexpressed in high grade malignant ERBB2-transformed tumor-derived cholangiocarcinoma cells as compared to less tumorigenic BDE-1 cells.[63] The reason for the discrepancy between our data and those of Merino-Azpitarte and coworkers who showed that overexpression of SOX17 in EGI-1 cells reduces proliferation, is not clear. We note that those authors generated stable lines which overexpressed SOX17 and which displayed heterogeneous survival rates. We therefore propose that the function of SOX17 should be investigated while considering that SOX17 levels and the heterogeneity of cellular context might determine pro-tumorigenic or tumor-suppressing properties.

Tensins are focal adhesion proteins involved in cell migration by bridging the extracellular matrix receptors and the actin skeleton. TNS4 is activated by KRAS signaling in cultured colorectal cancer cells and pancreatic adenocarcinoma cell lines in which it enhances the motility of cells.[64] In cultured hepatocellular cancer cells TNS4 is activated by EGF-ERK signaling and also enhances cell motility.[52] In our mouse model we detect TNS4 expression already in normal cholangiocytes, and monitor an increase in its expression during tumorigenesis. Moreover, TNS4 and SOX17 co-localized in IPNB/iCCA SOX17 positive cholangiocytes, and in transfection experiments we positioned TNS4 epistatically downstream of SOX17. TNS4 and SOX17 mRNA expression was positively correlated in iCCA, as well as in breast carcinoma and prostate cancer but not in other tumor types (Supplementary Figure 2B)

In conclusion, we have developed a novel mouse model which recapitulates formation of iCCA precursor lesions and their evolution toward malignancy, much like in humans. The carcinogenic process is associated with activation of a EGFR/ERBB2-KRAS-SOX17-TNS4 gene cascade that contributes to cell migration and proliferation.

Acknowledgements

The authors thank Sandra Ormenese, Mieke Dewerchin, Javed Manesia, Satdarshan Paul Monga and Michael Oertel for advice with laser-capture microdissection; Noémie Ammar-Khodja, Jean-Nicolas Lodewijckx and Mourad El Kaddouri for help.

Figure Legends

Figure 1. Development of biliary lesions in mice expressing oncogenic *Kras*^{G12D} and fed a cholangitis-inducing diet. (*A*) Timing of *Kras*^{G12D} expression and of 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC)-induced cholangitis. (*B*) Development of biliary lesions in livers collected after 18 weeks of DDC feeding. Control mice do not express *Kras*^{G12D} and did not receive DDC. Opn-Cre^{ER}/LSL-Kras^{G12D} mice expressing KRAS^{G12D} following tamoxifen administration (*Kras*^{G12D}) do not show any lesion, whereas DDC-treated mice in the absence (DDC) or presence of KRAS^{G12D} (*Kras*^{G12D}+DDC) develop ductular proliferations and intraductal papillary neoplasms of bile ducts (IPNB) surrounded by a fibrotic capsule. (*C*) Intrahepatic cholangiocarcinoma (iCCA) are detected only in the *Kras*^{G12D}+DDC mice and starting at 21 weeks of DDC treatment. bd, bile duct; H&E, hematoxylin and eosin; pv: portal vein; W, weeks. *, porphyrin plug.

Figure 2. Histomorphological features of biliary lesions in mice expressing oncogenic *Kras*^{G12D} and fed a cholangitis-inducing diet. *(A*) In Opn-Cre^{ER}/LSL-Kras^{G12D}/Rosa26^{eYFP} mice treated with tamoxifen to induce KRAS^{G12D} and eYFP in cholangiocytes (*Kras*^{G12D}) and fed a DDC diet for 21 weeks (*Kras*^{G12D}+DDC), cells of ductular proliferations, IPNB and iCCA proliferate; they express CK19, EpCAM and eYFP, demonstating their cholangiocyte originn and Ki67 in proliferating cells. (*B*) High and low magnification of iCCA in (*Kras*^{G12D}+DDC) mice, here at 25 weeks of DDC treatment. iCCA cells organize as glandular tissue, display cuboidal morphology and express mucins, as evidenced by alcian blue staining. (*C*) In *Kras*^{G12D}+DDC mice, here an individual at 30W of DDC treatment, showing transition zones (black arrows) between IPNB and early stage of invasive iCCA are detected. Abbreviations are as in Fig. 1.

Figure 3. Signaling pathways activated during tumor progression in mice. (*A*) p-ERK staining is detected in cholangiocytes following *Kras*^{G12D} activation and in IPNB and iCCA. p-AKT is detected in iPNB and iCCA, while HES1 is detected only in iCCA. p-SMAD2 is detected only in IPNB. ERBB2 is detected in biliary epithelial cells in all tested conditions. (*B*) Levels of ERBB2 protein are higher in iCCA than in IPNB and ERBB2 protein is active as evidenced by detection of p-ERBB2^{Y1248}. Expression of

Egfr and *Erbb2* mRNA was higher in IPNB and iCCA than in normal purified cholangiocytes. Data in dot plots are means +/- SEM. *, p<0.05.

Figure 4. Differentially expressed genes between IPNB and iCCA. (*A*) Principal component analysis of RNA sequencing data of control cholangiocytes, ductular proliferations, IPNB and iCCA. Control cholangiocytes are purified from tamoxifen-treated Opn-Cre^{ER}/Rosa26R^{eYFP} mice, the lesions were microdissected from Opn-Cre^{ER}/LSL-Kras^{G12D}/Rosa26R^{eYFP} mice treated with tamoxifen to induce Kras^{G12D} and fed a DDC diet for 25 weeks. (*B*) Table of the 19 genes differentially expressed between IPNB and iCCA. (*C*) The genes which are overexpressed in murine iCCA *versus* IPNB are also overexpressed in murine iCCA *versus* control cholangiocytes. Most of these genes are overexpressed in human iCCA versus control tissue. Two human iCCA cohorts are compared with mouse iCCA. (*D*) Correlation between SOX17 and TNS4 mRNA expression in human iCCA. The cohorts are described in TCGA and in Sia and coworkers.[33]

Figure 5. SOX17 and TNS4 expression in mouse and human iCCA. (*A*) *SOX17* and *TNS4* mRNA and protein are detected at higher levels in murine iCCA than in IPNB and control cholangiocytes. Control cholangiocytes are purified from tamoxifen-treated Opn-Cre^{ER}/Rosa26R^{eYFP} mice, IPNB and iCCA are microdissected from Opn-Cre^{ER}/LSL-Kras^{G12D}/Rosa26R^{eYFP} mice treated with tamoxifen to induce KRAS^{G12D} and fed a DDC diet for 25 weeks. Data in dot plots are means +/- SEM. * p<0.05; ** p<0.01. (*B*) Expression of SOX17 is detected in murine IPNB and iCCA, not in normal cholangiocytes. TNS4 is detected in biliary cells from normal mouse ducts, IPNB and iCCA. (C) Expression of SOX17 is detected in human iCCA, not in normal cholangiocytes. Dotted lines delineate bile ducts; white arrowheads, SOX17-expressing biliary cells; yellow arrowheads, SOX17-expressing CD31-positive endothelial cells. bd, bile duct; ns, non-specific labeling; pv, portal vein.

