Dissertation

submitted to the

Combined Faculties for Natural Sciences and for Mathemathics
of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by
Noujan Ganjian
M. Sc. Infection Biology
Oral-examination: January 30th, 2024

Identification of a plasticity driver of Combined Hepatocellular-Cholangiocarcinoma using functional interspecies comparison

Abstract

Primary liver cancer (PLC) comprises hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (iCCA), which may occur together as combined HCC-CCA (cHCC-CCA). For both tumor components, a single cell of origin is suggested. It is debated whether the cell of origin belongs to the stem/progenitor cell type or may represent a fully differentiated liver cell. It is believed that iCCA develops from biliary epithelial cells and that hepatocytes give rise to HCC. However, lineage tracing experiments in mice showed that iCCA may develop from hepatocytes while HCC can originate from ductular cells. These results suggest that differentiated liver epithelial cells can transdifferentiate based on their cellular plasticity.

Mutation spectra of both hepatocellular carcinoma and cholangiocarcinoma have been published: However, it remains to be elucidated, which of these alterations are driving the tumor phenotype or determine its biological behavior. Determining the underlying mechanisms will have profound impact on our understanding of cancer biology and treatment options.

In this project, an interspecies approach combining human, mouse, and *in vitro* data was performed to identify factors that affect the phenotype of liver cancer cells. Human primary liver cancer samples were selected by morphological analysis and subjected to genome-wide exomic and transcriptomic profiling. Next generation sequencing (NGS) data were integrated and candidate genes potentially affecting the phenotype of the tumor cells were identified. Between the two components of cHCC-CCA, 54 differentially expressed and/or differentially mutated genes were found. These were functionally validated by *in vivo* RNAi screening in mosaic mouse models of HCC (*MYC-AKT1* in wildtype mice) and iCCA (*KRAS^{G12V}* in p19-deficient mice), which were generated by hydrodynamic tail vein injection of transposon vectors.

Histological analysis of formalin-fixed paraffin-embedded individual tumor nodules followed by immunohistological assessment of hepatocellular (Hnf4α) and biliary (Sox9, Krt19) markers allowed for the identification of potential phenotype modulating genes.

Thrombospondin 3 (THBS3) was identified as a phenotypic driver: *THBS3* was mutated in the HCC compartment of human cHCC-CCA. Both its knockdown and expression of the synonymous mutation *Thbs3^{R102Q}* in the iCCA mouse model resulted in a cHCC-CCA phenotype. Functional analyses were conducted in a primary murine isogenic iCCA cell line *in vitro*. The findings were confirmed in a human iCCA cell line. In both murine iCCA and HCC

models, *Thbs3* wildtype maintains a cholangiocytic tumor phenotype (*Sox9+, Krt19+, Hnf4α-*) *in vitro*, while *Thbs3^{R102Q}* expression promoted a hepatoid phenotype (*Sox9-, Krt19-, Hnf4α+*). *Thbs3* knockdown largely phenocopied *Thbs3^{R102Q}* expression in the iCCA cell line. Also, the results from functional assays suggested a tumor suppressor role of Thbs3 wildtype, while at the same time Thbs3 wildtype seems to be essential for the survival of iCCA cells.

Furthermore, transcriptomic and ATAC sequencing analysis was performed on Thbs3 variant-expressing and knockdown cell lines in the iCCA model. NGS data integration and subsequent GSEA against the M2 mouse collection of the Molecular Signatures Database enabled the retrieval of four canonical pathways, including TGFβ signaling. Concluding, this study found that *THBS3* knockdown alters TGFβ signaling at transcriptional level.

Zusammenfassung

Primäre Leberkarzinome werden in hepatozelluläre (HCC), cholangiozelluläre (CCA) und gemischte hepatozelluläre und cholangiozelluläre Karzinome (cHCC-CCA) unterteilt. Die Ursprungszelle von cHCC-CCA steht gegenwärtig zur Debatte. Es ist strittig, ob cHCC-CCA aus Progenitorzellen oder differenzierten Leberzellen entstehen. Man geht allgemein davon aus, dass Cholangiokarzinome aus Cholangiozyten und hepatozelluläre Karzinome aus Hepatozyten entstehen. *Lineage-tracing*-Ansätze zeigten jedoch, dass Cholangiokarzinome auch aus Hepatozyten und hepatozelluläre Karzinome auch aus Cholangiozyten entstehen können. Diese Ergebnisse legen nahe, dass differenzierte Leberepithelzellen auf der Grundlage ihrer zellulären Plastizität die Fähigkeit zur Transdifferenzierung besitzen.

Die Mutationsspektren von hepatozellulären- und cholangiozellulären Karzinomen sind bekannt. Es bleibt jedoch zu klären, welche dieser Veränderungen den Tumorphänotyp oder das biologische Verhalten bestimmen. Die Identifikation der zugrundeliegenden Mechanismen wird tiefgreifende Auswirkungen auf unser Verständnis und der Behandlungsoptionen haben.

In diesem Projekt wurde ein speziesübergreifender Ansatz durchgeführt, bei dem humane, murine und *in vitro* Daten integriert wurden. Ziel war die Identifikation von Faktoren, die den Phänotyp von Leberkrebszellen bestimmen. Menschliche primäre Leberkrebsproben wurden Morphologie basiert ausgewählt und einem genomweiten exomischen und transkriptomischen Profiling unterzogen. Durch die Integration von NGS-Daten wurden 54 Gene selektiert, die möglicherweise den Phänotyp der Tumorzellen beeinflussen. Diese zeigten eine differenzielle Expression und/oder differenziell nachweisbare Mutation zwischen den beiden Komponenten von cHCC-CCA. Die Kandidaten wurden mittels eines multiplex RNA-Interferenzscreens in zwei Mosaik-Mausmodellen (HCC: *MYC-AKT1* in Wildtyp Mäusen; iCCA: *KRAS^{G12V}* in p19-defizienten Mäusen) funktionell validiert. Die Mausmodelle wurden mittels hydrodynamischer Schwanzveneninjektion von Transposonvektoren generiert.

Die histologische Analyse Formalin-fixierter, in Paraffin-eingebetteter individueller Tumorknoten, gefolgt von einer immunhistologischen Bewertung hepatozellulärer (Hnf4α) und biliärer (Sox9, Krt19) Marker ermöglichte die Identifizierung potenzieller Modulatoren des Phänotyps.

Thrombospondin 3 (*THBS3*) wurde als phänotypisches Treibergen identifiziert. Eine *THBS3*-Mutation war in der HCC-Komponente eines humanen cHCC-CCA nachweisbar. Im iCCA-

Mausmodell führten sowohl *Knockdown* als auch die Expression der synonymen Mutation *Thbs3^{R102Q}* zu einem cHCC-CCA-Phänotyp. Funktionelle Analysen wurden *in vitro* in einer primären murinen isogenen iCCA-Zelllinie und bestätigend in einer humanen iCCA-Zelllinie durchgeführt. In sowohl muriner iCCA als auch HCC-Zelllinie förderte *Thbs3* Expression den iCCA-Phänotyp (*Sox9+, Krt19+, Hnf4α-*) *in vitro*, wohingegen die *Thbs3^{R102Q}*-Expression zu einem eher HCC-ähnlichen Phänotypen (*Sox9-, Krt19-, Hnf4α+*) führte. *Thbs3-Knockdown* zeigte einen ähnlichen Effekt wie die *Thbs3^{R102Q}*-Expression. Schließlich wiesen die Ergebnisse der funktionellen Analysen auf eine Tumorsuppressor Funktion von wildtypischem Thbs3 hin, während Thbs3 gleichzeitig für das Überleben von iCCA-Zellen essenziell zu sein scheint.

Des Weiteren wurden Transkriptom- und ATAC-Sequenzierungsanalysen an Thbs3-Varianten-exprimierenden und Knockdown-Zelllinien im iCCA-Modell durchgeführt. Die Integration der NGS-Daten und anschließende GSEA gegen die M2-Maus-Sammlung der Molecular Signatures Database ermöglichte die Identifizierung von vier kanonischen Signalwegen, einschließlich des TGFβ-Signalwegs. Zusammenfassend zeigte diese Studie, dass *THBS3-Knockdown* den TGFβ-Signalweg auf transkriptioneller Ebene beeinflusst.

Table of Contents

List of abbreviations

AFP	Alpha-Fetoprotein
AKT1	AKT serine/threonine kinase 1
Alb	Albumin
ANNOVAR	ANNOtate VARiation
ANOVA	Analysis of variance
APS	Ammonium persulfate
aqua dest.	Distilled water
ATAC	Assay for Transposase-Accessible Chromatin
ATP	Adenosine triphosphate
BEC	Biliary epithelial cells
BSA	Bovine serum albumin
CAF	Cancer-associated fibroblast
CaKIG	pCaggs KRAS^{G12V}-IRES-GFP
CaMIA	pCaggs MYC-IRES-AKT1
CCA	Cholangiocarcinoma
cDNA	Complementary DNA
CO_2	Carbon dioxide
Ctrl	Control
DAA	Direct antiviral agents
DAPI	4,6-Diamidino-2-phenylindole
dH20	Distilled DNase, RNase free water
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DN	Dysplastic nodules
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
Dox	Doxycycline

DTT	Dithiothreitol
eCCA	Extrahepatic cholangiocarcinoma
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-mesenchymal transition
EtOH	Ethanol
FC	Fold change
FDR	False Discovery Rate
FFPE	formalin-fixed paraffin-embedded
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GEMM	Genetically engineered mouse model
GemOx	Gemcitabine/oxaliplatin
GFP	Green fluorescent protein
GSEA	Gene Set Enrichment Analysis
H&E	Hematoxylin and eosin
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDTVI	Hydrodynamic Tail Vein Injection
HEK293T	Human embryonic kidney 293T
HEPES	N-2-hydroxyethylpiperazine-N′‘-2-ethanesulfonic acid
HNF4α	Hepatocyte nuclear factor 4
HSC	Hepatic stellate cell
IAP	Integrin-associated protein
ICI	Immune checkpoint inhibitor
IR	Inverted repeats
KRAS	KRAS proto-oncogene, GTPase
KRT19	Keratin 19
KRT7	Keratin 7

lincRNA	Long non-coding RNA
logFC	Log fold-change
LPC	Liver progenitor cells
MAPK	Mitogen-activated protein kinase
MDR	Multidrug resistant
Mut	Mutant
MYC	Myc proto-oncogene protein
NaCl	Sodium chloride
NAFLD	Nonalcoholic fatty liver disease
NASH	Nonalcoholic steatohepatitis
NECD	Notch receptor extracellular domain
NGS	Next generation sequencing
NICD	Notch intracellular cytoplasmic domain
OD	Optical density
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
PFA	Paraformaldehyde
pH	Potential of hydrogen
PLC	Primary liver cancer
PMSF	Phenylmethylsulphonyl fluoride
PolyPhen-2	Polymorphism Phenotyping v2
PSC	Primary sclerosing cholangitis
RE	Restriction enzyme
RFA	Radiofrequency ablation
RNA	Ribonucleic acid
RNAi	RNA interference
RPMI	Roswell Park Memorial Institute
rSAP	Recombinant Shrimp Alkaline Phosphatase

RT	Room temperature
SB13	Sleeping Beauty 13
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	Standard error of the mean
shRNA	Short hairpin RNA
TACE	Trans-arterial chemoembolization
TAM	Tumor-associated macrophages
TCGA	The Cancer Genome Atlas
TE-Buffer	Tris-acetate-EDTA buffer
TEMED	Tetramethylethylenediamine
TGFβ	Transforming growth factor β
THBS	Thrombospondins
TKI	Tyrosine kinase inhibitor
TME	Tumor microenvironment
TNFα	Tumor Necrosis Factor α
Treg	Regulatory T cells
Tris	Trisaminomethane
WB	Western blot
WES	Whole exome sequencing
WT	Wildtype

List of tables

List of figures

1 Introduction

With approximately 900,000 new cases in 2020 [1], primary liver cancer (PLC) represents the sixth most frequent cancer worldwide causing the third most cancer-related deaths [2]. With increasing incidence of PLC, the WHO predicts more than a million deaths by 2030 [3-5]. PLC is a heterogeneous group of cancers. In 2018, hepatocellular carcinoma (HCC) was the most common subtype globally contributing 80% of all cases followed by intrahepatic cholangiocarcinoma (iCCA) (15%) and other entities such as combined HCC-CCA (cHCC-CCA) [6, 7].

1.1 Hepatocellular carcinoma

1.1.1 Epidemiology, etiology, and risk factors

East Asia and Africa harbor the highest incidence and mortality of HCC, but incidence rates are also rising in Europe and the USA [8]. In 90% of all HCC cases a chronic liver disease is causative [9-12].

Risk factors include chronic infection with the hepatitis B virus (HBV) causing approximately 60% of cases in Asia and Africa while only causing about 20% of cases in the industrialized world [13, 14]. In an effort to reduce the impact of HBV infection, successful preventive measures such as vaccination, interrupting mother-to-child transmission, and testing of blood products were taken [15]. Importantly, chronic HBV infection can exert direct oncogenic effects independent of the disease stage [13, 16]. Another cause of virus-induced chronic liver disease is chronic hepatitis C virus (HCV) infection, which may lead to complications such as liver cirrhosis and HCC development in 10%-20% of affected individuals [17]. However, the risk of HCV-related HCC has decreased by 50%-80% as a consequence of inventing highly effective direct antiviral agents resulting in sustained virologic response of patients [18]. Depending on disease duration and stage, HCC surveillance may be still required for the early diagnosis of HCC development after viral clearance [19].

Apart from chronic viral infections, other risk factors include alcoholic and nonalcoholic steatohepatitis (NASH) [19]. Chronic alcohol consumption accounts for about 15%-30% of HCC cases. In particular, the populations of the industrialized world reveal a tremendous increase in obesity. In Germany, NASH is estimated to increase from 3.33 million cases in 2016

to 4.74 million in 2030. In the USA, NAFLD incidence is predicted to rise by 122% between 2016-2030, resulting in 12,240 HCC cases annually [20]. Other sociodemographic factors further promoting the risk of HCC development include male sex (because of alcohol use and smoking [1]) and an age higher than 70 years in both genders [21]. Additional socioeconomic factors adding to the risk of HCC development include race or ethnicity of minorities [22]. More known risk factors include exposure to aflatoxin B1 [23], aristolochic acid [24] and smoking [25] while coffee and aspirin have been correlated with preventive functions [26].

1.1.2 Molecular hepatocarcinogenesis

Chronic liver damage results in parenchymal loss and progressive fibrotic remodeling, and finally leads to liver cirrhosis [27-29]. Hepatocarcinogenesis is a stepwise process that often results from cirrhosis and is associated with the accumulation and clonal selection of genetic and epigenetic aberrations. Consequently, foci of dysplastic hepatocytes may enlarge and become dysplastic nodules (DN) (> 1 mm). These lesions further accumulate molecular alterations and progress into early and advanced HCC (Fig. 1) [30-32].
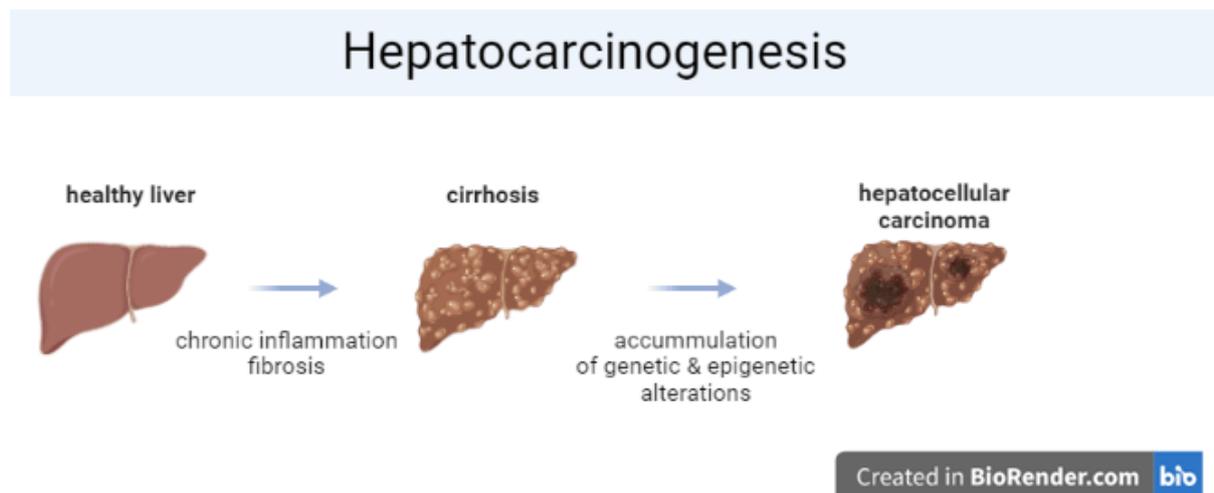
Fig. 1: Schematic of stepwise HCC development.
Hepatocarcinogenesis starts as a chronic liver disease at a certain stage leading to liver cirrhosis. Cirrhosis represents a premalignant condition, in which dysplastic hepatocellular lesions may develop that have the potential to progress into early and progressed HCC by selecting advantageous genomic abnormalities.

Mutations in the *TERT* promoter can be already detected in DN and represent the most frequent alteration in HCC (60% of cases) [33]. 30%-50% of cases comprise mutations in the Wnt-β-catenin signaling pathway, more specifically mutations in the *CTNNB1* gene (30% of cases) [24] and in the *AXIN1* gene (10% of cases) [34]. Mutations in *TP53* can be found in approximately 30% of cases, in particular in cases associated with aflatoxin B1 exposure and/or

HBV infection). Further mutations occur in genes related to chromatin remodeling, cell cycle control, epigenetic regulatory mechanisms, oxidative stress, and protumorigenic signaling cascades like AKT-mTOR and MAPK pathways [3, 35]. Apart from mutations, expression of *CCND1, FGF19, VEGFA, MYC* or *MET* due to recurrent focal chromosome amplifications activate oncogenic signaling pathways [36]. Overall, only approximately 25% of HCC tumors have actionable mutations. However, these represent low-frequency mutations (each detected in < 10% of cases), while the most frequent mutations are undruggable by current treatment options [24, 37-39].

1.1.3 Diagnosis and treatment

In cirrhotic patients HCC can be diagnosed by non-invasive imaging techniques, namely contrast-enhanced Computer Tomography or Magnetic Resonance Imaging, based on the characteristic arterialization that occurs during malignant transformation of preneoplastic lesions [40].

HCC treatment relies on liver function, patients´ performance status and tumor burden which is defined by the size, the number and localization of tumor nodules [41]. Early-stage tumors are eligible for resection, transplantation, and local ablation. [9, 10]. Ablation causes direct tumor cell necrosis by chemical or physical means [42]. Globally adopted as standard care, intermediate-stage tumors are treated by trans-arterial chemoembolization (TACE) meaning intraarterial infusion of cytotoxic agent and subsequent embolization of the tumor-feeding vessels. Patients with advanced disease receive systemic therapies including immune checkpoint inhibitors (ICI), tyrosine kinase inhibitors (TKI) and monoclonal antibodies [9, 37, 41, 43-47]. Current first-line treatment options include combination therapy with atezolizumab and bevacizumab targeting PDL1 and VEGF, respectively [48, 49] or dual immune checkpoint inhibition using durvalumab-tremelimumab [50]. Sorafenib and lenvatinib represent the most effective monotherapies [51, 52]. In general, prediction of treatment response based on certain genetic mutations is so far not possible and α-fetoprotein (Afp) is the only validated blood-based biomarker [4, 53-55].

1.2 Intrahepatic cholangiocarcinoma

1.2.1 Epidemiology, etiology, and risk factors

Intrahepatic cholangiocarcinoma is associated with poor prognosis due to clinical silence at early stages, aggressive disease progression, and limited treatment options [56]. The five-year overall survival rate is as low as 10% with a median survival of 24 months only [57, 58]. Although being a rare malignancy, incidence and mortality are steadily rising for 20 years [59, 60]. In a subset of cases, iCCA evolves from a cirrhotic liver and in general, any chronic liver disease infers an increased risk for the development of iCCA. Thus, the underlying etiologies are shared with HCC. Chronic biliary diseases like primary sclerosing cholangitis, Caroli's disease, or liver fluke infestation in Southeast Asia are among the additional risk factors for iCCA development [61-64]. Additionally, chemical agents (e. g. polychlorinated biphenyls [65]) and radionucleotides (e. g. Thorotrast [66]) have been associated with iCCA development. On the molecular level, mutations of genes involved in inflammation, DNA repair, carcinogen metabolism, and biliary transport can favor iCCA development [67-69].

1.2.2 Molecular intrahepatic cholangiocarcinogenesis

Cholangiocarcinogenesis is also a multistep process that is based on a pro-inflammatory environment and genetic and epigenetic factors such as chromosome aberrations and alterations in tumor suppressor genes and oncogenes leading to changes in proliferation, apoptosis, angiogenesis, invasion and metastasis [70-74].

Pro-inflammatory cytokines that are released into the biliary environment during inflammation are responsible for the malignant transformation of biliary epithelial cells (BEC) [70]. Bile acids have been found to induce cytokine expression and thus activate cyclooxygenase-2 and receptor tyrosine kinases such as EGFR [75, 76]. Growth factor receptors of the ERBB family activate the mitogen-activated protein kinase (MAPK) signaling pathway [77]. IL-6 has been found to promote survival by inducing mitogenic signals and is released by tumor and stromal inflammatory cells [78]. Oxidative stress activates the hedgehog signaling pathway, which promotes tumor development by increasing proliferation, migration, and invasive properties of cells [79, 80]. Further induced pathways include Notch signaling, which is responsible for lineage commitment of cholangiocytes and was associated with increased proliferation and

survival of iCCA cells. Furthermore, NOTCH1 was shown to enhance migration through RAC1 activation and epithelial-mesenchymal transition (EMT) induction [81-84].

Genes involved in chromatin regulation that are frequently mutated include *ARID1A, PBRM1* and *BAP1* [72]. Other genes frequently mutated in intrahepatic iCCA include *IDH1/2, BRAF* and *FGFR2*. *BRAF* mutations can be found in 1%-3% of iCCA and are associated with cell transformation through the MEK/ERK axis [85, 86]. *FGFR2* gene fusions appear in 10%-15% of iCCA and lead to constitutive tyrosine kinase activity [87, 88]. *TP53* mutations have been observed more frequently in liver fluke-related iCCA implying an etiological association [89]. *KRas* mutations have been found to cooperatively contribute to iCCA development together with *Tp53* or *Pten* mutations in genetic mouse experiments [90]. The PI3K/AKT pathway has also been implicated in iCCA. AKT is a serine/threonine kinase and exerts oncogenic activity by enhancement of cell survival [91].

1.2.3 Treatment

iCCA patients are frequently diagnosed with advanced disease, in which curative resection is no longer feasible. For more than a decade the combination of cisplatin plus gemcitabine (CISGEM) was the first-line systemic therapy [92]. A recent study showed that iCCA patients benefit from adding the PD-L1 inhibitor durvalumab to CISGEM treatment [93]. 5-fluorouracil/oxaliplatin was established as second-line treatment [94]. However, biomarkers for chemotherapy responders are lacking and cancers exhibit mechanisms of chemoresistance which help them to escape cytostatic drugs [95, 96]. Poor prognosis of iCCA patients calls for custom-made therapy based on the identification of common mutations in these cancers. The IDH1 inhibitor ivosidenib demonstrated a clinical benefit in previously treated, advanced IDH1-mutant CCA [97]. In addition, dual BRAF and MEK inhibition showed promising activity in patients with BRAF^{V600}-mutated biliary tract cancer in a phase 2 study. Another recurrent molecular feature of iCCA is the presence of gene fusions [98]. In particular, the high prevalence of FGFR2 gene fusions has become an attractive target and is still under active clinical evaluation [99-101]. Other gene rearrangements that are amendable for efficient drug targeting include fusions involving the NRG1 and NTRK genes [102, 103]. As more than 50% of iCCA contain potentially druggable alterations, early molecular profiling using large DNA and RNA panels may improve patients´ survival in clinical practice [104, 105].

1.3 Extracellular matrix and the tumor microenvironment

The extracellular matrix (ECM) is a complex and dynamic constituent of the liver that plays an important role in cancer development and progression. Apart from structural proteins like collagen it comprises matricellular proteins such as thrombospondins (THBS) [106]. ECM proteins are involved in many signaling pathways. The activation of integrin αv for example leads to the activation of transforming growth factor-β (TGFβ) [107]. TGFβ is a cytokine that plays a key role in generating a premalignant environment characterized by necrosis, inflammation, and fibrosis [108-111]. Fibrosis is not only collagen deposition, but also involves changes in many other ECM proteins that have been associated with HCC development [109]. TGFβ plays a complex role as it can have both cytostatic and tumor promoting effects through the TGFβ/SMAD axis and the activation of hepatic stellate cells [112].

The composition and organization of the ECM affect the stiffness of the microenvironment [113]. Matrix stiffness is a characteristic of inflammation and fibrosis and thus contributes to tumor development through integrin β1-mediated Fak, Erk, PKB/Akt, and Stat3 signaling pathways or the receptor-independent Hippo pathway which regulates cell differentiation and survival [114, 115] and is suggested to be involved in the transformation and dedifferentiation of hepatocytes [116]. Upon tumor development, the premalignant tumor environment becomes the tumor microenvironment (TME) which coevolves and interacts with tumor cells and impacts tumor growth, migration, invasion, and angiogenesis. The TME mainly includes the ECM, cancer-associated fibroblasts (CAFs), angiogenic cells and immune cells [117]. CAFs are responsible for producing the fibrous stroma that is associated with a more aggressive HCC phenotype [118, 119]. iCCA is characterized by a hypovascular, α-smooth muscle actin positive, desmoplastic stroma containing many CAFs. Activated CAFs as well as Kupffer cells, tumor-associated macrophages and iCCA cells produce IL-6. In iCCA, IL-6 predominantly targets epithelial cells through both IL-6/STAT3 and IL-6/p38 pathways, ultimately inhibiting cell death emphasizing the important role of the TME in cancer biology (Fig. 2) [78, 120-123].
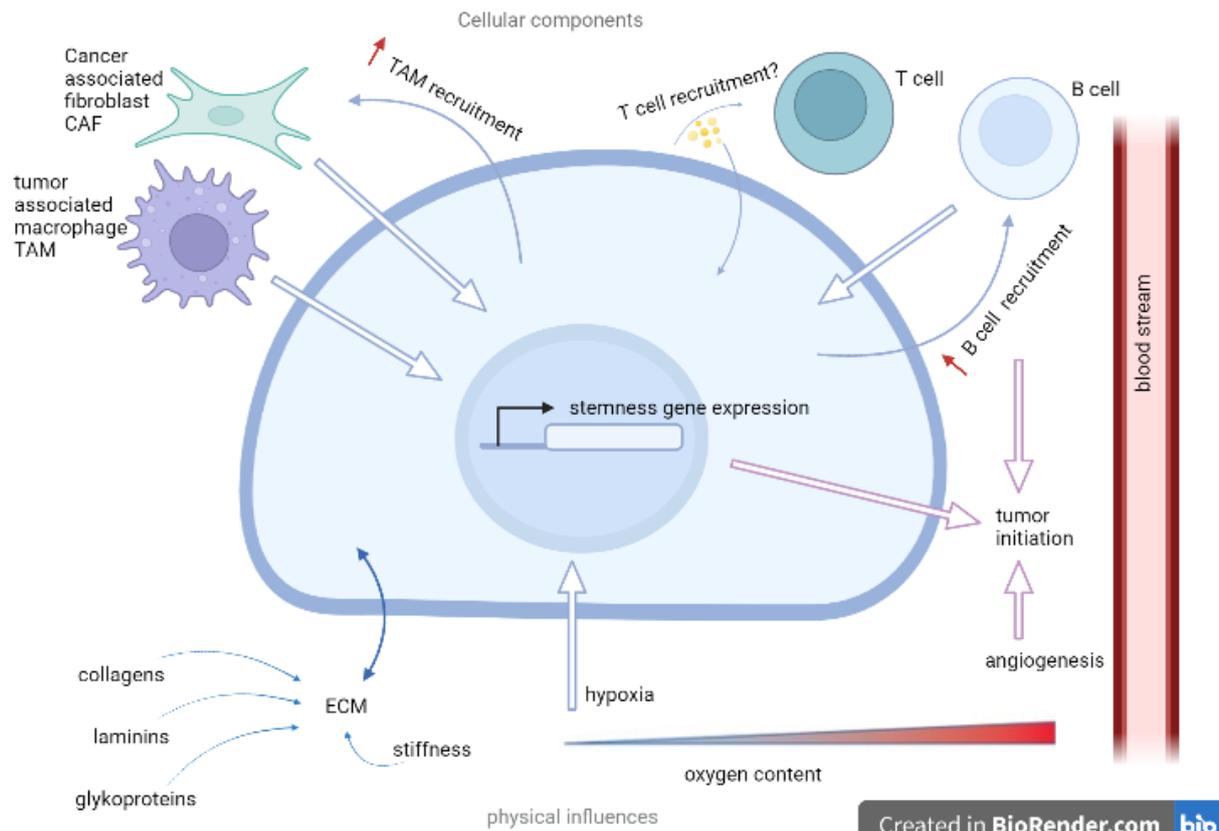
Fig. 2: The interplay of the tumor microenvironment with tumor epithelial cells.
The TME of iCCA plays an important role in disease initiation and progression. Many cell types such as CAFs and immune cells interact with tumor epithelial cells resulting in the activation of multiple protumorigenic signaling pathways (modified according to [124]).

1.4 Combined hepatocellular and cholangiocarcinoma

Incidence of cHCC-CCA ranges from 0.4%-14.2% with very limited treatment options except for surgical resection [125-127]. Compared to HCC, cHCC-CCA is more aggressive and has poorer outcomes similar to iCCA [128]. In cluster analysis, the two cHCC-CCA components were neither found to cluster with classic HCC nor classic iCCA, respectively, but did cluster with each other [128].

Despite their striking differences in histomorphology, the individual components of cHCC-CCA have a clonal origin as shown by the presence of shared mutations [128]. However, the cellular origin of cHCC-CCA remains a matter of debate. Studies show that cHCC-CCA may originate from both bipotent liver progenitor cells (LPC) and/or mature hepatocytes. Thus, the given tumor phenotype may be explained by cellular plasticity [6, 129-139]. Plasticity is defined as the ability to reversibly obtain phenotypes of other cell types of the same tissue and has been suggested as a rescue mechanism in liver disease occurring as an alternative to stem cell-mediated regeneration [140, 141]. Supporting this possibility, transdifferentiation of

hepatocytes into BEC was observed during chronic liver injury [142, 143]. Studies have also shown that biliary cells serve as facultative stem cells for hepatocytes restoring hepatic mass after hepatic tissue loss [144-146]. Fukuda *et al.* (2004) and other groups [134, 141, 147] suggested that hepatocytes transdifferentiate into Keratin 19 positive (Krt19+) biliary cells. Later, Fan *et al.* (2012) demonstrated in a lineage-tracing mouse experiment that iCCA can originate from hepatocytes with activation of Notch and Akt signaling [134] In accordance, a switch from iCCA to a hepatocellular adenoma-like lesion was observed upon *Notch2* deletion in mice, which was associated with downregulation of the biliary markers SRY-Box Transcription Factor 9 (Sox9) and epithelial cell adhesion molecule (EpCAM) [148].

Interestingly, Notch signaling was also activated in Akt/YAP-induced iCCA and was essential for tumor formation in this model. This may be explained by the finding that the biliary differentiation factor Sox9 acts as a YAP/TAZ regulated transcription factor during YAP induced hepatocarcinogenesis [143, 148, 149]. Notably, *Sox9* is also expressed by LPC [150]. In samples with high YAP activity, hepatocyte-derived iCCA and HCC could be distinguished in mouse and human livers based on intrinsic *Sox9* expression. Sox9 high mouse tumors were poorly differentiated carcinomas with high amount of fibrous stroma and expression of BEC lineage genes. In contrast, Sox9 low tumors showed upregulated expression of fetal hepatocyte genes and did not express *Krt19* [143, 149].

KRT19 is a marker for liver stem or progenitor cells (LSCs) and cholangiocytes but is not expressed by mature hepatocytes and most HCC [151, 152]. In contrast, HNF4α induces lineage commitment in LPC [153]. While forced expression of HNF4α results in less proliferating, more differentiated HCC with slow progression, Krt19+ HCC (10%-28%) are associated with a more aggressive phenotype and recurrence after resection or radiofrequency ablation [151, 152, 154-162].

In summary, PLC represents a heterogenous group of malignancies, which show partially overlapping risk factors but follow different routes of molecular carcinogenesis. The underlying mechanisms of cHCC-CCA initiation and progression remain to be elucidated.

1.5 Thrombospondins

THBS are a family of adhesive extracellular glycoproteins [163]. There are five THBS family members which are divided into trimer-forming subgroup A (THBS1 & 2) and pentamer-forming subgroup B (THBS3, 4 & 5) (Fig. 3) [164]. All family members share

THBS-type 2 and 3 repeats, which represent EGF-type repeats and a set of seven contiguous, E/F hand-type calcium binding repeats, respectively [165]. Furthermore, every THBS is a calcium-binding protein, and the presence of calcium plays a crucial role in preserving their tertiary structure [166].

In addition to their role in providing structural support, THBS also play a significant role in mediating various functions, including cell-cell and cell-ECM interactions, for example interactions with integrins which is supported by string analysis for THBS3 and THBS4. Integrins are transmembrane receptors through which cells sense changes in the ECM in response to external stimuli. In a mouse model, deletion of the integrin-linked kinase in hepatocytes resulted in cell death showing that the ECM may impact epithelial cell survival [167]. The best characterized function of THBS comprises their role in activating the Wnt/β-catenin signaling cascade. As agonists of the stem and progenitor cell receptors LGR4/5/6, they control stem cell regulation in multiple organs. [168-171]. THBS1 has been shown to regulate the stem cell niche in mammary cells in healthy mice in a Wnt4-mediated manner [172, 173].
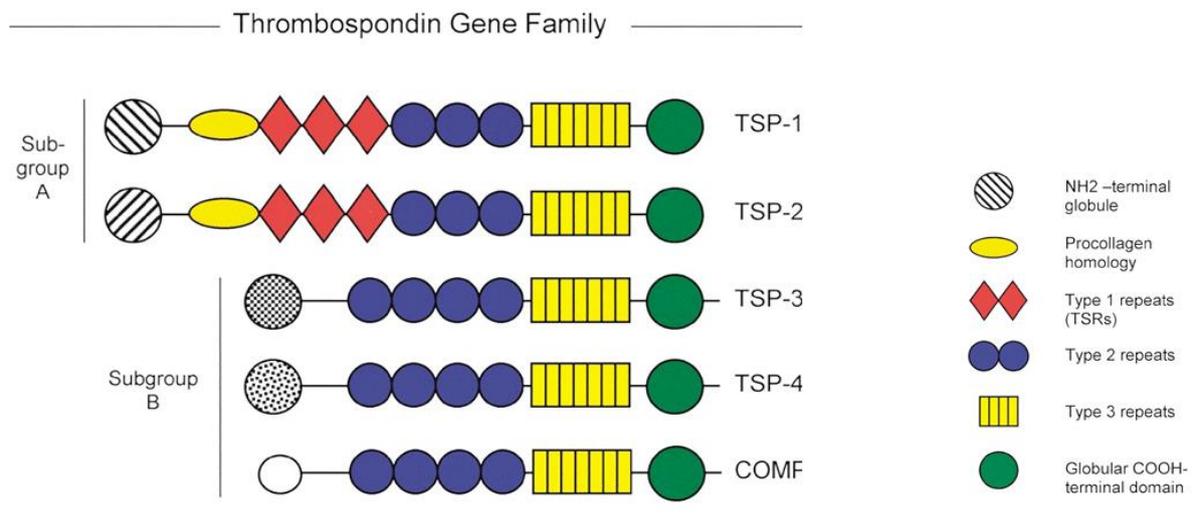
Fig. 3: Schematic depiction of thrombospondin gene family members.
The THBS gene family is subdivided into two structurally distinct groups A and B. Group A comprises THBS1 and THBS2. These members are longer and have additional type 1 repeats and a procollagen homology sequence compared to group B members, which represent shorter proteins. All members share a globular C-terminal domain and type 2 and type 3 repeats [174].

THBS dysregulation has increasingly been linked to carcinogenesis [175, 176]. In colorectal cancer, THBS2 promoted cell migration and invasion by activation of the Wnt/β-catenin axis and was negatively correlated to overall survival of patients [177]. In gastric cancer, THBS1-mediated activation of Wnt signaling was found *in vitro* and *in vivo*. Thbs1 depletion reduced

clonogenicity, migratory capacity, and stemness [178]. In iCCA, *THBS1* expression has been associated with hypovascularity and development of intrahepatic metastasis [179]. In addition, THBS1 has been shown to promote invasion by upregulation of metalloproteinases [179-182]. In HCC, THBS2 has been shown to activate the Wnt/β-catenin signaling cascade by binding to transmembrane receptors on stem and progenitor cells leading to their proliferation [183-187]. Furthermore, *THBS4* is upregulated in many cancers, has been implicated in regulating inflammation, and was associated with cancer growth and EMT [188-192].

THBS family members share a high homology of their C-terminal part both at the DNA and at the amino acid level, suggesting that the N-terminal part defines their different functions. In contrast to pro-angiogenic functions of THBS4 in HCC, THBS1 and THBS2 are characterized by anti-angiogenic effects [174, 193-195]. Another functional difference between family members is the pro-fibrotic role of THBS1 that was found in several organs [196-198]. THBS4 on the other hand was found to decrease ECM production *in vitro* and *in vivo* [199, 200]. The C-terminal domain of *THBS1* carries two binding sequences for CD47, which is a transmembrane integrin-associated protein (IAP) with elevated expression in cancers [201, 202]. Moreover, increased integrin expression was found in HCC models *in vivo* and in HCC patient samples and was associated with advanced disease, tumor encapsulation and prognosis [203-205].

To date, there are only few studies characterizing the role of THBS3 in cancer. *THBS3* expression has been negatively correlated to patients´ outcome in osteosarcoma and gastric cancer [206, 207]. Additionally, THBS3 has been proposed as an oncogene in breast cancer and a recent study by Ter Steege *et al.* (2022) established a relation between increased *THBS3* expression, aggressive tumor biology and poor differentiation in the latter type of cancer [208]. Interestingly, pan-cancer analysis observed reduced *THBS3* mRNA expression in 13 cancer types, while 11 cancer entities revealed upregulated *THBS3* expression, with the latter feature predicting worse outcome. Also, THBS3 is predicted to be involved in the activation of EMT and the infiltration of immune cells in human cancers. Thus, further studies are needed to explore and clarify the specific roles of THBS3 in these processes [209].

1.5 Mosaic mouse models

Mus musculus is considered a valuable tool for cancer research as it has several advantageous characteristics such as small size, lifespan of 3 years, and breeding features. Additionally, it

exhibits physiological and molecular similarities to the human system and is accessible for genetic modifications [210]. Importantly, mouse models allow the study of intermediate timepoints of tumor development, while human cancer is only amenable for cross sectional analysis [211, 212].

