THE UNIVERSITY
of EDINBURGH

Defining Functional Cooperation of Oncogenic KRAS and Rare Driver Mutations in Intrahepatic Cholangiocarcinoma

Mollie L Wilson MSci, B.Sc (Hons)

Submitted for the degree:

Doctor of Philosophy

The University of Edinburgh

2022

Declaration

I declare that this thesis is solely of my own composition, and unless stated otherwise, is comprised entirely of my own work. This work has not been submitted, in whole or in part, in any other application for a degree or professional qualification.

Mollie Wilson 09/05/2022

Abstract

Intrahepatic cholangiocarcinoma (iCCA) is an aggressive and lethal malignancy arising from the bile ducts. Patients with iCCA typically present with genetically heterogeneous tumours. There is an exception of a small minority of patients where actionable mutations are present, but the heterogeneity of this cancer has hampered the development of targeted therapeutics as the number of recurrent mutations between patients is limited. The majority of iCCA patients are not amenable to surgery, and standard of care chemotherapies act to extend life and are not considered as a curative treatment. As such, there is a significant requirement to explore the functional genetics of iCCA with the goal of developing more effective treatments and therapies. Deep sequencing studies have identified oncogenic mutations in *KRAS* in iCCA and these represent one of the most recurrent alterations in this cancer. Therefore I focussed on the role of *KRAS* mutations and their interactions with loss of function mutations.

During this work, I identified and validated *in vivo* which loss of function mutations were capable of interacting with oncogenic $KRAS^{G12D}$ mutations. To do this, I developed and carried out a CRISPR-*sp*Cas9 library screen informed by mutations that have been previously identified as pathogenic in patient iCCA. Using this approach, I identified the loss of cytoskeletal signalling protein Neurofibromin 2 (*NF2*) as a novel driver event in iCCA and that loss of this gene interacts with mutant *KRAS* to initiate cancer.

To investigate the role of *NF2* further, I explored the relationship between *Nf2*-loss, *Trp53*-loss and mutant KRAS, and I found that *Nf2* loss cooperated with *Trp53* loss to

lead to cancer with a severely accelerated phenotype and increased lethality. Additionally, these tumours were representative of a rare and aggressive sarcomatoid subtype of iCCA.

To understand how loss of both *Trp53* and *Nf2* leads to this aggressive phenotype, I used proteomic analysis of a number of signalling pathways, and RNA sequencing to define which mechanisms are altered and contribute to the aggressive cancer phenotype. These analyses identified the co-activation of Wnt/ β-catenin and Pi3K/AKT signalling as recurrently activated in iCCA. To test whether the aggressive nature of dual *Trp53* and *Nf2* loss was a result of changes in these signalling pathways, I inhibited them pharmacologically with inhibitors targeting Wnt/β-catenin and Pi3K/AKT signalling. Treatment with both of these inhibitors significantly improved survival of tumour bearing mice. Finally, I established a genetically engineered mouse model of iCCA not reliant on *Nf2*-loss for its progression to address whether inhibition of Wnt/β-catenin and Pi3K/AKT could represent a wide-spanning therapeutic and determine whether the treatment of a different pathological subtype of iCCA would be with these inhibitors would reduce disease progression. I showed this form of well-differentiated iCCA is also amenable to inhibition with this combination of therapies.

These data demonstrate that *Nf2* is a rare driver gene of iCCA that acts in a cooperative manner with oncogenic KRAS to accelerate tumourigenesis *in vivo*. Examination of these tumours highlighted a concurrent Wnt and PI3K signalling signature, and I demonstrate that pharmacological co-inhibition substantially impedes iCCA growth, highlighting the use of Wnt/β-catenin and Pi3K/AKT signalling inhibition as a broad treatment for iCCA patients.

Lay Summary

Intrahepatic cholangiocarcinoma (iCCA) is a rare cancer that occurs in the biliary tree in the liver – a network of tubes that transfers bile from the liver to the bowel to aid the breakdown of fats during digestion. The changes in DNA which allow this cancer to grow are complex, and so it has been difficult to develop medicines to treat this disease. Current treatments, chemotherapies, only extend life by 3-6 months unless patients have surgery on the liver to remove the cancer but this is only applicable to a small number of patients. For this reason my project aimed to examine the underlying causes of iCCA in the body with the intention of improving knowledge of the disease and contributing to the development of treatment options.