Figure 6. An EGF-stimulated gene cascade involving SOX17 and TNS4 active during iCCA tumorigenesis. (*A*) Expression of *KRAS*, *SOX17* and *TNS4* is upregulated in EGI-1 cells incubated with EGF. (*B*) Structure of the EGFR/ERBB2-KRAS-SOX17-TNS4 cascade involved in iCCA tumorigenesis. (*C*) EGI-cells stably expressing doxycyclin-inducible SOX17 (SOX17-EGI-1) or TNS4 (TNS4-EGI-1)

(upper panels) overexpress KRAS and TNS4, and ERBB2 and KRAS respectively (lower panels). (*D*) Doxycyclin-induced TNS4 accelerates migration of EGI-1 and NMC cells in a wound-healing assay. At each timepoint, the width of the wound is indicated as the means +/- SEM (µm) of 3 widths measured on 6 pictures covering the entire wound. (*E*) Doxycyclin-induced SOX17 or TNS4 stimulate proliferation of EGI-1 cells; 16 x 10^4 cells were plated at day 1 (upper panels). Cells illustrated in the upper panels were counted following crystal violet staining (lower left graph). In separate experiments, Hoechst fluorescence intensity of control, SOX17–EGI-1 and TNS4–EGI-1 cells was measured 4 days after plating (lower right graph). Data in graphs are represented +/- SEM. n= 3 or 4. Data in dot plot graphs represent fold inductions relative to control +/- SEM. * $p<0.05$ and *** $p<0.001$.

References

1. Macias RIR, Kornek M, Rodrigues PM, et al. Diagnostic and prognostic biomarkers in cholangiocarcinoma. Liver Int 2019;39 Suppl 1:108-122.
2. Banales JM, Cardinale V, Carpino G, et al. Expert consensus document: Cholangiocarcinoma: current knowledge and future perspectives consensus statement from the European Network for the Study of Cholangiocarcinoma (ENS-CCA). Nat Rev Gastroenterol Hepatol 2016;13:261-80.
3. Moeini A, Sia D, Bardeesy N, et al. Molecular Pathogenesis and Targeted Therapies for Intrahepatic Cholangiocarcinoma. Clin Cancer Res 2016;22:291-300.
4. Blechacz B. Cholangiocarcinoma: Current Knowledge and New Developments. Gut Liver 2017;11:13-26.
5. Sirica AE, Gores GJ, Groopman JD, et al. Intrahepatic Cholangiocarcinoma: Continuing Challenges and Translational Advances. Hepatology 2019;69:1803-1815.
6. Labib PL, Goodchild G, Pereira SP. Molecular Pathogenesis of Cholangiocarcinoma. BMC Cancer 2019;19:185.
7. Roos E, Soer EC, Klompmaker S, et al. Crossing borders: A systematic review with quantitative analysis of genetic mutations of carcinomas of the biliary tract. Crit Rev Oncol Hematol 2019;140:8-16.
8. Fouassier L, Marzioni M, Afonso MB, et al. Signalling networks in cholangiocarcinoma: Molecular pathogenesis, targeted therapies and drug resistance. Liver Int 2019;39 Suppl 1:43-62.

9. Nakagawa H, Suzuki N, Hirata Y, et al. Biliary epithelial injury-induced regenerative response by IL-33 promotes cholangiocarcinogenesis from peribiliary glands. Proc Natl Acad Sci U S A 2017;114:E3806-E3815.
10. Guest RV, Boulter L, Kendall TJ, et al. Cell Lineage Tracing Reveals a Biliary Origin of Intrahepatic Cholangiocarcinoma. Cancer Res 2014;74:1005-1010.
11. Sekiya S, Suzuki A. Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes. J Clin Invest 2012;122:3914-8.
12. Fan B, Malato Y, Calvisi DF, et al. Cholangiocarcinomas can originate from hepatocytes in mice. J Clin Invest 2012;122:2911-5.
13. Tschaharganeh DF, Xue W, Calvisi DF, et al. p53-Dependent Nestin Regulation Links Tumor Suppression to Cellular Plasticity in Liver Cancer. Cell 2016;165:1546-1547.
14. Benhamouche S, Curto M, Saotome I, et al. Nf2/Merlin controls progenitor homeostasis and tumorigenesis in the liver. Genes Dev 2010;24:1718-30.
15. Saha SK, Parachoniak CA, Ghanta KS, et al. Mutant IDH inhibits HNF-4alpha to block hepatocyte differentiation and promote biliary cancer. Nature 2014;513:110-4.
16. Komuta M, Govaere O, Vandecaveye V, et al. Histological diversity in cholangiocellular carcinoma reflects the different cholangiocyte phenotypes. Hepatology 2012;55:1876-1888.
17. Basturk O, Aishima S, Esposito I. Biliary intraepithelial neoplasia. WHO classification of tumours - 5th edition. Digestive system tumours: WHO, 2019:273-275.
18. Nakanuma Y, Basturk O, Esposito I, et al. Intraductal papillary neoplasms of the bile ducts. WHO classification of tumours - 5th edition. Digestive system tumours: WHO, 2019:279-282.
19. Kendall T, Verheij J, Gaudio E, et al. Anatomical, histomorphological and molecular classification of cholangiocarcinoma. Liver Int 2019;39 Suppl 1:7-18.
20. Aishima S, Kubo Y, Tanaka Y, et al. Histological features of precancerous and early cancerous lesions of biliary tract carcinoma. J Hepatobiliary Pancreat Sci 2014;21:448-52.
21. Ohtsuka M, Shimizu H, Kato A, et al. Intraductal papillary neoplasms of the bile duct. Int J Hepatol 2014;2014:459091.
22. Bhalla A, Mann SA, Chen S, et al. Histopathological evidence of neoplastic progression of von Meyenburg complex to intrahepatic cholangiocarcinoma. Hum Pathol 2017;67:217-224.
23. Ong CK, Subimerb C, Pairojkul C, et al. Exome sequencing of liver fluke-associated cholangiocarcinoma. Nat Genet 2012;44:690-3.
24. Churi CR, Shroff R, Wang Y, et al. Mutation profiling in cholangiocarcinoma: prognostic and therapeutic implications. PLoS One 2014;9:e115383.

25. Zou S, Li J, Zhou H, et al. Mutational landscape of intrahepatic cholangiocarcinoma. Nat Commun 2014;5:5696.
26. Nakamura H, Arai Y, Totoki Y, et al. Genomic spectra of biliary tract cancer. Nat Genet 2015;47:1003-10.
27. Javle M, Bekaii-Saab T, Jain A, et al. Biliary cancer: Utility of next-generation sequencing for clinical management. Cancer 2016;122:3838-3847.
28. Hsu M, Sasaki M, Igarashi S, et al. KRAS and GNAS mutations and p53 overexpression in biliary intraepithelial neoplasia and intrahepatic cholangiocarcinomas. Cancer 2013;119:1669-74.
29. Sasaki M, Matsubara T, Nitta T, et al. GNAS and KRAS Mutations are Common in Intraductal Papillary Neoplasms of the Bile Duct. PLoS ONE 2014;8:e81706.
30. Schlitter AM, Born D, Bettstetter M, et al. Intraductal papillary neoplasms of the bile duct: stepwise progression to carcinoma involves common molecular pathways. Mod Pathol 2014;27:73-86.
31. Khan SA, Tavolari S, Brandi G. Cholangiocarcinoma: Epidemiology and risk factors. Liver Int 2019;39 Suppl 1:19-31.
32. Andersen JB, Spee B, Blechacz BR, et al. Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. Gastroenterology 2012;142:1021-1031.
33. Sia D, Hoshida Y, Villanueva A, et al. Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. Gastroenterology 2013;144:829-40.
34. Loeuillard E, Fischbach SR, Gores GJ, et al. Animal models of cholangiocarcinoma. Biochim Biophys Acta 2018;1865:982-992.
35. Cadamuro M, Brivio S, Stecca T, et al. Animal models of cholangiocarcinoma: What they teach us about the human disease. Clin Res Hepatol Gastroenterol 2018;42:403-415.
36. Pirenne S, Lemaigre F. Genetically-engineered animal models of biliary tract cancers. Current Opinion in Gastroenterology 2020;220.
37. Espanol-Suner R, Carpentier R, Van Hul N, et al. Liver progenitor cells yield functional hepatocytes in response to chronic liver injury in mice. Gastroenterology 2012;143:1564-1575.
38. Lesaffer B, Verboven E, Van Huffel L, et al. Comparison of the Opn-CreER and Ck19-CreER Drivers in Bile Ducts of Normal and Injured Mouse Livers. Cells 2019;8:380.
39. Hingorani SR, Petricoin EF, Maitra A, et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 2003;4:437-50.