Initially, xenograft and orthotopic models were established in which cancer cells grown *in vitro* or intact tumor fragments from patients were implanted subcutaneously or transplanted into nude mice [213, 214]. However, many compounds that were successfully tested in both xenograft and orthotopic models were later shown to be ineffective in clinical trials. This emphasized that these models were not useful predictors of treatment response in human patients [212]. The growing evidence that the local tumor microenvironment is also important for the biological properties led to the development of more sophisticated animal models [210]. Genetically engineered mouse models mimic the different stages of tumor progression adequately regarding molecular features and the microenvironment and thus enable the study of pathophysiological features [210, 215]. Notably, similar genetic lesions may induce different pathologies in mice and humans [216]. Genetically engineered mouse model subcategories include transgenic, inducible, and conditional systems. The classic transgenic model is based on genetic manipulation of fertilized murine eggs or embryonic stem cells by means of microinjection or lentiviral gene transduction [215, 217]. Based on ectopic promoter and enhancer elements, homologous oncogene or tumor suppressor gene expression is activated or inactivated, respectively [210, 215]. Disadvantages include the absence of clinical features or exaggerated phenotypes resulting in infertility or embryonic lethality in cases where the targeted genes were essential for normal development [218, 219]. Furthermore, this system did not mirror spontaneous tumor initiation as it is based on homologous gene expression in all cells including those of the microenvironment [219, 220].

Establishment of mosaic mice gave rise to a model with cell-autonomous effects allowing to study tumor heterogeneity. Regarding the study of liver disease, they were either based on the re-transplantation of *ex vivo* genetically manipulated embryonic LPC into the livers of recipient mice [221, 222] or on hydrodynamic plasmid DNA delivery into the tail vein of mice. Hydrodynamic tail vein injection (HDTVI) is the easier and more advantageous method allowing for efficient delivery of vectors into hepatocytes. This cell-specificity results from the close spatial relationship between liver sinusoidal endothelial cells and hepatocytes. The hydrodynamic pressure arising from rapid injection of a large volume (8%-10% of mouse body weight) induces cardiac congestion and subsequently results in retrograde perfusion of the liver

[223-225]. This disrupts the endothelial barrier enabling DNA uptake by hepatocytes [223, 226, 227].

Somatic integration of the delivered DNA into liver epithelial cells is facilitated by the Sleeping Beauty (SB) transposon system. This powerful tool is based on an active transposase and a transposon consisting of a gene-specific DNA sequence flanked by inverted repeats (IR). The transposon is cut and pasted by the transposase [228, 229]. Consequently, the delivered DNA is stably integrated in liver epithelial cells, transcribed, and subsequently translated which may eventually promote tumorigenesis. For the study of tumor suppressor genes, the transposon system can be used to deliver stable ribonucleic acid interference (RNAi) constructs [222, 230]. This approach can also be used to deliver a short hairpin RNA (shRNA) library, which can be used for *in vivo* RNAi screening, thus highlighting the explorative power of such a model system [230].

1.6 Objectives

PLC is a deadly disease as it is mostly detected at an advanced stage, in which only palliative treatment options are available. The molecular landscape of both HCC and iCCA has been characterized. cHCC-CCA, however, is a heterogeneous tumor entity with characteristics of both HCC and iCCA. The protumorigenic mechanisms of the genetic and epigenetic alterations identified in cHCC-CCA remain mostly elusive.

In PLC, the cells of origin are still a matter of debate with two main theories being proposed (Fig. 4). One theory favors cellular plasticity as the key driving force, while the alternative is built on the assumption that LPC are the source of the tumor initiating cell.
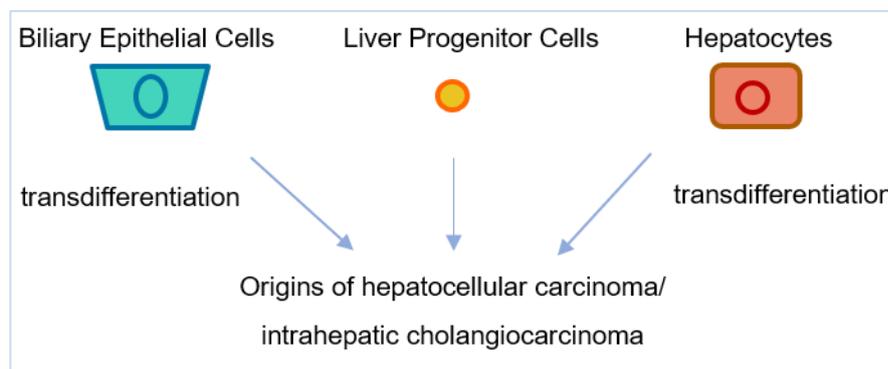
Fig. 4: Potential cells of origin of human PLC.

The aim of this study was

1.   to determine the differential genetic alterations between areas of hepatocellular and cholangiocellular differentiation in cHCC-CCA using exome and RNA sequencing.

2.   to validate the phenotype-driving potential of the differentially altered candidate genes using suitable mosaic mouse models.

3.   to identify and functionally characterize the mechanisms determining the tumor phenotype.

2 Materials

2.1 Antibiotics, chemicals, and mediums

Tab. 1: Antibiotics, chemicals and mediums

Chemicals, medium, antibiotics	Supplier
Acetic acid	Carl Roth, Karlsruhe, Germany
Agarose	Sigma-Aldrich, Taufkirchen, Germany
Albumin Fraction V, biotin-free, NZ-Origin	Carl Roth, Karlsruhe, Germany
Ammonium peroxydisulphate	Carl Roth, Karlsruhe, Germany
Ampicillin (100 mg/mL)	Sigma-Aldrich, Taufkirchen, Germany
Bacto™ Agar	BD Biosciences, Heidelberg, Germany
Blasticidin S HCl (10 mg/mL)	Gibco/Life Technologies, Darmstadt, Germany
BlueStar PLUS Prestained Protein Marker	NIPPON Genetics, Düren, Germany
Carbenicillin Dinatrium	Carl Roth, Karlsruhe, Germany
Coomassie® Brilliant Blue G 250	SERVA Electrophoresis, Heidelberg, Germany
Crystal Violet	Sigma-Aldrich, Taufkirchen, Germany
Dimethylsulfoxid (DMSO)	Carl Roth, Karlsruhe, Germany
dNTP Mix (10 mM)	Thermo Fisher Scientific, Waltham, USA
DTT	SERVA Electrophoresis, Heidelberg, Germany
Dulbecco's Modified Eagle Medium (DMEM)	Gibco/Life Technologies, Darmstadt, Germany
Ethanol	Merck Chemicals GmbH, Darmstadt, Germany
EZ-Run™ Prestained Rec Protein Ladder	Thermo Fisher Scientific, Waltham, USA
Fetal Bovine Serum (FBS)	Gibco/Life Technologies, Darmstadt, Germany
Gel Loading Dye, Purple (6×)	New England Biolabs, Frankfurt, Germany
Glycerol	Carl Roth, Karlsruhe, Germany
LB Agar	Carl Roth, Karlsruhe, Germany
LB Broth, granulated	Carl Roth, Karlsruhe, Germany
Methanol	VWR International S.A.S., Radnor, USA
Opti-MEM® I Reduced-Serum Medium	Gibco/Life Technologies, Darmstadt, Germany

Chemicals, medium, antibiotics	Supplier
Penicillin-Streptomycin	Gibco/Life Technologies, Darmstadt, Germany
Phenylmethylsulphonyl fluoride (PMSF)	Sigma-Aldrich, Taufkirchen, Germany
Polyethylenimine (PEI)	Polysciences, Warrington, USA
Powdered milk	Carl Roth, Karlsruhe, Germany
Puromycin Dihydrochloride	Gibco/Life Technologies, Darmstadt, Germany
Quick-Load® 1 kb Extend DNA Ladder	New England Biolabs, Frankfurt, Germany
RPMI-1640	Thermo Fisher Scientific, Waltham, USA
Sodium-Dodecyl-Sulfate (SDS)	SERVA Electrophoresis, Heidelberg, Germany
Sodium chloride (NaCl)	neoFroxx GmbH, Einhausen, Germany
Tetramethylethylenediamine (TEMED)	Carl Roth, Karlsruhe, Germany
Trisaminomethane (Tris)	Carl Roth, Karlsruhe, Germany
Tween® 20	Carl Roth, Karlsruhe, Germany

2.2 Antibodies

Tab. 2: Antibodies

Antigen	Species	Dilution	Ordering Number/ Company	Application
β-ACTIN	mouse	1/10,000	#A5441/ Sigma-Aldrich	WB
B220	mouse	1/50	#550539/ Biosciences	IHC
CD3	rabbit	1/150	#RM9107/Thermo Fisher Scientific	IHC
CD68	rat	1/50	#16676/ Thermo Fisher Scientific	IHC
FRA1	mouse	1/50–1/300	#376148/ Santa Cruz	IHC, WB
FRA1	rabbit	1/1000	#5281T/ Cell signaling	WB
GAPDH	mouse	1/5000	#MCA4739/ Bio-Rad	WB
H2A.X	rabbit	1/200	#14-047/ ProScie	IHC
HNF4α	rabbit	1/1000	#3113S/ Cell signaling	IHC
IRDye 680RD α-Mouse IgG	donkey	1/10,000	#925-68072/ LI-COR	WB
IRDye 800CW α-Mouse	goat	1/10,000	#926-32210/ LI-COR	WB
IRDye 800CW α-Rabbit	goat	1/10,000	#926-32211/ LI-COR	WB

Antigen	Species	Dilution	Ordering Number/ Company	Application
KERATIN19	rabbit	1/500	#ab133496/ Abcam	IHC
pH2A.X	rabbit	1/250	#9718S/ Cell signaling	IHC
SOX9	rabbit	1/1000	#AB5535/ Merck	IHC
THBS3	mouse	1/50-1/150	#25348/ Santa Cruz	IHC/ WB

2.3 Bacteria and cell lines

Tab. 3: Bacteria and cell lines

Bacteria and cell lines	Supplier
AMp19$^{-/-}$ (derived from murine HCC tissue)	Kindly provided by Prof. Dr. med. Lars Zender (University Hospital Tübingen, Germany)
Kp19$^{-/-}$ (derived from murine iCCA tissue)	Kindly provided by Prof. Dr. med. Lars Zender (Tübingen)
HEK293T/17 (human embryonic kidney)	ATCC, USA
Hepa1-6 (murine hepatoma)	ATCC, USA
Hep56 (murine hepatoma)	ATCC, USA
HUH28 (human iCCA)	Kindly provided by Prof. Dr. rer. nat. Stefanie Rössler (Heidelberg University Hospital, Heidelberg, Germany)
NEB® Stable Competent E. coli	New England Biolabs, Frankfurt, Germany
XL10-Gold Ultracompetent Cells	Agilent Technologies, Santa Clara, USA

2.4 Buffers and solutions

Tab. 4: Buffers and solutions

Buffers and solutions	Content	Application
2x Sample buffer	125 mM Tris HCl pH 6.8, 4% SDS, 20% glycerol, 0.02% bromophenol blue	SDS-PAGE
4x Loading buffer	250 mM Tris HCl pH 6.8, 8% SDS, 40% glycerol, 0.04% bromophenol blue, 100 mM DTT	SDS-PAGE
Blocking solution	5% milk powder or 5% BSA in TBST	WB

Buffers and solutions	Content	Application
Cell Lysis Buffer (10×)	Cell Signaling	WB
Crystal violet staining solution	1% Crystal violet, 25% Methanol	CFA
LB agar	1.5% agar in LB medium	Bacteria
LB medium	2% LB broth in H2O	Bacteria
PBS	140 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4, 1.8 mM KH2PO4	diverse
Running buffer (pH8.5-8.7)	0.25 M Tris base, 2 M glycine, 1% SDS	SDS-PAGE
Super Optimal broth with Catabolite repression	3.603 g/l dextrose, 0.186 g/l KCl, 4.8 g/l MgSO4, 20 g/L tryptone, 5 g/L yeast extract	Bacteria
T4 DNA Ligase Buffer	New England Biolabs, Frankfurt, Germany	Cloning
TAE buffer	40 mM Tris base, 10 mM EDTA pH 8.0, 6% acetic acid	AGE
TAE gel	1%-2% agarose in TAE buffer	AGE
TBST (pH 7.6)	25 mM Tris base, 140 mM NaCl, 0.02% Tween20	WB
Transfer buffer	25 mM Tris base, 200 mM glycine, 20% methanol	WB

2.5 Consumables

Tab. 5: Consumables

Consumables	Supplier
25 cm^2 Cell Culture flask with vented cap	Orange Scientific, Braine-l'Alleud, Belgium
Amersham Protran Nitrocellulose blotting membrane	GE Healthcare, Buckinghamshire, UK
Cryogenic vials	Greiner Bio-One, Frickenhausen, Germany
DISTRITIPS	Gilson, Berlin, Germany
Falcon® Cell culture plates	Corning, New York, USA
Falcon® round bottom 14 ml test tubes	Corning, New York, USA
Falcon tubes	Greiner Bio-One, Frickenhausen, Germany
Greiner Bio-One™ Pipette Tips	Greiner Bio-One, Frickenhausen, Germany
MicroAmp™ Fast Optical 96-Well Reaction Plate	Thermo Fisher Scientific, Waltham, USA
MicroAmp™ Optical Adhesive Film	Thermo Fisher Scientific, Waltham, USA

Consumables	Supplier
Microcentrifuge tubes	Eppendorf, Hamburg, Germany Sarstedt, Nümbrecht, Germany
Microscope cover glasses	Marienfeld, Lauda-Königshofen, Germany
Microscope slides "Menzel Gläser"	Thermo Fisher Scientific, Waltham, USA
neoCulture Cell scrapers ABS	neoLab, Heidelberg, Germany
PARAFILM® M	Thermo Fisher Scientific, Waltham, USA
PCR Tubes	Kisker Biotech, Steinfurt, Germany
Petri dishes	Sarstedt, Nümbrecht, Germany,
Pipette Tips	Sarstedt, Nümbrecht, Germany
SafeSeal Tips Professional 10 μl, Sterile	Biozym, Hessisch Oldendorf, Germany
Sterile stripettes®	Corning, New York, USA
Sterile syringe filters, pore size 0.45 μm	VWR International, Bruchsal, Germany
Syringes	BD Biosciences, Heidelberg, Germany
TipOne® Filter Tips	STARLAB, Hamburg, Germany

2.6   Enzymes and kits

Tab. 6: Enzymes and kits

Enzymes and kits	Supplier
ATAC-Seq Kit	Active Motif, Waterloo, Belgium
BamHI-HF	New England Biolabs, Frankfurt, Germany
CellTiter-Blue Cell Viability Assay	Promega, Wisconsin, USA
EcoRI-HF	New England Biolabs, Frankfurt, Germany
EndoFree Plasmid Maxi Kit	Qiagen, Hilden, Germany
PrimeScript™ Reverse Transcriptase	Takara Bio, Saint-Germain-en-Laye, France
Maxwell® 16 FFPE Plus LEV DNA Purification Kit	Promega, Wisconsin, USA
Mix & Go E. coli Transformation Kit	Zymo Research, Freiburg, Germany
Monarch® Plasmid Miniprep Kit	New England Biolabs, Frankfurt, Germany
MycoAlert™ Mycoplasma Detection Kit	Lonza, cologne, Germany

Enzymes and kits	Supplier
NEBuilder® HiFi DNA Assembly Master Mix	New England Biolabs, Frankfurt, Germany
NotI-HF	New England Biolabs, Frankfurt, Germany
NucleoSpin RNA Kit	MACHEREY-NAGEL, Düren, Germany
PhosStop	Roche, Mannheim, Germany
Phusion High-Fidelity DNA Polymerase	Thermo Fisher Scientific, Waltham, USA
PowerUp™ SYBR™ Green Master Mix	Thermo Fisher Scientific, Waltham, USA
PureYield™ Plasmid Midiprep System	Promega, Wisconsin, USA
Qubit™ dsDNA HS kit	Thermo Fisher Scientific, Waltham, USA
QuikChange Lightning Site-Directed Mutagenesis Kit	Agilent Technologies, Santa Clara, USA
rDNase	MACHEREY-NAGEL, Düren, Germany
RedTaq Ready Mix	Sigma-Aldrich, Taufkirchen, Germany
ReliaPrep™ DNA Clean-Up & Concentration System	Promega, Wisconsin, USA
Shrimp Alkaline Phosphatase (rSAP)	New England Biolabs, Frankfurt, Germany
T4 DNA Ligase	New England Biolabs, Frankfurt, Germany
Wizard® SV Gel and PCR Clean-Up System	Promega, Wisconsin, USA

2.7 Equipment

Tab. 7: Equipment

Equipment	Supplier
Axio Vert 40c	Zeiss, Oberkochen, Germany
Counting chamber BLAUBRAND® Neubauer	Brand, Frankfurt, Germany
D-DiGit® Gel Scanner	LI-COR, Bad Homburg, Germany
DISCOVERY COMFORT pipettes	Corning, New York, USA
Dri-Block® Heater	Thermo Fisher Scientific, Waltham, USA
EV265 & EV231 Electrophoresis Power Supplies	Consort, Turnhout, Belgium
FluorChem™ M system	ProteinSimple, San Jose, USA
FLUOstar Omega Microplate Reader	BMG Labtech, Ortenberg, Germany
Gel electrophoresis chamber	FEBIKON, Wermelskirchen, Germany

Equipment	Supplier
Heracell™ VIOS 250i CO_2 Incubator	Thermo Fisher Scientific, Waltham, USA
Heraeus Megafuge 1.0RS	Heraeus, Hanau, Germany
Heraeus Megafuge 16R Centrifuge	Thermo Fisher Scientific, Waltham, USA
HI-2210 Bench Top pH Meter	Hanna Instruments, Kehl, Germany
Incubator Hood TH 15	Edmund Bühler, Bodelshausen, Germany
JEM-1400Flash Electron Microscope	JEOL, Freising, Germany
KERN EW6000-1M	KERN & SOHN, Balingen, Germany
Leica EM TRIM 2.	Leica, Nussloch, Germany
Magnetic stirrer C-MAG MS 7	IKA®-Werke, Staufen, Germany
Maxwell 16 Research extraction system	Promega, Wisconsin, USA
Memmert 37°C Bacteria Incubator	Memmert, Schwabach, Germany
Fluorescence microscope BX53	OLYMPUS, Hamburg, Germany
MIKRO 200 R	Hettichlab, Tuttlingen, Germany
Mini Trans-Blot® Cell	Bio-Rad, Munich, Germany
Mini-PROTEAN® Tetra Cell	Bio-Rad, Munich, Germany
Mr. Frosty™ Freezing Container	Thermo Fisher Scientific, Waltham, USA
myFUGE™ Mini centrifuge	Benchmark Scientific, Sayreville, USA
NanoDrop™ ND-1000 UV/Vis Spectrophotometer	Thermo Fisher Scientific, Waltham, USA
Odyssey® DLx Imaging System	LI-COR, Bad Homburg, Germany
PIPETBOY acu 2	INTEGRA Biosciences, Hudson, USA
PowerPac™ Basic Power Supply	Bio-Rad, Munich, Germany
ProteinSimple FluorChem E Imaging	Alpha Innotech, Kasendorf, Germany
PTC 200 Peltier Thermal Cycler	Biozym, Hessisch Oldendorf, Germany
QuantStudio 3 Real-Time PCR System	Thermo Fisher Scientific, Waltham, USA
Revolver Rotator, digital (D-6050)	neoLab, Heidelberg, Germany
Roller mixer RN 5	CAT, Ballrechten-Dottingen, Germany
SW22 Shaking water bath	JULABO, Allentown, USA
Thermomixer Compact 5350 Mixer	Eppendorf, Hamburg, Germany
Transsonic T460/H ultrasonic bath	Elma, Singen, Germany

Equipment	Supplier
VACUSAFE vacuum pump	INTEGRA Biosciences, Hudson, USA
Vortex Mixer 7-2020	neoLab, Heidelberg, Germany

2.8   Oligonucleotides

2.8.1   Cloning primers (transposon vectors)

Tab. 8: Cloning primers (transposon vectors)

Primer	Sequence 5'-3'
CaKIG_shRNA_fwd	TAAGCAGGCGCGCCGCCACCTGCTGTTGACAGTGAGC
CaKIG_shRNA_rev	TGCTTACTCGAGTTCCGAGGCAGTAGGC
CaMIA_shRNA_fwd	TAAGCAACCGGTGCCACCTGCTGTTGACAGTGAGC
CaMIA_shRNA_rev	TGCTTAGCTAGCTTCCGAGGCAGTAGGC
pCaggs_mAdra1a_fwd	TAAGCAGGCGCGCCGCCACCATGGTGCTT
pCaggs_mAdra1a_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGT
pCaggs_mAnkrd1_fwd	TAAGCAGGCGCGCCGCCACCATGGTACTGAG
pCaggs_mAnkrd1_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGT
pCaggs_mAsb4_fwd	TAAGCAGGCGCGCCGCCACCATGGACGGC
pCaggs_mAsb4_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGT
pCaggs_mAsb15_fwd	AAGCAGGCGCGCCGCCACCATGGATATTAATGATGATTCTAACGA
pCaggs_mAsb15_rev	TGCTTACTCGAGGCTGTCCGTAGAGGTCAAAC
pCaggs_mDtx1_fwd	TAAGCAGGCGCGCCGCCACCATGTCACGG
pCaggs_mDtx1_rev	TGCTTACTCGAGGCTGTCCGTAGAGGTCAAAC
pCaggs_mFosl1_fwd	TAAGCAGGCGCGCCGCCACCATGTACCGAGACTACGGG
pCaggs_mFosl1_rev	TGCTTACCTCGAGCAAAGCCAGGAGTGTAGG
pCaggs_mFoxj1_fwd	TAAGCAGGCGCGCCGCCACCATGGCGG
pCaggs_mFoxj1_rev	TGCTTACCTCGAGCAAAGCCAGGAGTGTAGG
pCaggs_mGata1_fwd	TAAGCAGGCGCGCCGCCACCATGGATTTTCCT
pCaggs_mGata1_rev	TGCTTACCTCGAGCAAAGCCAGGAGTGTAGG

Primer	Sequence 5'-3'
pCaggs_mGli1_fwd	TAAGCAGGCGCGCCGCCACCATGTTCAATCCAATGACTCCACCACA AG
pCaggs_mGli1_rev	TGCTTACTCGAGGGCACTAGAGT
pCaggs_mGlis1_fwd	TAAGCAGGCGCGCCGCCACCGCCACCATGCATTGCGAGGTGGC
pCaggs_mGlis1_rev	TGCTTACCTCGAGCAAAGCCAGGAGTGTAGG
pCaggs_mKlf5_fwd	TAAGCAGGCGCGCCGCCACCATGCCCAC
pCaggs_mKlf5_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGT
pCaggs_mMacc1_fwd	TAAGCAGGCGCGCCGCCACCATGCTAATCAGT
pCaggs_mMacc1_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGT
pCaggs_mOrc1_fwd	TTTTGGCAAAGAATTCCCTCGATACCGGGGGCCACCATGCCATCCT AC
pCaggs_mOrc1_rev	AGCAGAGAGAAGTTTGTGGCGCCGCTGCCCGCTCTTCTTTGAGAGC AAAC
pCaggs_P2A_tRFP_fwd	TATGGCCACAACCATGAGCGAGCTGATCAAG
pCaggs_P2A_tRFP_rev	AAGTGTGATCAGTTAGAACCGGTGCTCATCTGTGCCCCAGTTTG
pCaggs_mThbs3_fwd	TTTTGGCAAAGAATTCCCTCGATACCGGGGGCCACCATGGAGAAG CCG
pCaggs_mThbs3_rev	AGCAGAGAGAAGTTTGTGGCGCCGCTGCCCGCACTCTTCCCTGGAG CAG
pCaggs_mTmprss4_fwd	TAAGCAGGCGCGCCGCCACCATGGAGTCAGACAGTG
pCaggs_mTmprss4_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGTTTGCTAG
pCaggs_Vtcn1_fwd	TAAGCAGGCGCGCCGCCACCATGGCTTCC
pCaggs_Vtcn1_rev	TGCTTAGCGGCCGCTTCATCTGTGCCCCAGT

2.8.2 Cloning primers (lentiviral vectors)

Tab. 9: Cloning primers (lentiviral vectors)

Primer	Sequence 5'-3'
MirE_fwd	TGAACTCGAGAAGGTATATTGCTGTTGACAGTGAGCG
MirE_rev	TCTCGAATTCTAGCCCCTTGAAGTCCGAGGCAGTAGGC
pLV_hTHBS3_fwd	TCCATTTCAGGTGTCGTGAGATGGAGAAGCCGGAACTTTG

Primer	Sequence 5'-3'
pLV_hTHBS3_rev	AGCAGAGAGAAGTTTGTGGCGCCGCTGCCCGGCACTCTTCCCTGGA GCAG
pLV_mThbs3_fwd	TTTTGGCAAAGAATTCCCTCGATACCGGGGGGCCACCATGGAGAAG CCG
pLV_mThbs3_rev	AGCAGAGAGAAGTTTGTGGCGCCGCTGCCCGCACTCTTCCCTGGAG CAG

2.8.3 Mutagenesis primers

Tab. 10: Mutagenesis primers

Primer	Sequence 5'-3'
hTHBS3_g305a_fwd	GGTGCGATACCAGCAGGAGGATGGCAAAG
hTHBS3_g305a_rev	CTTTGCCATCCTCCTGCTGGTATCGCACC
mAdra1a_t599g_fwd	TACATAACCAGGCTGATGGTCAGTGGCACGTAGAAA
mAdra1a_t599g_rev	TTTCTACGTGCCACTGACCATCAGCCTGGTTATGTA
mOrc1_c1334t_a1335g_fwd	GAGGAGTGGCACGATGAGGCATTCTAGGTTTAGGATTTTTCTG
mOrc1_c1334t_a1335g_rev	CAGAAAAATCCTAAACCTAGAATGCCTCATCGTGCCACTCCTC
mThbs3_g305a_fwd	GATGTTAGGCAGTATAATGTAGTCACTCTCTGAGCCAT
mThbs3_g305a_rev	ATGGCTCAGAGAGTGACTACATTATACTGCCTAACATC
mTmprss4_c1232t_fwd	TCCAGGAGTACTTAGGCCGCCGCATCC
mTmprss4_c1232t_rev	GGATGCGGCGGCCTAAGTACTCCTGGA
mVtcn1_g358c_fwd	CATCCGTGAGCTGCAGGTTTTTCAGTCTCAGGG
mVtcn1_g358c_rev	CCCTGAGACTGAAAAACCTGCAGCTCACGGATG

2.8.4 Semi-quantitative real-time PCR (qPCR) primers

Tab. 11: Semi-quantitative real-time PCR (qPCR) primers

Gene	Sequence 5'-3'	Accession number
hAFP_fwd	AGGGTGTTTAGAAAACCAGCTACC	NM_001134.3
hAFP_rev	TGCAGCAGTCTGAATGTCCG	

Gene	Sequence 5'-3'	Accession number
hALB_fwd	ATGCCCCGGAACTCCTTTTC	NM_000477.7
hALB_rev	CGAAGTTCATCGAGCTTTGGC	
hHNF4α_fwd	GGCAATGACACGTCCCCAT	NM_178849.3
hHNF4α_rev	CTCGAGGCACCGTAGTGTTT	
hKRT7_fwd	TGGGAGCCGTGAATATCTCTGT	NM_005556.4
hKRT7_rev	GAGAAGCTCAGGGCATTGCT	
hKRT19_fwd	ACAGCCACTACTACACGACC	NM_002276.5
hKRT19_rev	GTTCCGTCTCAAACTTGGTTCG	
hSOX9_fwd	CTCTGGAGACTTCTGAACGAGAG	NM_000346.4
hSOX9_rev	GTTCTTCACCGACTTCCTCCG	
hTHBS3_fwd	AATGAGCAATCCTACCCAGACAG	NM_007112.5
hTHBS3_rev	GTCCTTGGTGTCCTGATGCC	
mAdra1a_fwd	TGATGCCCATTGGGTCCTTC	NM_001271760.1
mAdra1a_rev	GGGTATATGATAGGGTTGATGCA	
mAfp_fwd	TGGTTACACGAGGAAAGCCC	NM_007423.4
mAfp_rev	GGAGGCAATGCTCACCATCT	
mAlb_fwd	TGTCAACCCCAACTCTCGTG	NM_009654.4
mAlb_rev	CAGACACACACGGTTCAGGA	
mAnkrd1_fwd	ACTGAGAGTAGAGGAGCTGGTAA	NM_014391.3
mAnkrd1_rev	CTGTTGGCCGGAAGTGTCT	
mAsb4_fwd	GCATCACTGCCCCTATCAGC	NM_023048.5
mAsb4_rev	AGCCTTCAGTTTTCCGAAGTCA	
mAsb15_fwd	GGCAACGTCCACTTGAGAGA	NM_080847.3
mAsb15_rev	CCGCCAAAGCAAACACATCA	
mDtx1_fwd	TGTGCCACCACATCGAGAAT	NM_008052.3
mDtx1_rev	TACCTGTGTCTTGCCGGAAC	
mFosl1_fwd	TGTACCGAGACTACGGGGAA	NM_010235.2
mFosl1_rev	AAGGTGGAACTTCTGCTGCT	
mFoxj1_fwd	GCTGGGGACAGAGAACCG	NM_008240.3

Gene	Sequence 5'-3'	Accession number
mFoxj1_rev	CGAATGTGAGGCCTGGCT	
mGata1_fwd	ATGGATTTTCCTGGTCTAGGGG	NM_001410603.1
mGata1_rev	CACCAGGGCAGAATCCACAA	
mGli1_fwd	CCATTGGTACCATGAGCCCTT	NM_010296.2
mGli1_rev	CACAGGGCTGGACTCCATAG	
mGlis1_fwd	CCCCTGTCTGTGAGAAGCTG	NM_147221.2
mGlis1_rev	CGGTAGGGCTTCTTGCTTGA	
mHnf4α_fwd	GAAAATGTGCAGGTGTTGACC	NM_008261.3
mHnf4α_rev	AGCTCGAGGCTCCGTAGTGTT	
mKlf5_fwd	GGATCTGGAGAAGCGACGTAT	NM_009769.4
mKlf5_rev	GGCTTCTCGCCCGTATGA	
mKrt7_fwd	GAACCGCTCTATCCAGAGGC	NM_033073.3
mKrt7_rev	CAGCTCCCCTTGTTCCTCAG	
mKrt19_fwd	TGAAGCCACCTACCTTG	NM_008471.3
mKrt19_rev	GATCTGCTCAGAGTGGACGG	
mMacc1_fwd	AGCCTATTGTTGGCCACGAG	NM_001163136.1
mMacc1_rev	GCCCTTGCCGTGCTGTAATA	
mOrc1_fwd	CGCTGGAACTCGATGGCTTA	NM_011015.2
mOrc1_rev	GGCTTTTTAGGAGGCGAGGT	
mSox9_fwd	TAAGTTCCCCGTGTGCATCC	NM_011448.4
mSox9_rev	GTGTGGCTTGTTCTTGCTGG	
mThbs3_fwd	TATGCTCGAGCCAGCAAACA	NM_013691.3
mThbs3_rev	TCACAGCCTCCGTTGTTACC	
mTmprss4_fwd	ACCCCTCAACAACCGTGATATT	NM_145403.3
mTmprss4_rev	GCCACAATCACGAGGGCTAT	
mVtcn1_fwd	GGGGCAGATCATCTTTTGGAGTATT	NM_178594.4
mVtcn1_rev	GCCTGAAATGCCAAAGCCAA	

2.8.5 Sequencing primers

Tab. 12: Sequencing primers

Primer	Sequence 5'-3'	Accession number
hTHBS3_seq1	CTAGGGGAGCAGACCAAG	
hTHBS3_seq2	GACACCCAACTCTGGGCAG	NM_007112.5
hTHBS3_seq3	CAGACCATGAACAGTGACC	
mAdra1a_seq1	CAAGAGAGAAAGCCGAG	NM_001271760.1
mAnkrd1_seq1	GGA AACGCAGATGTCCTG	NM_014391.3
mAsb4_seq1	GCGCTTCAAGGAGCAGGAG	NM_023048.5
mAsb15_seq1	GCTGAGTATGGCCACTGTGAC	NM_080847.3
mDtx1_seq1	GCCACCCGGAGGACCTCC	NM_008052.3
mDtx1_seq2	GGGAAGATGGAGTTTCACC	
mFosl1_seq1	TGTACCGAGACTACGGGGAA	NM_010235.2
mFoxj1_seq1	GCCTCCCAGGAACCTAGC	NM_008240.3
mGata1_seq1	CCTGTGCAATGCCTGTGGC	NM_001410603.1
mGli1_seq1	CGCTGGGATGGTTGCAGCC	
mGli1_seq2	GCCTGGAGAACCTTAGGC	NM_010296.2
mGli1_seq3	GCCTGGTCCACCAACCAAC	
mGli1_seq4	CCTGAGGTGGGCAGGTTAG	
mGlis1_seq1	GCCTCCTCCTCACCCTGTG	NM_147221.2
mGlis1_seq2	GCCCACTCAGCCAAAGAGC	
mKlf5_seq1	GGCAGGCCTTAACCCACACC	NM_009769.4
mMacc1_seq1	GGACATATAGCTGTGGGAG	
mMacc1_seq2	GCACTTGTTTGTTTTCCGGG	NM_001163136.1
mMacc1_seq3	GCCAAGAGTTAGTGGCAC	
mOrc1_seq1	CCTGCCCGGAACACAACA	NM_011015.2
mOrc1_seq2	GGCAGACCCCCTCCAA	
mThbs3_seq1	CTAGGGGAGCAGACCAAG	NM_013691.3
mThbs3_seq2	GGAACGTGTGTGGGCCT	

Primer	Sequence 5'-3'	Accession number
mThbs3_seq3	CGGGGACGATGACAATGACG	
mTmprss4_seq1	GCAGCCCACTGCTTCAGG	NM_145403.3
mVtcn1_seq1	GGGGCAGATCATCTTTTGGAGTATT	NM_178594.4

2.8.6 shRNA oligos

Tab. 13: shRNA oligos

oligo	Sequence 5'-3'
shAsb15.211	TGCTGTTGACAGTGAGCGACCAGGCTATATTTCATCCTAATAGTGAAGCCACAGATGTATTAGGATGAAATATAGCCTGGCTGCCTACTGCCTCGGA
shAsb15.1834	TGCTGTTGACAGTGAGCGCGCGAAGATACTTATTATTTAATAGTGAAGCCACAGATGTATTAAATAATAAGTATCTTCGCATGCCTACTGCCTCGGA
shFosl1.948	TGCTGTTGACAGTGAGCGATCCGAGTCTGGTTTTCACCTATAGTGAAGCCACAGATGTATAGGTGAAAACCAGACTCGGAGTGCCTACTGCCTCGGA
shFosl1.949	TGCTGTTGACAGTGAGCGCCCGAGTCTGGTTTTCACCTATTAGTGAAGCCACAGATGTAATAGGTGAAAACCAGACTCGGATGCCTACTGCCTCGGA
shOrc1.330	TGCTGTTGACAGTGAGCGCACTGATTGAATTATTTCAAAATAGTGAAGCCACAGATGTATTTTGAAATAATTCAATCAGTTTGCCTACTGCCTCGGA
shOrc1.2732	TGCTGTTGACAGTGAGCGAAGGATGGAAATTGTTGTTATATAGTGAAGCCACAGATGTATATAACAACAATTTCCATCCTGTGCCTACTGCCTCGGA
shThbs3.659	TGCTGTTGACAGTGAGCGACTGAGTGAATGTCCATTCCAATAGTGAAGCCACAGATGTATTGGAATGGACATTCACTCAGGTGCCTACTGCCTCGGA
shThbs3.922	TGCTGTTGACAGTGAGCGCCTGTATGGAAGTGTATGAGTATAGTGAAGCCACAGATGTATACTCATACACTTCCATACAGTTGCCTACTGCCTCGGA
shThbs3.1162	TGCTGTTGACAGTGAGCGACAATGACATTGATGAATGTAATAGTGAAGCCACAGATGTATTACATTCATCAATGTCATTGCTGCCTACTGCCTCGGA
shThbs3.1529	TGCTGTTGACAGTGAGCGACAAGAAGATGCTGATAACGATTAGTGAAGCCACAGATGTAATCGTTATCAGCATCTTCTTGCTGCCTACTGCCTCGGA

oligo	Sequence 5'-3'
shThbs3.2143	TGCTGTTGACAGTGAGCGCGGGTGATGTCTGTGAAGATGATAGTGA AGCCACAGATGTATCATCTTCACAGACATCACCCATGCCTACTGCC TCGGA

2.9 Plasmids

Tab. 14: Plasmids

Plasmids	origin	application
CaKIG-mirE	Gift from AG Zender (Tübingen)	Stable expression *in vivo*
CaKIG-shAsb15.211	This study	Stable expression *in vivo*
CaKIG-shAsb15.1834	This study	Stable expression *in vivo*
CaKIG-shFosl1.948	This study	Stable expression *in vivo*
CaKIG-shFosl1.949	This study	Stable expression *in vivo*
CaKIG-shOrc1.330	This study	Stable expression *in vivo*
CaKIG-shOrc1.2732	This study	Stable expression *in vivo*
CaKIG-shRNA_library	This study	Stable expression *in vivo*
CaKIG-shThbs3.659	This study	Stable expression *in vivo*
CaKIG-shThbs3.1529	This study	Stable expression *in vivo*
CaMIA-mirE	Gift from AG Zender (Tübingen)	Stable expression *in vivo*
CaMIA-shThbs3.659	This study	Stable expression *in vivo*
CaMIA-shThbs3.1529	This study	Stable expression *in vivo*
LT3-GEPIR-mirE_Renilla	Gift from AG Zender (Tübingen)	Stable inducible expression *in vitro*
LT3-GEPIR-mirE-shThbs3.659	This study	Stable inducible expression *in vitro*
LT3-GEPIR-mirE-shThbs3.922	This study	Stable expression *in vivo*
LT3-GEPIR-mirE-shThbs3.1162	This study	Stable expression *in vivo*
LT3-GEPIR-mirE-shThbs3.1529	This study	Stable inducible expression *in vitro*
LT3-GEPIR-mirE-shThbs3.2143	This study	Stable expression *in vivo*

Plasmids	origin	application
pCaggs-Adra1a_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Adra1a_mut_RFP	This study	Stable expression *in vivo*
pCaggs-Ankrd1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Asb4_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Asb15_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Dtx1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Fosl1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Foxj1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Gata1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Gli1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Glis1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Klf5_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Macc1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Orc1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Orc1_mut_RFP	This study	Stable expression *in vivo*
pCaggs-polyII	Gift from AG Zender (Tübingen)	Stable expression *in vivo*
pCaggs-Thbs3_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Thbs3_mut_RFP	This study	Stable expression *in vivo*
pCaggs-Tmprss4_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Tmprss4_mut_RFP	This study	Stable expression *in vivo*
pCaggs-Vtcn1_wt_RFP	This study	Stable expression *in vivo*
pCaggs-Vtcn1_mut_RFP	This study	Stable expression *in vivo*
pLV_EF1a_IRES_BLAST	Addgene/ #85133	Stable expression *in vitro*
pLV-Thbs3_wt	This study	Stable expression *in vitro*
pLV-Thbs3_mut	This study	Stable expression *in vitro*
pLV-hTHBS3_wt	This study	Stable expression *in vitro*
pLV-hTHBS3_mut	This study	Stable expression *in vitro*
pMD2.G	Gift from AG Tschaharganeh (Heidelberg University Hospital)	Lentiviral envelope plasmid

Plasmids	origin	application
psPax2	Gift from AG Tscharganeh (Heidelberg)	Lentiviral packaging plasmid
pT3-TRE-tRFP-mirE	This study	Stable expression *in vitro*

2.10 Software

Tab. 15: Software and Webtools

Software	Provider
Adobe Illustrator CS5	Adobe Systems, Munich, Germany
Adobe Photoshop CS6	Adobe Systems, Munich, Germany
Agilent QuikChange Primer Design	https://www.agilent.com/store/primerDesignProgram
ANNOVAR	http://www.openbioinformatics.org/annovar/
Aperio ImageScope v12.4.3.7001	Leica Biosystems, Nussloch, Germany
BioRender	https://biorender.com
cbioportal webpage	https://www.cbioportal.org/
cellSens Dimension	Olympus, Hamburg, Germany
COSMIC v88	https://cancer.sanger.ac.uk/cosmic
Endnote 21	ClarivateTM, Chandler, USA
FIJI/ImageG v1.46j	www.fiji.sc
Human Protein Atlas	https://www.proteinatlas.org
GraphPad Prism8	GraphPad Software, San Diego, USA
Image Studio v3.1.4	LI-COR Biosciences, Bad Homburg, Germany
Ingenuity® Pathway Analysis (IPA®)	QIAGEN
Microsoft Excel 2019	Microsoft
MutationTaster	https://www.mutationtaster.org
NanoDrop 1000 3.8.1	Thermo Fisher Scientific, Waltham, USA
NEB TM calculator	https://tmcalculator.neb.com/
Omega v3.00 R2 and MARS	BMG Labtech, Ortenberg, Germany
Polyphen-2	http://genetics.bwh.harvard.edu/pph2/

Software	Provider
QuantStudio™ Design & Analysis Software v1.4.3	Thermo Fisher Scientific, Waltham, USA
SnapGene Viewer	GSL Biotech, Chicago, USA
StepOne v2.3	Applied Biosystems, Darmstadt, Germany

3 Methods

3.1 Human samples

3.1.1 Human DNA and RNA extraction from cHCC-CCA

The cHCC-CCA cases were selected based on histological analyses of archived formalin-fixed paraffin-embedded (FFPE) human PLC samples. Tissue areas showing either hepatocellular or cholangiocellular differentiation of each sample were microdissected using each six consecutive 10 µm thick FFPE sections. The individual samples from each area were pooled, deparaffinized and digested with proteinase K overnight. DNA and RNA were extracted with an automated Maxwell® 16 Research extraction system (Promega) using the Maxwell® 16 FFPE Plus LEV DNA Purification Kit and Maxwell® RSC RNA FFPE Kit according to the manufacturer's instructions. DNA and RNA concentrations were measured fluorometrically using the Qubit™ dsDNA HS kit. Overall, 13 cHCC-CCA samples passed quality control and could be used for exome and transcriptome sequencing. The use of patient material for research purposes was approved by the ethics committee of the Medical Faculty Heidelberg (S-134/2014).