During my project I examined the DNA changes, or mutations, that cause iCCA by developing mice with this disease. A single piece of DNA, known as a gene, often mutated in some iCCA patients is called *KRAS*. However, as mutations in *KRAS* are insufficient to cause cancer in the liver alone, I aimed to look at what mutations in other genes work together with those in *KRAS* in order to cause cancer. I chose which genes to investigate their interaction with *KRAS* based on patient iCCA data. It was important to use animals as models of this cancer as they develop a tumour that is similar to what the cancer is like in humans, and that is contained in a similar environment in the body. Examining the underlying genetics of a disease is a flexible way in which to learn about its mechanism and how it causes illness.

My experiments involved using a system where I used the cancer formed in mice to understand how cells with different DNA mutations change the way that the cells sense their environment and receive messages. These messages are the cause of cells

growing in an out of control way - the cells becoming cancerous. I identified a specific set of these messages which are used by some tumours to grow. I then used two drugs together to target these messages in mice, and growth of tumours in the mice were reduced significantly. Importantly, this combination of drugs has not been used for this cancer type in humans before. Therefore this is a key step in the development of a novel treatment option for patients with iCCA, and my work shows that the blocking of these messages with drug treatments has the potential to slow down iCCA growth in patients.

Acknowledgements

This work has been published in part in *"In Vivo Modeling of Patient Genetic Heterogeneity Identifies New Ways to Target Cholangiocarcinoma"* by Younger and Wilson *et al* 2022. (1) It is included as Supplementary Material.

I dedicate this thesis to the powerful women who shaped me; SMH, KJR, SJW, my beloved ALR and of course SLR, who makes anything feel possible. You all have made me who I am, I learned the determination I needed to complete this work from you. Thank you for supporting me no matter what, you've championed my success.

To my friends, I'm still here because of you. Brenola, Timme and Jakke, I hope this work makes you happy, and proud also, you've made my PhD a possibility. To Michael, Vic, Posc, Kez and Fraser. Thank you all for helping me believe in myself and being so full of vitality you make the hard things possible.

Thank you to the Boulter lab members past and present, and everyone else in W3 for your guidance on my work, your support, and helping keep things in perspective with kindness, cakes and wee beers. Especially Nick, Ed, Anabel and Mia. Of course Scott, my other half, I couldn't have done it without you, this experience has been unforgettable because of you, thank you for making me laugh every day.

Thank you Alex, your support and love has healed me, and guided me through my PhD. I am so grateful to you for your compassion and intelligence.

Finally Luke, thank you for your advice, for always having my back and for your boundless kindness. You're an inspiring human and a pleasure to work with, I'm proud to have been a member of your lab.

Contents

Chapter 1

Introduction

The genetic and molecular diversity of intrahepatic cholangiocarcinomas (iCCAs), malignant tumours that arise from the epithelia of peripheral branches of the biliary tree, present longstanding challenges in their medical treatment. iCCA presents as a lethal disease with poor prognosis and high mortality. Increasing depth of understanding of this disease at the genetic and molecular level is paramount to therapeutic intervention and ultimately improvement of patient survival. To address this lack of knowledge, my PhD project utilises genetic screening and molecular approaches to identify and characterise phenotypic subtypes of iCCA in relation to specific genetic changes in murine models of iCCA. Firstly I address the liver as an organ and the importance of the biliary tree, then I explore cancer as a disease, then I introduce iCCA as a burden of global health, and its biological background and finally examine systems that are used to investigate the functional genetics of cancer, and in the case of my PhD, iCCA.