40. Kowarz E, Loscher D, Marschalek R. Optimized Sleeping Beauty transposons rapidly generate stable transgenic cell lines. Biotechnol J 2015;10:647-53.
41. Mates L, Chuah MK, Belay E, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nat Genet 2009;41:753-61.
42. Ueno Y, Alpini G, Yahagi K, et al. Evaluation of differential gene expression by microarray analysis in small and large cholangiocytes isolated from normal mice. Liver Int 2003;23:449-59.
43. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166-9.
44. von Ahlfen S, Missel A, Bendrat K, et al. Determinants of RNA quality from FFPE samples. PLoS One 2007;2:e1261.
45. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010;11:R106.
46. Fickert P, Stoger U, Fuchsbichler A, et al. A new xenobiotic-induced mouse model of sclerosing cholangitis and biliary fibrosis. Am J Pathol 2007;171:525-36.
47. Schmitz KJ, Lang H, Wohlschlaeger J, et al. AKT and ERK1/2 signaling in intrahepatic cholangiocarcinoma. World J Gastroenterol 2007;13:6470-7.
48. El Khatib M, Bozko P, Palagani V, et al. Activation of Notch signaling is required for cholangiocarcinoma progression and is enhanced by inactivation of p53 in vivo. PLoS One 2013;8:e77433.
49. Sirica AE. Role of ErbB family receptor tyrosine kinases in intrahepatic cholangiocarcinoma. World J Gastroenterol 2008;14:7033-58.
50. Zhou G, Cao P, Li Y. RNA over-editing leads to aggressiveness of intrahepatic cholangiocarcinoma. Gene Expression Omnibus, accession number GSE119336 2019.
51. Delgiorno KE, Hall JC, Takeuchi KK, et al. Identification and manipulation of biliary metaplasia in pancreatic tumors. Gastroenterology 2014;146:233-244
52. Chan LK, Chiu YT, Sze KM, et al. Tensin4 is up-regulated by EGF-induced ERK1/2 activity and promotes cell proliferation and migration in hepatocellular carcinoma. Oncotarget 2015;6:20964-76.
53. Corada M, Orsenigo F, Morini MF, et al. Sox17 is indispensable for acquisition and maintenance of arterial identity. Nat Commun 2013;4:2609.
54. Pellat A, Vaquero J, Fouassier L. Role of ErbB/HER family of receptor tyrosine kinases in cholangiocyte biology. Hepatology 2018;67:762-773.
55. Li J, Razumilava N, Gores GJ, et al. Biliary repair and carcinogenesis are mediated by IL-33-dependent cholangiocyte proliferation. J Clin Invest 2014;124:3241-51.

56. Yamada D, Rizvi S, Razumilava N, et al. IL-33 facilitates oncogene-induced cholangiocarcinoma in mice by an interleukin-6-sensitive mechanism. Hepatology 2015;61:1627-42.
57. Uemura M, Ozawa A, Nagata T, et al. Sox17 haploinsufficiency results in perinatal biliary atresia and hepatitis in C57BL/6 background mice. Development 2013;140:639-48.
58. Spence JR, Lange AW, Lin SC, et al. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. Dev Cell 2009;17:62-74.
59. Higashiyama H, Ozawa A, Sumitomo H, et al. Embryonic cholecystitis and defective gallbladder contraction in the Sox17-haploinsufficient mouse model of biliary atresia. Development 2017;144:1906-1917.
60. Uhlen M, Oksvold P, Fagerberg L, et al. Towards a knowledge-based Human Protein Atlas. Nat Biotechnol 2010;28:1248-50.
61. Merino-Azpitarte M, Lozano E, Perugorria MJ, et al. SOX17 regulates cholangiocyte differentiation and acts as a tumor suppressor in cholangiocarcinoma. J Hepatol 2017;67:72-83.
62. Goeppert B, Konermann C, Schmidt CR, et al. Global alterations of DNA methylation in cholangiocarcinoma target the Wnt signaling pathway. Hepatology 2014;59:544-54.
63. Dumur CI, Campbell DJ, DeWitt JL, et al. Differential gene expression profiling of cultured neu-transformed versus spontaneously-transformed rat cholangiocytes and of corresponding cholangiocarcinomas. Exp Mol Pathol 2010;89:227-35.
64. Al-Ghamdi S, Albasri A, Cachat J, et al. Cten is targeted by Kras signalling to regulate cell motility in the colon and pancreas. PLoS One 2011;6:e20919.