3.1.2 Whole exome sequencing

Whole exome sequencing was conducted by Dr. Robert Geffers (Genome Analytics, Helmholtz Centre for Infection Research, Braunschweig). In preparation of the library, the concentration and quality of the extracted genomic DNA (gDNA) was determined using a 2100 Bioanalyzer (Agilent Technologies). 200 ng of gDNA was fragmented in a Covaris microTUBE using a Covaris S2 ultrasonicator (Covaris) (200 cycles per burst, 80 seconds); fragments with an average length of 300 base pairs were generated (bp). Fragment size was verified with a 2100 Bioanalyzer (Agilent Technologies). Then, 200 ng of fragmented gDNA was used for the generation of DNA sequencing libraries using the TruSeq SBS Kit v3-HS (Illumina) according to the manufacturer's instructions. Libraries were sequenced on an Illumina HiSeq2500 platform (2×75 bp, paired end run). On average, the resulting read count was 12.5x106 per single exome and a mean coverage of 50X. Using the FastQC tool, the quality of FASTQ files was determined before and after trimming. Additionally, artificial Illumina adapter sequences and bad quality sequence reads were removed with TrimGalore!. Using the BWA aligner tool

({HYPERLINK "http://bio-bwa.sourceforge.net/"}), trimmed FASTQ files were then aligned to the human reference genome hg19. Subsequently, duplicated reads (PCR products) were tagged with MarkDuplicates (Picard tools). Variants were called based on the resulting BAM files that were processed using GenomeAnalysisTK-1.4 ({HYPERLINK "https://gatk.broadinstitute.org/hc/en-us"). Inclusion criteria of variants were read depth and a normalized variant call confidence >10 in all samples, predicted impact (SnpEff) at least moderate and that variants were not observed in the 1000 genomes project (variant allele frequency <0.05). Using MutationTaster, neutral and polymorphic variants were excluded. PolyPhen-2 (HDIV >0.452, Hvar >0.446) was used to remove non-functional variants.

3.1.3 Expression profiling and analysis

RNA sequencing was conducted by Dr. Robert Geffers (Braunschweig). RNA sequencing libraries were generated using the TruSeqRNA Access Library Preparation Kit according to the manufacturer's instructions. 100 ng of RNA sequencing libraries were sequenced on an Illumina HiSeq2500 platform. On average, the resulting read count was 3×10^7 per transcriptome. Using STAR, raw RNA sequencing reads were mapped to the reference sequence and with Picard tools, duplicates were marked and sorted. Reassignment of mapping quality was achieved with Split´N´Trim. Inclusion criteria for differentially expressed genes between HCC and iCCA areas of cHCC-CCA were defined as >100 counts per million reads and a false discovery rate <0.05. Furthermore, only genes annotated as transcriptional regulators or lincRNA were considered. Differential expression of genes was based on LogFC values which represent a fold change in expression between the 2 components of cHCC-CCA and a significant change was defined as LogFC values >1.5 or <-1.5 with a detection threshold between the areas of each tumor at p <0.001. A positive LogFC value means higher expression in the HCC area compared to the iCCA area of cHCC-CCA. Vice versa a negative LogFC value means higher expression in the iCCA area of cHCC-CCA. Notably, the LogFC values do not represent the mean LogFC values of a gene across all 13 samples but the LogFC value of one sample or the mean LogFC value of various samples that passed the significance threshold of p <0.001.

3.2 Mouse work

3.2.1 Mosaic mouse models using hydrodynamic tail vein injection

Mosaic mice were generated by hydrodynamic tail vein injection (HDTVI) of transposon vectors in collaboration with the research group of Prof. Dr. med. Lars Zender (Tübingen). This method was used to specifically deliver plasmids (shRNA, expression plasmids) into the livers of wildtype and p19$^{-/-}$ C57BL/6 mice, respectively. In this method, plasmid DNA dissolved in a saline solution equaling 10% of the mouse body weight is injected in less than 10 s into the tail vein of mice. In case of CaKIG (KRAS^{G12V}-IRES-GFP), 5 μg were used. For CaMIA (MYC-AKT1), 25 μg were injected. Irrespective of the model, 1 μg of Sleeping Beauty transposase was co-injected. All experiments were performed according to the institutional regulations of the institutional regulation of the animal facility in Tübingen.

3.2.2 Mouse tissue harvest

Following HDTVI, tumor formation was assessed by clinical examination such as abdominal palpation and ultrasound imaging in collaboration with the research group of Prof. Dr. med. Lars Zender (Tübingen). Upon tumor development, mice were sacrificed, and individual tumor nodules were dissected. Half of each nodule was formalin fixed overnight and subsequently paraffin embedded. FFPE samples were used for histological and immunohistochemical analysis. The other half of each nodule was freshly frozen for subsequent NGS analysis. In all experiments, FFPE sections were stained with hematoxylin and eosin (H&E) and histopathological evaluation and tumor typing was performed by Prof. Dr. med. Thomas Longerich (Heidelberg University Hospital).

3.2.3 Sanger sequencing

In collaboration with the research group of Prof. Dr. med. Lars Zender (Tübingen), enriched shRNA from the library screen experiment were identified by Sanger sequencing. For this, genomic DNA was isolated from fresh frozen tissue of each nodule. The integrated shRNA sequences were PCR-amplified from the flanking Illumina adaptor sequences. Sanger sequencing was performed at Microsynth SeqLab, Switzerland. Resulting sequences were matched with shRNA sequences.

3.3   Methods of molecular biology

3.3.1   RNA isolation and cDNA synthesis

Total RNA was isolated according to the manufacturer's instructions using the NucleoSpin® RNA II kit and stored at -80°C. cDNA was synthesized from 100 ng – 2000 ng of RNA according to the manufacturer's instructions using the PrimeScript RT Master Mix and subsequently stored at -20°C.

3.3.2   Semi-quantitative real-time PCR (qPCR)

Gene expression levels were assessed by means of qPCR. Reactions were based on the PowerUp™ SYBR™ Green Master Mix (Tab. 16). Cycling conditions were as follows: 95°C for 10 min, then 40x cycles of 95°C for 15 s and 60°C for 60 s followed by melting curve analysis for product specificity which was set up as follows: 95°C for 15 s, 60°C for 60s, 60-95°C with 0.15°C/second). Relative gene expression was calculated with the delta-delta Ct ($2^{-\Delta\Delta CT}$) method. Housekeeping gene GAPDH and Gapdh were used for normalization of human and mouse samples, respectively.

Tab. 16: qPCR master mix (1 rxn)

Reagent	Volume	Final concentration
SYBR Green Mix	5.0 µl	50%
Primer_fwd [10 µM]	0.4 µl	0.4 µM
Primer_rev [10 µM]	0.4 µl	0.4 µM
cDNA (1:25)	2.0 µl	
dH2O	2.2 µl	

3.3.3   Expression profiling and pathway analysis of Thbs3-expressing CaKIG cells

The purpose of RNA sequencing was the detection of differentially expressed genes and the identification of differentially altered signaling pathways affected by THBS3 expression. For this, CaKIG cells either stably expressing a wildtype or a mutant Thbs3 gene as well as CaKIG cells with Dox-inducible Thbs3 knockdown were used. The following eight clones were analyzed in triplicates: CaKIG Thbs3 wt1, wt4, mut2, mut5, ctrl, shRNA.Renilla ctrl,

shThbs3.659, and shThbs3.1529. Total RNA was isolated 48h after cell seeding and Dox induction. Quality control and RNA sequencing were performed by Dr. Robert Geffers (Braunschweig). Data analysis was performed by Klaus Kluck and Iordanis Ourailidis (research group Medical Bioinformatics headed by Prof. Dr. rer. nat. Jan Budczies, Institute of Pathology, Heidelberg University Hospital). Fastp [231] was used for adapter trimming and low-quality reads filtering. Salmon [232], along with the GRCm39 primary assembly, were used for the quantification of transcript expression. The pipeline can be openly accessed at {HYPERLINK "https://github.com/iouraili/SalmRNAseq"}. DESeq2 statistical package was used for differential gene expression. Benjamini-Hochberg (BH) corrected significance level was set at FDR < 0.1.

3.3.4 Assay for Transposase-Accessible Chromatin (ATAC) sequencing

The regulation of gene expression is based on chromatin accessibility. The Assay for Transposase Accessible Chromatin allows for high-throughput sequencing (ATAC-seq) revealing chromatin accessible regions at a genome-wide level. These analyses were performed using the same eight clones as in 3.3.3. All samples were prepared according to the manufacturer's instructions using the ATAC-Seq Kit (Active Motif). Quality control and RNA sequencing were performed by Dr. Robert (Braunschweig). Data analysis was performed by Klaus Kluck and Iordanis Ourailidis (Heidelberg). Samples were preprocessed which included read trimming and base-call quality checks. Then samples were aligned to mm10 with BWA. Subsequently, peaks were called with MAC2. This included peak annotation (intronic, exomic, intragenic, distance to closest transcription start site (TSS) or transcription termination site (TTS)) and the calculation of consensus peaks with corresponding peak heights. Each step was accompanied by quality checks. Then, two analyses were conducted for merged and non-merged replicates. Resulting peak tracks were visualized by the generation of IGV session files. The generation of PCA plots and correlation heat maps served for visualization of the results and for quality checks. The pipeline can be openly accessed at {HYPERLINK "https://nf-co.re/atacseq"} [233]. DESeq2 statistical package was used for differential accessibility analysis. BH corrected significance level was set at FDR < 0.1.

3.3.5 Cloning of mirE-based shRNA library

Cloning of the mirE-based shRNA library was performed by the research group of Prof. Dr. med Lars Zender (Tübingen). shRNA sequences were retrieved from the splashRNA

website {HYPERLINK "http://splashrna.mskcc.org/"} by entering the entrezID number which considers all mRNA variants. The 6 first predictions per candidate gene were synthesized and tested *in vitro* regarding their knockdown efficiency. The 5 most efficient shRNA per gene were used for subsequent *in vivo* experiments. This resulted in a total of 270 shRNA targeting the 54 selected candidate genes (Appendix Tab. A). First, the ordered oligos were dissolved to a stock concentration of 1 µg/µl. Then the stock solution was diluted to a final concentration of 0.05 ng/µl. The diluted solution served as a template for PCR amplification (Tab. 17). Thermocycling conditions are described in Tab. 17. Primers MirE_fwd and MirE_rev (2.8.1) were used to create overhanging sites for the restriction enzymes (RE) XhoI and EcoRI. The transposon vectors CaMIA (MYC-AKT1) and CaKIG (KRAS^{G12V}-IRES-GFP) were digested with XhoI and EcoRI followed by column purification using the Qiagen Quick Gel Extraction Kit (Qiagen) and subsequent standard calf-intestinal alkaline phosphatase reaction according to the manufacturer's instructions (NEB) to remove phosphate groups from the 5' end of DNA strands. Then, backbones and the library pool were purified using phenol/chloroform. For this, an equal volume of phenol/chloroform was added to the DNA and the solutions were mixed. The mixture was then transferred into a phase lock light tube and separated by centrifugation at 14000 rpm for 10 min. Subsequently, the top phase was transferred to a fresh Eppendorf tube. 10% volume of sodium acetate (3 M, pH 4.8), 3 volumes of 100% ethanol and 1 µl glycogen were added to the Eppendorf tube, which was then incubated at -20°C. After 2 h, the tube was centrifuged again at 14000 rpm for 20 min. Following centrifugation, the supernatant was aspirated. The pellet was then washed with 1 ml of 70% ethanol by centrifugation at 14000 rpm for 20 min. After aspirating the medium, the pellet was air dried and lastly re-suspended in dH2O. The purified backbone and library pool were ligated using Roche Rapid ligation Kit according to the manufacturer´s instructions (Roche). The resulting library comprised a 1000-fold overrepresentation of each shRNA. In a pre-chilled Eppendorf tube, 20 µl ElectroMax DH10B electrocompetent bacteria were transformed with the transposon vector carrying the shRNA library by electroporation using a Gene Pulser Xcell (Biorad) at 2 kV time constant <5 ms. 1 ml Super Optimal broth with catabolite repression medium was added and the tube was incubated at 37°C for 1 h while shaking. Then, the transformed cells were plated on agar plates and incubated overnight at 37°C. Maxipreps were prepared using the EndoFree Plasmid Maxi Kit by picking colonies from the plates according to the manufacturer's instructions (Qiagen).

Tab. 17: Thermocycling conditions for PCR

Step	Description	Temperature	Time	Cycles
1	Initial denaturation	95°C	120 s	1
2	Denaturation	95°C	20 s	
3	Annealing	54°C	20 s	30
4	Elongation	72°C	30 s	
5	Final elongation	72°C	180 s	1

3.3.6 Cloning of phenotype-driving genes

Cloning of potential phenotype-driving genes into a transposon vector was either conducted by PCR-amplification of fragments with overhangs carrying restriction enzyme (RE) sites or by high-fidelity (HIFI) assembly (New England Biolabs). The genetic information of interest was either commercially available as part of a pcDNA3.1 vector (GenScript) or in case of Asb15, Fosl1, Orc1, and Thbs3, PCR-amplified from gDNA of murine cell lines Hepa1-6, Hep56, Kp19$^{-/-}$, and AMp19$^{-/-}$.

Regarding the cloning strategy with RE sites, both backbone and insert (either in original vector or as PCR fragment) were digested with the same two REs (either AscI and NotI or AscI and XhoI) at 37°C for 4 h. Additionally, backbones were incubated with recombinant Shrimp Alkaline Phosphatase at 37°C for 2 h to remove phosphate groups from 5' and 3' ends of DNA. Subsequently, vector backbones and gene-of-interest coding DNA fragments were jointly loaded onto a 1% agarose gel. Respective cDNA and linear vector backbone were cut from the gel, purified using Wizard® SV Gel and PCR Clean-Up System before ligating them using T4 DNA ligase at 16°C overnight.

The HIFI assembly strategy using the NEBuilder® HiFi DNA Assembly Master Mix and NEBuilder Assembly Tool (New England Biolabs) differs from the RE cloning strategy in that the PCR-amplified gel-purified insert is not cut by REs, instead the gene-of-interest-containing DNA fragment is annealed to the linear vector backbone with large overlapping sequence stretches of 20 – 30 nucleotides and a proprietary high-fidelity polymerase fills in the missing bases on both strands. The vector backbone is processed in the same way as mentioned before.

In both cloning strategies, the ligation products were used to transform Mix & Go competent cells by heat shock at 42°C for 45 s and subsequent incubation on ice for 2 min. Then, cells

were plated on agar plates and incubated at 37°C overnight. Colonies were picked and minipreps were prepared using the Monarch® Plasmid Miniprep Kit according to the manufacturer's instructions (New England Biolabs). Sequences were verified by Sanger sequencing using the respective sequencing primers (2.8.5) by Microsynth Seqlab (Göttingen, Germany). Sequences were evaluated using SnapGene Viewer. Once sequences were verified, maxipreps were prepared for subsequent *in vivo* experiments. For *in vitro* experiments, midipreps were prepared according to the manufacturer's instructions using the PureYield™ Plasmid Midiprep System (Promega).

3.3.8 Site-directed mutagenesis

Variants of candidate genes were generated by introduction of nucleotide substitutions according to the manufacturer's instructions using the QuickChange Lightning Site-Directed Mutagenesis Kit with mutagenesis primers shown in 2.8.3.

3.4 Methods of cell biology

3.4.1 Cell cryopreservation

Long-term storage of cells was achieved by cryo-conservation. For this, cells of 80%-90% confluent 10 cm plates were trypsinized, pelleted by centrifugation at 1000 rpm for 5 min and subsequently re-suspended in medium supplemented with 10% DMSO. The cell suspension was transferred to cryogenic vials and incubated at -80°C for 24 h in freezing containers to cool down at a rate of 1°C/min. Afterwards, cells were stored inside the vapor phase of liquid nitrogen tanks at -196°C. For reuse of cells, cells were thawed and washed by centrifugation to remove DMSO residuals. Then, they were plated into 10 cm plates with pre-warmed medium and cultured at 37°C.

3.4.2 Cell cultivation

Cell lines used in this study were all adherently growing cells of either human or mouse origin as shown in 2.3. The human iCCA cell line HUH28 was cultured in RPMI supplemented with 10% FCS and 1% penicillin/streptomycin. Murine HCC cell lines Hepa1-6 and CaMIA and murine iCCA cell line CaKIG were cultured in DMEM supplemented with 10% FCS and 1% penicillin/streptomycin. Medium for human HEK293T cells was additionally supplemented

with 1% HEPES, 1% L-glutamine, and 1% sodium pyruvate. Cells were kept at 37°C and 5% CO_2 in a humidified incubator (Tab. 7) and routinely checked for mycoplasma contamination every two months. Cells were passaged twice a week by first washing with PBS, then trypsinizing with trypsin-EDTA and subsequent re-suspension in pre-warmed medium before seeding into new cell culture dishes.

3.4.3 Cell seeding

Cells were seeded at different concentrations prior to the start of experiments. For this, cells were washed with PBS, trypsinized and re-suspended at an adequate concentration for counting using a Neubauer chamber. Then, cells were plated at the required cell number per well.

3.4.4 Transient cell transfection

For verifying the successful cloning of candidate genes into the transposon vectors, transient transfection of Hepa1-6 cells was performed using transfection reagent Lipofectamine 3000 according to the manufacturer's instructions. For this, cells were seeded at a 50%-70% confluency 24 h prior to transfection. The most efficient ratio of transfection reagent to DNA was determined as 5:1 in pilot transfection experiments. Cells were harvested 48 h after transfection. RNA was isolated, cDNA transcribed, and overexpression of the respective candidate assessed by qPCR.

For virus particle production, 1.5×10^6 HEK293T cells were seeded into $25 cm^2$ cell culture flasks 24 h prior to transfection. The transfection mix included 4.17 µg of lentiviral plasmid DNA, 3.33 µg of psPAX2 packaging plasmid, 1.04 µg of pMD2.G envelope plasmid and transfection reagent polyethylenimine (PEI) at a concentration of 1 µg/µl diluted in 1 ml OptiMEM. The mix was incubated at RT for 30 min. Then, cells were washed and provided with new medium before the mix was added dropwise onto the cells.

3.4.5 Lentiviral transduction of cells

Lentiviral transduction was used to achieve either stable gene expression or an inducible expression knockdown of a gene-of-interest in eukaryotic cells. First, lentiviral plasmids (Tab. 14) were co-delivered into HEK293T cells as described in 3.4.5. On day 1 post transfection of HEK293 T cells, medium of cells was changed. Also, the cells to be finally infected with virus particles were seeded at a 50%-80% confluency. On day 2 post-infection,

virus supernatant was harvested at 48 h and 54 h post transfection and filtered using 0.45 μm Millex-HA filters. The resulting virus-containing suspension was either used to infect cells or was frozen at -80°C for later use. For infection, medium of previously seeded cells was changed and polybrene was added at a concentration of 5 μg/ml before adding 1 ml of the virus-containing suspension. Infection of cells resulted in the stable integration of the respective transgene or shRNA into the genome of the cells. On day 1 post-infection of the target cells, the medium of infected cells was replaced. On day 2 post-infection, cells were transferred to a 10 cm plate and selection of successfully transduced cells containing a resistance cassette was achieved by treatment with blasticidin at 1-5 μg/ml or puromycin at 1 μg/ml.

3.4.6 Generation of isogenic cell lines

After lentiviral transduction and antibiotic selection of cells stably expressing the respective gene, cells were seeded according to 3.4.3 at a very high dilution of approximately 1-5 cells/well into all wells of a 96-well plate. Cells were incubated at 37°C. After a week, cells were screened under the microscope and wells with one single isogenic cell colony were marked. Once wells were confluent, 12 wells were selected and their cells each transferred into a well of a 12-well plate. Subsequently, the 12 cell clones were transferred into 6-well plates and their RNA was isolated to confirm overexpression of the respective gene by qPCR.

3.4.7 Inducible gene expression systems

For controlling gene expression, the tetracycline-controlled system derived from the tetracycline-resistance operon in bacteria was used. The Tet-On system used in this study allows activation of gene expression by adding the tetracycline-derivative Dox, which binds to the Tet repressor protein (TetR) variant that functions in reverse fashion. Hereupon, TetR undergoes conformational change that allows *tet* operator binding. Subsequent activation of the *tet* promoter drives expression of the downstream-positioned gene. For the induction of shRNA in stably transduced cells, Dox was added at a concentration of 5 μg/ml after cell passaging. This concentration was determined in a concentration response curve experiment. 48 h after Dox induction, cells were harvested for RNA or protein isolation and subsequent experiments.

3.5 Functional assays

3.5.1 Cell viability assay

Cell viability assays were conducted using CellTiter-Blue® reagent which represents a fluorescent method for monitoring cell viability. Living cells convert the redox dye (resazurin) into a fluorescent end product (resorufin) while nonviable cells do not generate a fluorescent signal. Cells were seeded into 6-well plates at the following concentrations: AMp19$^{-/-}$ cells at 15K/well; Kp19$^{-/-}$ cells and HUH28 cells at 50K/well. At each time point, 2 ml CellTiter-Blue® reagent (1:10 dilution in cell culture medium) was added onto the cells and incubated at 37℃ for 4 h. Then, supernatant was transferred to 96-well plates (8 technical replicates) and fluorescence was measured using a FLUOstar Omega microplate reader. RNA was isolated for further analysis.

3.5.2 Colony formation assay

Clonogenicity of cells was assessed by seeding cells at a low concentration (50 - 2,000 cells/well) into 6-well plates (3 technical replicates). Cells were cultured for 8 d, then washed with PBS and subsequently stained with 0.5% crystal violet solution for 1 h. After staining, cells were washed with water to remove any crystal violet residues. Colony formation was analyzed using the ImageJ plugin ColonyArea[170] and was calculated as the product of area and signal intensity.

3.5.3 Migration assay

Migration capacity of cells was assessed by a transwell-based approach in 24-well plates using ThinCerts TC inserts with pore diameters of 8 μm. In preparation, inserts were incubated in FCS-free medium for 4 h. Then, cells were washed with PBS and trypsinized. Defined trypsin inhibitor (DTI) was used to stop trypsinization as cells were diluted in FCS-free medium. Medium supplemented with FCS was added to the bottom of wells of 24-well plates, FCS-free cell suspension at the required concentration was added into the inserts and inserts were placed into wells. Cells were left to migrate towards FCS as a chemoattractant at 37°C for 24 h. Then, inserts were removed carefully. The inside of inserts was cleaned with a cotton swab to remove cells that did not migrate to the outside of the transwell. Then, cells were fixed in 0.5 ml ice-cold MeOH for 10 min. Subsequently, cells were stained with 0.5 ml of 0.5% crystal violet

solution. Lastly, cells were washed with water to remove any crystal violet residues. Trans wells were then left to dry for 24 h before they were analyzed using the ImageJ plugin Migration Counting Test. Per transwell, 5 pictures (top, bottom, left, right, center) at 40x magnification were analyzed.

3.6 Methods of protein and RNA biochemistry

3.6.1 Protein isolation and quantification

Total protein was extracted from cells of confluent 10 cm plates. After washing of cells with PBS, protein was isolated with 1x Cell Lysis Buffer supplemented with 1x PhosStop, 1x Protease Inhibitor Mix G and 1x PMSF. Sonification (3x 20 s) in an ultrasound water bath was used to disrupt the cell membranes completely. Between sonification steps, samples were incubated on ice for 30 s. Cell debris was removed by centrifugation at 14,000 rpm at 4°C for 10 min. The protein concentration was determined by measuring the absorption at 280 nm using a NanoDrop device. Protein lysates were stored at -20°C.

3.6.2 SDS-Polyacrylamide gel electrophoresis (PAGE) and Western Blot (WB)

Proteins were separated by molecular weight using 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). In preparation, 4x protein sample buffer was added to protein lysates after which the lysates were boiled at 95°C for 5 min. Electrophoresis was conducted in running buffer at 120 V for 2 h. Separated proteins were then blotted onto a nitrocellulose membrane using ice-cold borate buffer at 130 V and 1000 mA/chamber for 1.5 h. Subsequently, the membranes were blocked in 5% BSA in TBS-T for 2 h and then incubated with primary antibody in blocking solution at 4°C overnight (2.2). Membranes were washed 3x for 15 min with TBS-T and incubated with an adequate secondary antibody in blocking solution at RT for 1 h. Washing was repeated as before and then fluorescence signals were detected and quantified using the Odyssey-CLx Infrared Imaging system with the ImageStudio Lite software. Loading controls were GAPDH or β-ACTIN.

3.6.3 Immunohistochemistry (IHC)

Immunohistochemical staining was performed on FFPE tissue and cell block sections by the Center for Model System and Comparative Pathology (CMCP, Institute of Pathology,

Heidelberg University Hospital). Initially, tissue samples were cut at 3 μm sections using a microtome. Then, sections were dried on microscope slides overnight and re-hydrated stepwise (xylene 3x 5 min, 100% ethanol 2x 2 min, 96% ethanol 2x 2 min, 70% ethanol 2x 2 min). Washing was concluded by rinsing the slides with aqua dest. Afterwards, antigens were retrieved using a pressure cooker or steamer with Target Retrieval Solution Citrate pH 6. Following cooldown for 30 min, slides were washed with TBS or TBS-T for 10 min and subsequently incubated with primary antibodies in a wet chamber for 1 h.

For detection, slides were washed twice with TBS or TBS-T for 5 min and then incubated with POLYVIEW® PLUS AP reagent for 45 min. Lastly, slides were washed twice with TBS or TBS-T for 5 min and developed using the Permanent AP Red Kit for 5 min. Antibody concentrations were applied according to the manufacturer's instructions.

3.7 Statistical analysis

Data is presented as mean ± standard deviation (SD). Using GraphPad Prism 9 software suitable statistical tests were performed. First, normality and lognormality tests were conducted. When data were normally distributed, parametric tests, namely t-test, Chi Square test or one/two-way ANOVA was conducted. If data were not normally distributed, non-parametric tests such as Fisher's Exact Test, Mann-Whitney U test or Kruskal-Wallis tests were applied. The significance level was defined as $p \leq 0.05$.

3.7.1 NGS analysis

RNA sequencing and ATAC sequencing analyses were performed with DESeq2 statistical package. Benjamini-Hochberg procedure was used to yield corrected FDR (FDR(BH)) < 0.1.

3.7.2 Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was performed by Klaus Kluck (Heidelberg). GSEA was performed on genes derived from expression profiling (3.3.3) and ATAC-seq (3.3.4) and on the intersection gene set between differential gene expression and promoter accessibility based on ATAC-seq of Thbs3 variant expressing and knockdown cell lines.

Genes were analyzed against M2 (curated), M3 (regulatory and target) and M5 (ontology) gene sets from {HYPERLINK "https://www.gsea-msigdb.org/gsea/msigdb/mouse/collections.jsp"} Significance was set at FDR(BH) < 0.1.

4   Results

4.1 Identification of plasticity driver genes of human primary liver cancer

In this study, I had the goal to identify genes which determine the phenotype of cHCC-CCA. 13 classical cHCC-CCA cases were selected based on histological analysis performed by Prof. Dr. med. Thomas Longerich (Heidelberg). In particular, they showed areas with clear-cut cholangiocellular and clear-cut hepatocellular differentiation, respectively, which were large enough to allow reliable microdissection. Subsequently, DNA and RNA were isolated from each of the two tumor areas and subjected to RNA and Whole Exome Sequencing (WES). Next generation sequencing data were processed by Dr. Robert Geffers (Braunschweig) using the ANNOVAR application [234] (Fig. 5).
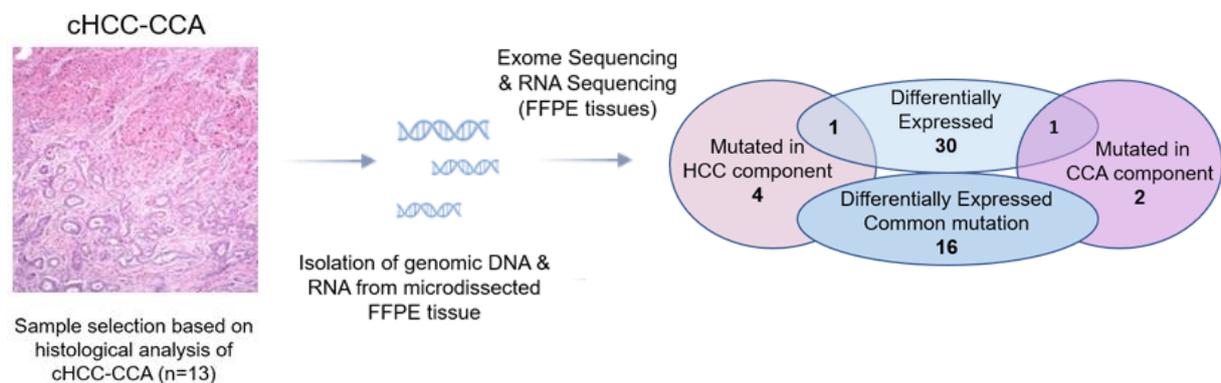
Fig. 5: Workflow for the identification of potentially phenotype driving genes.
Results are presented as a Venn diagram showing the number of differentially expressed and/or mutated genes between the two components of cHCC-CCA.

Of the total 32 differentially expressed genes that I found as a result of RNA sequencing analysis, 30 genes were differentially expressed between the two compartments of cHCC-CCA. The other two did not show any visible expression trend across the tumors. Furthermore, I found nine of the genes upregulated in the HCC compartment and 21 genes were higher expressed in the iCCA compartment. The final candidates that I derived from transcriptome profiling are listed in Tab. 18. For external validation, I matched the sequencing data with the TCGA data sets of human HCC and iCCA. TCGA data regarding the long intervening noncoding RNA (lincRNA) candidate PLUT was not available for validation.

Tab. 18: Expression status of candidate genes between the components of cHCC-CCA.

Gene	Expression and tumor component	Matching TCGA expression data [235, 236]	LogFC	p-value
ARID3C	No visible trend across tumors		-0.129	<0.001
CEBPA	No visible trend across tumors		0.159	<0.001
ANKRD1	Upregulated in iCCA	yes	-6.222	<0.001
ASB15	Upregulated in iCCA	no	-7.070	<0.001
ASB4	Upregulated in iCCA	no	-7.048	<0.001
FOSL1	Upregulated in iCCA	yes	-7.733	<0.001
FOXJ1	Upregulated in iCCA	yes	-7.259	<0.001
FOXN4	Upregulated in iCCA	no	-6.711	<0.001
GATA1	Upregulated in iCCA	no	-8.029	<0.001
GLI1	Upregulated in iCCA	yes	-6.898	<0.001
HR	Upregulated in iCCA	yes	-6.739	<0.001
KLF5	Upregulated in iCCA	yes	-6.289	<0.001
MACC1	Upregulated in iCCA	yes	-6.555	<0.001
MLIP	Upregulated in iCCA	no	-6.087	<0.001
MYCL	Upregulated in iCCA	no	-7.083	<0.001
PLUT	Upregulated in iCCA	Not available	-6.469	<0.001
POU6F2	Upregulated in iCCA	no	-6.701	<0.001
PRDM8	Upregulated in iCCA	yes	-7.835	<0.001
REXO4	Upregulated in iCCA	no	-6.974	<0.001
SPIB	Upregulated in iCCA	yes	-7.551	<0.001
TBX19	Upregulated in iCCA	yes	-7.001	<0.001
ZNF19	Upregulated in iCCA	yes	-6.310	<0.001
ZNF92	Upregulated in iCCA	yes	-6.800	<0.001
DTX1	Upregulated in HCC	yes	8.148	<0.001
ISX	Upregulated in HCC	yes	9.370	<0.001
MNX1	Upregulated in HCC	no	7.629	<0.001
NAT8	Upregulated in HCC	yes	6.888	<0.001

Gene	Expression and tumor component	Matching TCGA expression data [235, 236]	LogFC	p-value
SALL4	Upregulated in HCC	no	6.145	<0.001
SIM2	Upregulated in HCC	no	7.320	<0.001
SOWAHB	Upregulated in HCC	yes	6.952	<0.001
TBX3	Upregulated in HCC	yes	6.748	<0.001
TCEA3	Upregulated in HCC	yes	7.410	<0.001

In terms of mutational status, I identified five mutations specific to HCC and three mutations exclusively present in the iCCA region of cHCC-CCA. Furthermore, I found 16 mutations that were present in both compartments but showed differential RNA expression levels (Tab. 19).

By combining the results from both analyses, I obtained a list of 54 potential phenotype driver genes of PLC.

Tab. 19: Mutation candidate genes of cHCC-CCA as a result of WES.

Gene	Mutation	SNV	Detected in	Expression status	LogFC	p-value
ATP5F1	I150V	Att/Gtt	CCA	Upregulated in HCC		
CRTAP	V214A	gTg/gCg	CCA	Expressed		
MND1	K85M	aAg/aTg	CCA	Expressed		
ARHGAP22	R628C	Cgc/Tgc	HCC	Expressed		
KANK4	H171R	cAc/cGc	HCC	Expressed		
SETSIP	D263G	gAt/gGt	HCC	Upregulated in iCCA		
THBS3	R102Q	cGg/cAg	HCC	Expressed		
ZNF316	S123C	tCc/tGc	HCC	Expressed		
ABCA7	I1690T	aTc/aCc	HCC & iCCA	Upregulated in iCCA	-0.847	0.565
ADRA1A	I200S	aTc/aGc	HCC & iCCA	Upregulated in iCCA	-3.411	0.033
DNAH2	Y516H	Tac/Cac	HCC & iCCA	Upregulated in iCCA	-2.947	0.061
ORC1	T466M	aCg/aTg	HCC & iCCA	Upregulated in iCCA	-0.943	0.533
RAVER1	P29Q	cCg/cAg	HCC & iCCA	Upregulated in iCCA	-0.794	0.589
SLC35F2	L100V	Ctt/Gtt	HCC & iCCA	Upregulated in iCCA	-2.605	0.102
VTCN1	V120L	Gtg/Ctg	HCC & iCCA	Upregulated in iCCA	-9.723	0.000

Gene	Mutation	SNV	Detected in	Expression status	LogFC	p-value
GLIS1	UTR_5'	5239	HCC & iCCA	Upregulated in HCC	1.309	0.377
KCNK13	A402T	Gca/Aca	HCC & iCCA	Upregulated in HCC	1.542	0.301
MBL2	R52C	Cgt/Tgt	HCC & iCCA	Upregulated in HCC	2.175	0.152
RILP	G284S	Ggc/Agc	HCC & iCCA	Upregulated in HCC	1.222	0.417
TMPRSS4	P413L	cCg/cTg	HCC & iCCA	Upregulated in HCC	3.786	0.034
TRIP10	R222H	cGc/cAc	HCC & iCCA	Upregulated in HCC	0.888	0.545
TUSC5	E34D	gaG/gaC	HCC & iCCA	Upregulated in HCC	1.925	0.335
ZC3H12D	K106R	aAa/aGa	HCC & iCCA	Upregulated in HCC	0.138	0.295
MUC4	I3701S	aTc/aGc	HCC & iCCA	Upregulated in both	-1.795	0.346

4.2 Identification of loss-of-function candidates of PLC

To validate the phenotype driving potential of the selected candidates, I applied RNAi technology in transposon-based mosaic mice. In collaboration with Prof. Dr. med. Lars Zender and colleagues (Tübingen), two mouse models were generated by hydrodynamic tail vein (HDTV) injection of a Sleeping Beauty transposon system (SB13) (Fig. 6). Murine HCC formation was induced by a *MYC-AKT1* transposon (CaMIA) [131] in wildtype (C57BL/6) mice, while the injection of a *KRAS^{G12V}* transposon (CaKIG) into p19-deficient (p19$^{-/-}$) C57BL/6 mice was used for the induction of iCCA development.
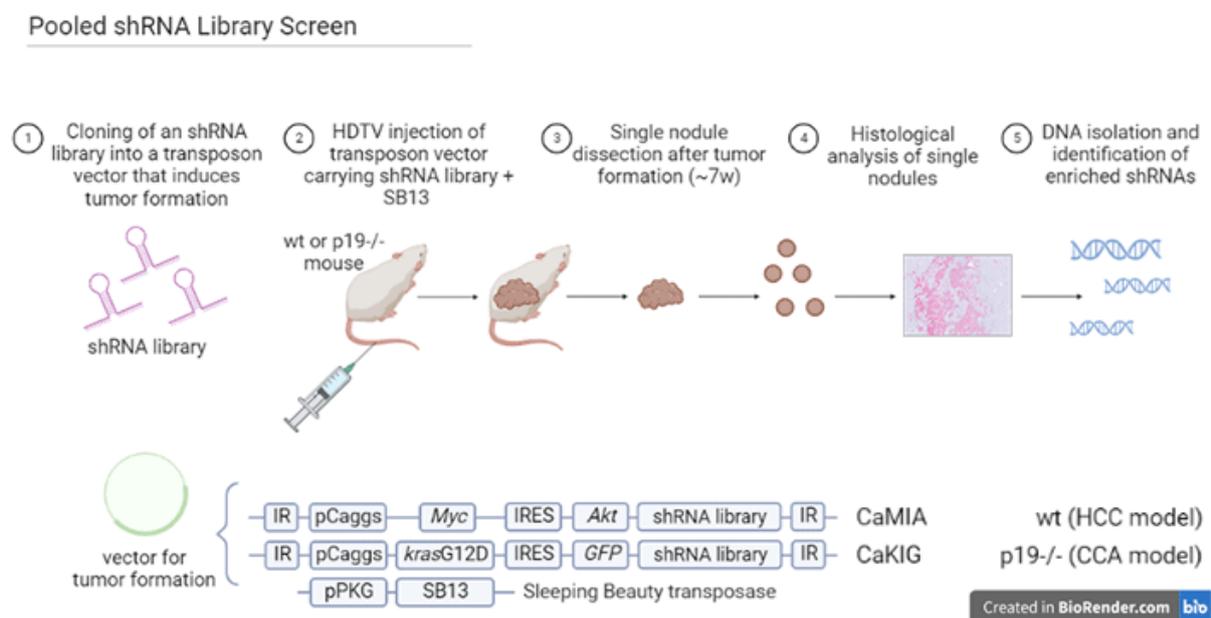
Fig. 6: Workflow of the *in vivo* RNAi screen for the identification of loss-of-function candidates. The schematic organization of the transposon vectors used is shown in the lower part.