1.1 Biliary and hepatic physiology and function

iCCA arises in the biliary tree of the liver, a critical metabolically active organ with myriad functions. The liver is responsible for homeostatic regulation of the blood and is a site of metabolism for many molecules, including that of drugs or other damaging substances such as urea. In addition to the liver's detoxifying role, it is also a key player in the digestive process. More specifically, the liver regulates chemical, hormone and amino acid levels in the blood, as well as producing proteins for blood plasma and regulating blood clotting. Its protective role in bodily immunity involves production of immune factors, and the removal of bacteria from the blood. The liver produces cholesterol and other fat carrying proteins and it functions to convert excess glucose to

glycogen where it is then stored. (2) With regard to its role in the digestive system, the liver synergises with the biliary system, which consists of the gallbladder and the hepatic, cystic and bile ducts, to excrete bile and subsequently digest and absorb lipids that have come from the gastrointestinal tract. The bile ducts form a complex 3D network known as the biliary tree that functions to secrete and move bile. Bile is vital in the excretion, emulsion, encouragement of absorption and elimination of compounds that other organs are not capable of handling, as well as aiding hepatic circulation and transference of hormonal signals that contribute to intestinal growth and development. (3) Bile consists of bile acids, bile pigments such as bilirubin, bicarbonate, cholesterol, phospholipids and water.

1.2 Hepatic physiology

The liver is the largest organ of the body, comprising 2-5% of adult human body weight. Its cellular composition consists mostly of hepatocytes, parenchymal cells which make up 60% of the total cells and 80% of the liver volume. (4) The liver parenchyma consists mostly of hepatocytes organised around a portal triad consisting of a portal vein, hepatic artery and bile duct. The liver has a unique vasculature whereby it receives blood from the hepatic artery that brings oxygenated blood, and the portal vein that runs in parallel and brings deoxygenated blood. (5) Sinusoids are channels lined by endothelial cells that are located at the periphery of the liver lobules, and are where the oxygenated and deoxygenated blood mixes. The space between the sinusoidal lumen and the basolateral hepatocyte membranes is known as the space of Disse. In this area there are microvilli that extend from the hepatocytic basolateral membrane to communicate with the capillaries and thus provide blood to

hepatocytes. The space of Disse is made up of collagens and proteoglycans, meaning it acts as an extracellular matrix scaffolding to support the hepatocytes and helps maintain the structure of hexagonal functional units of the liver known as lobules. The central vein is located at the centre of the lobules. Hepatocytes making up the lobules have varying functions, and are divided into zones based on these functions and their perfusion ability. Zone I denotes the periportal hepatocyte region that are closest to the oxygenated blood and nutrients, and therefore the site of oxidative metabolisms. Zone II forms the pericentral region and zone III, being the furthest distance from the portal triad, has the lowest perfusion and has the most significant role in detoxification.

1.3 Non-parenchymal cells

The liver also contains non-parenchymal cells, including liver sinusoidal endothelial cells which make up approximately 50% of the non-parenchymal cells in the liver. They are responsible for the clearance of protein, lipids, and polysaccharide from the blood and are organised into a unique fenestrated endothelium morphology. (6–8) The resident hepatic macrophage, Kupffer cells, comprise around 20% of the non-parenchymal cells, and they act as filtering cells to extrude pathogenic substances such as endotoxins, bacteria, dead cells, and debris from the circulating blood via phagocytosis. Kupffer cells also release pro-inflammatory cytokines such as interleukin 1β, tumour necrosis factor, and transforming growth factor during liver injury or infection that recruits other immune cells. (9) Despite their usual role in immune surveillance and protecting the liver, dysregulated Kupffer cells cause chronic inflammation that perpetuate non-alcoholic fatty liver disease and liver fibrosis. (7)

Liver lymphocytes are also commonly found throughout the parenchyma and in the portal tract, particularly cells of the innate immune system including natural killer T cells, γδ T cells and natural killer (NK) cells. (9) Hepatic stellate cells, the non-parenchymal resident mesenchymal cells, retain the feature of resident portal fibroblasts in that they are embedded in stromal matrix and are attached to capillary endothelial cells but with the additional ability to communicate between neighbouring cell types. Hepatic stellate cells exist in the space of Disse where they have a key role in response to liver injury. Where they are usually quiescent in the healthy liver, they become activated and differentiate into myofibroblasts when they lose retinol and express collagen upon liver injury. This can lead to secretion of extracellular matrix proteins that accumulate into scar tissue due to the proliferative and contractile nature of myofibroblasts. (10) Upon consideration of the fibrogenic nature of the injured liver, portal fibroblasts are also present in the liver. These are a cell population discrete from hepatic stellate cells yet they act in a similar way upon chronic injury where they also differentiate into fibrogenic myofibroblasts. They are located adjacent to the bile duct epithelia and thus have an important role in biliary fibrosis specifically, as well as in general scar formation in the liver. (11)