Figure 1

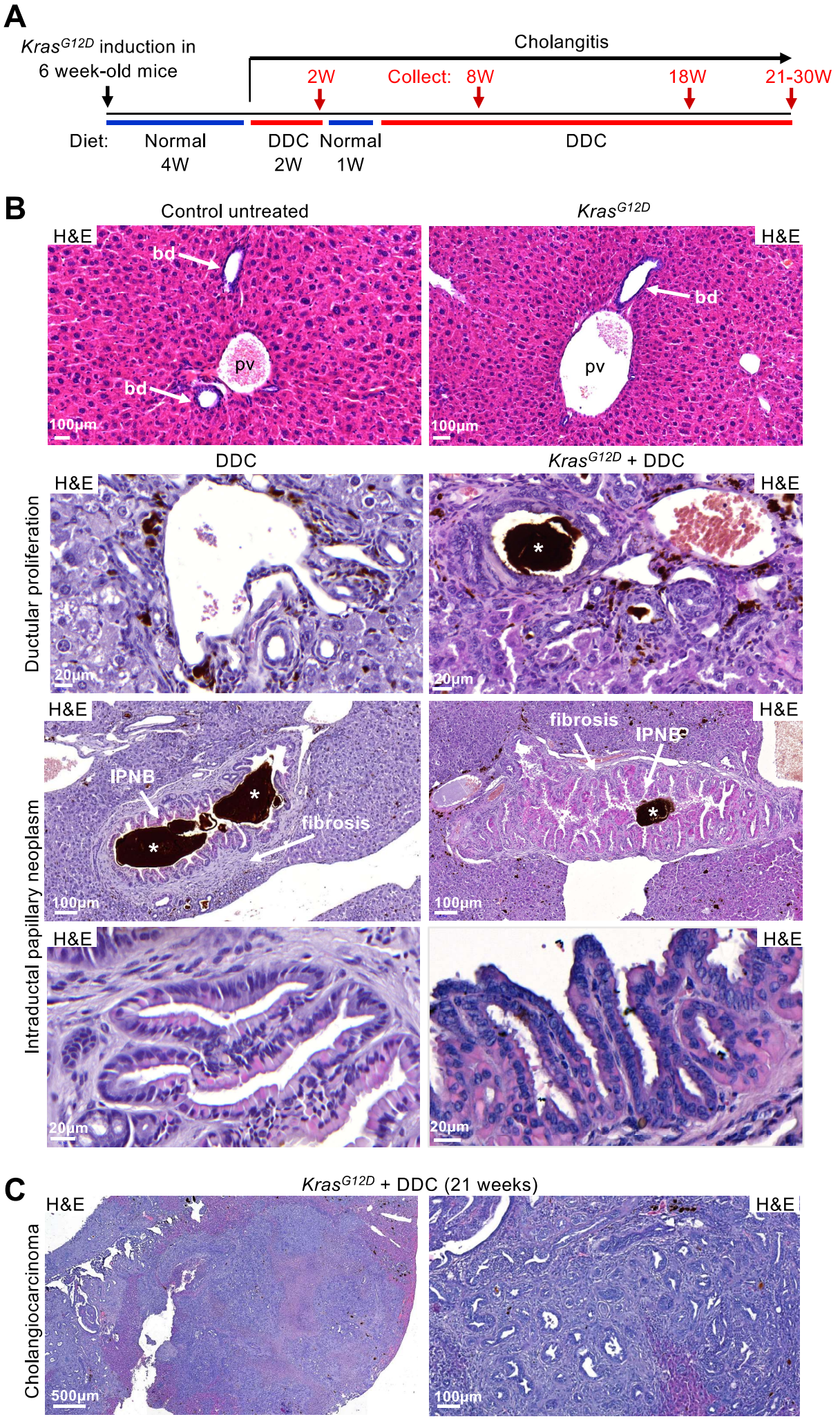
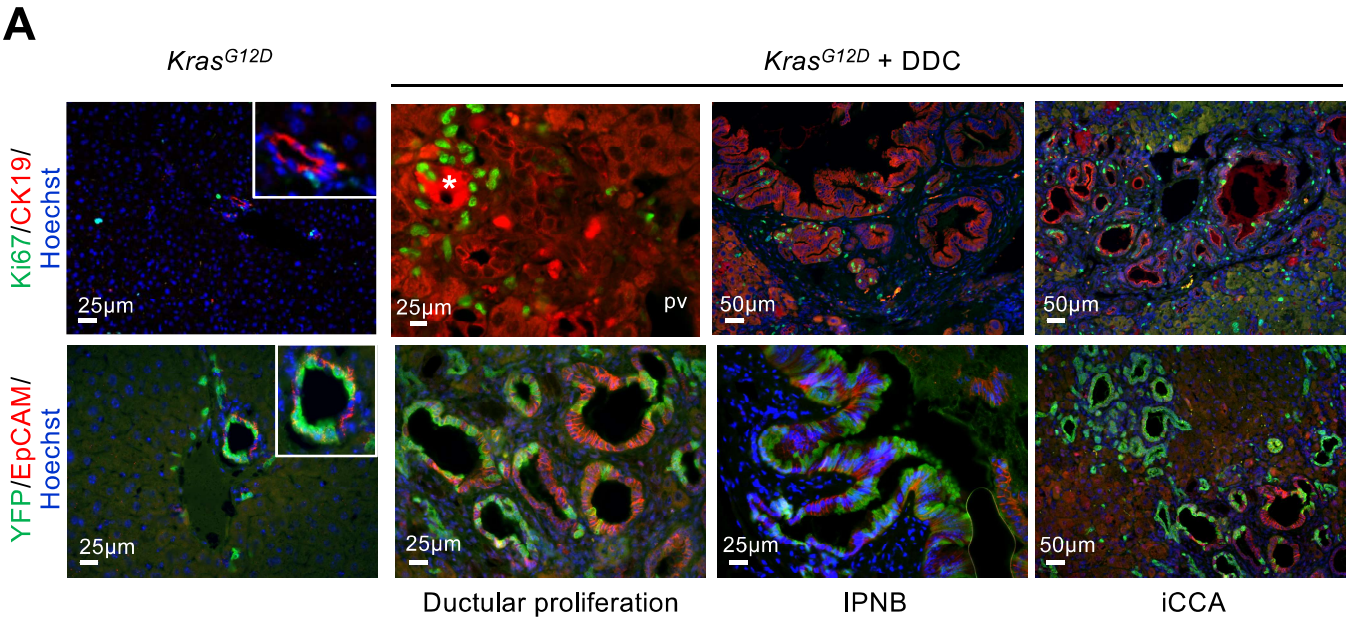
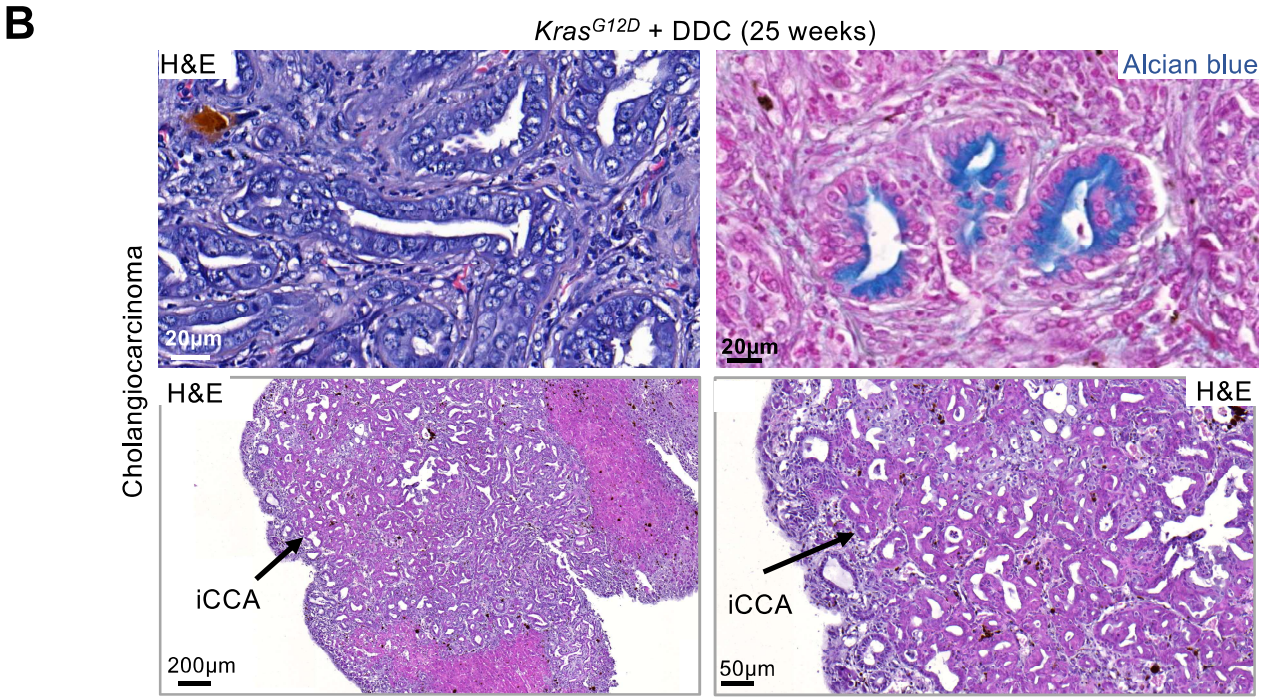