In preparation for the *in vivo* screening, a mirE-based shRNA library was generated using the five most efficient shRNA from *in vitro* testing per gene (Appendix Tab. A). An shRNA targeting *Renilla* was used as control. Overall, 271 shRNA (54 candidates x 5 shRNA + 1 control) were cloned into a transposon vector to induce PLC formation either in the HCC or in the iCCA model. A positive experimental readout was defined as a phenotypic switch to iCCA or cHCC-CCA in the HCC model or to HCC or cHCC-CCA in the iCCA model. Altogether, screening was performed in each ten mice of the HCC and the iCCA model, respectively. After co-injection of the library together with SB13, the mice were monitored up to seven weeks until tumor development occurred, and the experiment was terminated. In case of the iCCA model, expression of the transposon vector in harvested liver tumors could be visualized by GFP detection (Fig. 7). In total, I dissected and processed 96 individual liver tumor nodules for
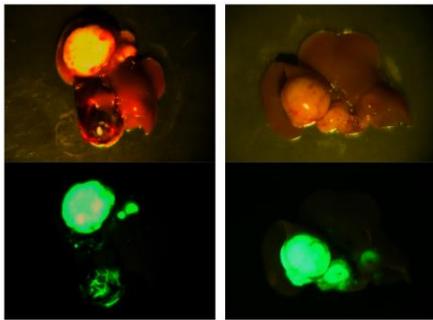
Fig. 7: Detection of GFP expression in murine liver cancer following hepatectomy during autopsy.

histological evaluation by Prof. Dr. med. Thomas Longerich (Heidelberg). These included 58 lesions from the HCC model and 38 tumors from the iCCA model. A phenotypic switch was noted in 17 nodules, of which six were dissected from the HCC and eleven from the iCCA model. I isolated DNA from the 17 nodules and subsequently performed sequencing. In this way I identified the expressed shRNA, respectively. As a final step, I integrated the available human data on the remaining candidates. In case of twelve nodules, the phenotypic switch of the identified shRNA did not match the human data, thus I did not consider the associated candidate genes for further evaluation. The remaining five nodules exhibited a phenotypic switch, which was consistent with the publicly available human data. I identified four candidate genes based on the expression of shRNA in these five nodules, namely *ASB15, FOSL1, ORC1* and *THBS3*. For *ASB15*, *ORC1* and *THBS3*, I found one of the five shRNA per gene that I had injected enriched in each one nodule, namely shAsb15.211, shOrc1.330 and shThbs3.1529. In case of *FOSL1*, I found shFosl1.948 enriched in two nodules. I retrieved *ORC1* and *THBS3* initially from the WES part of the sequencing analysis, while I derived *ASB15* and *FOSL1* from the RNA sequencing analysis as they were differentially expressed between the HCC and the iCCA areas of cHCC-CCA (Tab. 18 & 19).

4.3 Validation of THBS3 as a phenotype driver gene in PLC

4.3.1 *In vivo* validation in transposon-based mosaic mouse models

To individually validate the four remaining candidates, they were expressed in the mouse model in which they did not show intratumoral shRNA enrichment. Thus, *ASB15, FOSL1* and *ORC1* were expressed in the HCC model, while I validated *THBS3* in the iCCA model. Before testing them in the respective PLC model, I generated the variants $Orc1^{T466M}$ and $Thbs3^{R102Q}$ by site-directed mutagenesis.
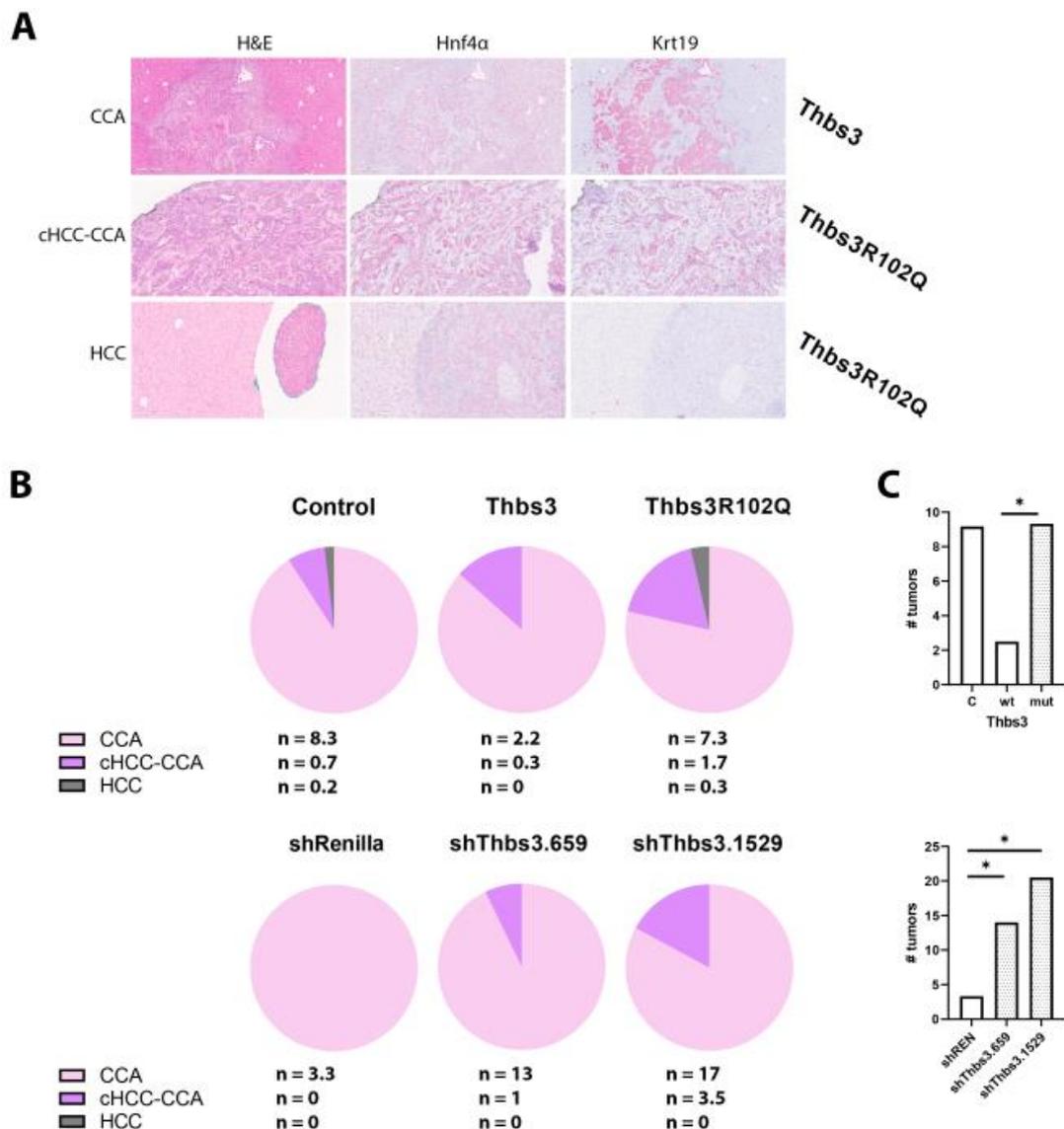
Fig. 8: Histological evaluation of mosaic mouse tumors with altered *Thbs3* gene expression.
A) Representative histological sections of murine liver tissues with mosaic expression of *Thbs3* variants. Hnf4α indicates hepatoid differentiation, while Krt19 is a marker of biliary or progenitor cell differentiation.
B) Distribution of tumor types in the respective models shown as average number of nodules per mouse.

C) Frequency of PLC formation in the respective models calculated as average number of nodules per mouse. Statistical test: One-way ANOVA, $P \leq 0.05$.

Overall, I dissected and processed 214 individual tumor nodules which were subsequently typed by histological assessment by Prof. Dr. med. Thomas Longerich (Heidelberg). Positive readouts were observed for *Asb15*, *Fosl1* and *Thbs3*. Expression of *Asb15* resulted in the expected formation of HCC in 89% of nodules (n = 16/18) in the HCC model, while the remaining tumors were cHCC-CCA (11%, n = 2/18). Induction of a fibrotic tumor stroma was evident in the latter two cases.

Concerning *Fosl1*, almost half of tumors were cHCC-CCA (47%, n = 18/38) with induction of a fibrotic stroma in 39% (n = 7/18) of them (data not shown).

Expression of *Thbs3^{R102Q}* resulted in a higher cHCC-CCA frequency compared to the expression of the wildtype gene (18% vs. 13%). Furthermore. expression of the *Thbs3* wildtype gene significantly reduced the likelihood of tumor formation compared to *Thbs3^{R102Q}* (2.5 nodules/*Thbs3*-wildtype mouse vs. 9.3 nodules/*Thbs3^{R102Q}* mouse, p $\leq$0.05, Fig. 8).

In parallel, I validated the four candidate genes in a second RNAi experiment, in which each two individual shRNA (the one initially scoring during the RNAi screen and a second with high knockdown efficacy; Appendix Tab. A) were used to knockdown candidate gene expression in p19$^{-/-}$ C57BL/6 mice. I used the iCCA model because both *ASB15* and *FOSL1* were found upregulated in the iCCA component of human cHCC-CCA and both mutated candidates (*ORC1*, *THBS3*) were detected in the HCC component of cHCC-CCA.

Based on histological analysis of 163 individual tumor nodules performed by Prof. Dr. med. Thomas Longerich (Heidelberg), I only validated *THBS3* as a phenotype driving gene. I did not consider the other three candidates for further evaluation as shRNA-mediated knockdown of *Asb15* and *Orc1* did not result in a consistent phenotypic change compared to the control and none of the mice with *Fosl1* knockdown showed any tumor development. Conversely, shRNA-mediated knockdown of *Thbs3* resulted in an increased frequency of cHCC-CCA formation (7% with shThbs3.659 and 17% with shThbs3.1529 vs. 0% with shRenilla). Additionally, I found that the shRNA-mediated *Thbs3* knockdown accelerated PLC development in both shThbs3.659- (4-fold) and shThbs3.1529- (7-fold) injected mice compared to the shRenilla control (14 nodules/shThbs3.659 mouse and 24 nodules/shThbs3.1529 mouse vs. 3.3 nodules/shRenilla mouse). I observed that this mimicked the results from the overexpression experiments, in which *Thbs3^{R102Q}* expression resulted in an about four times

higher tumor number compared to *Thbs3* wildtype-expressing mosaic mice. Similarly. As mentioned above, I saw that about 18% of *Thbs3^{R102Q}*-expressing tumors revealed a cHCC-CCA phenotype in the same model. Thus, I confirmed that the wildtype *THBS3* gene has a tumor suppressor function and seems to be important to maintain a cholangiocytic phenotype. In contrast, I found that both loss-of-expression and loss-of-function shift the cellular phenotype towards hepatocytic differentiation.

4.3.2 *In vitro* validation in primary murine isogenic cells

To confirm *in vivo* findings regarding the tumor suppressor and phenotype-modulating functions of THBS3 in PLC, I generated stable clones from primary murine isogenic cells. I derived Kp19^{-/-} cells (kindly provided by AG Zender, Tübingen) from the CaKIG-induced iCCA mouse model (KRAS^{G12V} transposon in a p19^{-/-} mouse background). From these cells, I generated and selected each two cell clones with stable expression of either the *Thbs3* wildtype or the *Thbs3^{R102Q}* variant (Thbs3 wt1 & wt4; mut2 & mut5, Fig. 9A). Furthermore, I generated stable knockdown cell clones using the shRNA shThbs3.659, shThbs3.1529 and shRenilla (control). I achieved stable *Thbs3* mRNA overexpression compared to the control (wt1: 200-fold; wt4, 136-fold; mut2, 216-fold; mut5, 162-fold; each p ≤0.05); the efficiency of shRNA-mediated *Thbs3* knockdown compared to the shRenilla control was 93% for shThbs3.659 and 92% for shThbs3.1529 (each p ≤0.05).

In these cell lines, I determined the mRNA expression of cholangiocytic (keratin7 (*Krt7*), *Krtl9*, and *Sox9*) and hepatocytic (*Hnf4α*, α-fetoprotein (*Afp*), and albumin (*Alb*)) marker genes by quantitative RT-PCR. I found that expression of the *Thbs3^{R102Q}* mutant resulted in a significantly decreased expression of cholangiocytic marker gene expression in the Kp19^{-/-}-derived clones *in vitro* (*Sox9*: 91% and 88% reduction; *Krt7*: 86% and 89% reduction; *Krt19*: 81% and 84% reduction for the Thbs3^{R102Q} clones mut2 and mut5, respectively; each p ≤0.05, Fig. 9B). Also, I observed that *Thbs3* knockdown reduced *Sox9* and *Krt19* expression (*Sox9*: 46% and 70%; *Krt19:* 60% and 69% for shThbs3.659 and shThbs3.1529, respectively; each p ≤0.05). However, I saw that *Krt7* expression was increased upon knockdown of *Thbs3* expression (12-fold for shThbs3.659 and 13-fold for shThbs3.1529, respectively; each p ≤0.05.

Furthermore, I observed that the expression of mutated *Thbs3^{R102Q}* upregulated hepatocytic marker genes (*Hnf4α*: 8-fold for both clone; *Afp*: 3-fold and 2-fold; *Alb*: 15-fold and 10-fold for the Thbs3^{R102Q} clones mut2 and mut5, respectively; each p ≤0.05, Fig. 9C). Similarly, I could

see that shRNA-mediated knockdown of *Thbs3* expression increased *Hnf4α* (11-fold and 15-fold for shThbs3.659 and shThbs3.1529, respectively; each p ≤0.05) and *Afp* mRNA levels (16-fold and 15-fold for shThbs3.659 and shThbs3.1529, respectively; each p ≤0.05). However, I found that the knockdown of *Thbs3* expression did not induce the *Alb* expression (5% and 1% for shThbs3.659 and shThbs3.1529, respectively; each p ≤0.05) compared to shRNA control.
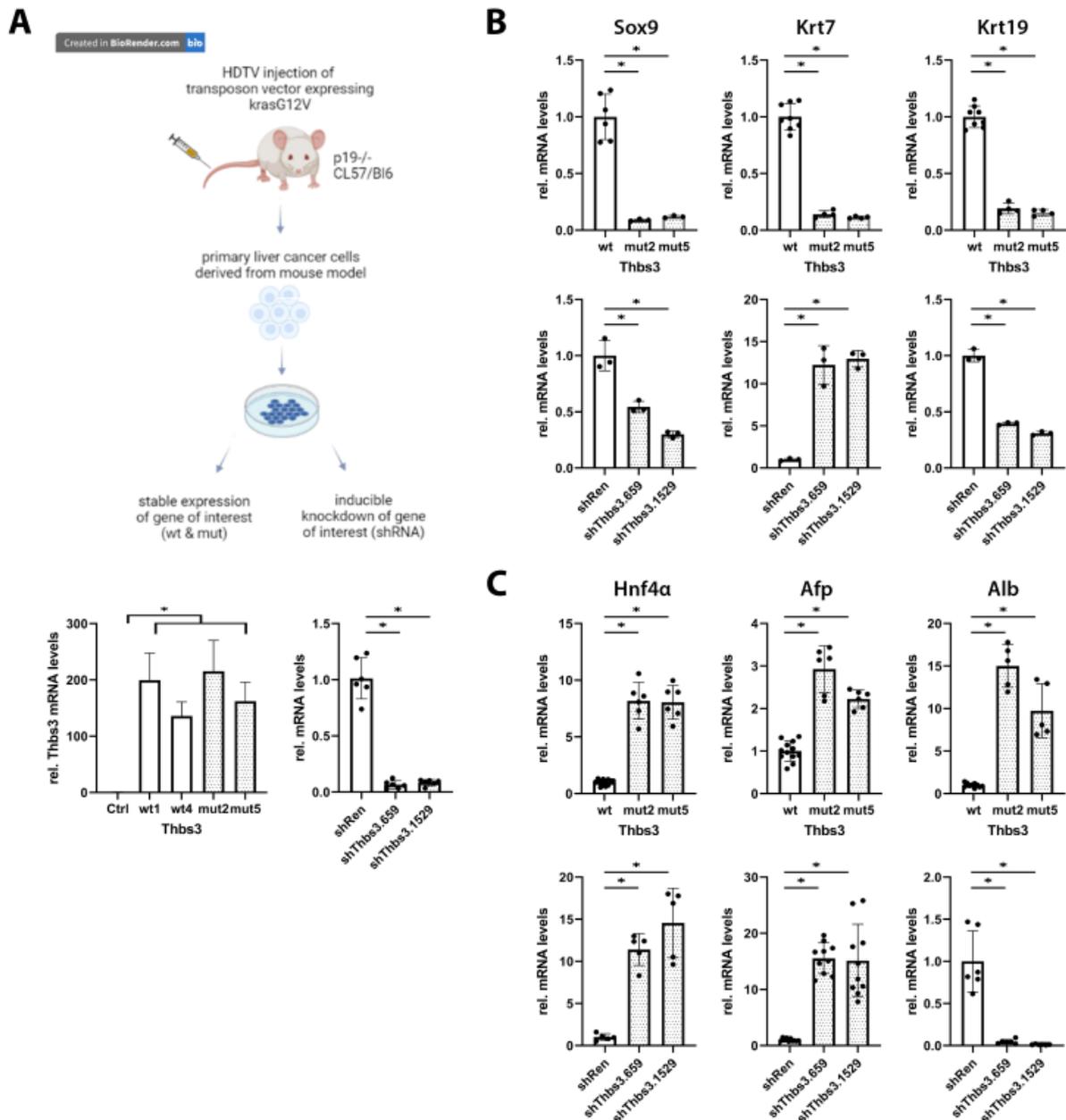
Fig. 9: *In vitro* validation of THBS3 as a driver of cellular plasticity.
A) Workflow of stable primary murine cell line generation. Relative *Thbs3* mRNA levels were determined by qPCR analysis normalized to control (Ctrl). Statistical test: t-test, P ≤0.05. **B)** qPCR analysis of cholangiocytic marker expression at mRNA level normalized to the mean expression of the wt clones 1 &4. Statistical test: t-test, P ≤0.05. **C)** qPCR analysis of hepatocytic marker expression at mRNA level normalized to the mean expression of the wt clones 1 &4. Statistical test: t-test, P ≤0.05.

In summary and in concordance with the previous results of this study, I observed that expression of the *Thbs3* wildtype gene maintained the cholangiocytic phenotype of Kp19$^{-/-}$ cells *in vitro* (*Sox9+, Krt19+, Hnf4α-*) while *Thbs3^{R102Q}* expression resulted in a phenotypic switch towards a hepatoid differentiation (*Sox9-, Krt19-, Hnf4α+*). Furthermore, I found that shRNA-mediated *Thbs3* knockdown only phenocopied the effect of *Thbs3^{R102Q}* expression in Kp19$^{-/-}$ cells with respect to *Hnf4α, Afp, Sox9*, and *Krt19* expression. However, *Krt7* and *Alb* mRNA levels were changed in the opposite direction.

Additionally, I performed functional characterization (proliferation, clonogenicity and migratory capacity) of these stable cell lines (Fig. 10). I assessed cell viability over a period of 48 h and measured at time points 24 h and 48 h. Compared to the control, I observed that *Thbs3^{R102Q}* expression increased the cell viability at 24 h (131% (mut2) and 159% (mut5); each p ≤0.05) and 48 h (130% (mut2) and 136% (mut5); each p ≤0.05), respectively. With *Thbs3* knockdown, on the contrary, I could see decreased cell viability by 38% (shThbs3.659) and 51% (shThbs3.1529) at 24 h and 41% for both clones at 48 h, respectively (each p ≤0.05, Fig. 10A). I found both clonogenicity (9-fold for mut2 and 6-fold for mut5; each p ≤0.05) and migration (19-fold for mut2 & 26-fold for mut5; each p ≤0.05) increased in *Thbs3^{R102Q}*-expressing cells, while *Thbs3* knockdown subtotally blocked the clonogenic (1% & 5%) and migratory (each < 1%) capacities of shThbs3.659- and shThbs3.1529-expressing Kp19$^{-/-}$ cells, respectively (each p ≤0.05, Fig. 10B&C). I also observed, that *Thbs3* knockdown did not phenocopy the functional effects observed after *Thbs3^{R102Q}* expression. Instead, I saw a decrease in proliferation and almost no clonogenic or migratory capacities in shThbs3.659- and shThbs3.1529-expressing cells *in vitro* implying that *Thbs3* expression may be essential for survival of Kp19$^{-/-}$ cells.
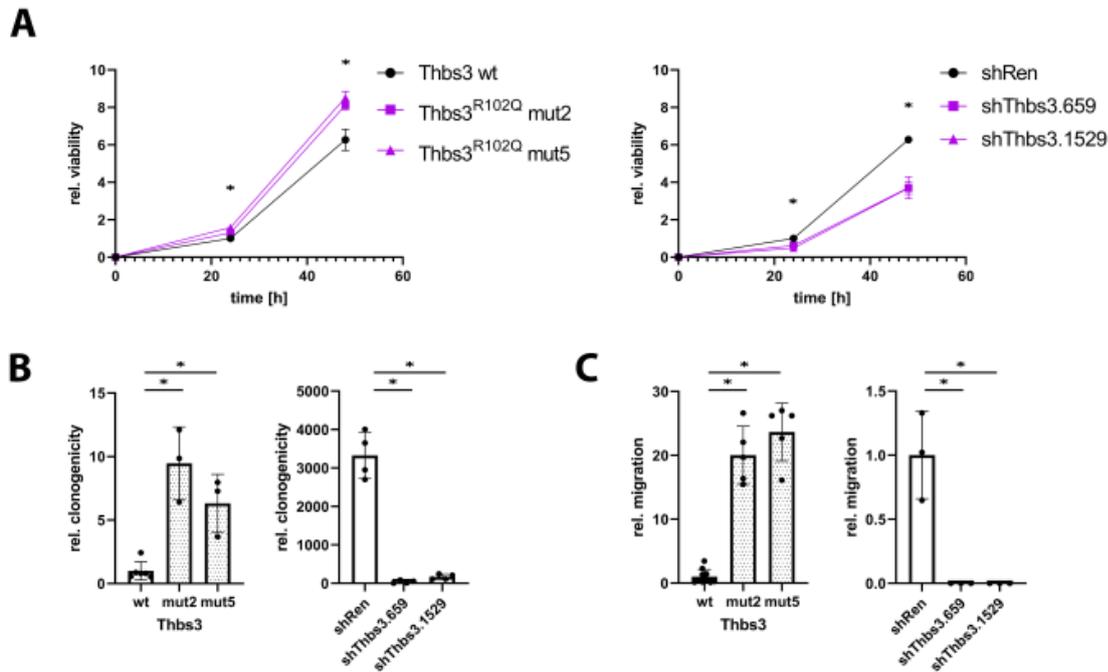
Fig. 10: Functional characterization of THBS3 in Kp19-/- cells.
A) Cell Titer Blue-based cell viability assay at different time points. The graph summarizes the data from 4 independent experiments. Statistical test: t-test, P ≤0.05 at each time point. **B)** Colony formation assay over the period of 8 d. Statistical test: t-test, P ≤0.05. **C)** Insert-based migration assay. Statistical test: t-test, P ≤0.05.

To characterize the effect of *THBS3* gene expression in isogenic murine HCC cells, I stably transfected AMp19$^{-/-}$ cells (derived from the CaMIA HCC mouse model and kindly provided by AG Zender, Tübingen) with *Thbs3* wildtype, *Thbs3^{R102Q}* mutant, or gene-specific shRNA (shThbs3.659, shThbs3.1529, shRenilla control). I observed stable overexpression of *Thbs3 wt* (192-fold) and the *Thbs3^{R102Q}* variant (83-fold) (Fig. 11C). In addition, I achieved shRNA-mediated knockdown of *Thbs3* expression (shThbs3.659: 3%, shThbs3.1529: 8%; each p ≤0.05, Fig. 11C) compared to the shRenilla control. Notably, the generation of a primary HCC cell line induced by AKT1 and Myc was only successful in the p19$^{-/-}$ mouse background, thus the *in vitro* model I used did not fully match the *in vivo* approach, in which the *p19* gene was expressed. As before, I analyzed the expression of cholangiocytic and hepatocytic marker genes at the mRNA level. I normalized all data to the respective controls.

I observed that the expression of mutated *Thbs3^{R102Q}* significantly decreased cholangiocytic marker expression in Amp19$^{-/-}$ cells (*Sox9*: 30%; *Krt7*: 68%; *Krt19*: 66% compared to control; each p ≤0.05, Fig. 11A). Furthermore, I saw that *Thbs3* knockdown also reduced *Sox9, Krt7*, and *Krt19* mRNA expression in shThbs3.659- and shThbs3.1529-expressing cells compared to shRenilla (*Sox9*: 46% & 51%; *Krt7*: 54% & 66%; *Krt19*: 45% & 42%, respectively; each p ≤0.05). Additionally, I could observe that the expression of the hepatocytic marker genes

Hnf4α and *Alb* was significantly increased in *Thbs3^{R102Q}*-expressing cells (*Hnf4α*: 7-fold; *Alb*: 5-fold induction compared to *Thbs3* wt, p ≤0.05, Fig. 11B) and following shRNA-mediated *Thbs3* knockdown (*Hnf4α*: 3-fold & 2-fold; *Alb*: 2-fold for both shThbs3.659 & shThbs3.1529, respectively; each p ≤0.05).

Regarding the functional experiments, I saw that clonogenicity was reduced by 44% (*Thbs3^{R102Q}* variant) and 72% & 70% (shThbs3.659 & shThbs3.1529) compared to the respective controls (each p ≤0.05, Fig. 11D). Furthermore, I observed that the expression of the *Thbs3^{R102Q}* variant reduced cell viability (56% at t = 24 h, 74% at t = 48 h, 76% at t = 72 h, and 63% at t = 96 h) compared to the control (each p ≤0.05). Similarly, I could see that the knockdown of *Thbs3* reduced cell viability (51% & 43% at t = 24 h, 33% & 43% at t = 48 h, 28% for both shThbs3.659 & shThbs3.1529 at t = 72 h, respectively; each p ≤0.05; Fig. 11E).
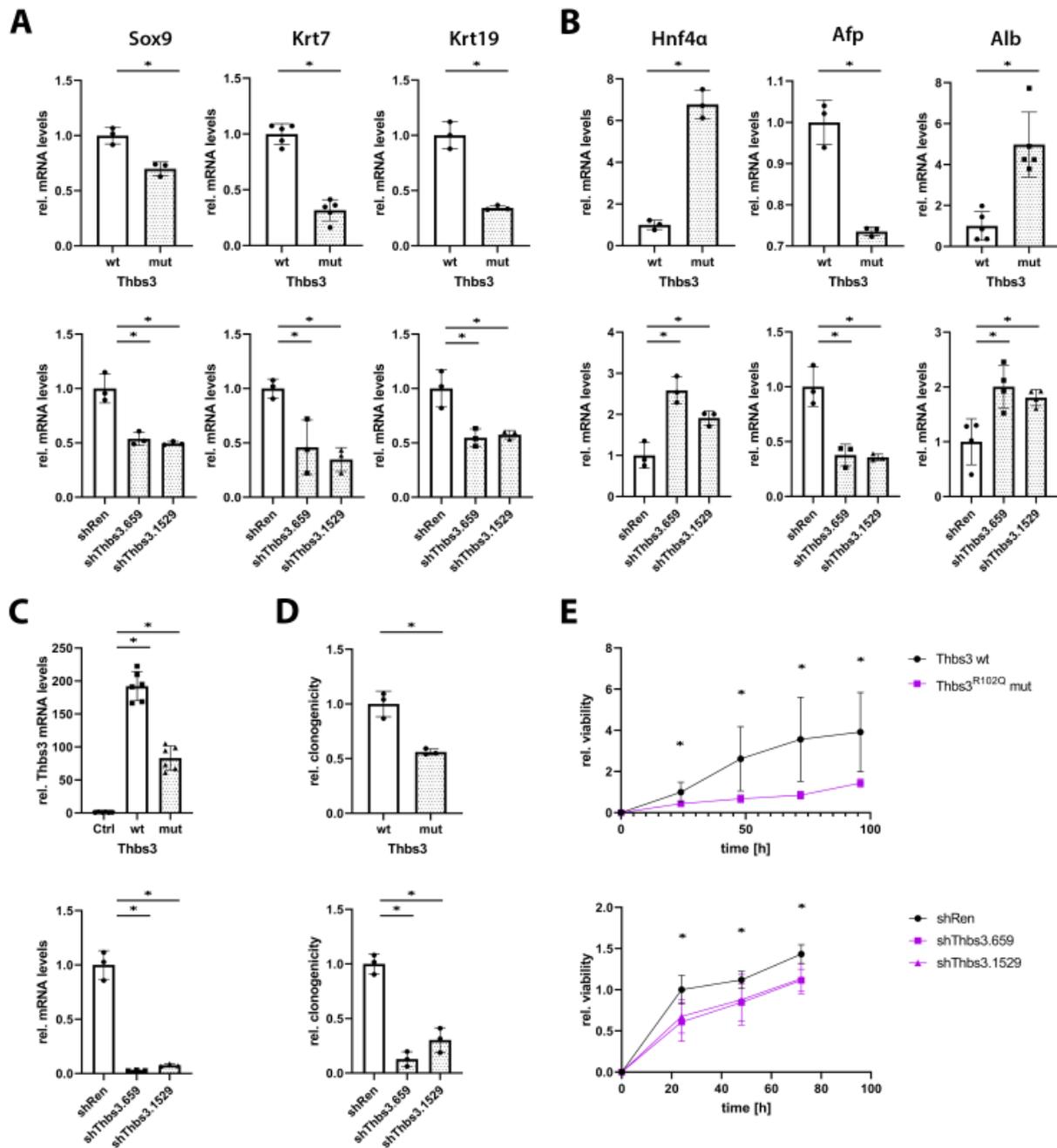
Fig. 11: Functional characterization of Thbs3 in AMp19-/- cells.
A) qPCR analysis of cholangiocytic marker expression at the mRNA level. Statistical test: t-test, $P \leq 0.05$. **B)** qPCR analysis of hepatocytic marker expression at the mRNA level. Statistical test: t-test, $P \leq 0.05$. **C)** *Thbs3* overexpression and knockdown were assessed at mRNA level by qPCR analysis. Statistical test: t-test, $P \leq 0.05$. **D)** Colony formation over a period of 8 d. Statistical test: t-test, $P \leq 0.05$. **E)** Cell Titer Blue-based cell viability assay at different time points. The graph summarizes results from 4 independent experiments. Statistical test: t-test, $P \leq 0.05$ at each time point.

Supporting the previous findings, I could also validate the phenotype-modulating function of Thbs3 in an HCC cell model *in vitro*. I observed that the knockdown of *Thbs3* expression completely phenocopied the functional effects observed after *Thbs3*[R102Q] overexpression. I furthermore noted that the loss of the Thbs3 wildtype function either through *Thbs3*[R102Q]

overexpression or *Thbs3* depletion reduced *Afp* expression compared to the respective control. Furthermore, I observed that both *Thbs3^{R102Q}* overexpression and knockdown of *Thbs3* expression decreased proliferation and clonogenicity compared to the respective control, again emphasizing a role of Thbs3 for cell survival.

4.3.3 *In vitro* validation in a human iCCA cell line

To compare the function of Thbs3 in the murine and the human system, I also analyzed the phenotype-modulating and tumor suppressive functions of the wildtype *THBS3* gene and its variant *THBS3^{R102Q}* in the human iCCA cell line HUH28. First, I generated the human *THBS3^{R102Q}* mutant by site-directed-mutagenesis. Subsequently, I generated HUH28 cells with stable gene expression. Then, I assessed stable overexpression of *THBS3* wt (70-fold) & *THBS3^{R102Q}* variant (74-fold) compared to the control in independent experiments (Fig. 12C). Analogously to before, I analyzed the cellular differentiation at mRNA level by qRT-PCR. I observed that the expression of *THBS3^{R102Q}* variant decreased the expression of the cholangiocytic markers *SOX9 (*59%*), KRT7 (*80%), and *KRT19 (*75%) and increased the expression of the hepatocytic markers *HNF4α (2-fold)* and *ALB* (7-fold) compared to *THBS3* wildtype expressing control cells (p ≤0.05, Fig. 12A&B). Both clonogenicity and cell viability were strongly increased compared to the wildtype control (p ≤0.05, Fig. 12D&E).

Thus, I confirmed the previous results in murine iCCA cells in the human iCCA model *in vitro*.
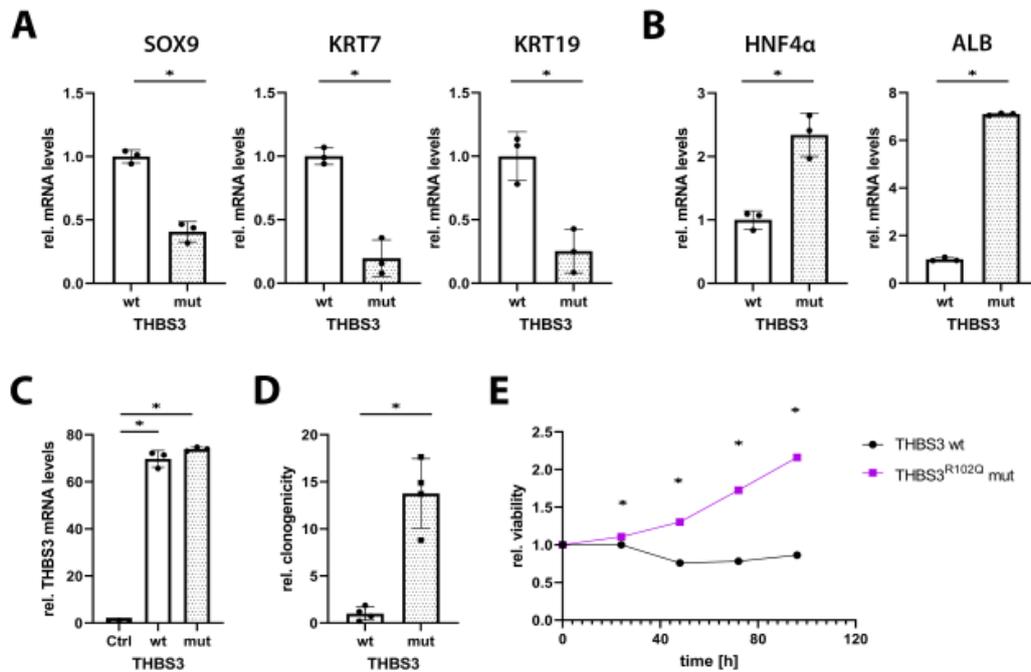
Fig. 12: *In vitro* validation of THBS3 functions in the human iCCA cell line HUH28.
A) qPCR analysis of cholangiocytic marker expression at the mRNA level. Statistical test: t-test, P ≤0.05. **B)** qPCR analysis of hepatocytic marker expression at the mRNA level. Statistical test: t-test, P ≤0.05. **C)** *THBS3* overexpression was confirmed by qPCR analysis. Statistical test: t-test, P ≤0.05. **D)** Colony formation assay over a period of 8 d. Statistical test: t-test, P ≤0.05. **E)** Cell Titer Blue-based cell viability assay at different time points. The graph summarizes results from 4 independent experiments. Statistical test: t-test, P ≤0.05 at each time point.

4.3.4 *THBS3* knockdown alters TGFβ signaling at transcriptional level

To identify the underlying mechanism, by which THBS3 affects the cellular phenotype of liver cancer cells, I analyzed isogenic *Thbs3* variant-expressing (*Thbs3* wt/*Thbs3*[R102Q] mutant) and *Thbs3* knockdown (shThbs3.659; shThbs3.1529; shRenilla control) cell lines by transcriptomic and ATAC sequencing and subsequent GSEA. NGS data processing and GSEA were kindly performed by Klaus Kluck and Iordanis Ourailidis (Institute of Pathology, Heidelberg).
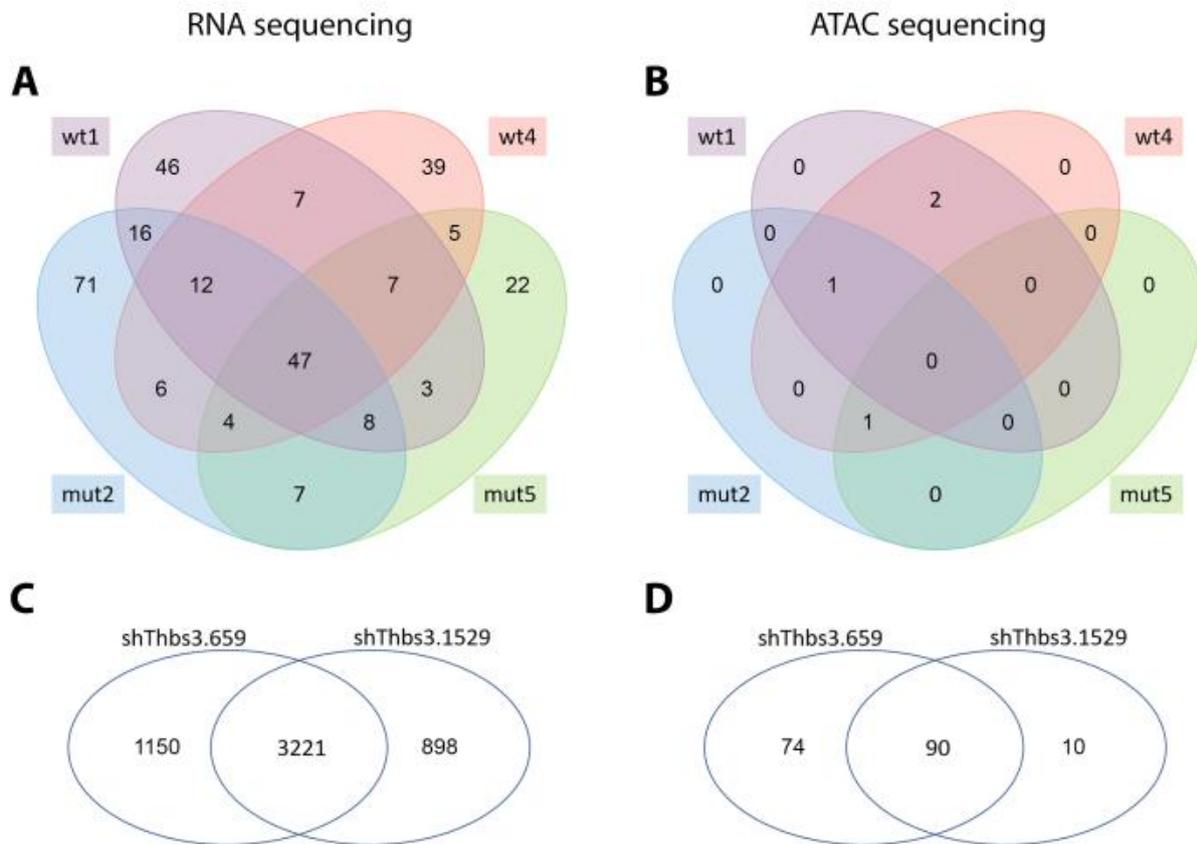
Fig. 13: NGS analyses of *Thbs3* knockdown and variant-expressing cell lines.
(A) Venn diagram depicting the number of differentially expressed genes between Thbs3 wt (wt1 & wt4) and mutant (mut2 & mut5)-clones compared to control. **(B)** Venn diagram showing the number of differentially accessible promoter sites between these clones and control. Statistical package DESeq2, FDR(BH) < 0.1. **(C)** Venn diagram showing the number of differentially expressed genes between the two Thbs3 knockdown cells clones (shThbs3.659 & shThbs3.1529) and control (shRenilla). **(D)** Venn diagram showing the number of differentially accessible promoter sites between these cells. Statistical package DESeq2, FDR(BH) < 0.1.