Finally, the cholangiocytes, epithelial cells that line the complex three-dimensional network of interconnecting ductules and ducts known collectively as the biliary tree, are highly dynamic cells that function to modify bile as it passes through the biliary tree. Cholangiocytes are classified as being small or large, generally based on whether they are lining small or large ducts, respectively. Duct morphologic heterogenetiy itself will be discussed later in this chapter. Small and large cholangiocytes also differ in their nuclear to cytoplasmic ratio and thus their plasticity, as the ratio is greater in small

cholangiocytes, and these are less differentiated. (12) Cholangiocytes have a luminal and basolateral plasma membrane and are joined to adjacent cells by *zonula occludens* that establish and maintain epithelial cell polarity, and gap junctions that allow cell-cell communication. (13) Due to their key role in modification and movement of bile, cholangiocytes have many microvilli extending from their apical membrane which act to increase surface area. Their increased surface area means there is more interaction with the bile. Cholangiocytes also extend a single primary cilium that is composed of a microtubules and a centriole-derived basal body. (12) The cholangiocyte actin cytoskeleton has important structural and functional roles as it supports the plasma membrane and controls cell polarity, vesicle trafficking and distribution of membrane proteins. (14)

1.4 The biliary system

Considering the biliary tree system as a whole, it originates from bile canaliculi, formed by adjacent hepatocyte apical membranes joined by tight junctions, which facilitate bile transport from hepatocytes where bile is produced. (3,13) The high surface area formed by lattice patterning of the canaliculi result in increased bile flow, concurrently to the flow of blood entering the liver for perfusion, as bile exits the liver to administer its many functions. Once secreted by hepatocytes, bile travels through the canaliculi to the canals of Herring which are the first components of the ductal tree, lined by bile duct epithelium containing secretory and absorptive enzymes and transporter proteins that refine bile. (15) The presence of bipotential progenitor cells at the canals of Herring means cells can develop to hepatocytes or cholangiocytes. The system of ductules and ducts become progressively larger until they eventually delta into the

common bile duct. From here, bile is transported to the gallbladder to be concentrated and stored, or it is moved further along the common bile duct to the duodenum as required.

The biliary tree itself can be divided into intra- and extrahepatic parts. Intrahepatic ducts make up the peripheral branches of the biliary tree located in the liver parenchyma. They are considered as small (<15 µm diameter), interlobular (15–100 µm diameter), septal (100–300 µm diameter), area (300–400 µm diameter), segmental (400-800 µm), and hepatic ducts (>800 µm). (13) The right and left hepatic bile ducts are lined by cylindrical cholangiocytes that produce mucin, as are the large intrahepatic bile ducts (including segmental, area and septal ducts) whereas the small interlobular bile ducts and ductules are lined with cuboidal cholangiocytes that are mucin negative. (16) Ductules also contain hepatic progenitor cells that are capable of differentiation into both cholangiocytes and hepatocytes. When the biological systems in place to maintain cellular homeostasis fail, this is when cells can proliferate in an uncontrollable manner and become cancerous.

1.5 Cancer as a disease

Cancer is hugely complex in terms of its biology and its causes. Normal cells usually exist under a system of homeostasis whereby the amount of cell division and proliferation is equal to that of differentiation and apoptosis. This was described by Cairns in 1975 where a single stem cell will divide into two daughter cells, one of which will remain in the stem cell niche while the other is committed to a differentiation pathway. (17) Therefore the multiplication of cells is controlled by signals that act in a positive or negative way, and these signals themselves are controlled by genetic

alterations. (18) It is the disruption of the genetic controls of these signals which cause cells to lose their homeostatic state and become capable of uncontrolled division and invasion: cells become cancerous.