A
Kras^G12D induction in
6 week-old mice
Cholangitis
2W
Collect: 8W
18W
21-30W
Diet:
Normal
4W
DDC
2W
Normal
1W
DDC
B
Control untreated
Kras^G12D
H&E
bd
pv
100µm
DDC
Kras^G12D + DDC
Ductular proliferation
20µm
IPNB
fibrosis
Intraductal papillary neoplasm
C
Kras^G12D + DDC (21 weeks)
Cholangiocarcinoma
500µm

Figure 2


A
Kras G12D
Kras G12D + DDC
Ki67/CK19/
Hoechst
25µm
25µm
pv
50µm
50µm
YFP/EpCAM/
Hoechst
25µm
25µm
25µm
50µm
Ductular proliferation
IPNB
iCCA


B
Kras G12D + DDC (25 weeks)
Cholangiocarcinoma
H&E
20µm
Alcian blue
20µm
H&E
iCCA
200µm
H&E
iCCA
50µm

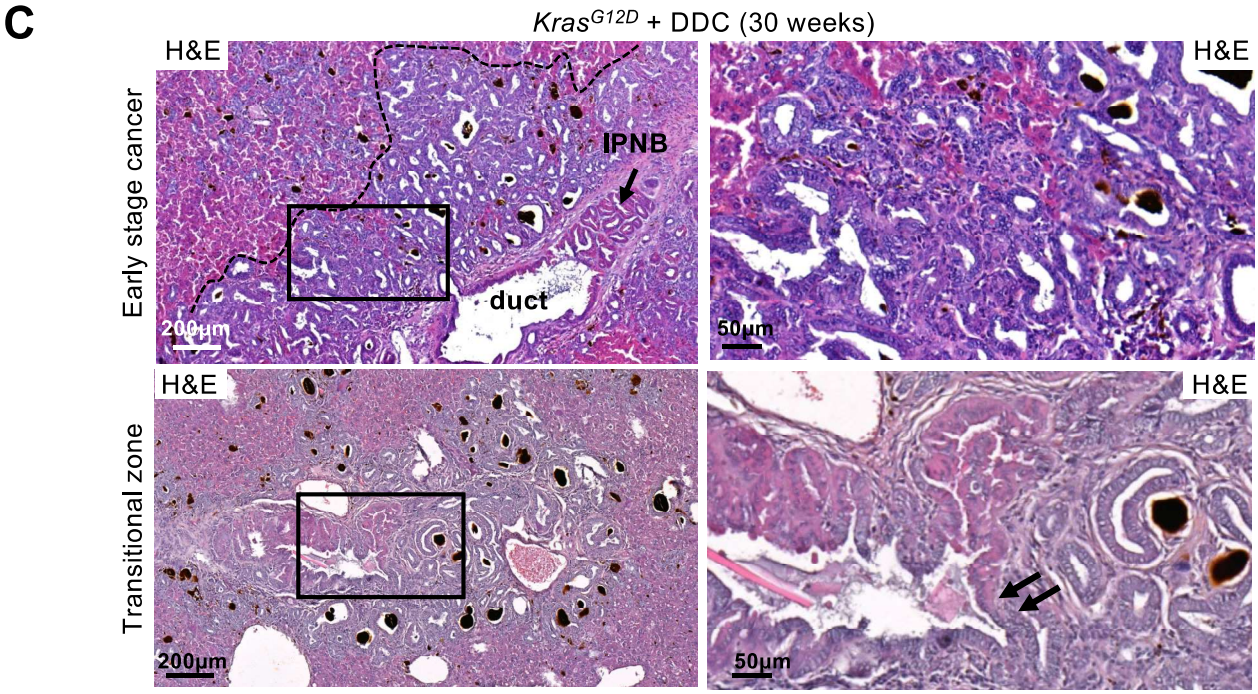
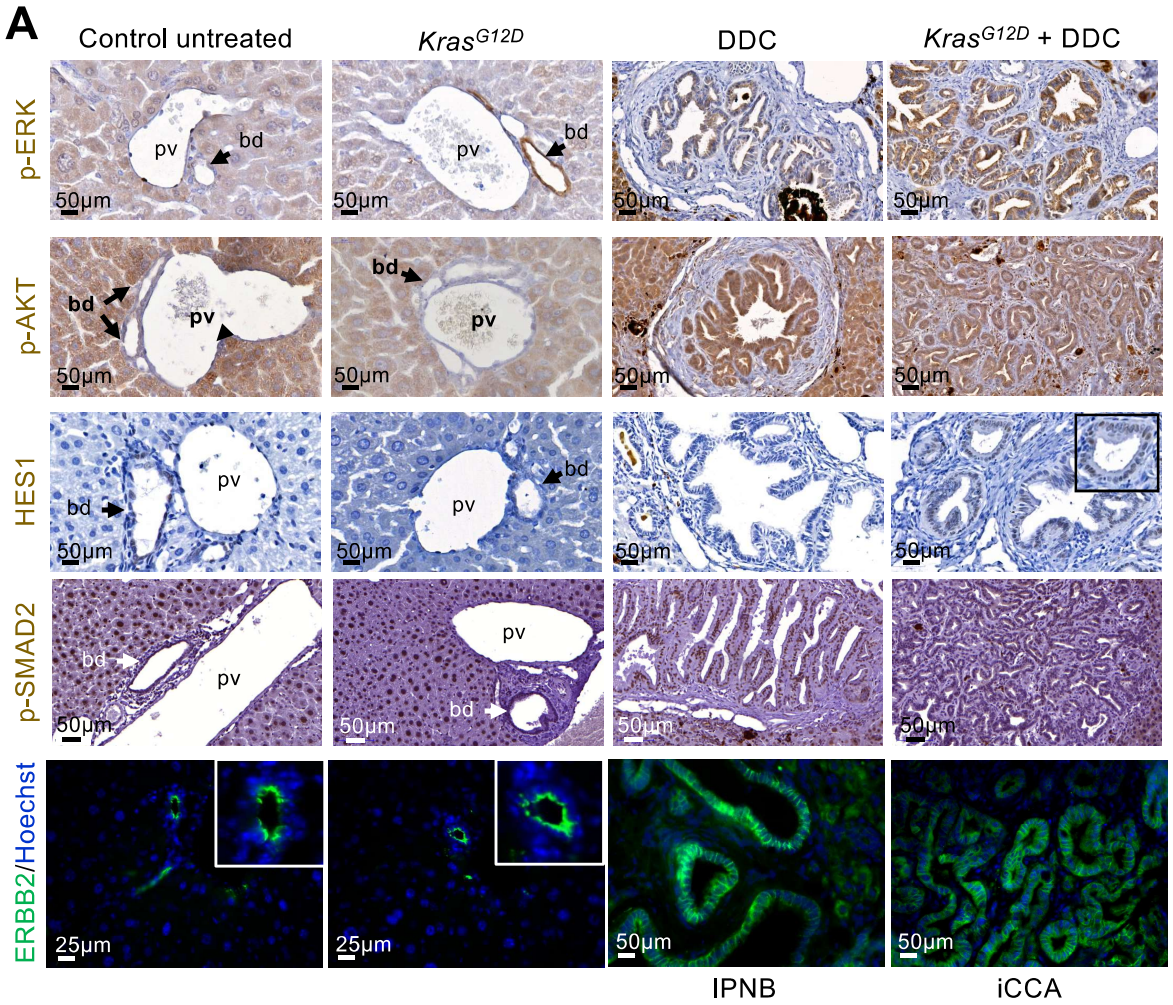
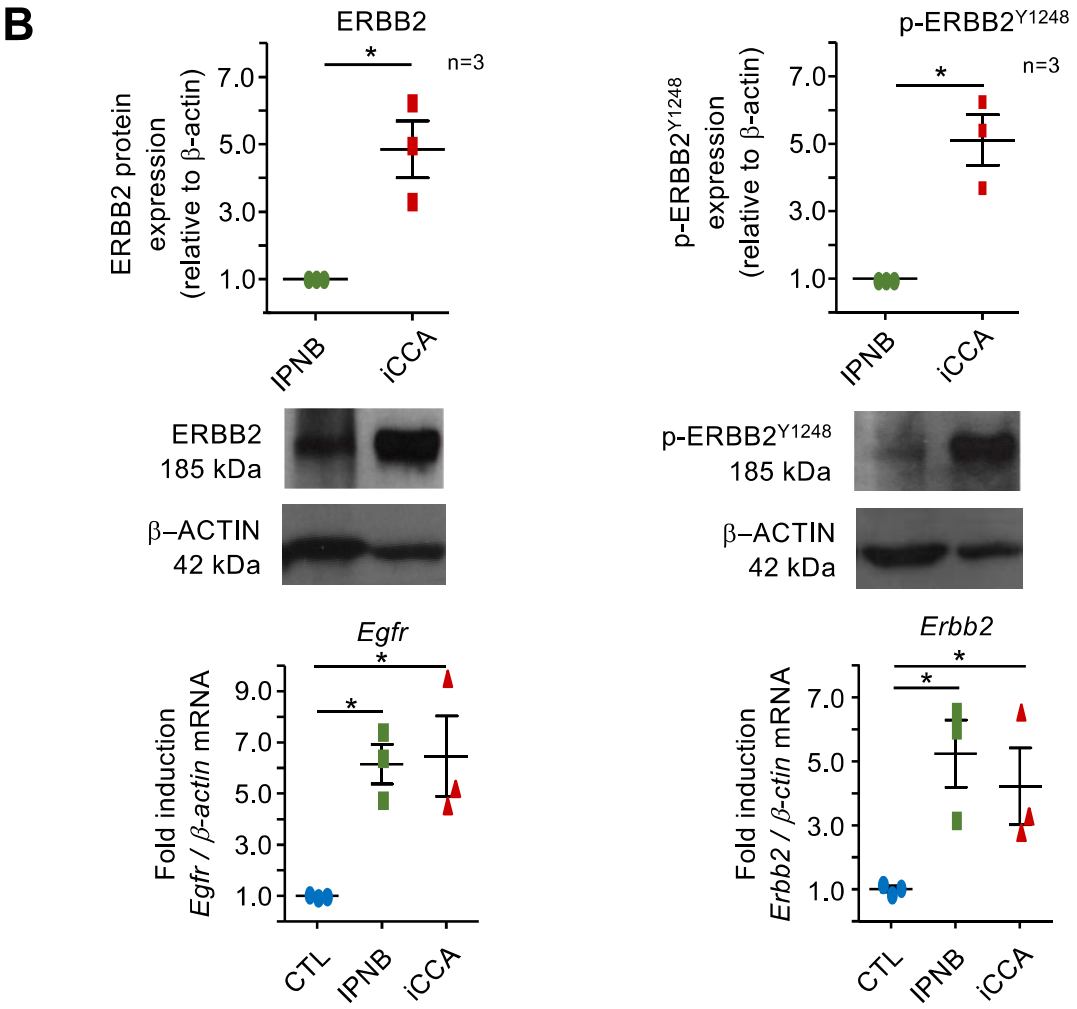
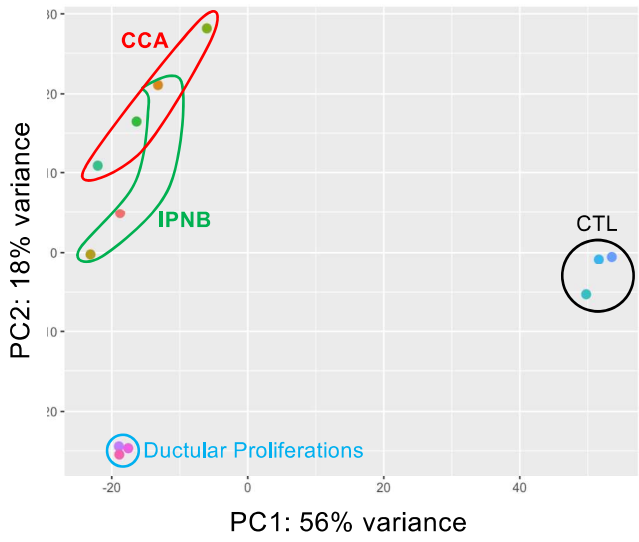