In Fig. 13, I present the results from transcriptomic (A&C) and ATAC (B&D) sequencing analysis. In both Thbs3 wildtype and in both Thbs3^{R102Q} mutant clones, I detected that each seven genes were differentially expressed compared to control, respectively (FDR(BH) < 0.1, Fig. 13A). I derived from this that wildtype Thbs3 signaling resulted in the differential expression of seven genes which were not differentially expressed upon loss of Thbs3 function by *Thbs3^{R102Q}* loss-of-function mutation. On the other hand, seven distinct genes were differentially expressed as a result of Thbs3^{R102Q} signaling. Furthermore, I observed two genes with differential chromatin accessibility at promoter sites in both wildtype clones compared to control (FDR(BH) < 0.1, Fig. 13B). Analogously, I derived from this that wildtype Thbs3 signaling resulted in differential chromatin accessibility at promoter sites for two genes. After integration of both datasets from ATAC and transcriptomic sequencing, I identified two merged hit genes that were both differentially expressed and were accessible at promoter sites based on

the combined data. First, I saw that the *Pmp22* gene showed reduced promoter accessibility and was downregulated in *Thbs3^{R102Q}*-expressing cells (FDR(BH) < 0.1). While I found that accessibility at promoter sites of the *Pmp22* gene was not unique to *Thbs3^{R102Q}*-expressing cell clones (mut2 & mut5), but was also exhibited by Thbs3 wildtype clone 4, I observed this characteristic becoming exclusive to *Thbs3^{R102Q}*-expressing cell clones when considering the integration of ATAC and transcriptomic sequencing data. I understood that this exclusivity arose because the gene expression of *Pmp22* was not significantly downregulated in Thbs3 wildtype clone 4. From this, I concluded that the loss of wildtype Thbs3 function led to a transcriptional decrease in a signaling pathway involving the gene *Pmp22*.

Second, I saw that the *Mpzl2* gene revealed reduced promoter accessibility and downregulation in Thbs3 wildtype cells (FDR(BH) < 0.1). Based on this, I inferred that wildtype Thbs3 signaling led to an augmentation in a signaling pathway that included *Mpzl2*. The GSEA of the integrated NGS datasets did not reveal any statistically significant gene set enrichment.

Regarding the Thbs3 knockdown cell lines, I detected 3221 differentially expressed genes (C) and 90 genes with differential chromatin accessibility at promoter sites (D) compared to control (FDR(BH) < 0.1) based on NGS analysis. After integration of both datasets, I found two genes that showed increased expression and increased chromatin accessibility at the gene promoter site and 55 genes with decreased expression and reduced chromatin accessibility at promoter sites (FDR(BH) < 0.1, Tab. 20). These genes were analyzed by GSEA against the M2 mouse collection of the Molecular Signatures Database. I retrieved four canonical pathways (FDR(BH) < 0.1, Fig. 14). Of the genes analyzed, I found *Fmo5, Igfbp6, Maoa, Pkhd1l1, Sh3tc2, Slc43a3, Trib2*, and *Upk1b* to be downregulated in murine high grade large cell neuroendocrine lung carcinoma compared to normal lung tissue (enrichment FC = 5.9). Furthermore, I saw that *Col6a1, Maoa, Mtus1, Pkhd1l1, Tbc1d2b*, and *Wt1* were downregulated (via TGFβ1R) by TGFβ1 in murine embryonic fibroblasts (enrichment FC = 9.6). Additionally, I could see that *Carns1, Foxred2, Medag*, and *Sulf1* were upregulated in cells expressing the fusion proteins MLL-AF4 and/or AF4-MLL (enrichment FC = 15.4). Lastly, I found that *Col6a1, Fbln1*, and *Fxyd5* were selectively expressed by meningeal cells in embryonic day 14.5 mouse telencephalons (enrichment FC = 29.0).

Tab. 20: Integrated NGS data analysis of Thbs3 knockdown cells

Gene	Gene expression and chromatin accessibility at promoter sites	FDR(BH)
Cyp2c68	Upregulated expression and increased accessibility	<0.1
D630003M21Rik	Upregulated expression and increased accessibility	<0.1
4930461G14Rik	Downregulated expression and decreased accessibility	<0.1
5730409E04Rik	Downregulated expression and decreased accessibility	<0.1
Ak1	Downregulated expression and decreased accessibility	<0.1
Akt2	Downregulated expression and decreased accessibility	<0.1
Arhgap32	Downregulated expression and decreased accessibility	<0.1
C9	Downregulated expression and decreased accessibility	<0.1
Carns1	Downregulated expression and decreased accessibility	<0.1
Cdkn2a	Downregulated expression and decreased accessibility	<0.1
Cfap45	Downregulated expression and decreased accessibility	<0.1
Col6a1	Downregulated expression and decreased accessibility	<0.1
Dclk1	Downregulated expression and decreased accessibility	<0.1
Ddr1	Downregulated expression and decreased accessibility	<0.1
Efemp1	Downregulated expression and decreased accessibility	<0.1
Ephx2	Downregulated expression and decreased accessibility	<0.1
Fbln1	Downregulated expression and decreased accessibility	<0.1
Fcgrt	Downregulated expression and decreased accessibility	<0.1
Fetub	Downregulated expression and decreased accessibility	<0.1
Fggy	Downregulated expression and decreased accessibility	<0.1
Fmo5	Downregulated expression and decreased accessibility	<0.1
Foxred2	Downregulated expression and decreased accessibility	<0.1
Fxyd5	Downregulated expression and decreased accessibility	<0.1
Gm13293	Downregulated expression and decreased accessibility	<0.1
Hoxaas2	Downregulated expression and decreased accessibility	<0.1
Igfbp6	Downregulated expression and decreased accessibility	<0.1
Ipo4	Downregulated expression and decreased accessibility	<0.1
Klk10	Downregulated expression and decreased accessibility	<0.1

Gene	Gene expression and chromatin accessibility at promoter sites	FDR(BH)
Lbp	Downregulated expression and decreased accessibility	<0.1
Maoa	Downregulated expression and decreased accessibility	<0.1
Medag	Downregulated expression and decreased accessibility	<0.1
Mtus1	Downregulated expression and decreased accessibility	<0.1
Mxra8	Downregulated expression and decreased accessibility	<0.1
Nfe2l3	Downregulated expression and decreased accessibility	<0.1
Opcml	Downregulated expression and decreased accessibility	<0.1
Pdk4	Downregulated expression and decreased accessibility	<0.1
Pkhd1l1	Downregulated expression and decreased accessibility	<0.1
Plcl1	Downregulated expression and decreased accessibility	<0.1
Pstpip1	Downregulated expression and decreased accessibility	<0.1
Ptgs2os2	Downregulated expression and decreased accessibility	<0.1
Rorc	Downregulated expression and decreased accessibility	<0.1
Sardh	Downregulated expression and decreased accessibility	<0.1
Sbsn	Downregulated expression and decreased accessibility	<0.1
Sh3rf2	Downregulated expression and decreased accessibility	<0.1
Sh3tc2	Downregulated expression and decreased accessibility	<0.1
Slc17a9	Downregulated expression and decreased accessibility	<0.1
Slc43a3	Downregulated expression and decreased accessibility	<0.1
Slc9b2	Downregulated expression and decreased accessibility	<0.1
Sorbs3	Downregulated expression and decreased accessibility	<0.1
Srgap3	Downregulated expression and decreased accessibility	<0.1
Sulf1	Downregulated expression and less chromatin accessibility	<0.1
Tbc1d2b	Downregulated expression and less chromatin accessibility	<0.1
Trib2	Downregulated expression and less chromatin accessibility	<0.1
Upk1b	Downregulated expression and less chromatin accessibility	<0.1
Usp54	Downregulated expression and less chromatin accessibility	<0.1
Wt1	Downregulated expression and less chromatin accessibility	<0.1
Zfp760	Downregulated expression and less chromatin accessibility	<0.1
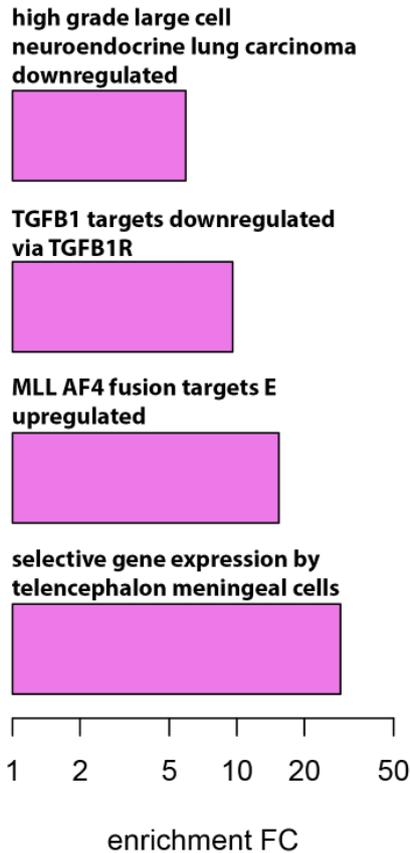
Fig. 14: GSEA of integrated NGS datasets.
The differentially expressed gene set detected upon knockdown of *Thbs3* expression shows and enrichment in four pathways. GSEA was performed against the M2 (curated) gene set from {HYPERLINK "https://www.gsea-msigdb.org/gsea/msigdb/mouse/collections.jsp"}

4.4 Identification of phenotype-driving gain-of-function candidates

In contrast to phenotype-driving loss-of-function candidates, which can be detected by RNAi screening, gain-of-function candidates will be missed by such an experimental design. Therefore, I used *in silico* analysis to select potential oncogenic candidates.

First, I used the bioinformatics software IPA. By inserting expression data, I could retrieve insights into potential biological functions and disease associations. Second, the computational tool OpenCRAVAT was kindly employed by Dr. Aurelie Tomczak (Institute of Pathology, Heidelberg University Hospital).) for variant analysis. On this basis, genetic variants and their potential impact on protein function and disease associations was analyzed. Lastly, I incorporated TCGA data into the analysis to examine correlations between genetic alterations, gene expression patterns, and clinical characteristics. I combined the complex biological information from the different tools and selected candidate genes based on their predicted variant consequences, and their implications in disease, particularly in cancer.

Using this approach, I considered *ADRA1A, ANKRD1, ASB4, DTX1, FOXJ1, GATA1, GLI1, GLIS1, KLF5, MACC1, TMPRSS4*, and *VTCN1* as potential gain-of-function candidates. As

concerns the candidates detected as differentially mutated between the HCC and the iCCA component of cHCC-CCA, I introduced the synonymous murine mutations by site-directed mutagenesis ($ADRA1A^{I200S}$, $TMPRSS4^{P413L}$, $VTCN1^{V120L}$). Then, I cloned the twelve wildtype genes and their variants into transposon vectors which were co-delivered by HDTVI in the suitable mouse model reflecting the human constellation in which they were detected. As *DTX1*, *GLIS1* and *TMPRSS4* were upregulated in the human HCC compartment, these were injected into the iCCA mouse model (Tab. 18 & 19). The other nine genes were injected into the HCC mouse model, as they were upregulated in the iCCA compartment of cHCC-CCA. As in the *in vivo* RNAi experiments, I sacrificed mice after tumor development and dissected 219 individual nodules for subsequent histological analysis for tumor typing. Positive readouts were observed for *DTX1* and *TMPRSS4* in the iCCA model. Expression of *Dtx1* resulted in an increased cHCC-CCA frequency (20%, n = 2/10 with *Dtx1*) compared to the control (11%, n = 2/19). Expression of mutated $Tmprss4^{P413L}$ resulted in a higher frequency of cHCC-CCA development compared to the expression of the wildtype *Tmprss4* gene ($TMPRSS4^{P413L}$: n = 3/12 vs, *TMPRSS4* wt: n = 0/7).

In summary, I validated two (*DTX1* and *TMPRSS4*) out of twelve potential gain-of-function candidates (*ADRA1A, ANKRD1, ASB4, DTX1, FOXJ1, GATA1, GLI1, GLIS1, KLF5, MACC1, TMPRSS4*, and *VTCN1*) *in vivo* which require further evaluation.

5 Discussion

5.1 The cellular origins of PLC

According to text book knowledge, it is believed that BEC are the origin of iCCA [237], but there is compelling experimental evidence from murine liver cancer models that the different epithelial cell lineages present in the liver may be the source of either HCC, or iCCA, and/or cHCC-CCA, respectively [138, 238-240]. Additionally, there are data from human patients pointing into the same direction. For instance, both HBV and HCV infect hepatocytes but not BEC. Nevertheless, (besides the associated, quantitatively dominant hepatocarcinogenesis) iCCA may develop as a result of chronic viral hepatitis, thus highlighting the cellular plasticity of adult human hepatocytes [143]. Transdifferentiation processes and subsequent cancer development have been shown in other organs as well. Pancreatic ductal adenocarcinoma (PDAC), for example, was first believed to originate from ductal pancreatic cells, but this view was later challenged by the finding that acinar cells may transdifferentiate into ductal cells and may form preneoplastic lesions that eventually progress to PDAC [241, 242].

Furthermore, a set of identical mutations has been observed in the two components of cHCC-CCA and certain genetic alterations have been detected in HCC with progenitor-like features and iCCA, respectively, again supporting the concept that the cell of origin can be within the same lineage in both cancer types [243-250]. Interestingly, a recent study showed that the individual components of cHCC-CCA do neither cluster with HCC nor iCCA, respectively, but instead form their own cluster. Thus, both components of cHCC-CCA are more similar to each other than to other tumors of the same histomorphology [128].

Other studies suggested that cHCC-CCA may originate from LPC, as these mixed tumors were found to express stem cell markers like c-Kit, CD133, EpCam, SALL4, or Nestin [128, 129, 251-255]. LPC are located in the stem cell niche within the canals of Hering, which represent the connection between the bile canaliculi of the liver acinus and the draining portal ductular system. Their bipotential nature allows them to undergo differentiation into both hepatocytic and cholangiocytic lineages. In the present study, only classical cHCC-CCA cases were included. Thus, the human tumors used to select candidate genes did not express stem cell markers. However, the experimental approach, which I used in the murine system, did not principally exclude the option of an LPC-derived tumorigenesis.

I detected 16 common mutations between the two compartments of human cHCC-CCA. Of these, I found *ABCA7, ADRA1A, DNAH2, ORC1, RAVER1, SLC35F2,* and *VTCN1* to be upregulated in the iCCA component and *GLIS1, KCNK13, MBL2, RILP, TMPRSS4, TRIP10, TUSC5*, and *ZC3H12D* to be upregulated in the HCC component. Based on my analysis, I inferred that these biphenotypic cancers may have originated from the same cell of origin. However, the initiating cell type remains unclear. Notably, I also observed that the *MUC4* gene can be mutated and upregulated in either the HCC or the CCA compartment (4.1, Tab. 19). Thus, *MUC4* mutations might exert a protumorigenic function that facilitates cancer initiation and/or progression *per se* without affecting the cellular differentiation. However, it could also be that different mutations (affecting different protein domains) may have different functional effects depending on the affected domain, the given cellular state and the surrounding environment.

5.2 Candidates modulating cellular differentiation in liver cancer

5.2.1 Loss-of-function candidates from pooled RNAi screen

Pooled RNAi technology *in vivo* is a very useful tool for finding loss-of-function candidates because thousands of genes can be screened simultaneously [256, 257]. By silencing the expression of specific genes, their loss-of-function effects can be identified [258]. Thus, the applied RNAi screen in transposon-based mosaic mice allowed to screen for the loss-of-function of all 54 candidate genes in parallel. Histological analysis of the individual tumor nodules developing in these mosaic mice showed a phenotypic switch in 17 nodules. Four potential loss-of-function candidate genes were identified by sequencing-based identification of the expressed shRNA sequences in the respective nodules, namely *ASB15, FOSL1, ORC1*, and *THBS3*. A subsequent single gene RNAi validation experiment *in vivo*, however, excluded *ASB15*, *FOSL1*, and *ORC1* as potential phenotype-modulating loss-of-function candidate genes because the phenotypic switch observed in the pooled screen could not be reproduced at the single gene level. *THBS3* will be discussed in more detail below (paragraph 5.3).

5.2.2 Gain-of-function candidates from *in silico* analysis in liver cancer

In order to identify potential gain-of-function candidates that may impact cellular differentiation, the remaining 50 candidate genes that were not considered as potential loss-of-

function candidates were assessed by *in silico* analysis. Of the genes that were differentially expressed between the two components of human cHCC-CCA, based on a combination of pathway analysis (IPA) and clinical characteristics attributable to cellular differentiation retrieved from TCGA data, candidates with potential biological functions and disease associations connected to cellular differentiation were selected. Using this approach and their upregulation in the respective tumor compartment, I considered *ANKRD1, ASB4, FOXJ1, GATA1, GLI1, KLF5,* and *MACC1* as candidate genes potentially promoting an iCCA phenotype and *DTX1* as a candidate gene potentially promoting an HCC phenotype.

Besides changes in gene expression, changes in the DNA sequence may eventually enhance the activity or function of a gene respectively its coded protein, thus constituting the concept of gain-of-function mutations [259]. One example is the proto-oncogene *KRAS* which encodes a GTPase involved in cell signaling regulating cell growth, differentiation, and apoptosis [260]. Gain-of-function mutations in *KRAS* are the most frequent genetic alterations found in human cancers, including iCCA [90, 260]. The result of *KRAS* gain-of-function mutations is constitutive activation of the KRAS protein that then persistently stimulates downstream signaling pathways, including cell proliferation and survival, thereby leading to tumorigenesis [90].

Considering this concept, candidate genes that were exclusively mutated in one of the compartments of human cHCC-CCA were assessed by *in silico* analysis. Genetic variants and their potential impact on protein function and disease associations regarding cellular differentiation were selected based on variant analysis (OpenCRAVAT). Then, clinical characteristics retrieved from TCGA data of genetic alterations of candidate genes attributable to cellular differentiation were incorporated. *ADRA1A, GLIS1, TMPRSS4*, and *VTCN1* were identified as additional potential gain-of-function candidates. *ADRA1A*^{I200S} and *VTCN1*^{V120L} variants were detected in the HCC component of human cHCC-CCA, whereas *GLIS1* (UTR_5' at position 5239) and *TMPRSS4*^{P413L} variants were detected in the iCCA component, respectively. Therefore, the presumption was that *ADRA1A* and *VTCN1* variants had gained a function promoting HCC differentiation, explaining their restriction to the HCC component. Analogously, the respective *GLIS1* and *TMPRSS4* variants may have gained a function, which either promotes cholangiocytic or blocks hepatocellular differentiation.

From the twelve candidate genes, namely *ADRA1A, ANKRD1, ASB4, DTX1, FOXJ1, GATA1, GLI1, GLIS1, KLF5, MACC1, TMPRSS4*, and *VTCN1*, initially considered as potential gain-of-

function variants, only the expression of *DTX1* and *TMPRSS4* increased the rate of cHCC-CCA formation in the iCCA model. Thus, *DTX1* upregulation and the *TMPRSS4^{P413L}* variant were validated as phenotypic driver genes promoting hepatocytic differentiation *in vivo*.

In the healthy liver, lineage commitment of LPC was shown to rely on a positive feedback loop involving EGFR/NOTCH1 signaling for the cholangiocytic and MET signaling for the hepatocytic lineage [261]. DTX1 is a regulator of the Notch pathway and was shown to play a role in cancer stem cell-mediated tumor formation in human hepatoma cell lines (Sphere-forming cell subpopulations with cancer stem cell properties in human hepatoma cell lines). However, my study represents the first evaluation of Dtx1 in a liver cancer model *in vivo*. Based on my findings that *Dtx1* expression nearly doubled the rate of cHCC-CCA formation compared to control, -*Dtx1* seems to be a promising gain-of-function candidate in liver cancer. Therefore, further studies are warranted characterizing the role of DTX1 in liver cancer and liver cell differentiation. Of note, *DTX1* was found upregulated in other cancer entities such as thymic tumors and glioblastoma [262, 263], while reduced *DTX1* levels were shown to promote proliferation and migration of breast cancer cells [264] suggesting that the function of DTX1 may vary between different cell types.

TMPRSS4, the other gain-of-function candidate that I could validate *in vivo*, is a positive regulator of the Raf/MEK/ERK1/2 pathway. In line with its potential gain-of-function role, TMPRSS4 has also been negatively correlated to patient outcome by promoting EMT and angiogenesis and is suggested as a prognostic marker in HCC [265]. Thus, TMPRSS4 may also be an interesting candidate for further functional characterization as a driver gene in liver cancer development.

5.3 Molecular and functional characterization of THBS3 in cHCC-CCA

5.3.1 Impact of THBS3 on cellular differentiation of PLC

I detected the *THBS3^{R102Q}* variant in the HCC compartment of human cHCC-CCA, and knockdown of *Thbs3* scored in the iCCA screening model *in vivo* leading to a more HCC-like tumor phenotype. Although the frequency of cHCC-CCA formation was higher in murine liver cancer nodules induced by *Thbs3^{R102Q}* expression, the single gene RNAi validation experiment did not confirm THBS3 as a phenotypic driver *in vivo* at the statistical level. However, the results of the *in vivo* validation experiments may be biased for various reasons. First, unintentional mouse loss in the *Thbs3* knockdown groups, which will be discussed in more

detail below (paragraph 5.5), reduced the number of available tumor nodules for histological analysis. Second, the use of two vector systems in parallel (instead of a single vector, in which all genes required for the model were coded in series) affected the identification of informative tumor nodules expressing the gene of interest. While the vector used for overexpression of mutant $KRas^{G12V}$ carried a GFP reporter, the vector carrying either wildtype $Thbs3$ or variant $Thbs3^{R102Q}$ lacked a comparable marker gene. Therefore, their expression could not be directly verified through marker gene detection. To address this issue, it was necessary to confirm mRNA expression of either wildtype $Thbs3$ or mutant $Thbs3^{R102Q}$ in the respective nodules from freshly frozen tissue. Future experiments should address this obstacle by either establishing a single-vector approach, as was done with the shRNA vectors, where the respective shRNA was cloned into the backbone carrying the oncogene for induction of tumorigenesis, or by adding a marker gene to the vector used for overexpression of the gene-of-interest in order to monitor gene expression.

Nevertheless, when I expressed the loss-of-function variant $Thbs3^{R102Q}$ in the hepatocyte-derived mutant $KRas^{G12V}$-induced iCCA cells *in vitro*, I observed a more HCC-like phenotype arguing for a phenotype-switching function of $THBS3^{R102Q}$. In particular, I observed that the loss of Thbs3 function in the *in vitro* iCCA model resulted in a loss of biliary differentiation and a phenotypic switch towards a more hepatocytic differentiation as suggested by decreased $Krt19$ and increased $Hnf4\alpha$ mRNA expression. In conclusion, $Thbs3^{R102Q}$ seems to be a loss-of-function variant promoting an HCC-like phenotype of cholangiocarcinoma cells.

Other studies have also implied that THBS family members play a role in cellular plasticity, e. g. Zhang *et al.* found that all THBS family members show strong activation of EMT through pan-cancer analysis [209]. Furthermore, they showed that *THBS2* promoted EMT and tumor metastasis in colorectal and gastric cancer cells. Based on correlation analyses using TCGA data, they further suggest that the THBS family might function in tumorigenesis together because they could significantly positively correlate *THBS2* with *THBS1, THBS3, THBS4*, and *THBS5* in colon cancer and with *THBS1, THBS3*, and *THBS4* in gastric cancer [209]. Another study showed that increased *THBS1* promotes colorectal cancer liver metastasis by increasing EMT [266]. Since it is probable that the other THBS genes such as *THBS3, THBS4*, and *THBS5*, also play a role in regulating cancer cell EMT, it is necessary to conduct additional experiments to address their involvement in cellular plasticity pathways.

5.3.2 The tumor suppressive function of THBS3

The expression of the wildtype *Thbs3* gene decreased the formation of tumor nodules in the iCCA model *in vivo* suggesting that the wildtype *THBS3* gene may act as a tumor suppressor gene in iCCA. In line with this, both shRNA-mediated *Thbs3* knockdown cell lines increased the frequency of iCCA formation compared to control cells.

Supporting this study's finding, Zhang *et al.* found that low *THBS3* expression was associated with shorter overall patient survival in HCC, esophageal carcinoma, and pancreatic adenocarcinoma. Furthermore, they showed that *THBS3* was downregulated in 13 kinds of cancers, including colon adenocarcinoma, lung squamous carcinoma and lung adenocarcinoma [209]. Downregulation of other THBS family members was also correlated to shorter patient survival in cancer, including lower *THBS4* expression in cervical cancer, thymus cancer, and tenosynovial giant cell tumors and lower *THBS5* expression in lung adenocarcinoma [209]. Furthermore, *THBS1* was shown to inhibit tumor growth, metastatic capacities, and angiogenesis in human breast carcinoma cell lines [267]. There are many studies, however, that also emphasize a tumor-promoting role of members of the THBS family. A recent study, e. g. found that *THBS3* expression was increased in clear cell renal cell carcinoma in patients and associated to poorer overall survival [268]. This emphasizes the complex and multifaceted roles of members of the THBS family and calls for further assessment of their roles in different kinds of tumor entities.

5.4 Potential implications of targeting THBS3 in liver cancer

PLC comprises a group of aggressive cancer entities (HCC, iCCA, cHCC-CCA) with increasing frequency around the world [1, 2]. Among these, cHCC-CCA is rare (0.4%-14.2%) but due to its biphenotypical differentiation very interesting for the evaluation of mechanisms determining the tumor phenotype and its associated biological functions [125-127]. Elucidating the initiation processes behind the development of cHCC-CCA would be of great medical value. As mentioned above, based on shared mutations it was determined that both areas of these tumors share a clonal origin [128]. Given that iCCA generally presents a more aggressive tumor phenotype than HCC [128], characterized by late-stage detection and limited therapeutic options [56], understanding the factors driving lineage commitment in cHCC-CCA becomes crucial. By elucidating the decisive factors for lineage commitment in cHCC-CCA, it may be possible to target specific aspects of signaling pathways associated with cHCC-CCA, aiming

to induce a slower, more mature, and more treatable HCC-like phenotype. Ultimately, this approach would contribute to the development of customized, individualized, targeted molecular therapies.

While I was able to show in this study that disabling Thbs3 function induced a more HCC-like phenotype in the iCCA model *in vitro*, the proliferation and clonogenicity was increased in these cells suggesting a more aggressive tumor phenotype. However, although seeming counterintuitive, in some cases, increased proliferation of tumor cells has led to a better patient outcome due to a better response to chemotherapy [269, 270]. Essentially, rapidly dividing cells can be more susceptible to the effects of these therapies [271]. As a result, a higher proliferation rate might correlate with a better response to treatment and, consequently, a more favorable patient outcome [272]. However, this concept is not universally applicable and varies depending on the specific cancer type, its genetic characteristics, the microenvironment, and the treatment approach. In many cases, increased tumor cell proliferation is associated with aggressiveness and poorer prognosis, as it can lead to faster disease progression and increased resistance to treatment [273]. Thus, the proliferation rate of tumor cells needs to be interpreted in the given clinical and tumor biological context. Apoptosis represents another example that can be viewed as a process with dual outcomes: firstly, as a suppressor of tumors by eliminating malignant or pre-malignant cells, and secondly, as a promoter of tumorigenesis by triggering reparative and regenerative reactions within the tumor microenvironment [274]. In conclusion, the faster proliferating more HCC-like phenotype induced by $Thbs3^{R102Q}$ expression *in vitro* needs to be evaluated in context of treatment sensitivity to be able to draw potentially clinically relevant conclusions.

As my data suggest that Thbs3 alters the TGFβ pathway at the transcriptional level, it would be interesting to analyze whether targeting TGFβ signaling may block protumorigenic functions induced by the loss of Thbs3 function. TGFβ signaling is involved in liver development, tissue regeneration, immune responses, fibrosis, but also hepatocellular carcinoma progression [109, 275]. It can be characterized by early and late expression signatures representing different functional outcomes in response to TGFβ signaling activation in the liver [276]. While the early signature is represented by genes that are rapidly and transiently activated or repressed in response to TGFβ signaling, the late signature includes genes that show sustained or delayed changes in expression [277]. The former encompasses genes involved in immediate molecular events, such as receptor activation, phosphorylation of downstream effectors (e.g., SMAD proteins), and their translocation to the nucleus. The latter, however, includes genes involved

in cellular and physiological processes requiring a longer period of time to be executed, such as extracellular matrix production, epithelial-mesenchymal transition, immune modulation, and tissue remodeling [277]. Regarding this study, phenotypic changes of cells and tumor tissue are likely represented by late signature genes, which have been implicated in liver cancer progression [278, 279]. Interestingly, the TGFβ signaling pathway has also been connected to THBS1, another THBS family member [280], arguing for a more general function relationship between THBS family members and TGFβ signaling.

5.5 Challenges and limitations of the experimental approach

The *in vivo* experiments performed in this study relied on the histological analysis of tumor nodules, in which a positive readout was defined as a phenotypic switch in tumor morphology compared to the basic model system.

While I observed a trend suggesting that Thbs3^{R102Q} promotes a more HCC-like tumor phenotype, the reliability of the *in vivo* results was biased by the missing information whether *Thbs3^{R102Q}* was indeed expressed in the tumors analyzed by histology. In addition, the tumor suppressive effects of the wildtype *Thbs3* gene negatively impacted on the size of the control group [281, 282]. Considering the experimental setup, which relied on the histological analysis of individual tumor nodules, a reliable evaluation depends on a sufficient number of informative tumor nodules to reach statistical significance. In the case of a powerful tumor suppressor, it might be unattainable to generate a sufficient number of tumor nodules for this purpose, even though the observation itself could be highly meaningful [281, 283].

Furthermore, there were instances of unintentional mouse loss in the *Thbs3* knockdown groups. Although the reasons remain unknown, potential attributable factors may include experimental errors, technical issues, or unanticipated variables [284]. The occurrence of unexpected mouse loss during experiments posed a challenge to this study as it compromised the integrity of the collected data compared to the study plan, thereby potentially leading to a bias in terms of data reliability and interpretability [281]. Consequently, this hindered me in drawing precise final conclusions, as the tumor phenotype data were normalized per mouse in a given experimental group to allow any informative statistical evaluation.

5.6 Conclusions and future prospective

The purpose of this dissertation was to identify genes that would help to shed light on the mechanism leading to the development of biphenotypical liver tumors. Based on vertical integration of human and murine liver cancer data, I identified *THBS3* as a likely plasticity driving gene with its variant p.R102Q promoting a more HCC-like phenotype in cholangiocarcinoma cells. Furthermore, overexpression of the wildtype *Thbs3* impaired murine liver tumor development compared to control suggesting that *THBS3* has tumor suppressive functions.

Unfortunately, the unintentional loss of mice challenged this study, not only due to ethical and financial concerns but particularly due to time constraints. Within the available funding period and timeframe, it became practically unfeasible for me to repeat the experiment. This lack of sufficient informative samples potentially impeded on the scientific progress in this specific area of my study.

Time limitations also limited the characterization of the Thbs3 function. The elucidation of the intricate molecular mechanisms underlying THBS3's (loss-of-) function in liver cancer is imperative to gain a comprehensive understanding of its role. Unraveling the precise downstream targets and molecular interactions associated with the action of THBS3 variants will provide valuable insights into the complex signaling networks driving lineage commitment in primary liver cancer. Future experiments could aim at determining whether the cellular function of THBS3 and the potentially associated cellular differentiation of liver cancer cells truly hinge on the interaction with the TGFβ signaling pathway.

Furthermore, integrating THBS3's functionality with other genetic mutations and molecular alterations commonly observed in primary liver cancer could yield a more holistic perspective on the disease. Exploring potential crosstalk between *THBS3* and other genes, such as oncogenes, holds the potential to unveil novel therapeutic targets and predictive markers for personalized treatment strategies.

Subsequently, it seems crucial to investigate the prevalence and clinical implications of *THBS3* alterations in primary liver cancer. By assessing *THBS3* expression levels and mutation status in patient samples and correlating them with relevant clinical parameters such as tumor type, proliferation status, prognosis, and treatment response, valuable prognostic information may be

obtained. Such knowledge may facilitate the identification of patients who are more likely to benefit from precision oncology therapies targeting either THBS3 or its downstream effectors.

In summary, the identification of THBS3 as a plasticity driver and tumor suppressor in primary liver cancer emphasizes its significance as a potential therapeutic target and underscores the need for in-depth exploration of its molecular functions. By characterizing the precise mechanism by which THBS3 affects the lineage of liver cancer cells, including the decipherment of its intricate interactions with other molecular components, we can expand our understanding of the disease and pave the way for innovative and effective therapeutic approaches for individuals affected by primary liver cancer.

6 Acknowledgements

Very importantly, I thank my childhood friends among others Annabelle Eilingsfeld, Delphina Hennig and Negar Sharafi from my home city of Frankfurt/Main for their lifelong companionship, for all we have been through, especially puberty, for all the laughing, singing, playing, partying, travelling, crying, hangovers, sleepovers, makeovers, all the memories we have made so far and the ones to come.

Special heartfelt thanks and huge gratitude goes to my BFF Nasli Malek for truly feeling my train of emotions and rescuing me whenever my train derails.

Also, enormous thanks to Dr. Adrian Kovač whose love and friendship are truly unconditional. In addition to his infinitely valuable friendship, I thank him for his professional help during my studies.

Many thanks to my Hamburg pearls Johanna von Berlepsch, Andreas Blusch and Olivia Drożdż.

In beautiful Hamburg, we spent THE best years together. Times were crazy! Special heartfelt thanks goes to Juju for being an amazing backpack travel partner. Thank you so much for sharing countless beautiful and some strange experiences during our trips.

Lastly, and importantly, a big thank you to my "new" friends in Heidelberg.

Very special thanks goes to Dr. Thorben Huth, Dr. Fabian Rose, Dr. Lena Thiess and Dr. Lilija Wehling for all the exciting and amazing fun (sporty) activities we experienced and times we spent together (partying). Looking forward to our next super fun challenge!

Furthermore, I would like to thank Eva Eiteneuer, Damaris Greule, Aslihan Inal and Ariane Neumann, for their everyday friendship lightening up lab life for me. Without you, the last 4 years at work would have been less colorful! Many thanks to all other lab members for providing a pleasant and helpful environment. A big thanks to Alphonse Charbel for his warm friendship and our amazing vibe.

Also, a big thank you to my gym friends Houda, Jelena, Michi, and Vlado and to my inspiring fitness coach Nicole, for keeping me not only physically but mentally balanced for the last years by sharing hard workouts and rewarding wellness. Very special thanks to Melanie Doley for her truly valuable friendship and great sense of humor.

Last thank you goes to my great flat mates Adrian, Dominik, Fanny and Vani for our fun, pleasant and comfy WG that I always felt happy to come home to. Goodbye Heidelberg.