C
Kras G12D + DDC (30 weeks)
Early stage cancer
H&E
IPNB
duct
200µm
H&E
50µm
Transitional zone
H&E
200µm
H&E
50µm

Figure 3


A
Control untreated
Kras G12D
DDC
Kras G12D + DDC
p-ERK
p-AKT
HES1
p-SMAD2
ERBB2/Hoechst
pv
bd
50µm
25µm
IPNB
iCCA


B
ERBB2
ERBB2 protein expression (relative to β-actin)
n=3
IPNB
iCCA
ERBB2
185 kDa
β–ACTIN
42 kDa
p-ERBB2Y1248
p-ERBB2Y1248 expression (relative to β-actin)
Egfr
Fold induction Egfr / β-actin mRNA
CTL
Erbb2
Fold induction Erbb2 / β-ctin mRNA

A



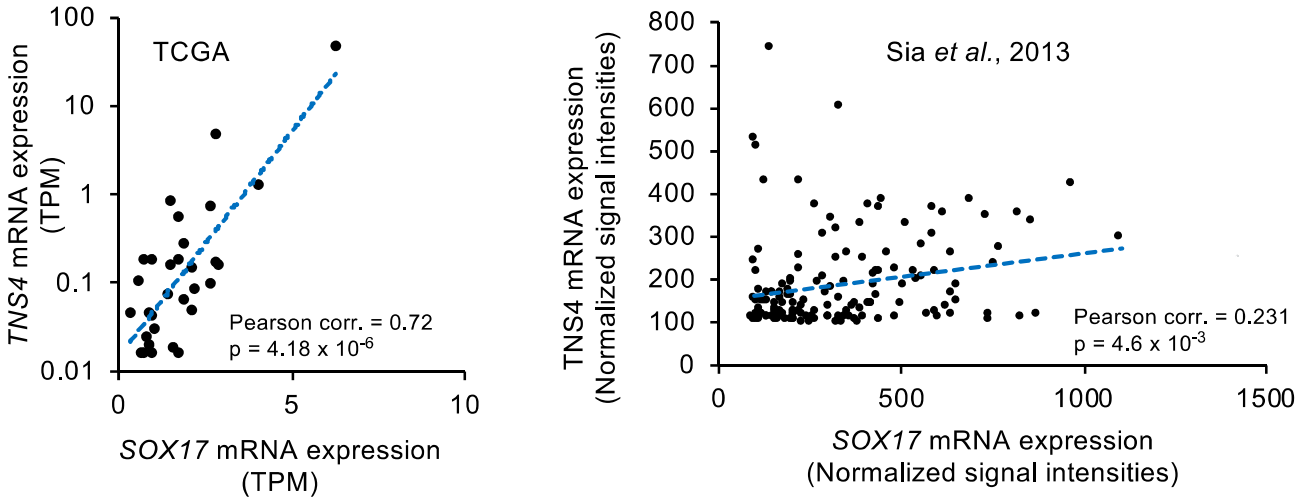
B

Symbol	Fold Change iCCA/IPNB	p_{adj}
Pax9	33.3955802	0.04872871
Dpcr1	20.7365393	0.03500676
Atg9b	11.0400319	4.8521E-05
Naip1	6.71720132	0.02432953
Aldh1a3	6.60987806	0.01341438
Tns4	**6.4800418**	**0.00072911**
Sox17	**6.16912873**	**0.02973183**
Aspa	5.21554345	0.01767884
Pmm1	3.42577373	0.02973183
Wdr89	2.93720232	0.02973183
Epha3	-2.32462465	0.04865337
Plet1	-3.50171381	0.00028544
Trpv6	-4.59693332	0.01341438
Lcn2	-4.68625364	0.04872871
S1pr4	-6.15198467	0.04561029
Nos2	-7.24047163	0.03360769