Appendix

A   Complete shRNA library sequences

Ren.713	TGCTGTTGACAGTGAGCGCAGGAATTATAATGCTTATCTATAGTGAAGCCAC AGATGTATAGATAAGCATTATAATTCCTATGCCTACTGCCTCGGA
Abca7.1372	TGCTGTTGACAGTGAGCGAACCTACGATTCAAGATTCGAATAGTGAAGCCAC AGATGTATTCGAATCTTGAATCGTAGGTGTGCCTACTGCCTCGGA
Abca7.2594	TGCTGTTGACAGTGAGCGCGGGCCATGATGTACAAACCAATAGTGAAGCCA CAGATGTATTGGTTTGTACATCATGGCCCATGCCTACTGCCTCGGA
Abca7.5089	TGCTGTTGACAGTGAGCGCTCCTGAAACAAGTGTTTCTTATAGTGAAGCCAC AGATGTATAAGAAACACTTGTTTCAGGATTGCCTACTGCCTCGGA
Abca7.5090	TGCTGTTGACAGTGAGCGCCCTGAAACAAGTGTTTCTTATTAGTGAAGCCAC AGATGTAATAAGAAACACTTGTTTCAGGATGCCTACTGCCTCGGA
Abca7.6515	TGCTGTTGACAGTGAGCGAACAGAGGATCTCTGTACCATATAGTGAAGCCAC AGATGTATATGGTACAGAGATCCTCTGTCTGCCTACTGCCTCGGA
Abca7.989	TGCTGTTGACAGTGAGCGACCAGATCTTCAACTTCATGAATAGTGAAGCCAC AGATGTATTCATGAAGTTGAAGATCTGGGTGCCTACTGCCTCGGA
Slc35f2.1683	TGCTGTTGACAGTGAGCGATGGGATGTTGGTGCAATAAGATAGTGAAGCCAC AGATGTATCTTATTGCACCAACATCCCAGTGCCTACTGCCTCGGA
Slc35f2.2143	TGCTGTTGACAGTGAGCGCAGCCACAGTTTAGAAAATTAATAGTGAAGCCAC AGATGTATTAATTTTCTAAACTGTGGCTTTGCCTACTGCCTCGGA
Slc35f2.2485	TGCTGTTGACAGTGAGCGCATGTAAATTGTTTTTGTATAATAGTGAAGCCAC AGATGTATTATACAAAAACAATTTACATATGCCTACTGCCTCGGA
Slc35f2.2525	TGCTGTTGACAGTGAGCGAGGTAAGTGTTCGTATTATTAATAGTGAAGCCAC AGATGTATTAATAATACGAACACTTACCCTGCCTACTGCCTCGGA
Slc35f2.2526	TGCTGTTGACAGTGAGCGAGTAAGTGTTCGTATTATTAAATAGTGAAGCCAC AGATGTATTTAATAATACGAACACTTACCTGCCTACTGCCTCGGA
Slc35f2.549	TGCTGTTGACAGTGAGCGCCCTGGAGATTTTGAGAAGAAATAGTGAAGCCAC AGATGTATTTCTTCTCAAAATCTCCAGGATGCCTACTGCCTCGGA
Adra1a.1067	TGCTGTTGACAGTGAGCGCGACCATCATCCTGGTTATGTATAGTGAAGCCAC AGATGTATACATAACCAGGATGATGGTCATGCCTACTGCCTCGGA
Adra1a.530	TGCTGTTGACAGTGAGCGAAGCACAGGTGAACATTTCTAATAGTGAAGCCAC AGATGTATTAGAAATGTTCACCTGTGCTGTGCCTACTGCCTCGGA
Adra1a.992	TGCTGTTGACAGTGAGCGCTGAGACCATCTGCCAAATCAATAGTGAAGCCAC AGATGTATTGATTTGGCAGATGGTCTCATTGCCTACTGCCTCGGA
Adra1a.529	TGCTGTTGACAGTGAGCGACAGCACAGGTGAACATTTCTATAGTGAAGCCAC AGATGTATAGAAATGTTCACCTGTGCTGGTGCCTACTGCCTCGGA
Adra1a.655	TGCTGTTGACAGTGAGCGATGACTCACTACTACATTGTCATAGTGAAGCCAC AGATGTATGACAATGTAGTAGTGAGTCACTGCCTACTGCCTCGGA

Adra1a.1223	TGCTGTTGACAGTGAGCGACAGTGCCAAGAATAAGACTCATAGTGAAGCCA CAGATGTATGAGTCTTATTCTTGGCACTGCTGCCTACTGCCTCGGA
Ankrd1.1664	TGCTGTTGACAGTGAGCGCTCGATGAAAGAAGAAATCAAATAGTGAAGCCA CAGATGTATTTGATTTCTTCTTTCATCGAATGCCTACTGCCTCGGA
Ankrd1.1722	TGCTGTTGACAGTGAGCGCTGTATGTATATTTTATATTTATAGTGAAGCCACA GATGTATAAATATAAAATATACATACAATGCCTACTGCCTCGGA
Ankrd1.1194	TGCTGTTGACAGTGAGCGCCACCACTTTTATTTATTCCTATAGTGAAGCCACA GATGTATAGGAATAAATAAAGTGGTGATGCCTACTGCCTCGGA
Ankrd1.323	TGCTGTTGACAGTGAGCGCTGAAATAATTGTTCAACTGAATAGTGAAGCCAC AGATGTATTCAGTTGAACAATTATTTCAATGCCTACTGCCTCGGA
Ankrd1.1693	TGCTGTTGACAGTGAGCGACGCTGTAAATGTATAAATGTATAGTGAAGCCAC AGATGTATACATTTATACATTTACAGCGCTGCCTACTGCCTCGGA
Ankrd1.1158	TGCTGTTGACAGTGAGCGATGGGTTGTAGGTGTTAAAATATAGTGAAGCCAC AGATGTATATTTTAACACCTACAACCCAGTGCCTACTGCCTCGGA
Arhgap22.2560	TGCTGTTGACAGTGAGCGCGGCCACCAAGAACTTTATTAATAGTGAAGCCAC AGATGTATTAATAAAGTTCTTGGTGGCCTTGCCTACTGCCTCGGA
Arhgap22.1035	TGCTGTTGACAGTGAGCGACCCACTCAGATGTCAATAAGATAGTGAAGCCAC AGATGTATCTTATTGACATCTGAGTGGGCTGCCTACTGCCTCGGA
Arhgap22.2561	TGCTGTTGACAGTGAGCGAGCCACCAAGAACTTTATTAAATAGTGAAGCCAC AGATGTATTTAATAAAGTTCTTGGTGGCCTGCCTACTGCCTCGGA
Arhgap22.2429	TGCTGTTGACAGTGAGCGCCTGGATACTTTTCCAAACATATAGTGAAGCCAC AGATGTATATGTTTGGAAAAGTATCCAGATGCCTACTGCCTCGGA
Arhgap22.2420	TGCTGTTGACAGTGAGCGAAGGCAGGATCTGGATACTTTTTAGTGAAGCCAC AGATGTAAAAAGTATCCAGATCCTGCCTGTGCCTACTGCCTCGGA
Arhgap22.2559	TGCTGTTGACAGTGAGCGAAGGCCACCAAGAACTTTATTATAGTGAAGCCAC AGATGTATAATAAAGTTCTTGGTGGCCTGTGCCTACTGCCTCGGA
Arid3c.593	TGCTGTTGACAGTGAGCGATGGTGGAAGTCATCAACCGAATAGTGAAGCCA CAGATGTATTCGGTTGATGACTTCCACCAGTGCCTACTGCCTCGGA
Arid3c.463	TGCTGTTGACAGTGAGCGAAAGGAGTTTCTGGATGACCTATAGTGAAGCCAC AGATGTATAGGTCATCCAGAAACTCCTTCTGCCTACTGCCTCGGA
Arid3c.693	TGCTGTTGACAGTGAGCGCCATGAAGTATTTGTACCCATATAGTGAAGCCAC AGATGTATATGGGTACAAATACTTCATGTTGCCTACTGCCTCGGA
Arid3c.376	TGCTGTTGACAGTGAGCGAAGCCAGTCTCCTGGAATCCAATAGTGAAGCCAC AGATGTATTGGATTCCAGGAGACTGGCTGTGCCTACTGCCTCGGA
Arid3c.913	TGCTGTTGACAGTGAGCGACAGGGTTCTGCTTCTGGTTTATAGTGAAGCCAC AGATGTATAAACCAGAAGCAGAACCCTGGTGCCTACTGCCTCGGA
Arid3c.479	TGCTGTTGACAGTGAGCGAACCTATTTAGCTTCATGCAGATAGTGAAGCCAC AGATGTATCTGCATGAAGCTAAATAGGTCTGCCTACTGCCTCGGA
Asb15.304	TGCTGTTGACAGTGAGCGCCCAGGAGTATGTGCAATATAATAGTGAAGCCAC AGATGTATTATATTGCACATACTCCTGGATGCCTACTGCCTCGGA

Asb15.1834	TGCTGTTGACAGTGAGCGCGCGAAGATACTTATTATTTAATAGTGAAGCCAC AGATGTATTAAATAATAAGTATCTTCGCATGCCTACTGCCTCGGA
Asb15.130	TGCTGTTGACAGTGAGCGCGGATATTAATGATGATTCTAATAGTGAAGCCAC AGATGTATTAGAATCATCATTAATATCCATGCCTACTGCCTCGGA
Asb15.211	TGCTGTTGACAGTGAGCGACCAGGCTATATTTCATCCTAATAGTGAAGCCAC AGATGTATTAGGATGAAATATAGCCTGGCTGCCTACTGCCTCGGA
Asb15.1833	TGCTGTTGACAGTGAGCGCTGCGAAGATACTTATTATTTATAGTGAAGCCAC AGATGTATAAATAATAAGTATCTTCGCATTGCCTACTGCCTCGGA
Asb15.305	TGCTGTTGACAGTGAGCGACAGGAGTATGTGCAATATAAATAGTGAAGCCA CAGATGTATTTATATTGCACATACTCCTGGTGCCTACTGCCTCGGA
Asb4.2958	TGCTGTTGACAGTGAGCGAACATAGAAAGTAGTTTGCTTATAGTGAAGCCAC AGATGTATAAGCAAACTACTTTCTATGTGTGCCTACTGCCTCGGA
Asb4.1525	TGCTGTTGACAGTGAGCGAAAGGCAGATACTATTATAGTATAGTGAAGCCAC AGATGTATACTATAATAGTATCTGCCTTCTGCCTACTGCCTCGGA
Asb4.2855	TGCTGTTGACAGTGAGCGCGTATTTGATGTATCAATTTAATAGTGAAGCCAC AGATGTATTAAATTGATACATCAAATACATGCCTACTGCCTCGGA
Asb4.2637	TGCTGTTGACAGTGAGCGCTGGTACCAGTTAAAATTGTAATAGTGAAGCCAC AGATGTATTACAATTTTAACTGGTACCAATGCCTACTGCCTCGGA
Asb4.1290	TGCTGTTGACAGTGAGCGACCAGAGGGAATCATTTACTAATAGTGAAGCCAC AGATGTATTAGTAAATGATTCCCTCTGGCTGCCTACTGCCTCGGA
Asb4.1912	TGCTGTTGACAGTGAGCGCGCCTAGGAATAGACATTTCAATAGTGAAGCCAC AGATGTATTGAAATGTCTATTCCTAGGCTTGCCTACTGCCTCGGA
Atp5f1.1272	TGCTGTTGACAGTGAGCGACAAGTGCATTGAAGATCTAAATAGTGAAGCCAC AGATGTATTTAGATCTTCAATGCACTTGGTGCCTACTGCCTCGGA
Atp5f1.885	TGCTGTTGACAGTGAGCGCGATAGTCTATGTGATTAAGAATAGTGAAGCCAC AGATGTATTCTTAATCACATAGACTATCATGCCTACTGCCTCGGA
Atp5f1.1545	TGCTGTTGACAGTGAGCGACCAACTGAAATTACCAAGTTATAGTGAAGCCAC AGATGTATAACTTGGTAATTTCAGTTGGGTGCCTACTGCCTCGGA
Atp5f1.1553	TGCTGTTGACAGTGAGCGCAATTACCAAGTTATAATTTAATAGTGAAGCCAC AGATGTATTAAATTATAACTTGGTAATTTTGCCTACTGCCTCGGA
Atp5f1.923	TGCTGTTGACAGTGAGCGAGAGAATTTATTGACAAACTTATAGTGAAGCCAC AGATGTATAAGTTTGTCAATAAATTCTCCTGCCTACTGCCTCGGA
Atp5f1.811	TGCTGTTGACAGTGAGCGCAGCTTGTATTTTCTATCCAAATAGTGAAGCCAC AGATGTATTTGGATAGAAAATACAAGCTATGCCTACTGCCTCGGA
Cebpa.1466	TGCTGTTGACAGTGAGCGAGGCAGTACTAGTATTAAGGAATAGTGAAGCCA CAGATGTATTCCTTAATACTAGTACTGCCGTGCCTACTGCCTCGGA
Cebpa.2083	TGCTGTTGACAGTGAGCGCCCCAAATATTTTGCTTTATCATAGTGAAGCCAC AGATGTATGATAAAGCAAAATATTTGGGATGCCTACTGCCTCGGA
Cebpa.1552	TGCTGTTGACAGTGAGCGATGCCTTGATATTTTATTTGGATAGTGAAGCCAC AGATGTATCCAAATAAAATATCAAGGCACTGCCTACTGCCTCGGA

Cebpa.1462	TGCTGTTGACAGTGAGCGAGCCCGGCAGTACTAGTATTAATAGTGAAGCCAC AGATGTATTAATACTAGTACTGCCGGGCCTGCCTACTGCCTCGGA
Cebpa.1463	TGCTGTTGACAGTGAGCGACCCGGCAGTACTAGTATTAAGTAGTGAAGCCAC AGATGTACTTAATACTAGTACTGCCGGGCTGCCTACTGCCTCGGA
Cebpa.1847	TGCTGTTGACAGTGAGCGCCCTGGGTGAGTTCATGGAGAATAGTGAAGCCAC AGATGTATTCTCCATGAACTCACCCAGGATGCCTACTGCCTCGGA
Crtap.1182	TGCTGTTGACAGTGAGCGACAGTTCAGTTCTTTAATGTGATAGTGAAGCCAC AGATGTATCACATTAAAGAACTGAACTGCTGCCTACTGCCTCGGA
Crtap.655	TGCTGTTGACAGTGAGCGCGAGAAACATGGAGTATTATAATAGTGAAGCCA CAGATGTATTATAATACTCCATGTTTCTCTTGCCTACTGCCTCGGA
Crtap.634	TGCTGTTGACAGTGAGCGCTCCAGATGACGAGATGATGAATAGTGAAGCCA CAGATGTATTCATCATCTCGTCATCTGGATTGCCTACTGCCTCGGA
Crtap.640	TGCTGTTGACAGTGAGCGCTGACGAGATGATGAAGAGAAATAGTGAAGCCA CAGATGTATTTCTCTTCATCATCTCGTCATTGCCTACTGCCTCGGA
Crtap.994	TGCTGTTGACAGTGAGCGAGACCATGTACCACTATTTACATAGTGAAGCCAC AGATGTATGTAAATAGTGGTACATGGTCGTGCCTACTGCCTCGGA
Crtap.922	TGCTGTTGACAGTGAGCGCGGAGTGTAAGATTCGTTGTGATAGTGAAGCCAC AGATGTATCACAACGAATCTTACACTCCATGCCTACTGCCTCGGA
Dnah2.8126	TGCTGTTGACAGTGAGCGAAAGGACTTTCATGATACCAAATAGTGAAGCCAC AGATGTATTTGGTATCATGAAAGTCCTTGTGCCTACTGCCTCGGA
Dnah2.10220	TGCTGTTGACAGTGAGCGACGGGATGAGATTATCAATCAATAGTGAAGCCAC AGATGTATTGATTGATAATCTCATCCCGGTGCCTACTGCCTCGGA
Dnah2.775	TGCTGTTGACAGTGAGCGCCCCCCCAGATTTTTATGAATAATAGTGAAGCCAC AGATGTATTATTCATAAAAATCTGGGGGATGCCTACTGCCTCGGA
Dnah2.12409	TGCTGTTGACAGTGAGCGCCACCACCTATATCAATGATTATAGTGAAGCCAC AGATGTATAATCATTGATATAGGTGGTGATGCCTACTGCCTCGGA
Dnah2.10221	TGCTGTTGACAGTGAGCGAGGGATGAGATTATCAATCAAATAGTGAAGCCA CAGATGTATTTGATTGATAATCTCATCCCGTGCCTACTGCCTCGGA
Dnah2.12817	TGCTGTTGACAGTGAGCGCGAGATATAACAAGTTGATGAATAGTGAAGCCA CAGATGTATTCATCAACTTGTTATATCTCTTGCCTACTGCCTCGGA
Dtx1.984	TGCTGTTGACAGTGAGCGCCCACATCGAGAATGTTCTTAATAGTGAAGCCAC AGATGTATTAAGAACATTCTCGATGTGGTTGCCTACTGCCTCGGA
Dtx1.3100	TGCTGTTGACAGTGAGCGATGTCTCCATTACATCTGTATATAGTGAAGCCAC AGATGTATATACAGATGTAATGGAGACACTGCCTACTGCCTCGGA
Dtx1.3098	TGCTGTTGACAGTGAGCGATGTGTCTCCATTACATCTGTATAGTGAAGCCAC AGATGTATACAGATGTAATGGAGACACAGTGCCTACTGCCTCGGA
Dtx1.3279	TGCTGTTGACAGTGAGCGACCAGGTTATTAAGTAGCTTTTTAGTGAAGCCAC AGATGTAAAAAGCTACTTAATAACCTGGGTGCCTACTGCCTCGGA
Dtx1.2613	TGCTGTTGACAGTGAGCGCCAAGACGGAGTTTGGTCCAATAGTGAAGCCAC AGATGTATTGGAACCAAACTCCGTCTTGTTGCCTACTGCCTCGGA

Dtx1.2011	TGCTGTTGACAGTGAGCGAAAGAAGAAACACCTCAAGAAATAGTGAAGCCACAGATGTATTTCTTGAGGTGTTTCTTCTTGTGCCTACTGCCTCGGA
Fosl1.948	TGCTGTTGACAGTGAGCGATCCGAGTCTGGTTTTCACCTATAGTGAAGCCACAGATGTATAGGTGAAAACCAGACTCGGAGTGCCTACTGCCTCGGA
Fosl1.926	TGCTGTTGACAGTGAGCGCCACCCTCTCTGACTCCTTTTATAGTGAAGCCACAGATGTATAAAAGGAGTCAGAGAGGGTGTTGCCTACTGCCTCGGA
Fosl1.699	TGCTGTTGACAGTGAGCGAAGAGATTGAAGAGCTGCAGAATAGTGAAGCCACAGATGTATTCTGCAGCTCTTCAATCTCTCTGCCTACTGCCTCGGA
Fosl1.949	TGCTGTTGACAGTGAGCGCCCGAGTCTGGTTTTCACCTATTAGTGAAGCCACAGATGTAATAGGTGAAAACCAGACTCGGATGCCTACTGCCTCGGA
Fosl1.688	TGCTGTTGACAGTGAGCGAGGGCTGCAGCGAGAGATTGAATAGTGAAGCCACAGATGTATTCAATCTCTCGCTGCAGCCCCTGCCTACTGCCTCGGA
Fosl1.624	TGCTGTTGACAGTGAGCGCCCGAAGAAAGGAGCTGACAGATAGTGAAGCCACAGATGTATCTGTCAGCTCCTTTCTTCGGTTGCCTACTGCCTCGGA
Foxj1.2336	TGCTGTTGACAGTGAGCGCAGTGAATGTAGTTATAGCTAATAGTGAAGCCACAGATGTATTAGCTATAACTACATTCACTATGCCTACTGCCTCGGA
Foxj1.2523	TGCTGTTGACAGTGAGCGCTGTAAAGTTCTTTACAATAAATAGTGAAGCCACAGATGTATTTATTGTAAAGAACTTTACATTGCCTACTGCCTCGGA
Foxj1.2008	TGCTGTTGACAGTGAGCGCCGCTCATTAGATGATAACAAATAGTGAAGCCACAGATGTATTTGTTATCATCTAATGAGCGTTGCCTACTGCCTCGGA
Foxj1.2524	TGCTGTTGACAGTGAGCGCGTAAAGTTCTTTACAATAAAATAGTGAAGCCACAGATGTATTTTATTGTAAAGAACTTTACATGCCTACTGCCTCGGA
Foxj1.2522	TGCTGTTGACAGTGAGCGAATGTAAAGTTCTTTACAATAATAGTGAAGCCACAGATGTATTATTGTAAAGAACTTTACATGTGCCTACTGCCTCGGA
Foxj1.2012	TGCTGTTGACAGTGAGCGCCATTAGATGATAACAAATTAATAGTGAAGCCACAGATGTATTAATTTGTTATCATCTAATGATGCCTACTGCCTCGGA
Foxn4.2096	TGCTGTTGACAGTGAGCGCCTGTGTGTACATTTATTTATATAGTGAAGCCACAGATGTATATAAATAAATGTACACACAGTTGCCTACTGCCTCGGA
Foxn4.2100	TGCTGTTGACAGTGAGCGCGTGTACATTTATTTATATTTATAGTGAAGCCACAGATGTATAAATATAAATAAATGTACACATGCCTACTGCCTCGGA
Foxn4.99	TGCTGTTGACAGTGAGCGCCCAGAATGTCAGAAATGATCATAGTGAAGCCACAGATGTATGATCATTTCTGACATTCTGGATGCCTACTGCCTCGGA
Foxn4.2097	TGCTGTTGACAGTGAGCGATGTGTGTACATTTATTTATATTAGTGAAGCCACAGATGTAATATAAATAAATGTACACACAGTGCCTACTGCCTCGGA
Foxn4.2587	TGCTGTTGACAGTGAGCGATCCTGTGACTGATTTTTCCAATAGTGAAGCCACAGATGTATTGGAAAAATCAGTCACAGGACTGCCTACTGCCTCGGA
Foxn4.2450	TGCTGTTGACAGTGAGCGATGCCAGGTTAGAGACATGAAATAGTGAAGCCACAGATGTATTTCATGTCTCTAACCTGGCACTGCCTACTGCCTCGGA
Gata1.846	TGCTGTTGACAGTGAGCGATGGCTTGTATCACAAGATGAATAGTGAAGCCACAGATGTATTCATCTTGTGATACAAGCCACTGCCTACTGCCTCGGA

Gata1.842	TGCTGTTGACAGTGAGCGACCTGTGGCTTGTATCACAAGATAGTGAAGCCAC AGATGTATCTTGTGATACAAGCCACAGGCTGCCTACTGCCTCGGA
Gata1.929	TGCTGTTGACAGTGAGCGCCCCAATGCACTAACTGTCAAATAGTGAAGCCAC AGATGTATTTGACAGTTAGTGCATTGGGTTGCCTACTGCCTCGGA
Gata1.668	TGCTGTTGACAGTGAGCGACCAGTCCTTTCTTCTCTCCCATAGTGAAGCCACA GATGTATGGGAGAGAAGAAAGGACTGGGTGCCTACTGCCTCGGA
Gata1.1280	TGCTGTTGACAGTGAGCGCCAGGGCCTGTCAGCCATCTTATAGTGAAGCCAC AGATGTATAAGATGGCTGACAGGCCCTGATGCCTACTGCCTCGGA
Gata1.552	TGCTGTTGACAGTGAGCGCCACGTTCTTGGACACCTTGAATAGTGAAGCCAC AGATGTATTCAAGGTGTCCAAGAACGTGTTGCCTACTGCCTCGGA
Gli1.381	TGCTGTTGACAGTGAGCGACTCGGAGTTCAGTCAAATTAATAGTGAAGCCAC AGATGTATTAATTTGACTGAACTCCGAGGTGCCTACTGCCTCGGA
Gli1.3625	TGCTGTTGACAGTGAGCGAACACAGTTTCCTGACAATAAATAGTGAAGCCAC AGATGTATTTATTGTCAGGAAACTGTGTCTGCCTACTGCCTCGGA
Gli1.576	TGCTGTTGACAGTGAGCGACACCTCAGATGAGTCATCAAATAGTGAAGCCAC AGATGTATTTGATGACTCATCTGAGGTGGTGCCTACTGCCTCGGA
Gli1.380	TGCTGTTGACAGTGAGCGCCCTCGGAGTTCAGTCAAATTATAGTGAAGCCAC AGATGTATAATTTGACTGAACTCCGAGGATGCCTACTGCCTCGGA
Gli1.3512	TGCTGTTGACAGTGAGCGACAAGCAGATGGTATTTCCTAATAGTGAAGCCAC AGATGTATTAGGAAATACCATCTGCTTGGTGCCTACTGCCTCGGA
Gli1.3574	TGCTGTTGACAGTGAGCGAGGGCTGTATTTAGTCTATGTATAGTGAAGCCAC AGATGTATACATAGACTAAATACAGCCCCTGCCTACTGCCTCGGA
Glis1.1017	TGCTGTTGACAGTGAGCGCCAGCTCTAGTTAGCTGTGTAATAGTGAAGCCAC AGATGTATTACACAGCTAACTAGAGCTGTTGCCTACTGCCTCGGA
Glis1.2181	TGCTGTTGACAGTGAGCGACAGGGGGACAGTCATTCTCTATAGTGAAGCCAC AGATGTATAGAGAATGACTGTCCCCCTGGTGCCTACTGCCTCGGA
Glis1.1016	TGCTGTTGACAGTGAGCGAACAGCTCTAGTTAGCTGTGTATAGTGAAGCCAC AGATGTATACACAGCTAACTAGAGCTGTCTGCCTACTGCCTCGGA
Glis1.685	TGCTGTTGACAGTGAGCGCCCCAGCCGCTCAGCACATCAATAGTGAAGCCAC AGATGTATTGATGTGCTGAGCGGCTGGGATGCCTACTGCCTCGGA
Glis1.1583	TGCTGTTGACAGTGAGCGACGGAGCCACACAGGCGAGAAATAGTGAAGCCA CAGATGTATTTCTCGCCTGTGTGGCTCCGCTGCCTACTGCCTCGGA
Glis1.973	TGCTGTTGACAGTGAGCGACTGTGCCTCCTCAGATATCAATAGTGAAGCCAC AGATGTATTGATATCTGAGGAGGCACAGGTGCCTACTGCCTCGGA
Hr.5442	TGCTGTTGACAGTGAGCGCTGGGGATTTTTTTTAATTAAATAGTGAAGCCAC AGATGTATTTAATTAAAAAAAATCCCCAATGCCTACTGCCTCGGA
Hr.5404	TGCTGTTGACAGTGAGCGAGCCAGCTCTTGTTATACTTAATAGTGAAGCCAC AGATGTATTAAGTATAACAAGAGCTGGCGTGCCTACTGCCTCGGA
Hr.5024	TGCTGTTGACAGTGAGCGATGGGACTGAAGATATTCTTGATAGTGAAGCCAC AGATGTATCAAGAATATCTTCAGTCCCAGTGCCTACTGCCTCGGA

Hr.5403	TGCTGTTGACAGTGAGCGACGCCAGCTCTTGTTATACTTATAGTGAAGCCAC AGATGTATAAGTATAACAAGAGCTGGCGCTGCCTACTGCCTCGGA
Hr.5028	TGCTGTTGACAGTGAGCGAACTGAAGATATTCTTGACTGATAGTGAAGCCAC AGATGTATCAGTCAAGAATATCTTCAGTCTGCCTACTGCCTCGGA
Hr.4895	TGCTGTTGACAGTGAGCGAAAGGCGAATGCAGAAATCTGATAGTGAAGCCA CAGATGTATCAGATTCTGCATTCGCCTTCTGCCTACTGCCTCGGA
Isx.594	TGCTGTTGACAGTGAGCGACCAGACCACAAAGCATTTGTATAGTGAAGCCAC AGATGTATACAAATGCTTTGTGGTCTGGGTGCCTACTGCCTCGGA
Isx.596	TGCTGTTGACAGTGAGCGAAGACCACAAAGCATTTGTAAATAGTGAAGCCA CAGATGTATTTACAAATGCTTTGTGGTCTGTGCCTACTGCCTCGGA
Isx.595	TGCTGTTGACAGTGAGCGACAGACCACAAAGCATTTGTAATAGTGAAGCCAC AGATGTATTACAAATGCTTTGTGGTCTGGTGCCTACTGCCTCGGA
Isx.1122	TGCTGTTGACAGTGAGCGAGGCAGCATCTGTGCAACTTCATAGTGAAGCCAC AGATGTATGAAGTTGCACAGATGCTGCCCTGCCTACTGCCTCGGA
Isx.493	TGCTGTTGACAGTGAGCGCCATGGAGAAGAGTTCAGGATATAGTGAAGCCA CAGATGTATATCCTGAACTCTTCTCCATGTTGCCTACTGCCTCGGA
Isx.988	TGCTGTTGACAGTGAGCGATCCCTTCAAGCTGGTTCCCTATAGTGAAGCCAC AGATGTATAGGGAACCAGCTTGAAGGGAGTGCCTACTGCCTCGGA
Kank4.4755	TGCTGTTGACAGTGAGCGACCTGTCCTTTTTGTTTATAAATAGTGAAGCCACA GATGTATTTATAAACAAAAAGGACAGGCTGCCTACTGCCTCGGA
Kank4.4355	TGCTGTTGACAGTGAGCGCCCATAAGAAACAGCTTATAAATAGTGAAGCCAC AGATGTATTTATAAGCTGTTTCTTATGGATGCCTACTGCCTCGGA
Kank4.3339	TGCTGTTGACAGTGAGCGCCAGTACATTCCCATCATATAATAGTGAAGCCAC AGATGTATTATATGATGGGAATGTACTGTTGCCTACTGCCTCGGA
Kank4.3764	TGCTGTTGACAGTGAGCGCTGGAAGTAAGCAGAAAACAAATAGTGAAGCCA CAGATGTATTTGTTTTCTGCTTACTTCCATTGCCTACTGCCTCGGA
Kank4.4579	TGCTGTTGACAGTGAGCGAACACACTTTTAATAAACGTTATAGTGAAGCCAC AGATGTATAACGTTTATTAAAAGTGTGTGTGCCTACTGCCTCGGA
Kank4.4754	TGCTGTTGACAGTGAGCGAGCCTGTCCTTTTTGTTTATAATAGTGAAGCCACA GATGTATTATAAACAAAAAGGACAGGCCTGCCTACTGCCTCGGA
Kcnk13.1414	TGCTGTTGACAGTGAGCGAGAAGATTTTTCTGATCTTTTATAGTGAAGCCAC AGATGTATAAAAGATCAGAAAAATCTTCCTGCCTACTGCCTCGGA
Kcnk13.1888	TGCTGTTGACAGTGAGCGCCCTGATCAAACAGACTGTGAATAGTGAAGCCAC AGATGTATTCACAGTCTGTTTGATCAGGATGCCTACTGCCTCGGA
Kcnk13.1855	TGCTGTTGACAGTGAGCGACTGCATCTACTCTTTGTTTAATAGTGAAGCCAC AGATGTATTAAACAAAGAGTAGATGCAGCTGCCTACTGCCTCGGA
Kcnk13.2209	TGCTGTTGACAGTGAGCGCGGGAGCCTTTGCAGTAATGAATAGTGAAGCCAC AGATGTATTCATTACTGCAAAGGCTCCCATGCCTACTGCCTCGGA
Kcnk13.1453	TGCTGTTGACAGTGAGCGAAAGTACCATCCTCTTCTTCAATAGTGAAGCCAC AGATGTATTGAAGAAGAGGATGGTACTTGTGCCTACTGCCTCGGA

107

Kcnk13.1407	TGCTGTTGACAGTGAGCGCCGGGAGGGAAGATTTTTCTGATAGTGAAGCCAC AGATGTATCAGAAAAATCTTCCCTCCCGTTGCCTACTGCCTCGGA
Klf5.1334	TGCTGTTGACAGTGAGCGACACCTGTCAGATACAACAGAATAGTGAAGCCA CAGATGTATTCTGTTGTATCTGACAGGTGGTGCCTACTGCCTCGGA
Klf5.889	TGCTGTTGACAGTGAGCGAGTGAACAATATCTTCATCAAATAGTGAAGCCAC AGATGTATTTGATGAAGATATTGTTCACCTGCCTACTGCCTCGGA
Klf5.1407	TGCTGTTGACAGTGAGCGATTGCACAAAAGTTTATACAAATAGTGAAGCCAC AGATGTATTTGTATAAACTTTTGTGCAACTGCCTACTGCCTCGGA
Klf5.581	TGCTGTTGACAGTGAGCGACGATAATTTCAGAGCATAAAATAGTGAAGCCAC AGATGTATTTTATGCTCTGAAATTATCGGTGCCTACTGCCTCGGA
Klf5.579	TGCTGTTGACAGTGAGCGCTCCGATAATTTCAGAGCATAATAGTGAAGCCAC AGATGTATTATGCTCTGAAATTATCGGAATGCCTACTGCCTCGGA
Klf5.888	TGCTGTTGACAGTGAGCGCGGTGAACAATATCTTCATCAATAGTGAAGCCAC AGATGTATTGATGAAGATATTGTTCACCTTGCCTACTGCCTCGGA
Macc1.2483	TGCTGTTGACAGTGAGCGATGGAACTTTAATACTTATAAATAGTGAAGCCAC AGATGTATTTATAAGTATTAAAGTTCCAGTGCCTACTGCCTCGGA
Macc1.1314	TGCTGTTGACAGTGAGCGCTCGTGTGATCCTGATTTTGAATAGTGAAGCCAC AGATGTATTCAAAATCAGGATCACACGAATGCCTACTGCCTCGGA
Macc1.3324	TGCTGTTGACAGTGAGCGAGCATACAGTTACTTGATATAATAGTGAAGCCAC AGATGTATTATATCAAGTAACTGTATGCCTGCCTACTGCCTCGGA
Macc1.3172	TGCTGTTGACAGTGAGCGCGACCATGAATCAGTATATATATAGTGAAGCCAC AGATGTATATATATACTGATTCATGGTCTTGCCTACTGCCTCGGA
Macc1.634	TGCTGTTGACAGTGAGCGCCCATTGCAGTTTGTAAATTAATAGTGAAGCCAC AGATGTATTAATTTACAAACTGCAATGGTTGCCTACTGCCTCGGA
Macc1.2025	TGCTGTTGACAGTGAGCGAGCACAGGAAGATTTTAATAGATAGTGAAGCCA CAGATGTATCTATTAAAATCTTCCTGTGCCTGCCTACTGCCTCGGA
MBL2.476	TGCTGTTGACAGTGAGCGCACTAGCGAAATTGATTCAGAATAGTGAAGCCAC AGATGTATTCTGAATCAATTTCGCTAGTATGCCTACTGCCTCGGA
MBL2.477	TGCTGTTGACAGTGAGCGCCTAGCGAAATTGATTCAGAAATAGTGAAGCCAC AGATGTATTTCTGAATCAATTTCGCTAGTTGCCTACTGCCTCGGA
MBL2.691	TGCTGTTGACAGTGAGCGCAGTGGCCAAAGATATTGCCTATAGTGAAGCCAC AGATGTATAGGCAATATCTTTGGCCACTTTGCCTACTGCCTCGGA
MBL2.769	TGCTGTTGACAGTGAGCGAAGTGCGCTATACTAATTGGAATAGTGAAGCCAC AGATGTATTCCAATTAGTATAGCGCACTCTGCCTACTGCCTCGGA
MBL2.479	TGCTGTTGACAGTGAGCGCAGCGAAATTGATTCAGAAATTTAGTGAAGCCAC AGATGTAAATTTCTGAATCAATTTCGCTATGCCTACTGCCTCGGA
MBL2.540	TGCTGTTGACAGTGAGCGATGCTCTTCTCTCTGAGTGAAATAGTGAAGCCAC AGATGTATTTCACTCAGAGAGAAGAGCACTGCCTACTGCCTCGGA
Mlip.1135	TGCTGTTGACAGTGAGCGCGCCTATAATGCCTTCTATTAATAGTGAAGCCAC AGATGTATTAATAGAAGGCATTATAGGCTTGCCTACTGCCTCGGA

Mlip.368	TGCTGTTGACAGTGAGCGAGAAGATGAAGCTACATGCAGATAGTGAAGCCACAGATGTATCTGCATGTAGCTTCATCTTCCTGCCTACTGCCTCGGA
Mlip.24	TGCTGTTGACAGTGAGCGACATGTCAAATCAGTTTCTAGATAGTGAAGCCACAGATGTATCTAGAAACTGATTTGACATGCTGCCTACTGCCTCGGA
Mlip.1177	TGCTGTTGACAGTGAGCGCCCCCTATATTTAGCAGTCAAATAGTGAAGCCACAGATGTATTTGACTGCTAAATATAGGGGATGCCTACTGCCTCGGA
Mlip.974	TGCTGTTGACAGTGAGCGAGCCAACTACTTGCTAAACTTATAGTGAAGCCACAGATGTATAAGTTTAGCAAGTAGTTGGCCTGCCTACTGCCTCGGA
Mlip.1168	TGCTGTTGACAGTGAGCGCACTAATGTTCCCCTATATTTATAGTGAAGCCACAGATGTATAAATATAGGGGAACATTAGTATGCCTACTGCCTCGGA
Mnd1.163	TGCTGTTGACAGTGAGCGACCGGATGATGGAGATATTTTTTAGTGAAGCCACAGATGTAAAAAATATCTCCATCATCCGGGTGCCTACTGCCTCGGA
Mnd1.713	TGCTGTTGACAGTGAGCGCCCAGAAGACTTTGACTACATATAGTGAAGCCACAGATGTATATGTAGTCAAAGTCTTCTGGATGCCTACTGCCTCGGA
Mnd1.783	TGCTGTTGACAGTGAGCGAAAGGTTTAAACAGCTAACTAATAGTGAAGCCACAGATGTATTAGTTAGCTGTTTAAACCTTGTGCCTACTGCCTCGGA
Mnd1.840	TGCTGTTGACAGTGAGCGACGTTTTGAAGCTTTAAATAAATAGTGAAGCCACAGATGTATTTATTTAAAGCTTCAAAACGGTGCCTACTGCCTCGGA
Mnd1.680	TGCTGTTGACAGTGAGCGCGAAGAAAGTAAAATTGATAAATAGTGAAGCCACAGATGTATTTATCAATTTTACTTTCTTCATGCCTACTGCCTCGGA
Mnd1.845	TGCTGTTGACAGTGAGCGCTGAAGCTTTAAATAAAGTATATAGTGAAGCCACAGATGTATATACTTTATTTAAAGCTTCAATGCCTACTGCCTCGGA
Mnx1.968	TGCTGTTGACAGTGAGCGCCCAGTTCAAGCTCAACAAGTATAGTGAAGCCACAGATGTATACTTGTTGAGCTTGAACTGGTTGCCTACTGCCTCGGA
Mnx1.678	TGCTGTTGACAGTGAGCGACCGGTCTACAGTTATTCGGCATAGTGAAGCCACAGATGTATGCCGAATAACTGTAGACCGGGTGCCTACTGCCTCGGA
Mnx1.731	TGCTGTTGACAGTGAGCGCCCCGGCGCTTTCCTACTCATATAGTGAAGCCACAGATGTATATGAGTAGGAAAGCGCCGGGTTGCCTACTGCCTCGGA
Mnx1.1268	TGCTGTTGACAGTGAGCGCGGACGAGGATGATGAAGAAGATAGTGAAGCCACAGATGTATCTTCTTCATCATCCTCGTCCTTGCCTACTGCCTCGGA
Mnx1.1271	TGCTGTTGACAGTGAGCGCCGAGGATGATGAAGAAGAGGATAGTGAAGCCACAGATGTATCCTCTTCTTCATCATCCTCGTTGCCTACTGCCTCGGA
Mnx1.1164	TGCTGTTGACAGTGAGCGAGGCAGTGAGGAGAAGACGGAATAGTGAAGCCACAGATGTATTCCGTCTTCTCCTCACTGCCGTGCCTACTGCCTCGGA
Muc4.9783	TGCTGTTGACAGTGAGCGAAGCAGTGTTTGTACAATGAAATAGTGAAGCCACAGATGTATTTCATTGTACAAACACTGCTCTGCCTACTGCCTCGGA
Muc4.1126	TGCTGTTGACAGTGAGCGACAGCAGCAAATCTCAAGTCAATAGTGAAGCCACAGATGTATTGACTTGAGATTTGCTGCTGGTGCCTACTGCCTCGGA
Muc4.9782	TGCTGTTGACAGTGAGCGAGAGCAGTGTTTGTACAATGAATAGTGAAGCCACAGATGTATTCATTGTACAAACACTGCTCCTGCCTACTGCCTCGGA

Muc4.1210	TGCTGTTGACAGTGAGCGCAGTGACATCATCTTCAGATAATAGTGAAGCCAC AGATGTATTATCTGAAGATGATGTCACTTTGCCTACTGCCTCGGA
Muc4.2384	TGCTGTTGACAGTGAGCGCCCAAGTACAACACATATGTTATAGTGAAGCCAC AGATGTATAACATATGTGTTGTACTTGGATGCCTACTGCCTCGGA
Muc4.9538	TGCTGTTGACAGTGAGCGACAGCAGCAAGATTCAAGCCTATAGTGAAGCCA CAGATGTATAGGCTTGAATCTTGCTGCTGCTGCCTACTGCCTCGGA
Mycl.1095	TGCTGTTGACAGTGAGCGCCCATAACTTCTTGGAACGAAATAGTGAAGCCAC AGATGTATTTCGTTCCAAGAAGTTATGGTTGCCTACTGCCTCGGA
Mycl.247	TGCTGTTGACAGTGAGCGCCAGCACTATTTCTACGACTATTAGTGAAGCCAC AGATGTAATAGTCGTAGAAATAGTGCTGATGCCTACTGCCTCGGA
Mycl.1090	TGCTGTTGACAGTGAGCGAAAGAACCATAACTTCTTGGAATAGTGAAGCCAC AGATGTATTCCAAGAAGTTATGGTTCTTCTGCCTACTGCCTCGGA
Mycl.246	TGCTGTTGACAGTGAGCGCTCAGCACTATTTCTACGACTATAGTGAAGCCAC AGATGTATAGTCGTAGAAATAGTGCTGATTGCCTACTGCCTCGGA
Mycl.1206	TGCTGTTGACAGTGAGCGCCCTCAGCAAGGCGTTAGAATATAGTGAAGCCAC AGATGTATATTCTAACGCCTTGCTGAGGATGCCTACTGCCTCGGA
Mycl.1315	TGCTGTTGACAGTGAGCGAGCGTACCTCAGTGGCTACTAATAGTGAAGCCAC AGATGTATTAGTAGCCACTGAGGTACGCGTGCCTACTGCCTCGGA
Nat8.790	TGCTGTTGACAGTGAGCGCGTGGATGTTTCTCTAATTCATTAGTGAAGCCAC AGATGTAATGAATTAGAGAAACATCCACATGCCTACTGCCTCGGA
Nat8.789	TGCTGTTGACAGTGAGCGCTGTGGATGTTTCTCTAATTCATAGTGAAGCCAC AGATGTATGAATTAGAGAAACATCCACAATGCCTACTGCCTCGGA
Nat8.412	TGCTGTTGACAGTGAGCGATGGAAGAATTATGTGTCCAAATAGTGAAGCCAC AGATGTATTTGGACACATAATTCTTCCAGTGCCTACTGCCTCGGA
Nat8.791	TGCTGTTGACAGTGAGCGATGGATGTTTCTCTAATTCATTTAGTGAAGCCAC AGATGTAAATGAATTAGAGAAACATCCACTGCCTACTGCCTCGGA
Nat8.677	TGCTGTTGACAGTGAGCGCACAGTGATGTTGTCCTTGTGATAGTGAAGCCAC AGATGTATCACAAGGACAACATCACTGTATGCCTACTGCCTCGGA
Nat8.411	TGCTGTTGACAGTGAGCGACTGGAAGAATTATGTGTCCAATAGTGAAGCCAC AGATGTATTGGACACATAATTCTTCCAGGTGCCTACTGCCTCGGA
Orc1.2732	TGCTGTTGACAGTGAGCGAAGGATGGAAATTGTTGTTATATAGTGAAGCCAC AGATGTATATAACAACAATTTCCATCCTGTGCCTACTGCCTCGGA
Orc1.2808	TGCTGTTGACAGTGAGCGAGGTAACATACTATGTATATAATAGTGAAGCCAC AGATGTATTATATACATAGTATGTTACCGTGCCTACTGCCTCGGA
Orc1.2962	TGCTGTTGACAGTGAGCGATATGATGTATTGGATTTTAAATAGTGAAGCCAC AGATGTATTTAAAATCCAATACATCATAGTGCCTACTGCCTCGGA
Orc1.2807	TGCTGTTGACAGTGAGCGCCGGTAACATACTATGTATATAGTGAAGCCAC AGATGTATATATACATAGTATGTTACCGTTGCCTACTGCCTCGGA
Orc1.699	TGCTGTTGACAGTGAGCGCAGCCAAGATTAAGAATGTAAATAGTGAAGCCA CAGATGTATTTACATTCTTAATCTTGGCTTTGCCTACTGCCTCGGA

Orc1.330	TGCTGTTGACAGTGAGCGCACTGATTGAATTATTTCAAAATAGTGAAGCCAC AGATGTATTTTGAAATAATTCAATCAGTTTGCCTACTGCCTCGGA
Pou6f2.827	TGCTGTTGACAGTGAGCGCCACAGGTCAACTAGTTACTAATAGTGAAGCCAC AGATGTATTAGTAACTAGTTGACCTGTGATGCCTACTGCCTCGGA
Pou6f2.828	TGCTGTTGACAGTGAGCGAACAGGTCAACTAGTTACTAATTAGTGAAGCCAC AGATGTAATTAGTAACTAGTTGACCTGTGTGCCTACTGCCTCGGA
Pou6f2.1603	TGCTGTTGACAGTGAGCGAGAGCAGGTATGCAGAATCTCATAGTGAAGCCA CAGATGTATGAGATTCTGCATACCTGCTCGTGCCTACTGCCTCGGA
Pou6f2.1004	TGCTGTTGACAGTGAGCGCCCAGATCTTGCCTGTGATCAATAGTGAAGCCAC AGATGTATTGATCACAGGCAAGATCTGGTTGCCTACTGCCTCGGA
Pou6f2.1883	TGCTGTTGACAGTGAGCGCAGCAGATTCTCCAGAAGAAAATAGTGAAGCCA CAGATGTATTTTCTTCTGGAGAATCTGCTATGCCTACTGCCTCGGA
Pou6f2.834	TGCTGTTGACAGTGAGCGCCAACTAGTTACTAATGCACAATAGTGAAGCCAC AGATGTATTGTGCATTAGTAACTAGTTGATGCCTACTGCCTCGGA
Prdm8.2747	TGCTGTTGACAGTGAGCGCAGATGTAATAATTATATGTAATAGTGAAGCCAC AGATGTATTACATATAATTATTACATCTATGCCTACTGCCTCGGA
Prdm8.2766	TGCTGTTGACAGTGAGCGCAAGGACAAAGTTTATTTTAGATAGTGAAGCCAC AGATGTATCTAAAATAAACTTTGTCCTTATGCCTACTGCCTCGGA
Prdm8.2835	TGCTGTTGACAGTGAGCGAAGCTGTGTTCTTACAAAAATATAGTGAAGCCAC AGATGTATATTTTTGTAAGAACACAGCTGTGCCTACTGCCTCGGA
Prdm8.2772	TGCTGTTGACAGTGAGCGAAAAGTTTATTTTAGATATAAATAGTGAAGCCAC AGATGTATTTATATCTAAAATAAACTTTGTGCCTACTGCCTCGGA
Prdm8.2832	TGCTGTTGACAGTGAGCGCGACAGCTGTGTTCTTACAAAATAGTGAAGCCAC AGATGTATTTTGTAAGAACACAGCTGTCATGCCTACTGCCTCGGA
Prdm8.2712	TGCTGTTGACAGTGAGCGCACGTTTCTACTTAAATTATTATAGTGAAGCCAC AGATGTATAATAATTTAAGTAGAAACGTATGCCTACTGCCTCGGA
Raver1.3399	TGCTGTTGACAGTGAGCGAATAGACCTTTGCTGAATTAAATAGTGAAGCCAC AGATGTATTTAATTCAGCAAAGGTCTATGTGCCTACTGCCTCGGA
Raver1.3411	TGCTGTTGACAGTGAGCGATGAATTAAAGTTTGTAAGTAATAGTGAAGCCAC AGATGTATTACTTACAAACTTTAATTCAGTGCCTACTGCCTCGGA
Raver1.2695	TGCTGTTGACAGTGAGCGACAGGATGATTATCACATCTACTAGTGAAGCCAC AGATGTAGTAGATGTGATAATCATCCTGGTGCCTACTGCCTCGGA
Raver1.3198	TGCTGTTGACAGTGAGCGATACCATGTTCTTTCAAACCCATAGTGAAGCCAC AGATGTATGGGTTTGAAAGAACATGGTAGTGCCTACTGCCTCGGA
Raver1.1780	TGCTGTTGACAGTGAGCGATGACAGCTATGGCTTTGATTATAGTGAAGCCAC AGATGTATAATCAAAGCCATAGCTGTCAGTGCCTACTGCCTCGGA
Raver1.2237	TGCTGTTGACAGTGAGCGAAAGCGGAAGAGGATTTTCTAATAGTGAAGCCA CAGATGTATTAGAAAATCCTCTTCCGCTTCTGCCTACTGCCTCGGA
Rexo4.2246	TGCTGTTGACAGTGAGCGCTCCTAGAACTGTCAAAAATAATAGTGAAGCCAC AGATGTATTATTTTTGACAGTTCTAGGAATGCCTACTGCCTCGGA

Rexo4.2247	TGCTGTTGACAGTGAGCGCCCTAGAACTGTCAAAAATAAATAGTGAAGCCAC AGATGTATTTATTTTTGACAGTTCTAGGATGCCTACTGCCTCGGA
Rexo4.1129	TGCTGTTGACAGTGAGCGCCAGGGACACTCAGAAATTCAATAGTGAAGCCA CAGATGTATTGAATTTCTGAGTGTCCCTGATGCCTACTGCCTCGGA
Rexo4.1130	TGCTGTTGACAGTGAGCGAAGGGACACTCAGAAATTCAAATAGTGAAGCCA CAGATGTATTTGAATTTCTGAGTGTCCCTGTGCCTACTGCCTCGGA
Rexo4.2204	TGCTGTTGACAGTGAGCGCCAGAGGACACTTGTTCACTAATAGTGAAGCCAC AGATGTATTAGTGAACAAGTGTCCTCTGATGCCTACTGCCTCGGA
Rexo4.295	TGCTGTTGACAGTGAGCGACCCCGAGAACTTTTCTCAAAATAGTGAAGCCAC AGATGTATTTTGAGAAAAGTTCTCGGGGGTGCCTACTGCCTCGGA
Rilp.1350	TGCTGTTGACAGTGAGCGCACCACACTTTACAGACATAAATAGTGAAGCCAC AGATGTATTTATGTCTGTAAAGTGTGGTATGCCTACTGCCTCGGA
Rilp.1451	TGCTGTTGACAGTGAGCGCGGTGACCAGTGCAATAAATAATAGTGAAGCCA CAGATGTATTATTTATTGCACTGGTCACCTTGCCTACTGCCTCGGA
Rilp.1349	TGCTGTTGACAGTGAGCGATACCACACTTTACAGACATAATAGTGAAGCCAC AGATGTATTATGTCTGTAAAGTGTGGTACTGCCTACTGCCTCGGA
Rilp.1165	TGCTGTTGACAGTGAGCGATAGTCCCCTGCTTTCAATGAATAGTGAAGCCAC AGATGTATTCATTGAAAGCAGGGGACTAGTGCCTACTGCCTCGGA
Rilp.1450	TGCTGTTGACAGTGAGCGAAGGTGACCAGTGCAATAAATATAGTGAAGCCA CAGATGTATATTTATTGCACTGGTCACCTGTGCCTACTGCCTCGGA
Rilp.1448	TGCTGTTGACAGTGAGCGCGCAGGTGACCAGTGCAATAAATAGTGAAGCCA CAGATGTATTTATTGCACTGGTCACCTGCATGCCTACTGCCTCGGA
Sall4.33	TGCTGTTGACAGTGAGCGAGCAATGCATTATTGATAATAATAGTGAAGCCAC AGATGTATTATTATCAATAATGCATTGCGTGCCTACTGCCTCGGA
Sall4.2669	TGCTGTTGACAGTGAGCGCAAGGTCTGAAAATAAAGTGAATAGTGAAGCCA CAGATGTATTCACTTTATTTTCAGACCTTATGCCTACTGCCTCGGA
Sall4.1343	TGCTGTTGACAGTGAGCGCTGGGACAAATTATATTCTAGATAGTGAAGCCAC AGATGTATCTAGAATATAATTTGTCCCAATGCCTACTGCCTCGGA
Sall4.2434	TGCTGTTGACAGTGAGCGACACCTGATTTTAGGATCCAAATAGTGAAGCCAC AGATGTATTTGGATCCTAAAATCAGGTGCTGCCTACTGCCTCGGA
Sall4.2664	TGCTGTTGACAGTGAGCGCACTGTAAGGTCTGAAAATAAATAGTGAAGCCAC AGATGTATTTATTTTCAGACCTTACAGTATGCCTACTGCCTCGGA
Sall4.9	TGCTGTTGACAGTGAGCGAAGCTGTCAGAGCTCATGATAATAGTGAAGCCAC AGATGTATTATCATGAGCTCTGACAGCTGTGCCTACTGCCTCGGA
Sim2.3119	TGCTGTTGACAGTGAGCGCGCATCTCATTTTTAAATTTAATAGTGAAGCCAC AGATGTATTAAATTTAAAAATGAGATGCATGCCTACTGCCTCGGA
Sim2.2648	TGCTGTTGACAGTGAGCGACACGATATAAATTAATTTATATAGTGAAGCCAC AGATGTATATAAATTAATTTATATCGTGCTGCCTACTGCCTCGGA
Sim2.913	TGCTGTTGACAGTGAGCGAACAGGCAACAGTATTTATGAATAGTGAAGCCAC AGATGTATTCATAAATACTGTTGCCTGTGTGCCTACTGCCTCGGA

Sim2.3626	TGCTGTTGACAGTGAGCGCACCAGCAGTGTTTAAAAATAATAGTGAAGCCAC AGATGTATTATTTTTAAACACTGCTGGTATGCCTACTGCCTCGGA
Sim2.3606	TGCTGTTGACAGTGAGCGAGAGATCCTTTTTATAAACTTATAGTGAAGCCAC AGATGTATAAGTTTATAAAAAGGATCTCCTGCCTACTGCCTCGGA
Sim2.2668	TGCTGTTGACAGTGAGCGCACAGAGACAACTATTTTAATATAGTGAAGCCAC AGATGTATATTAAAATAGTTGTCTCTGTATGCCTACTGCCTCGGA
Sowahb.3440	TGCTGTTGACAGTGAGCGAAAGCTGCTAAGTATAACTTTATAGTGAAGCCAC AGATGTATAAAGTTATACTTAGCAGCTTCTGCCTACTGCCTCGGA
Sowahb.3776	TGCTGTTGACAGTGAGCGCTTGTATCAAGCTGTTATTATATAGTGAAGCCAC AGATGTATATAATAACAGCTTGATACAATTGCCTACTGCCTCGGA
Sowahb.3750	TGCTGTTGACAGTGAGCGAAGCAGATTGATTGTTATGTTATAGTGAAGCCAC AGATGTATAACATAACAATCAATCTGCTGTGCCTACTGCCTCGGA
Sowahb.3605	TGCTGTTGACAGTGAGCGCTGTGTAAATTCTGAATGTAAATAGTGAAGCCAC AGATGTATTTACATTCAGAATTTACACATTGCCTACTGCCTCGGA
Sowahb.3501	TGCTGTTGACAGTGAGCGACCTGAAGTACTGAATGTCTAATAGTGAAGCCAC AGATGTATTAGACATTCAGTACTTCAGGGTGCCTACTGCCTCGGA
Sowahb.3058	TGCTGTTGACAGTGAGCGCACAGTGCTTGGTGTAAACTGATAGTGAAGCCAC AGATGTATCAGTTTACACCAAGCACTGTATGCCTACTGCCTCGGA
Spib.233	TGCTGTTGACAGTGAGCGACCAAACTGTTCAGCTCTGTTATAGTGAAGCCAC AGATGTATAACAGAGCTGAACAGTTTGGGTGCCTACTGCCTCGGA
Spib.98	TGCTGTTGACAGTGAGCGCCAGCTGCAAGCCCTTCAGTTATAGTGAAGCCAC AGATGTATAACTGAAGGGCTTGCAGCTGTTGCCTACTGCCTCGGA
Spib.689	TGCTGTTGACAGTGAGCGACAAGCGCATGACGTATCAGAATAGTGAAGCCA CAGATGTATTCTGATACGTCATGCGCTTGCTGCCTACTGCCTCGGA
Spib.417	TGCTGTTGACAGTGAGCGCTCAGAGGAAGAAGACATTATGTAGTGAAGCCA CAGATGTACATAATGTCTTCTTCCTCTGATTGCCTACTGCCTCGGA
Spib.743	TGCTGTTGACAGTGAGCGAAGGCGAAATCCGCAAGGTCAATAGTGAAGCCA CAGATGTATTGACCTTGCGGATTTCGCCTGTGCCTACTGCCTCGGA
Spib.527	TGCTGTTGACAGTGAGCGAACGCAAGAAGCTGCGCCTGTATAGTGAAGCCA CAGATGTATACAGGCGCAGCTTCTTGCGTGTGCCTACTGCCTCGGA
Tbx19.309	TGCTGTTGACAGTGAGCGCCCCTACAGAGAAGCAACTTCATAGTGAAGCCAC AGATGTATGAAGTTGCTTCTCTGTAGGGTTGCCTACTGCCTCGGA
Tbx19.1574	TGCTGTTGACAGTGAGCGACAGTAGAGAACTGTCTGTATTTAGTGAAGCCAC AGATGTAAATACAGACAGTTCTCTACTGGTGCCTACTGCCTCGGA
Tbx19.365	TGCTGTTGACAGTGAGCGCTCAAGGAAGTCACTAATGAAATAGTGAAGCCA CAGATGTATTTCATTAGTGACTTCCTTGAATGCCTACTGCCTCGGA
Tbx19.1768	TGCTGTTGACAGTGAGCGATGCTTTCTCTGAGTTGATTTATAGTGAAGCCAC AGATGTATAAATCAACTCAGAGAAAGCAGTGCCTACTGCCTCGGA
Tbx19.813	TGCTGTTGACAGTGAGCGAAGCTCTCAAAATCAAGTACAATAGTGAAGCCAC AGATGTATTGTACTTGATTTTGAGAGCTGTGCCTACTGCCTCGGA

Tbx19.843	TGCTGTTGACAGTGAGCGACAAGGCCTTCTTGGATGCCAATAGTGAAGCCAC AGATGTATTGGCATCCAAGAAGGCCTTGGTGCCTACTGCCTCGGA
Tbx3.691	TGCTGTTGACAGTGAGCGCCGCTGTTGAATTTTTTAATTATAGTGAAGCCAC AGATGTATAATTAAAAAATTCAACAGCGATGCCTACTGCCTCGGA
Tbx3.3013	TGCTGTTGACAGTGAGCGACAGTGCACTTTGTTAGATGTATAGTGAAGCCAC AGATGTATACATCTAACAAAGTGCACTGGTGCCTACTGCCTCGGA
Tbx3.657	TGCTGTTGACAGTGAGCGCTGCCCTCTTTTTTATTTTAAATAGTGAAGCCACA GATGTATTTAAAATAAAAAAGAGGGCAATGCCTACTGCCTCGGA
Tbx3.694	TGCTGTTGACAGTGAGCGATGTTGAATTTTTTAATTATTATAGTGAAGCCACA GATGTATAATAATTAAAAAATTCAACAGTGCCTACTGCCTCGGA
Tbx3.3014	TGCTGTTGACAGTGAGCGAAGTGCACTTTGTTAGATGTAATAGTGAAGCCAC AGATGTATTACATCTAACAAAGTGCACTGTGCCTACTGCCTCGGA
Tbx3.3781	TGCTGTTGACAGTGAGCGATTGGACCATTAGTTCTTTTAATAGTGAAGCCAC AGATGTATTAAAAGAACTAATGGTCCAACTGCCTACTGCCTCGGA
Tcea3.1084	TGCTGTTGACAGTGAGCGCGACCACTTTTGTCTTATGCAATAGTGAAGCCAC AGATGTATTGCATAAGACAAAAGTGGTCATGCCTACTGCCTCGGA
Tcea3.1027	TGCTGTTGACAGTGAGCGACAAGAAGAAGAATTGTACCTATAGTGAAGCCA CAGATGTATAGGTACAATTCTTCTTCTTGCTGCCTACTGCCTCGGA
Tcea3.1030	TGCTGTTGACAGTGAGCGCGAAGAAGAATTGTACCTATAATAGTGAAGCCAC AGATGTATTATAGGTACAATTCTTCTTCTTGCCTACTGCCTCGGA
Tcea3.1029	TGCTGTTGACAGTGAGCGCAGAAGAAGAATTGTACCTATATAGTGAAGCCAC AGATGTATATAGGTACAATTCTTCTTCTTTGCCTACTGCCTCGGA
Tcea3.1023	TGCTGTTGACAGTGAGCGCAGTGCAAGAAGAAGAATTGTATAGTGAAGCCA CAGATGTATACAATTCTTCTTCTTGCACTTTGCCTACTGCCTCGGA
Tcea3.879	TGCTGTTGACAGTGAGCGACAGAGCTCATTGCCAAGATGATAGTGAAGCCAC AGATGTATCATCTTGGCAATGAGCTCTGGTGCCTACTGCCTCGGA
Thbs3.1162	TGCTGTTGACAGTGAGCGACAATGACATTGATGAATGTAATAGTGAAGCCAC AGATGTATTACATTCATCAATGTCATTGCTGCCTACTGCCTCGGA
Thbs3.659	TGCTGTTGACAGTGAGCGACTGAGTGAATGTCCATTCCAATAGTGAAGCCAC AGATGTATTGGAATGGACATTCACTCAGGTGCCTACTGCCTCGGA
Thbs3.922	TGCTGTTGACAGTGAGCGCCTGTATGGAAGTGTATGAGTATAGTGAAGCCAC AGATGTATACTCATACACTTCCATACAGTTGCCTACTGCCTCGGA
Thbs3.2143	TGCTGTTGACAGTGAGCGCGGGTGATGTCTGTGAAGATGATAGTGAAGCCAC AGATGTATCATCTTCACAGACATCACCCATGCCTACTGCCTCGGA
Thbs3.1529	TGCTGTTGACAGTGAGCGACAAGAAGATGCTGATAACGATTAGTGAAGCCA CAGATGTAATCGTTATCAGCATCTTCTTGCTGCCTACTGCCTCGGA
Thbs3.2100	TGCTGTTGACAGTGAGCGATGGTACCCAATCCTAATCAGATAGTGAAGCCAC AGATGTATCTGATTAGGATTGGGTACCAGTGCCTACTGCCTCGGA
Tmprss4.1996	TGCTGTTGACAGTGAGCGCGACCAGTCAGTTCTAATGTAATAGTGAAGCCAC AGATGTATTACATTAGAACTGACTGGTCTTGCCTACTGCCTCGGA

Tmprss4.2055	TGCTGTTGACAGTGAGCGATCTGTGGTCAATATCATTAAATAGTGAAGCCAC AGATGTATTTAATGATATTGACCACAGAGTGCCTACTGCCTCGGA
Tmprss4.440	TGCTGTTGACAGTGAGCGCCAAGGTGATTCTGGATAAATATAGTGAAGCCAC AGATGTATATTTATCCAGAATCACCTTGATGCCTACTGCCTCGGA
Tmprss4.2015	TGCTGTTGACAGTGAGCGCAAGGTGTATATTTTAGTGTCATAGTGAAGCCAC AGATGTATGACACTAAAATATACACCTTATGCCTACTGCCTCGGA
Tmprss4.1995	TGCTGTTGACAGTGAGCGAAGACCAGTCAGTTCTAATGTATAGTGAAGCCAC AGATGTATACATTAGAACTGACTGGTCTCTGCCTACTGCCTCGGA
Tmprss4.2054	TGCTGTTGACAGTGAGCGACTCTGTGGTCAATATCATTAATAGTGAAGCCAC AGATGTATTAATGATATTGACCACAGAGGTGCCTACTGCCTCGGA
Trip10.1924	TGCTGTTGACAGTGAGCGATTGTACATATTTGTCATTTAATAGTGAAGCCAC AGATGTATTAAATGACAAATATGTACAACTGCCTACTGCCTCGGA
Trip10.1973	TGCTGTTGACAGTGAGCGACCAAGTTTTGTTTTATATTAATAGTGAAGCCAC AGATGTATTAATATAAAACAAAACTTGGCTGCCTACTGCCTCGGA
Trip10.1974	TGCTGTTGACAGTGAGCGACAAGTTTTGTTTTATATTAAATAGTGAAGCCAC AGATGTATTTAATATAAAACAAAACTTGGTGCCTACTGCCTCGGA
Trip10.1203	TGCTGTTGACAGTGAGCGCGAAGATGAAAGATGTATATGATAGTGAAGCCA CAGATGTATCATATACATCTTTCATCTTCTTGCCTACTGCCTCGGA
Trip10.1972	TGCTGTTGACAGTGAGCGCGCCAAGTTTTGTTTTATATTATAGTGAAGCCAC AGATGTATAATATAAAACAAAACTTGGCTTGCCTACTGCCTCGGA
Trip10.1200	TGCTGTTGACAGTGAGCGCGAAGAAGATGAAAGATGTATATAGTGAAGCCA CAGATGTATATACATCTTTCATCTTCTTCATGCCTACTGCCTCGGA
Vtcn1.558	TGCTGTTGACAGTGAGCGAAGAGATAAATGTGGACTATAATAGTGAAGCCA CAGATGTATTATAGTCCACATTTATCTCTGTGCCTACTGCCTCGGA
Vtcn1.763	TGCTGTTGACAGTGAGCGAACATACTCCTGTATGATTGAATAGTGAAGCCAC AGATGTATTCAATCATACAGGAGTATGTGTGCCTACTGCCTCGGA
Vtcn1.690	TGCTGTTGACAGTGAGCGACAGCTTTGAGTTGAACTCTGATAGTGAAGCCAC AGATGTATCAGAGTTCAACTCAAAGCTGGTGCCTACTGCCTCGGA
Vtcn1.557	TGCTGTTGACAGTGAGCGACAGAGATAAATGTGGACTATATAGTGAAGCCA CAGATGTATATAGTCCACATTTATCTCTGGTGCCTACTGCCTCGGA
Vtcn1.373	TGCTGTTGACAGTGAGCGCCAGCAGCATGAGATGTTCAGATAGTGAAGCCAC AGATGTATCTGAACATCTCATGCTGCTGTTGCCTACTGCCTCGGA
Vtcn1.310	TGCTGTTGACAGTGAGCGATGGCTGAAAGAAGGCATCAAATAGTGAAGCCA CAGATGTATTTGATGCCTTCTTTCAGCCACTGCCTACTGCCTCGGA
Zc3h12d.1515	TGCTGTTGACAGTGAGCGACAGCTCCTATTGTTATCTGAATAGTGAAGCCAC AGATGTATTCAGATAACAATAGGAGCTGGTGCCTACTGCCTCGGA
Zc3h12d.890	TGCTGTTGACAGTGAGCGACAACACTAAGCAACTTCCTAATAGTGAAGCCAC AGATGTATTAGGAAGTTGCTTAGTGTTGGTGCCTACTGCCTCGGA
Zc3h12d.1514	TGCTGTTGACAGTGAGCGACCAGCTCCTATTGTTATCTGATAGTGAAGCCAC AGATGTATCAGATAACAATAGGAGCTGGGTGCCTACTGCCTCGGA

Zc3h12d.177	TGCTGTTGACAGTGAGCGACAAGATGGAATTTTTCCAGAATAGTGAAGCCAC AGATGTATTCTGGAAAAATTCCATCTTGCTGCCTACTGCCTCGGA
Zc3h12d.178	TGCTGTTGACAGTGAGCGAAAGATGGAATTTTTCCAGAAATAGTGAAGCCAC AGATGTATTTCTGGAAAAATTCCATCTTGTGCCTACTGCCTCGGA
Zc3h12d.171	TGCTGTTGACAGTGAGCGCTCGGAGCAAGATGGAATTTTTTAGTGAAGCCAC AGATGTAAAAAATTCCATCTTGCTCCGATTGCCTACTGCCTCGGA

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorgelegte Dissertation "Identification and validation of a plasticity driver of Combined Hepatocellular-Cholangiocarcinoma using functional interspecies comparison" selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Des Weiteren bestätige ich, dass ich an keiner anderen Stelle ein Prüfungsverfahren beantragt bzw. die Dissertation in dieser oder einer anderen Form bereits als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.

Heidelberg, den 06.10.2023

Noujan Ganjian

Literature

1.   Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA Cancer J Clin, 2018. 68(6): p. 394-424.

2.   Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. 71(3): p. 209-249.

3.   Llovet, J.M., et al., *Hepatocellular carcinoma.* Nat Rev Dis Primers, 2016. 2: p. 16018.

4.   Villanueva, A., *Hepatocellular Carcinoma. Reply.* N Engl J Med, 2019. 381(1): p. e2.

5.   (IARC), T.W.-T.I.A.f.R.o.C., *World Cancer Report: Cancer Research for Cancer Prevention on the occasion of the 20th anniversary of World Cancer Day.* 2020.

6.   Sia, D., et al., *Identification of an Immune-specific Class of Hepatocellular Carcinoma, Based on Molecular Features.* Gastroenterology, 2017. 153(3): p. 812-826.

7.   Rumgay, H., et al., *Global, regional and national burden of primary liver cancer by subtype.* Eur J Cancer, 2022. 161: p. 108-118.

8.   McGlynn, K.A., J.L. Petrick, and W.T. London, *Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability.* Clin Liver Dis, 2015. 19(2): p. 223-38.

9.   Liver, E.A.f.t.S.o.t., *EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma.* JOURNAL OF HEPATOLOGY, 2018. 69: p. 182–236.

10.   Marrero, J.A., et al., *Diagnosis, Staging, and Management of Hepatocellular Carcinoma: 2018 Practice Guidance by the American Association for the Study of Liver Diseases.* Hepatology, 2018. 68(2): p. 723-750.

11.   Singal, A.G. and H.B. El-Serag, *Hepatocellular Carcinoma From Epidemiology to Prevention: Translating Knowledge into Practice.* Clin Gastroenterol Hepatol, 2015. 13(12): p. 2140-51.

12.   El-Serag, H.B., *Hepatocellular carcinoma.* N Engl J Med, 2011. 365(12): p. 1118-27.

13.   Global Burden of Disease Liver Cancer, C., et al., *The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level: Results From the Global Burden of Disease Study 2015.* JAMA Oncol, 2017. 3(12): p. 1683-1691.

14.   Yuen, M.F., et al., *Hepatitis B virus infection.* Nat Rev Dis Primers, 2018. 4: p. 18035.

15.   Shi, J.F., et al., *Is it possible to halve the incidence of liver cancer in China by 2050?* Int J Cancer, 2021. 148(5): p. 1051-1065.

16.   Levrero, M. and J. Zucman-Rossi, *Mechanisms of HBV-induced hepatocellular carcinoma.* J Hepatol, 2016. 64(1 Suppl): p. S84-S101.

17.   Roudot-Thoraval, F., *Epidemiology of hepatitis C virus infection.* Clin Res Hepatol Gastroenterol, 2021. 45(3): p. 101596.

18.   Kanwal, F., et al., *Risk of Hepatocellular Cancer in HCV Patients Treated With Direct-Acting Antiviral Agents.* Gastroenterology, 2017. 153(4): p. 996-1005 e1.

19.   Estes, C., et al., *Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease.* Hepatology, 2018. 67(1): p. 123-133.

20.   Estes, C., et al., *Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030.* J Hepatol, 2018. 69(4): p. 896-904.

21.   Rich, N.E., A.C. Yopp, A.G. Singal, and C.C. Murphy, *Hepatocellular Carcinoma Incidence Is Decreasing Among Younger Adults in the United States.* Clin Gastroenterol Hepatol, 2020. 18(1): p. 242-248 e5.

22. Rich, N.E., et al., *Racial and Ethnic Differences in Presentation and Outcomes of Hepatocellular Carcinoma.* Clin Gastroenterol Hepatol, 2019. 17(3): p. 551-559 e1.

23. Kensler, T.W., B.D. Roebuck, G.N. Wogan, and J.D. Groopman, *Aflatoxin: a 50-year odyssey of mechanistic and translational toxicology.* Toxicol Sci, 2011. 120 Suppl 1(Suppl 1): p. S28-48.

24. Schulze, K., et al., *Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets.* Nat Genet, 2015. 47(5): p. 505-511.

25. Lee, Y.C., et al., *Meta-analysis of epidemiologic studies on cigarette smoking and liver cancer.* Int J Epidemiol, 2009. 38(6): p. 1497-511.

26. Simon, T.G., et al., *Association of Aspirin with Hepatocellular Carcinoma and Liver-Related Mortality.* N Engl J Med, 2020. 382(11): p. 1018-1028.

27. Hernandez-Gea, V. and S.L. Friedman, *Pathogenesis of liver fibrosis.* Annu Rev Pathol, 2011. 6: p. 425-56.

28. Capece, D., et al., *The inflammatory microenvironment in hepatocellular carcinoma: a pivotal role for tumor-associated macrophages.* Biomed Res Int, 2013. 2013: p. 187204.

29. Seki, E. and R.F. Schwabe, *Hepatic inflammation and fibrosis: functional links and key pathways.* Hepatology, 2015. 61(3): p. 1066-79.

30. Roncalli, M., et al., *Liver precancerous lesions and hepatocellular carcinoma: the histology report.* Dig Liver Dis, 2011. 43 Suppl 4: p. S361-72.

31. Torrecilla, S., et al., *Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma.* J Hepatol, 2017. 67(6): p. 1222-1231.

32. Karagozian, R., Z. Derdak, and G. Baffy, *Obesity-associated mechanisms of hepatocarcinogenesis.* Metabolism, 2014. 63(5): p. 607-17.

33. Schulze, K., J.C. Nault, and A. Villanueva, *Genetic profiling of hepatocellular carcinoma using next-generation sequencing.* J Hepatol, 2016. 65(5): p. 1031-1042.

34. Guichard, C., et al., *Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma.* Nat Genet, 2012. 44(6): p. 694-8.

35. Zehir, A., et al., *Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.* Nat Med, 2017. 23(6): p. 703-713.

36. Chiang, D.Y., et al., *Focal gains of VEGFA and molecular classification of hepatocellular carcinoma.* Cancer Res, 2008. 68(16): p. 6779-88.

37. Llovet, J.M., R. Montal, D. Sia, and R.S. Finn, *Molecular therapies and precision medicine for hepatocellular carcinoma.* Nat Rev Clin Oncol, 2018. 15(10): p. 599-616.

38. Zucman-Rossi, J., A. Villanueva, J.C. Nault, and J.M. Llovet, *Genetic Landscape and Biomarkers of Hepatocellular Carcinoma.* Gastroenterology, 2015. 149(5): p. 1226-1239 e4.

39. Hyman, D.M., B.S. Taylor, and J. Baselga, *Implementing Genome-Driven Oncology.* Cell, 2017. 168(4): p. 584-599.

40. Matsui, O., et al., *Hepatocelluar nodules in liver cirrhosis: hemodynamic evaluation (angiography-assisted CT) with special reference to multi-step hepatocarcinogenesis.* Abdom Imaging, 2011. 36(3): p. 264-72.

41. Llovet, J.M., C. Bru, and J. Bruix, *Prognosis of hepatocellular carcinoma: the BCLC staging classification.* Semin Liver Dis, 1999. 19(3): p. 329-38.

42. Llovet, J.M., et al., *Locoregional therapies in the era of molecular and immune treatments for hepatocellular carcinoma.* Nat Rev Gastroenterol Hepatol, 2021. 18(5): p. 293-313.

43. Llovet, J.M. and J. Bruix, *Systematic review of randomized trials for unresectable hepatocellular carcinoma: Chemoembolization improves survival.* Hepatology, 2003. 37(2): p. 429-42.

44. Salem, R., et al., *Y90 Radioembolization Significantly Prolongs Time to Progression Compared With Chemoembolization in Patients With Hepatocellular Carcinoma.* Gastroenterology, 2016. 151(6): p. 1155-1163 e2.

45. Heimbach, J.K., et al., *AASLD guidelines for the treatment of hepatocellular carcinoma.* Hepatology, 2018. 67(1): p. 358-380.

46. Vogel, A., et al., *Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up.* Ann Oncol, 2018. 29(Suppl 4): p. iv238-iv255.

47. Llovet, J.M., et al., *Trial Design and Endpoints in Hepatocellular Carcinoma: AASLD Consensus Conference.* Hepatology, 2021. 73 Suppl 1: p. 158-191.

48. Finn, R.S., et al., *Atezolizumab plus Bevacizumab in Unresectable Hepatocellular Carcinoma.* N Engl J Med, 2020. 382(20): p. 1894-1905.

49. Finn, R.S., et al., *IMbrave150: Updated overall survival (OS) data from a global, randomized, open-label phase III study of atezolizumab (atezo) plus bevacizumab (bev) versus sorafenib (sor) in patients (pts) with unresectable hepatocellular carcinoma (HCC).* Journal of Clinical Oncology, 2021. 39(3).

50. Wong, K.M., G.G. King, and W.P. Harris, *The Treatment Landscape of Advanced Hepatocellular Carcinoma.* Curr Oncol Rep, 2022. 24(7): p. 917-927.

51. Llovet, J.M., et al., *Sorafenib in advanced hepatocellular carcinoma.* N Engl J Med, 2008. 359(4): p. 378-90.

52. Kudo, M., et al., *Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial.* Lancet, 2018. 391(10126): p. 1163-1173.

53. El-Khoueiry, A.B., et al., *Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial.* Lancet, 2017. 389(10088): p. 2492-2502.

54. Finn, R.S., et al., *Pembrolizumab As Second-Line Therapy in Patients With Advanced Hepatocellular Carcinoma in KEYNOTE-240: A Randomized, Double-Blind, Phase III Trial.* J Clin Oncol, 2020. 38(3): p. 193-202.

55. Singal, A.G., A. Pillai, and J. Tiro, *Early detection, curative treatment, and survival rates for hepatocellular carcinoma surveillance in patients with cirrhosis: a meta-analysis.* PLoS Med, 2014. 11(4): p. e1001624.

56. Banales, J.M., 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(5): p. 261-80.

57. Diggs, L.P., et al., *CD40-mediated immune cell activation enhances response to anti-PD-1 in murine intrahepatic cholangiocarcinoma.* J Hepatol, 2021. 74(5): p. 1145-1154.

58. de Jong, M.C., et al., *Intrahepatic cholangiocarcinoma: an international multi-institutional analysis of prognostic factors and lymph node assessment.* J Clin Oncol, 2011. 29(23): p. 3140-5.

59. Rizvi, S., et al., *Cholangiocarcinoma - evolving concepts and therapeutic strategies.* Nat Rev Clin Oncol, 2018. 15(2): p. 95-111.

60. Chong, D.Q. and A.X. Zhu, *The landscape of targeted therapies for cholangiocarcinoma: current status and emerging targets.* Oncotarget, 2016. 7(29): p. 46750-46767.

61. Sripa, B., et al., *Liver fluke induces cholangiocarcinoma.* PLoS Med, 2007. 4(7): p. e201.

62. Ong, C.K., et al., *Exome sequencing of liver fluke-associated cholangiocarcinoma.* Nat Genet, 2012. 44(6): p. 690-3.

63. Rizvi, S. and G.J. Gores, *Pathogenesis, diagnosis, and management of cholangiocarcinoma.* Gastroenterology, 2013. 145(6): p. 1215-29.

64. Lipsett, P.A., et al., *Choledochal cyst disease. A changing pattern of presentation.* Ann Surg, 1994. 220(5): p. 644-52.

65. Adenugba, A., et al., *Polychlorinated biphenyls in bile of patients with biliary tract cancer.* Chemosphere, 2009. 76(6): p. 841-6.

66. Lipshutz, G.S., T.V. Brennan, and R.S. Warren, *Thorotrast-induced liver neoplasia: a collective review.* J Am Coll Surg, 2002. 195(5): p. 713-8.

67. Tyson, G.L. and H.B. El-Serag, *Risk factors for cholangiocarcinoma.* Hepatology, 2011. 54(1): p. 173-84.

68. Bergquist, A., H. Glaumann, B. Persson, and U. Broome, *Risk factors and clinical presentation of hepatobiliary carcinoma in patients with primary sclerosing cholangitis: a case-control study.* Hepatology, 1998. 27(2): p. 311-6.

69. Charbotel, B., B. Fervers, and J.P. Droz, *Occupational exposures in rare cancers: A critical review of the literature.* Crit Rev Oncol Hematol, 2014. 90(2): p. 99-134.

70. Fava, G., et al., *Molecular pathology of biliary tract cancers.* Cancer Lett, 2007. 250(2): p. 155-67.

71. Sirica, A.E., *Cholangiocarcinoma: molecular targeting strategies for chemoprevention and therapy.* Hepatology, 2005. 41(1): p. 5-15.

72. Jiao, Y., et al., *Exome sequencing identifies frequent inactivating mutations in BAP1, ARID1A and PBRM1 in intrahepatic cholangiocarcinomas.* Nat Genet, 2013. 45(12): p. 1470-1473.

73. Rizvi, S., M.J. Borad, T. Patel, and G.J. Gores, *Cholangiocarcinoma: molecular pathways and therapeutic opportunities.* Semin Liver Dis, 2014. 34(4): p. 456-64.

74. Razumilava, N. and G.J. Gores, *Cholangiocarcinoma.* Lancet, 2014. 383(9935): p. 2168-79.

75. Yoon, J.H., et al., *Bile acids induce cyclooxygenase-2 expression via the epidermal growth factor receptor in a human cholangiocarcinoma cell line.* Gastroenterology, 2002. 122(4): p. 985-93.