Ighv7-3	-21.0920687	0.04800234
Muc13	-56.7170907	4.8521E-05
Igkv4-80	-100.699452	0.0046736

C

Symbol	Mouse iCCA/CTL	p_{adj}	Human TCGA Cohort iCCA/CTL (n=34)	p_{adj}	Human Cohort GSE119336 iCCA/CTL (n=15)	p_{adj}
Aldh1a3	119.0455614	3.43E-27	6.347926756	0.0001758	0.746366038	3.28E-07
Epha3	71.66947934	1.15E-40	2.93108534	0.000796322	1.531730952	2.91E-07
Atg9b	51.6336212	8.32E-18	27.44873263	1.53E-05	3.417971186	4.78E-07
Pax9	44.77763278	9.47E-06	210.783572	1.65E-04	67.98491672	1.15E-05
Dpcr1	44.30280819	1.20E-07	37.65092673	0.324689899	1.914066514	4.59E-07
Tns4	**39.95113205**	**7.69E-25**	**184.73067**	**0.146014049**	**13.52238906**	**1.14E-06**
Naip1	22.96508145	9.91E-13	6.583127377	4.04E-09	1.114004562	5.34E-07
Sox17	**20.97516544**	**0.0001775**	**1.92208856**	**0.004239907**	**1.652591106**	**3.27E-08**
Pmm1	12.4122368	1.52E-15	1.748436572	0.000360451	0.779193434	1.09E-07
Wdr89	2.139602879	0.00384384	2.217813442	9.24E-09	1.412188956	6.27E-08

D



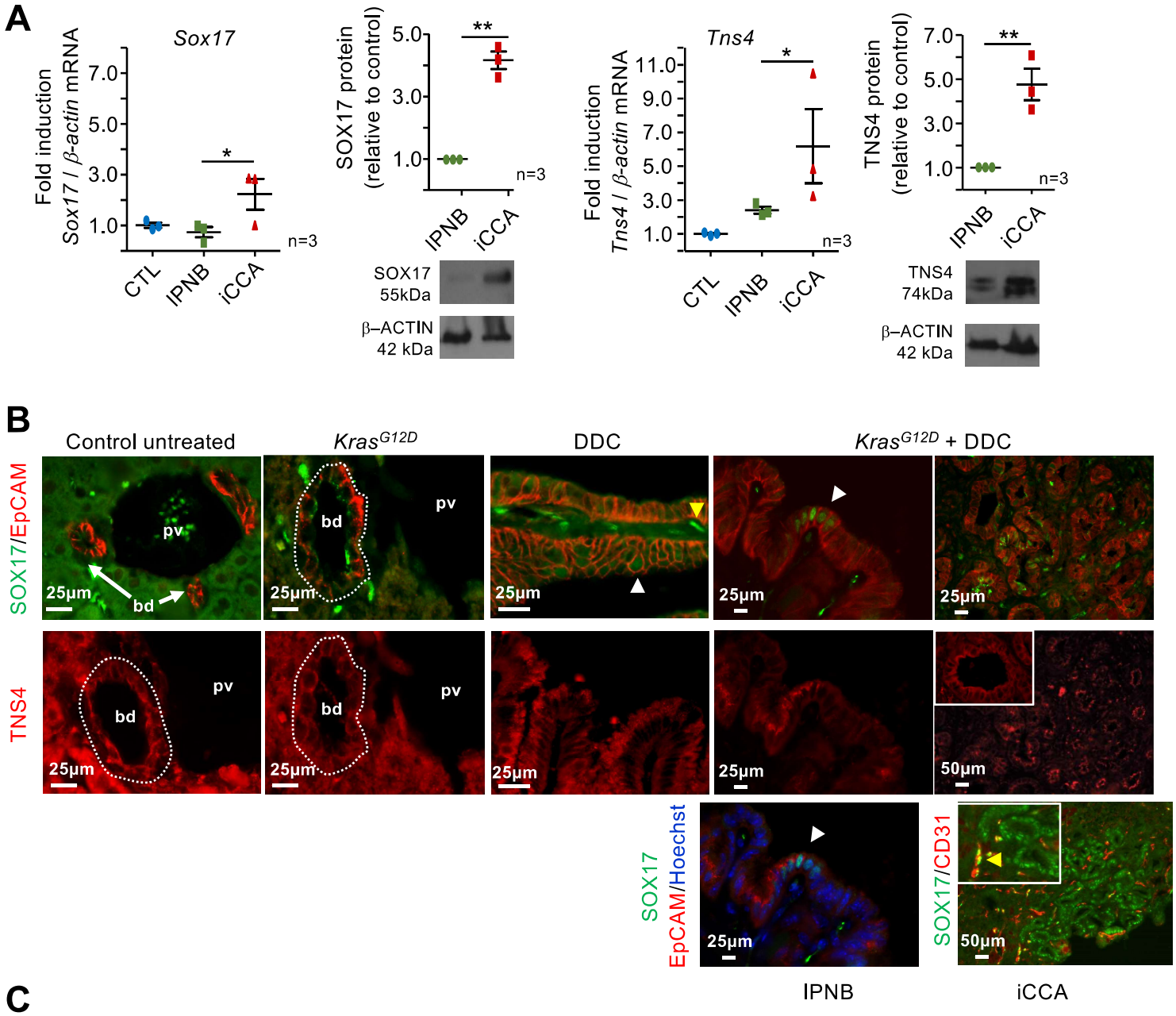
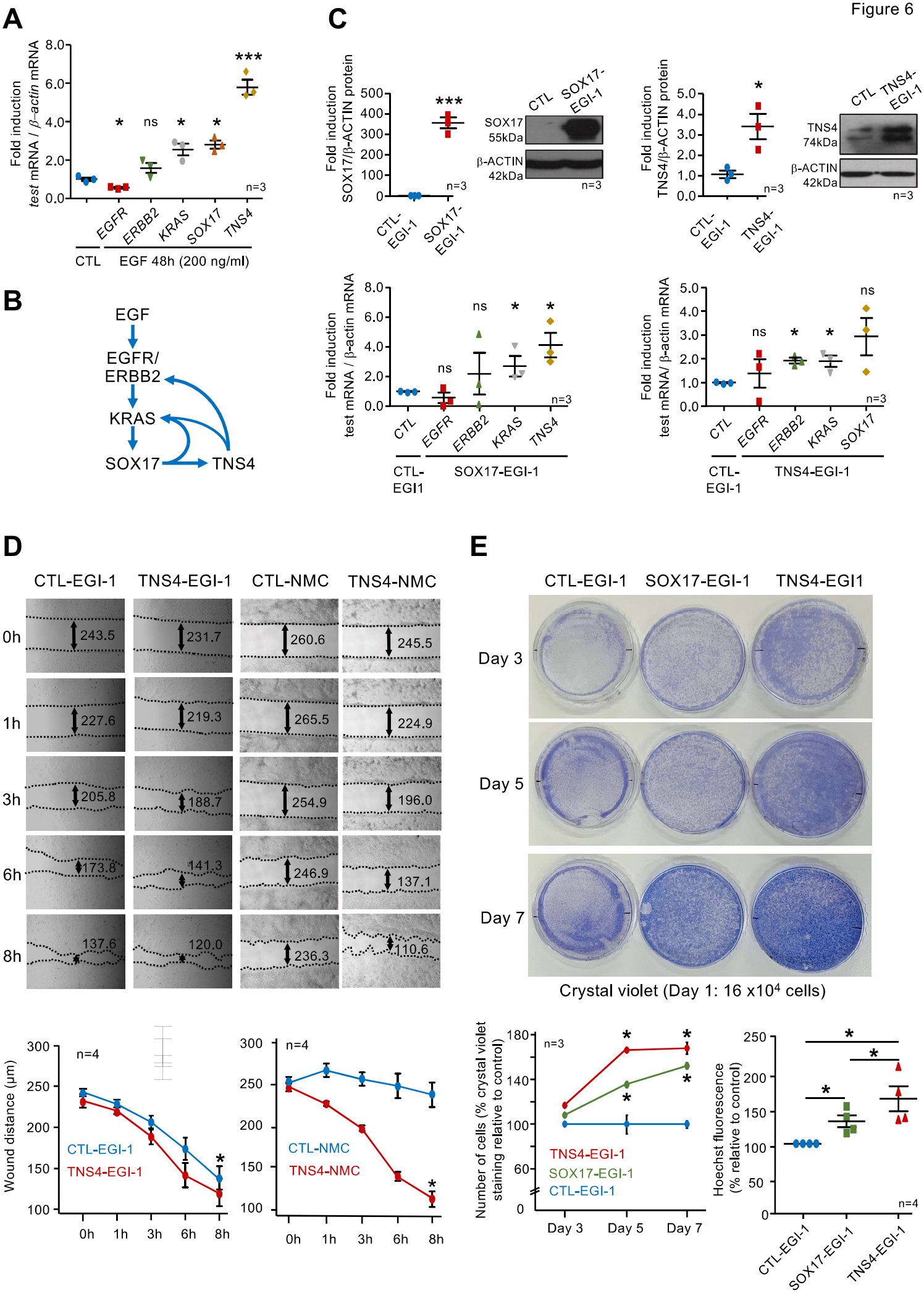