76. Yoon, J.H., et al., *Oxysterols induce cyclooxygenase-2 expression in cholangiocytes: implications for biliary tract carcinogenesis.* Hepatology, 2004. 39(3): p. 732-8.

77. Kiguchi, K., et al., *Constitutive expression of ErbB-2 in gallbladder epithelium results in development of adenocarcinoma.* Cancer Res, 2001. 61(19): p. 6971-6.

78. Park, J., L. Tadlock, G.J. Gores, and T. Patel, *Inhibition of interleukin 6-mediated mitogen-activated protein kinase activation attenuates growth of a cholangiocarcinoma cell line.* Hepatology, 1999. 30(5): p. 1128-33.

79. Dwyer, J.R., et al., *Oxysterols are novel activators of the hedgehog signaling pathway in pluripotent mesenchymal cells.* J Biol Chem, 2007. 282(12): p. 8959-68.

80. Nachtergaele, S., et al., *Oxysterols are allosteric activators of the oncoprotein Smoothened.* Nat Chem Biol, 2012. 8(2): p. 211-20.

81. Geisler, F., et al., *Liver-specific inactivation of Notch2, but not Notch1, compromises intrahepatic bile duct development in mice.* Hepatology, 2008. 48(2): p. 607-16.

82. Zong, Y., et al., *Notch signaling controls liver development by regulating biliary differentiation.* Development, 2009. 136(10): p. 1727-39.

83. Singrang, N., et al., *NOTCH1 regulates the viability of cholangiocarcinoma cells via 14-3-3 theta.* J Cell Commun Signal, 2019. 13(2): p. 245-254.

84. Zhou, Q., et al., *The roles of Notch1 expression in the migration of intrahepatic cholangiocarcinoma.* BMC Cancer, 2013. 13: p. 244.

85. Goeppert, B., et al., *BRAF V600E-specific immunohistochemistry reveals low mutation rates in biliary tract cancer and restriction to intrahepatic cholangiocarcinoma.* Mod Pathol, 2014. 27(7): p. 1028-34.

86. Dankner, M., et al., *Classifying BRAF alterations in cancer: new rational therapeutic strategies for actionable mutations.* Oncogene, 2018. 37(24): p. 3183-3199.

87. Rizvi, S., et al., *A Hippo and Fibroblast Growth Factor Receptor Autocrine Pathway in Cholangiocarcinoma.* J Biol Chem, 2016. 291(15): p. 8031-47.

88. Arai, Y., et al., *Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma.* Hepatology, 2014. 59(4): p. 1427-34.

89. Chan-On, W., et al., *Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers.* Nat Genet, 2013. 45(12): p. 1474-8.

90. Saborowski, A., et al., *Mouse model of intrahepatic cholangiocarcinoma validates FIG-ROS as a potent fusion oncogene and therapeutic target.* Proc Natl Acad Sci U S A, 2013. 110(48): p. 19513-8.

91. Morales-Ruiz, M., A. Santel, J. Ribera, and W. Jimenez, *The Role of Akt in Chronic Liver Disease and Liver Regeneration.* Semin Liver Dis, 2017. 37(1): p. 11-16.

92. Valle, J.W., et al., *Cisplatin and gemcitabine for advanced biliary tract cancer: a meta-analysis of two randomised trials.* Ann Oncol, 2014. 25(2): p. 391-8.

93. Oh, D.Y., et al., *Gemcitabine and cisplatin plus durvalumab with or without tremelimumab in chemotherapy-naive patients with advanced biliary tract cancer: an open-label, single-centre, phase 2 study.* Lancet Gastroenterol Hepatol, 2022. 7(6): p. 522-532.

94. Lamarca, A., et al., *Second-line FOLFOX chemotherapy versus active symptom control for advanced biliary tract cancer (ABC-06): a phase 3, open-label, randomised, controlled trial.* Lancet Oncol, 2021. 22(5): p. 690-701.

95. Oscar Briz, M.J.P., Jose J. G. Marin, *Further understanding of mechanisms involved in liver cancer chemoresistance* Hepatoma Research, 2017.

96. Marin, J.J.G., et al., *Chemoresistance and chemosensitization in cholangiocarcinoma.* Biochim Biophys Acta Mol Basis Dis, 2018. 1864(4 Pt B): p. 1444-1453.

97. Abou-Alfa, G.K., et al., *Ivosidenib in IDH1-mutant, chemotherapy-refractory cholangiocarcinoma (ClarIDHy): a multicentre, randomised, double-blind, placebo-controlled, phase 3 study.* Lancet Oncol, 2020. 21(6): p. 796-807.

98. Farshidfar, F., et al., *Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles.* Cell Rep, 2017. 18(11): p. 2780-2794.

99. Abou-Alfa, G.K., et al., *Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study.* Lancet Oncol, 2020. 21(5): p. 671-684.

100. Javle, M., et al., *Infigratinib (BGJ398) in previously treated patients with advanced or metastatic cholangiocarcinoma with FGFR2 fusions or rearrangements: mature results from a multicentre, open-label, single-arm, phase 2 study.* Lancet Gastroenterol Hepatol, 2021. 6(10): p. 803-815.

101. Tooke, S.M., H.C. Amstutz, and A.K. Hedley, *Results of transtrochanteric rotational osteotomy for femoral head osteonecrosis.* Clin Orthop Relat Res, 1987(224): p. 150-7.

102. Pranzatelli, M.R., *The comparative pharmacology of the behavioral syndromes induced by TRH and by 5-HT in the rat.* Gen Pharmacol, 1988. 19(2): p. 205-11.

103. Kelley, R.K., J. Bridgewater, G.J. Gores, and A.X. Zhu, *Systemic therapies for intrahepatic cholangiocarcinoma.* J Hepatol, 2020. 72(2): p. 353-363.

104. Verlingue, L., et al., *Precision medicine for patients with advanced biliary tract cancers: An effective strategy within the prospective MOSCATO-01 trial.* Eur J Cancer, 2017. 87: p. 122-130.

105. Tomczak, A., et al., *Precision oncology for intrahepatic cholangiocarcinoma in clinical practice.* Br J Cancer, 2022. 127(9): p. 1701-1708.

106. Friedman, S.L., *Signalling pathways in liver disease.* Signaling Pathways in Liver Diseases, 2015: p. 85–96.

107. Henderson, N.C., et al., *Targeting of alphav integrin identifies a core molecular pathway that regulates fibrosis in several organs.* Nat Med, 2013. 19(12): p. 1617-24.

108. Chen, R.J., H.H. Wu, and Y.J. Wang, *Strategies to prevent and reverse liver fibrosis in humans and laboratory animals.* Arch Toxicol, 2015. 89(10): p. 1727-50.

109. Affo, S., L.X. Yu, and R.F. Schwabe, *The Role of Cancer-Associated Fibroblasts and Fibrosis in Liver Cancer.* Annu Rev Pathol, 2017. 12: p. 153-186.

110. Stover, D.G., B. Bierie, and H.L. Moses, *A delicate balance: TGF-beta and the tumor microenvironment.* J Cell Biochem, 2007. 101(4): p. 851-61.

111. Meindl-Beinker, N.M., K. Matsuzaki, and S. Dooley, *TGF-beta signaling in onset and progression of hepatocellular carcinoma.* Dig Dis, 2012. 30(5): p. 514-23.

112. Dooley, S. and P. ten Dijke, *TGF-beta in progression of liver disease.* Cell Tissue Res, 2012. 347(1): p. 245-56.

113. Wang, H.M., et al., *Liver stiffness measurement as an alternative to fibrotic stage in risk assessment of hepatocellular carcinoma incidence for chronic hepatitis C patients.* Liver Int, 2013. 33(5): p. 756-61.

114. Schrader, J., et al., *Matrix stiffness modulates proliferation, chemotherapeutic response, and dormancy in hepatocellular carcinoma cells.* Hepatology, 2011. 53(4): p. 1192-205.

115. Dupont, S., et al., *Role of YAP/TAZ in mechanotransduction.* Nature, 2011. 474(7350): p. 179-83.

116. Williams, M.J., A.D. Clouston, and S.J. Forbes, *Links between hepatic fibrosis, ductular reaction, and progenitor cell expansion.* Gastroenterology, 2014. 146(2): p. 349-56.

117. Hernandez-Gea, V., S. Toffanin, S.L. Friedman, and J.M. Llovet, *Role of the microenvironment in the pathogenesis and treatment of hepatocellular carcinoma.* Gastroenterology, 2013. 144(3): p. 512-27.

118. Gascard, P. and T.D. Tlsty, *Carcinoma-associated fibroblasts: orchestrating the composition of malignancy.* Genes Dev, 2016. 30(9): p. 1002-19.

119. Sirica, A.E., *The role of cancer-associated myofibroblasts in intrahepatic cholangiocarcinoma.* Nat Rev Gastroenterol Hepatol, 2011. 9(1): p. 44-54.

120. Haga, H., et al., *Tumour cell-derived extracellular vesicles interact with mesenchymal stem cells to modulate the microenvironment and enhance cholangiocarcinoma growth.* J Extracell Vesicles, 2015. 4: p. 24900.

121. Yang, X.W., et al., *STAT3 overexpression promotes metastasis in intrahepatic cholangiocarcinoma and correlates negatively with surgical outcome.* Oncotarget, 2017. 8(5): p. 7710-7721.

122. Isomoto, H., et al., *Interleukin 6 upregulates myeloid cell leukemia-1 expression through a STAT3 pathway in cholangiocarcinoma cells.* Hepatology, 2005. 42(6): p. 1329-38.

123. Huyen, N.T., V. Prachayasittikul, and W. Chan-On, *Anoikis-resistant cholangiocarcinoma cells display aggressive characteristics and increase STAT3 activation.* J Hepatobiliary Pancreat Sci, 2016. 23(7): p. 397-405.

124. Shinmura, R., et al., *Cirrhotic nodules: association between MR imaging signal intensity and intranodular blood supply.* Radiology, 2005. 237(2): p. 512-9.
125. Connell, L.C., J.J. Harding, J. Shia, and G.K. Abou-Alfa, *Combined intrahepatic cholangiocarcinoma and hepatocellular carcinoma.* Chin Clin Oncol, 2016. 5(5): p. 66.
126. Maximin, S., et al., *Current update on combined hepatocellular-cholangiocarcinoma.* Eur J Radiol Open, 2014. 1: p. 40-8.
127. Lee, J.H., et al., *Long-term prognosis of combined hepatocellular and cholangiocarcinoma after curative resection comparison with hepatocellular carcinoma and cholangiocarcinoma.* J Clin Gastroenterol, 2011. 45(1): p. 69-75.
128. Xue, R., et al., *Genomic and Transcriptomic Profiling of Combined Hepatocellular and Intrahepatic Cholangiocarcinoma Reveals Distinct Molecular Subtypes.* Cancer Cell, 2019. 35(6): p. 932-947 e8.
129. Roskams, T., *Liver stem cells and their implication in hepatocellular and cholangiocarcinoma.* Oncogene, 2006. 25(27): p. 3818-22.
130. Pikarsky, E., *Neighbourhood deaths cause a switch in cancer subtype.* Nature, 2018. 562(7725): p. 45-46.
131. Seehawer, M., et al., *Necroptosis microenvironment directs lineage commitment in liver cancer.* Nature, 2018. 562(7725): p. 69-75.
132. Kowalik, M.A., et al., *Cytokeratin-19 positivity is acquired along cancer progression and does not predict cell origin in rat hepatocarcinogenesis.* Oncotarget, 2015. 6(36): p. 38749-63.
133. Mu, X., et al., *Hepatocellular carcinoma originates from hepatocytes and not from the progenitor/biliary compartment.* J Clin Invest, 2015. 125(10): p. 3891-903.
134. Fan, B., et al., *Cholangiocarcinomas can originate from hepatocytes in mice.* J Clin Invest, 2012. 122(8): p. 2911-5.
135. Malato, Y., et al., *Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration.* J Clin Invest, 2011. 121(12): p. 4850-60.
136. Espejel, S., et al., *Induced pluripotent stem cell-derived hepatocytes have the functional and proliferative capabilities needed for liver regeneration in mice.* J Clin Invest, 2010. 120(9): p. 3120-6.
137. Sekiya, S. and A. Suzuki, *Intrahepatic cholangiocarcinoma can arise from Notch-mediated conversion of hepatocytes.* J Clin Invest, 2012. 122(11): p. 3914-8.
138. Holczbauer, A., et al., *Modeling pathogenesis of primary liver cancer in lineage-specific mouse cell types.* Gastroenterology, 2013. 145(1): p. 221-231.
139. Mu, X., et al., *Epithelial Transforming Growth Factor-beta Signaling Does Not Contribute to Liver Fibrosis but Protects Mice From Cholangiocarcinoma.* Gastroenterology, 2016. 150(3): p. 720-33.
140. Fabregat, I. and D. Caballero-Diaz, *Transforming Growth Factor-beta-Induced Cell Plasticity in Liver Fibrosis and Hepatocarcinogenesis.* Front Oncol, 2018. 8: p. 357.
141. Michalopoulos, G.K., L. Barua, and W.C. Bowen, *Transdifferentiation of rat hepatocytes into biliary cells after bile duct ligation and toxic biliary injury.* Hepatology, 2005. 41(3): p. 535-44.
142. Yanger, K., et al., *Robust cellular reprogramming occurs spontaneously during liver regeneration.* Genes Dev, 2013. 27(7): p. 719-24.
143. Tarlow, B.D., et al., *Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes.* Cell Stem Cell, 2014. 15(5): p. 605-18.
144. Petersen, B.E., V.F. Zajac, and G.K. Michalopoulos, *Bile ductular damage induced by methylene dianiline inhibits oval cell activation.* Am J Pathol, 1997. 151(4): p. 905-9.

145. Petersen, B.E., V.F. Zajac, and G.K. Michalopoulos, *Hepatic oval cell activation in response to injury following chemically induced periportal or pericentral damage in rats.* Hepatology, 1998. 27(4): p. 1030-8.

146. Strain, A.J., et al., *Human liver-derived stem cells.* Semin Liver Dis, 2003. 23(4): p. 373-84.

147. Fukuda, K., et al., *The origin of biliary ductular cells that appear in the spleen after transplantation of hepatocytes.* Cell Transplant, 2004. 13(1): p. 27-33.

148. Wang, J., et al., *Notch2 controls hepatocyte-derived cholangiocarcinoma formation in mice.* Oncogene, 2018. 37(24): p. 3229-3242.

149. Liu, Y., et al., *Yap-Sox9 signaling determines hepatocyte plasticity and lineage-specific hepatocarcinogenesis.* J Hepatol, 2022. 76(3): p. 652-664.

150. Yimlamai, D., et al., *Hippo pathway activity influences liver cell fate.* Cell, 2014. 157(6): p. 1324-1338.

151. Durnez, A., et al., *The clinicopathological and prognostic relevance of cytokeratin 7 and 19 expression in hepatocellular carcinoma. A possible progenitor cell origin.* Histopathology, 2006. 49(2): p. 138-51.

152. Kim, H., et al., *Human hepatocellular carcinomas with "Stemness"-related marker expression: keratin 19 expression and a poor prognosis.* Hepatology, 2011. 54(5): p. 1707-17.

153. Hu, X., et al., *Direct induction of hepatocyte-like cells from immortalized human bone marrow mesenchymal stem cells by overexpression of HNF4alpha.* Biochem Biophys Res Commun, 2016. 478(2): p. 791-7.

154. Bonzo, J.A., et al., *Suppression of hepatocyte proliferation by hepatocyte nuclear factor 4alpha in adult mice.* J Biol Chem, 2012. 287(10): p. 7345-56.

155. Walesky, C., et al., *Hepatocyte-specific deletion of hepatocyte nuclear factor-4alpha in adult mice results in increased hepatocyte proliferation.* Am J Physiol Gastrointest Liver Physiol, 2013. 304(1): p. G26-37.

156. Yin, C., et al., *Differentiation therapy of hepatocellular carcinoma in mice with recombinant adenovirus carrying hepatocyte nuclear factor-4alpha gene.* Hepatology, 2008. 48(5): p. 1528-39.

157. Wei, L., et al., *Oroxylin A activates PKM1/HNF4 alpha to induce hepatoma differentiation and block cancer progression.* Cell Death Dis, 2017. 8(7): p. e2944.

158. Uenishi, T., et al., *Cytokeratin 19 expression in hepatocellular carcinoma predicts early postoperative recurrence.* Cancer Sci, 2003. 94(10): p. 851-7.

159. Govaere, O., et al., *Keratin 19: a key role player in the invasion of human hepatocellular carcinomas.* Gut, 2014. 63(4): p. 674-85.

160. Tsuchiya, K., et al., *Expression of keratin 19 is related to high recurrence of hepatocellular carcinoma after radiofrequency ablation.* Oncology, 2011. 80(3-4): p. 278-88.

161. Rhee, H., et al., *Poor outcome of hepatocellular carcinoma with stemness marker under hypoxia: resistance to transarterial chemoembolization.* Mod Pathol, 2016. 29(9): p. 1038-49.

162. Miltiadous, O., et al., *Progenitor cell markers predict outcome of patients with hepatocellular carcinoma beyond Milan criteria undergoing liver transplantation.* J Hepatol, 2015. 63(6): p. 1368-77.

163. Lawler, J., et al., *Identification and characterization of thrombospondin-4, a new member of the thrombospondin gene family.* J Cell Biol, 1993. 120(4): p. 1059-67.

164. Del Pozo Martin, Y., et al., *Mesenchymal Cancer Cell-Stroma Crosstalk Promotes Niche Activation, Epithelial Reversion, and Metastatic Colonization.* Cell Rep, 2015. 13(11): p. 2456-2469.

165. Lawler, J. and R.O. Hynes, *The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.* J Cell Biol, 1986. 103(5): p. 1635-48.

166. Stenina, O.I., et al., *Polymorphisms A387P in thrombospondin-4 and N700S in thrombospondin-1 perturb calcium binding sites.* FASEB J, 2005. 19(13): p. 1893-5.

167. Gkretsi, V., et al., *Loss of integrin linked kinase from mouse hepatocytes in vitro and in vivo results in apoptosis and hepatitis.* Hepatology, 2007. 45(4): p. 1025-34.

168. Kazanskaya, O., et al., *R-Spondin2 is a secreted activator of Wnt/beta-catenin signaling and is required for Xenopus myogenesis.* Dev Cell, 2004. 7(4): p. 525-34.

169. Clevers, H., *Wnt/beta-catenin signaling in development and disease.* Cell, 2006. 127(3): p. 469-80.

170. Koo, B.K., et al., *Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors.* Nature, 2012. 488(7413): p. 665-9.

171. Hao, H.X., et al., *ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner.* Nature, 2012. 485(7397): p. 195-200.

172. Cai, C., et al., *R-spondin1 is a novel hormone mediator for mammary stem cell self-renewal.* Genes Dev, 2014. 28(20): p. 2205-18.

173. Joshi, P.A., et al., *RANK Signaling Amplifies WNT-Responsive Mammary Progenitors through R-SPONDIN1.* Stem Cell Reports, 2015. 5(1): p. 31-44.

174. Adams, J.C. and R.P. Tucker, *The thrombospondin type 1 repeat (TSR) superfamily: diverse proteins with related roles in neuronal development.* Dev Dyn, 2000. 218(2): p. 280-99.

175. Yan, K.S., et al., *Non-equivalence of Wnt and R-spondin ligands during Lgr5(+) intestinal stem-cell self-renewal.* Nature, 2017. 545(7653): p. 238-242.

176. Ter Steege, E.J. and E.R.M. Bakker, *The role of R-spondin proteins in cancer biology.* Oncogene, 2021. 40(47): p. 6469-6478.

177. Qu, H.L., G.W. Hasen, Y.Y. Hou, and C.X. Zhang, *THBS2 promotes cell migration and invasion in colorectal cancer via modulating Wnt/beta-catenin signaling pathway.* Kaohsiung J Med Sci, 2022. 38(5): p. 469-478.

178. Wang, C., et al., *Rspondin-1 contributes to the progression and stemness of gastric cancer by LGR5.* Biochem Biophys Res Commun, 2022. 627: p. 91-96.

179. Tang, D., et al., *Angiogenesis in cholangiocellular carcinoma: expression of vascular endothelial growth factor, angiopoietin-1/2, thrombospondin-1 and clinicopathological significance.* Oncol Rep, 2006. 15(3): p. 525-32.

180. Kawahara, N., et al., *Enhanced expression of thrombospondin-1 and hypovascularity in human cholangiocarcinoma.* Hepatology, 1998. 28(6): p. 1512-7.

181. Sargiannidou, I., J. Zhou, and G.P. Tuszynski, *The role of thrombospondin-1 in tumor progression.* Exp Biol Med (Maywood), 2001. 226(8): p. 726-33.

182. Distler, J.H., et al., *Hypoxia-induced increase in the production of extracellular matrix proteins in systemic sclerosis.* Arthritis Rheum, 2007. 56(12): p. 4203-15.

183. de Lau, W., et al., *Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling.* Nature, 2011. 476(7360): p. 293-7.

184. de Lau, W.B., B. Snel, and H.C. Clevers, *The R-spondin protein family.* Genome Biol, 2012. 13(3): p. 242.

185. Ahn, S.M., et al., *Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification.* Hepatology, 2014. 60(6): p. 1972-82.

186. Conboy, C.B., et al., *R-spondin 2 Drives Liver Tumor Development in a Yes-Associated Protein-Dependent Manner.* Hepatol Commun, 2019. 3(11): p. 1496-1509.

187. Kan, Z., et al., *Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma.* Genome Res, 2013. 23(9): p. 1422-33.

188. Lu, X., et al., *Predicting features of breast cancer with gene expression patterns.* Breast Cancer Res Treat, 2008. 108(2): p. 191-201.

189. Frolova, E.G., et al., *Thrombospondin-4 regulates vascular inflammation and atherogenesis.* Circ Res, 2010. 107(11): p. 1313-25.

190. Muppala, S., et al., *Proangiogenic Properties of Thrombospondin-4.* Arterioscler Thromb Vasc Biol, 2015. 35(9): p. 1975-86.

191. Muppala, S., et al., *Thrombospondin-4 mediates TGF-beta-induced angiogenesis.* Oncogene, 2017. 36(36): p. 5189-5198.

192. Guo, D., et al., *THBS4 promotes HCC progression by regulating ITGB1 via FAK/PI3K/AKT pathway.* FASEB J, 2020. 34(8): p. 10668-10681.

193. Adams, J.C. and J. Lawler, *The thrombospondins.* Int J Biochem Cell Biol, 2004. 36(6): p. 961-8.

194. Adams, J.C. and J. Lawler, *The thrombospondins.* Cold Spring Harb Perspect Biol, 2011. 3(10): p. a009712.

195. Lawler, P.R. and J. Lawler, *Molecular basis for the regulation of angiogenesis by thrombospondin-1 and -2.* Cold Spring Harb Perspect Med, 2012. 2(5): p. a006627.

196. Murphy-Ullrich, J.E. and M.J. Suto, *Thrombospondin-1 regulation of latent TGF-beta activation: A therapeutic target for fibrotic disease.* Matrix Biol, 2018. 68-69: p. 28-43.

197. Crawford, S.E., et al., *Thrombospondin-1 is a major activator of TGF-beta1 in vivo.* Cell, 1998. 93(7): p. 1159-70.

198. Margagliotti, S., et al., *The Onecut transcription factors HNF-6/OC-1 and OC-2 regulate early liver expansion by controlling hepatoblast migration.* Dev Biol, 2007. 311(2): p. 579-89.

199. Frolova, E.G., et al., *Control of organization and function of muscle and tendon by thrombospondin-4.* Matrix Biol, 2014. 37: p. 35-48.

200. Frolova, E.G., et al., *Thrombospondin-4 regulates fibrosis and remodeling of the myocardium in response to pressure overload.* FASEB J, 2012. 26(6): p. 2363-73.

201. McDonald, J.F., J.M. Dimitry, and W.A. Frazier, *An amyloid-like C-terminal domain of thrombospondin-1 displays CD47 agonist activity requiring both VVM motifs.* Biochemistry, 2003. 42(33): p. 10001-11.

202. Barclay, A.N. and T.K. Van den Berg, *The interaction between signal regulatory protein alpha (SIRPalpha) and CD47: structure, function, and therapeutic target.* Annu Rev Immunol, 2014. 32: p. 25-50.

203. Lai, K.K., et al., *Extracellular matrix dynamics in hepatocarcinogenesis: a comparative proteomics study of PDGFC transgenic and Pten null mouse models.* PLoS Genet, 2011. 7(6): p. e1002147.

204. Begum, N.A., et al., *Differential display and integrin alpha 6 messenger RNA overexpression in hepatocellular carcinoma.* Hepatology, 1995. 22(5): p. 1447-55.

205. Zhao, G., et al., *Mechanical stiffness of liver tissues in relation to integrin beta1 expression may influence the development of hepatic cirrhosis and hepatocellular carcinoma.* J Surg Oncol, 2010. 102(5): p. 482-9.

206. Dalla-Torre, C.A., et al., *Effects of THBS3, SPARC and SPP1 expression on biological behavior and survival in patients with osteosarcoma.* BMC Cancer, 2006. 6: p. 237.

207. Deng, L.Y., et al., *Expression and prognostic significance of thrombospondin gene family in gastric cancer.* J Gastrointest Oncol, 2021. 12(2): p. 355-364.

208. Ter Steege, E.J., et al., *R-spondin-3 is an oncogenic driver of poorly differentiated invasive breast cancer.* J Pathol, 2022. 258(3): p. 289-299.

209. Zhang, C., et al., *The Integrative Analysis of Thrombospondin Family Genes in Pan-Cancer Reveals that THBS2 Facilitates Gastrointestinal Cancer Metastasis.* J Oncol, 2021. 2021: p. 4405491.

210. Frese, K.K. and D.A. Tuveson, *Maximizing mouse cancer models.* Nat Rev Cancer, 2007. 7(9): p. 645-58.

211. Van Dyke, T. and T. Jacks, *Cancer modeling in the modern era: progress and challenges.* Cell, 2002. 108(2): p. 135-44.

212. Sharpless, N.E. and R.A. Depinho, *The mighty mouse: genetically engineered mouse models in cancer drug development.* Nat Rev Drug Discov, 2006. 5(9): p. 741-54.

213. Rygaard, J. and C.O. Povlsen, *Heterotransplantation of a human malignant tumour to "Nude" mice.* Acta Pathol Microbiol Scand, 1969. 77(4): p. 758-60.

214. Hoffman, R.M., *Orthotopic metastatic (MetaMouse) models for discovery and development of novel chemotherapy.* Methods Mol Med, 2005. 111: p. 297-322.

215. Tuveson, D.A. and T. Jacks, *Technologically advanced cancer modeling in mice.* Curr Opin Genet Dev, 2002. 12(1): p. 105-10.

216. Maddison, K. and A.R. Clarke, *New approaches for modelling cancer mechanisms in the mouse.* J Pathol, 2005. 205(2): p. 181-93.

217. Jaenisch, R., *Transgenic animals.* Science, 1988. 240(4858): p. 1468-74.

218. Macleod, K.F. and T. Jacks, *Insights into cancer from transgenic mouse models.* J Pathol, 1999. 187(1): p. 43-60.

219. Zender, L., et al., *Cancer gene discovery in hepatocellular carcinoma.* J Hepatol, 2010. 52(6): p. 921-9.

220. Jonkers, J. and A. Berns, *Conditional mouse models of sporadic cancer.* Nat Rev Cancer, 2002. 2(4): p. 251-65.

221. Zender, L., et al., *Generation and analysis of genetically defined liver carcinomas derived from bipotential liver progenitors.* Cold Spring Harb Symp Quant Biol, 2005. 70: p. 251-61.

222. Zender, L., et al., *Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach.* Cell, 2006. 125(7): p. 1253-67.

223. Zhang, G., et al., *Hydroporation as the mechanism of hydrodynamic delivery.* Gene Ther, 2004. 11(8): p. 675-82.

224. Crespo, A., et al., *Hydrodynamic liver gene transfer mechanism involves transient sinusoidal blood stasis and massive hepatocyte endocytic vesicles.* Gene Ther, 2005. 12(11): p. 927-35.

225. Suda, T., X. Gao, D.B. Stolz, and D. Liu, *Structural impact of hydrodynamic injection on mouse liver.* Gene Ther, 2007. 14(2): p. 129-37.

226. Kobayashi, N., M. Nishikawa, K. Hirata, and Y. Takakura, *Hydrodynamics-based procedure involves transient hyperpermeability in the hepatic cellular membrane: implication of a nonspecific process in efficient intracellular gene delivery.* J Gene Med, 2004. 6(5): p. 584-92.

227. Andrianaivo, F., et al., *Hydrodynamics-based transfection of the liver: entrance into hepatocytes of DNA that causes expression takes place very early after injection.* J Gene Med, 2004. 6(8): p. 877-83.

228. Ivics, Z., Z. Izsvak, A. Minter, and P.B. Hackett, *Identification of functional domains and evolution of Tc1-like transposable elements.* Proc Natl Acad Sci U S A, 1996. 93(10): p. 5008-13.

229. Ivics, Z., P.B. Hackett, R.H. Plasterk, and Z. Izsvak, *Molecular reconstruction of Sleeping beauty, a Tc1-like transposon from fish, and its transposition in human cells.* Cell, 1997. 91(4): p. 501-510.

230. Zender, L., et al., *An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer.* Cell, 2008. 135(5): p. 852-64.

231. Chen, S., Y. Zhou, Y. Chen, and J. Gu, *fastp: an ultra-fast all-in-one FASTQ preprocessor.* Bioinformatics, 2018. 34(17): p. i884-i890.

232. Patro, R., et al., *Salmon provides fast and bias-aware quantification of transcript expression.* Nat Methods, 2017. 14(4): p. 417-419.

233. Ewels, P.A., et al., *The nf-core framework for community-curated bioinformatics pipelines.* Nat Biotechnol, 2020. 38(3): p. 276-278.

234. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic Acids Res, 2010. 38(16): p. e164.

235. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. 2(5): p. 401-4.

236. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. 6(269): p. pl1.

237. Guest, R.V., et al., *Cell lineage tracing reveals a biliary origin of intrahepatic cholangiocarcinoma.* Cancer Res, 2014. 74(4): p. 1005-10.

238. Tokumoto, N., et al., *Immunohistochemical and mutational analyses of Wnt signaling components and target genes in intrahepatic cholangiocarcinomas.* Int J Oncol, 2005. 27(4): p. 973-80.

239. Sugimachi, K., et al., *Altered expression of beta-catenin without genetic mutation in intrahepatic cholangiocarcinoma.* Mod Pathol, 2001. 14(9): p. 900-5.

240. Raggi, C., P. Invernizzi, and J.B. Andersen, *Impact of microenvironment and stem-like plasticity in cholangiocarcinoma: molecular networks and biological concepts.* J Hepatol, 2015. 62(1): p. 198-207.

241. Storz, P., *Acinar cell plasticity and development of pancreatic ductal adenocarcinoma.* Nat Rev Gastroenterol Hepatol, 2017. 14(5): p. 296-304.

242. Wong, C.H., Y.J. Li, and Y.C. Chen, *Therapeutic potential of targeting acinar cell reprogramming in pancreatic cancer.* World J Gastroenterol, 2016. 22(31): p. 7046-57.

243. Andersen, J.B., et al., *Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors.* Gastroenterology, 2012. 142(4): p. 1021-1031 e15.

244. Sia, D., V. Tovar, A. Moeini, and J.M. Llovet, *Intrahepatic cholangiocarcinoma: pathogenesis and rationale for molecular therapies.* Oncogene, 2013. 32(41): p. 4861-70.

245. Sia, D., et al., *Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes.* Gastroenterology, 2013. 144(4): p. 829-40.

246. Seok, J.Y., et al., *A fibrous stromal component in hepatocellular carcinoma reveals a cholangiocarcinoma-like gene expression trait and epithelial-mesenchymal transition.* Hepatology, 2012. 55(6): p. 1776-86.

247. Woo, H.G., et al., *Gene expression-based recurrence prediction of hepatitis B virus-related human hepatocellular carcinoma.* Clin Cancer Res, 2008. 14(7): p. 2056-64.

248. Cazals-Hatem, D., et al., *Clinical and molecular analysis of combined hepatocellular-cholangiocarcinomas.* J Hepatol, 2004. 41(2): p. 292-8.

249. Fujii, H., et al., *Genetic classification of combined hepatocellular-cholangiocarcinoma.* Hum Pathol, 2000. 31(9): p. 1011-7.

250. Murata, M., et al., *Combined hepatocellular/cholangiocellular carcinoma with sarcomatoid features: genetic analysis for histogenesis.* Hepatol Res, 2001. 21(3): p. 220-227.

251. Komuta, M., et al., *Clinicopathological study on cholangiolocellular carcinoma suggesting hepatic progenitor cell origin.* Hepatology, 2008. 47(5): p. 1544-56.

252. Roskams, T., *Different types of liver progenitor cells and their niches.* J Hepatol, 2006. 45(1): p. 1-4.

253. Kim, H., et al., *Primary liver carcinoma of intermediate (hepatocyte-cholangiocyte) phenotype.* J Hepatol, 2004. 40(2): p. 298-304.

254. Woo, H.G., et al., *Identification of a cholangiocarcinoma-like gene expression trait in hepatocellular carcinoma.* Cancer Res, 2010. 70(8): p. 3034-41.

255. Tschaharganeh, D.F., et al., *p53-dependent Nestin regulation links tumor suppression to cellular plasticity in liver cancer.* Cell, 2014. 158(3): p. 579-92.

256. Dahn, M.L. and P. Marcato, *In Vivo Genome-Wide Pooled RNAi Screens in Cancer Cells to Identify Determinants of Chemotherapy/Drug Response.* Methods Mol Biol, 2021. 2381: p. 189-200.

257. Noorani, I., A. Bradley, and J. de la Rosa, *CRISPR and transposon in vivo screens for cancer drivers and therapeutic targets.* Genome Biol, 2020. 21(1): p. 204.

258. Seyhan, A.A. and T.E. Rya, *RNAi screening for the discovery of novel modulators of human disease.* Curr Pharm Biotechnol, 2010. 11(7): p. 735-56.

259. Li, Y., et al., *Gain-of-Function Mutations: An Emerging Advantage for Cancer Biology.* Trends Biochem Sci, 2019. 44(8): p. 659-674.

260. Merz, V., et al., *Targeting KRAS: The Elephant in the Room of Epithelial Cancers.* Front Oncol, 2021. 11: p. 638360.

261. Kitade, M., et al., *Specific fate decisions in adult hepatic progenitor cells driven by MET and EGFR signaling.* Genes Dev, 2013. 27(15): p. 1706-17.

262. Huber, R.M., et al., *Deltex-1 activates mitotic signaling and proliferation and increases the clonogenic and invasive potential of U373 and LN18 glioblastoma cells and correlates with patient survival.* PLoS One, 2013. 8(2): p. e57793.

263. Lin, Y.W., et al., *OLIG2 (BHLHB1), a bHLH transcription factor, contributes to leukemogenesis in concert with LMO1.* Cancer Res, 2005. 65(16): p. 7151-8.

264. Liu, X., et al., *The associations between Deltex1 and clinical characteristics of breast cancer.* Gland Surg, 2021. 10(11): p. 3116-3127.

265. Wang, C.H., et al., *TMPRSS4 facilitates epithelial-mesenchymal transition of hepatocellular carcinoma and is a predictive marker for poor prognosis of patients after curative resection.* Sci Rep, 2015. 5: p. 12366.

266. Liu, X., et al., *THBS1 facilitates colorectal liver metastasis through enhancing epithelial-mesenchymal transition.* Clin Transl Oncol, 2020. 22(10): p. 1730-1740.

267. Weinstat-Saslow, D.L., et al., *Transfection of thrombospondin 1 complementary DNA into a human breast carcinoma cell line reduces primary tumor growth, metastatic potential, and angiogenesis.* Cancer Res, 1994. 54(24): p. 6504-11.

268. Wang, G., et al., *Elevated THBS3 predicts poor overall survival for clear cell renal cell carcinoma and identifies LncRNA/RBP/THBS3 mRNA networks.* Cell Cycle, 2023. 22(3): p. 316-330.

269. Stover, D.G., et al., *The Role of Proliferation in Determining Response to Neoadjuvant Chemotherapy in Breast Cancer: A Gene Expression-Based Meta-Analysis.* Clin Cancer Res, 2016. 22(24): p. 6039-6050.

270. Yan, M., et al., *miR-146b promotes cell proliferation and increases chemosensitivity, but attenuates cell migration and invasion via FBXL10 in ovarian cancer.* Cell Death Dis, 2018. 9(11): p. 1123.

271. Forkasiewicz, A., et al., *The usefulness of lactate dehydrogenase measurements in current oncological practice.* Cell Mol Biol Lett, 2020. 25: p. 35.

272. Ibrahim, A.M., S. Nady, M.W. Shafaa, and M.M. Khalil, *Radiation and chemotherapy variable response induced by tumor cell hypoxia: impact of radiation dose, anticancer drug, and type of cancer.* Radiat Environ Biophys, 2022. 61(2): p. 263-277.

273. Wu, C., et al., *Low expression of KIF20A suppresses cell proliferation, promotes chemosensitivity and is associated with better prognosis in HCC.* Aging (Albany NY), 2021. 13(18): p. 22148-22163.

274. Morana, O., W. Wood, and C.D. Gregory, *The Apoptosis Paradox in Cancer.* Int J Mol Sci, 2022. 23(3).

275. Abarca-Buis, R.F., E.A. Mandujano-Tinoco, A. Cabrera-Wrooman, and E. Krotzsch, *The complexity of TGFbeta/activin signaling in regeneration.* J Cell Commun Signal, 2021. 15(1): p. 7-23.

276. Yoo, J.E., et al., *The dual role of transforming growth factor-beta signatures in human B viral multistep hepatocarcinogenesis: early and late responsive genes.* J Liver Cancer, 2022. 22(2): p. 115-124.

277. Coulouarn, C., V.M. Factor, and S.S. Thorgeirsson, *Transforming growth factor-beta gene expression signature in mouse hepatocytes predicts clinical outcome in human cancer.* Hepatology, 2008. 47(6): p. 2059-67.

278. Yoo, J.E., et al., *Progressive Enrichment of Stemness Features and Tumor Stromal Alterations in Multistep Hepatocarcinogenesis.* PLoS One, 2017. 12(1): p. e0170465.

279. Wan, S., et al., *Tumor-associated macrophages produce interleukin 6 and signal via STAT3 to promote expansion of human hepatocellular carcinoma stem cells.* Gastroenterology, 2014. 147(6): p. 1393-404.

280. Liu, Y. and X. Cao, *Characteristics and Significance of the Pre-metastatic Niche.* Cancer Cell, 2016. 30(5): p. 668-681.

281. Garcia-Perez, M.A., *Statistical conclusion validity: some common threats and simple remedies.* Front Psychol, 2012. 3: p. 325.

282. Bate, S.T. and R.A. Clark, *The design and statistical analysis of animal experiments.* 2014, Cambridge, United Kingdom: Cambridge University Press. xv, 310 pages.

283. Bate, S. and N.A. Karp, *A common control group - optimising the experiment design to maximise sensitivity.* PLoS One, 2014. 9(12): p. e114872.

284. Heather M. Schellinck, D.P.C., Richard E. Brown, *Chapter 7 – how many ways can mouse behavior experiments go wrong? Confounding variables in mouse models of neurodegenerative diseases and how to control them.* Vol. 41. 2010. 255 - 366.