A
Sox17
Fold induction
Sox17 / β-actin mRNA
7.0
5.0
3.0
1.0
*
n=3
CTL
IPNB
iCCA
SOX17 protein
(relative to control)
5.0
3.0
1.0
**
n=3
IPNB
iCCA
SOX17
55kDa
β–ACTIN
42 kDa
Tns4
Fold induction
Tns4 / β-actin mRNA
11.0
9.0
7.0
5.0
3.0
1.0
*
n=3
CTL
IPNB
iCCA
TNS4 protein
(relative to control)
7.0
5.0
3.0
1.0
**
n=3
IPNB
iCCA
TNS4
74kDa
β–ACTIN
42 kDa
B
Control untreated
KrasG12D
DDC
KrasG12D + DDC
SOX17/EpCAM
TNS4
pv
bd
25µm
50µm
SOX17
EpCAM/Hoechst
SOX17/CD31
IPNB
iCCA
C
Control human liver
Human iCCA
SOX17
TNS4
SOX17/
TNS4/Hoechst
bd
ns
pv
25µm
10µm
50µm

Figure 6


A
Fold induction
test mRNA / β-actin mRNA
8.0
6.0
4.0
2.0
0.0
*
ns
*
*
***
n=3
EGFR
ERBB2
KRAS
SOX17
TNS4
CTL
EGF 48h (200 ng/ml)
B
EGF
EGFR/
ERBB2
KRAS
SOX17
TNS4
C
Fold induction
SOX17/β-ACTIN protein
500
400
300
200
100
0
***
n=3
CTL-
EGI-1
SOX17-
EGI-1
CTL
SOX17-
EGI-1
SOX17
55kDa
β-ACTIN
42kDa
n=3
Fold induction
TNS4/β-ACTIN protein
5.0
4.0
3.0
2.0
1.0
0.0
*
n=3
CTL-
EGI-1
TNS4-
EGI-1
CTL
TNS4-
EGI-1
TNS4
74kDa
β-ACTIN
42kDa
n=3
Fold induction
test mRNA / β-actin mRNA
8.0
6.0
4.0
2.0
0.0
ns
ns
*
*
n=3
CTL
EGFR
ERBB2
KRAS
TNS4
CTL-
EGI1
SOX17-EGI-1
Fold induction
test mRNA/ β-actin mRNA
5.0
4.0
3.0
2.0
1.0
0.0
ns
ns
*
*
ns
n=3
CTL
EGFR
ERBB2
KRAS
SOX17
CTL-
EGI-1
TNS4-EGI-1
D
CTL-EGI-1
TNS4-EGI-1
CTL-NMC
TNS4-NMC
0h
243.5
231.7
260.6
245.5
1h
227.6
219.3
265.5
224.9
3h
205.8
188.7
254.9
196.0
6h
173.8
141.3
246.9
137.1
8h
137.6
120.0
236.3
110.6
Wound distance (µm)
300
250
200
150
100
n=4
CTL-EGI-1
TNS4-EGI-1
*
0h
1h
3h
6h
8h
300
250
200
150
100
n=4
CTL-NMC
TNS4-NMC
*
0h
1h
3h
6h
8h
E
CTL-EGI-1
SOX17-EGI-1
TNS4-EGI1
Day 3
Day 5
Day 7
Crystal violet (Day 1: 16 x10⁴ cells)
Number of cells (% crystal violet staining relative to control)
180
140
100
60
0
n=3
*
*
*
*
TNS4-EGI-1
SOX17-EGI-1
CTL-EGI-1
Day 3
Day 5
Day 7
Hoechst fluorescence (% relative to control)
250
200
150
100
50
0
*
*
*
n=4
CTL-EGI-1
SOX17-EGI-1
TNS4-EGI-1