Cancer as a disease is a significant global health burden, being the first or second leading cause of death before 70 years old in 112 out of 183 countries. (19) As the world populations age and grow, cancer as a premature cause of death has also increased. According to the global cancer statistics report, in 2020 liver cancer was the third most common leading cause of cancer death, forming 8.3% of cancer deaths and the sixth most commonly diagnosed cancer. (19)

To be more specific in the definition of cancer, Hanahan and Weinberg distilled the complexities of cancer into a concise list of six so called “hallmarks of cancer” in 2000 (20), and have since added an additional two. (21) These hallmarks include: the ability to produce own growth signals, insensitivity to anti-growth signals, apoptosis avoidance, unlimited ability to replicate, metastasis, tissue invasion, ongoing angiogenesis, genomic mutations and instability, and finally, inflammation that promotes tumourigenesis. (21) As several of these characteristics are shared by both benign and malignant growths, it has been argued that even these detailed hallmarks are not specific enough to define this disease. (A benign growth, such as a papilloma for example, does not invade or transform surrounding cells, and it does not metastasise to other tissue types. (22) Additionally, invasion is a property of non-cancerous diseases such as endometriosis where cells do migrate to invade new tissue. (21)) However, the criteria are generally well accepted aside from the few exceptions and give a comprehensive definition to a nebulous disease.

Consideration of the underlying genetics of cancer is key to my project, and equally to consider the prolific ability of cancer cells to divide into more cells that are also cancerous (and subsequent cells rarely if ever being normal cells thereafter). This means that the heritability of genetic mutations is stable and transferred through cell division. Often each genetic change expands the population of neoplastic cells and thus creates an even bigger pool for subsequent genetic alterations, meaning cancer progresses from bad to worse in terms of accumulation of mutation which adds another dimension to its lethality.

1.5.1 Genetic alterations in cancer

As mentioned above, there are positive and negative signals that control cell multiplication, and these can either be genetic or epigenetic. (23) Cancer related genes are classed as either proto-oncogenes, those genes with the capacity for becoming oncogenic, or tumour suppressor genes. Proto-oncogene activation by mutation or epigenetic alteration results in positive stimulation of growth, and once activated they are known just as oncogenes. Usually these oncogenic mutations are dominant in nature and result in increased or altered function. (24) The other class of cancer related genes, tumour suppressor genes, are most commonly point mutations and deletions that generally result in a loss of gene function, and are therefore most likely to disrupt the mechanisms that prevent cells from proliferating in an uncontrolled manner. (25,26) As they are dominant compared to the wild type allele of a gene, a single mutation in a proto-oncogene allele is sufficient to activate it and there is no reason for cells to select for a homozygous mutation in this same gene as the cells are already cancerous. (24) It does, however, mean there is little to no chance of

oncogenic mutations being germline inherited, as they would be so disruptive to development to render the foetus nonviable. (27) In contrast, tumour suppressor genes require two independent mutational events across two alleles before a cell becomes cancerous. (26) For this reason, it's important to note that despite cancer being a somatic disease, there are heritable predispositions in tumour suppressor genes mutation in almost every type of cancer. (18) Germline mutation of a tumour suppressor in one allele is not sufficient for carcinogenesis to occur, and further somatic mutations must be accumulated to form cancer. For example, through environmental factors or by exposure to airborne carcinogenic agents. This outlines the basis of Knudson's two-hit hypothesis that explains predisposition to cancer, which remains true for the majority of tumour suppressor genes. (25) The hypothesis states that the germline mutation of one allele of a tumour suppressor gene acts as a predisposition to cancer because of the likelihood of the second allele being somatically mutated. (27)

Since genetic stability and the progressive inheritance of non-mutated cells is key in prevention of cancer, it is argued that another group of genes, in addition to oncogenes and tumour suppressor genes, are also responsible for tumourigenesis; stability genes. (28) Genetic instability is well known to play a role in cancer due to the many defects in chromosomes, gene rearrangements and amplifications. (29) Specifically, stability genes include genes that are involved in repair mechanisms including mismatch, nucleotide-excision, and base-excision repair, all of which aim to correct base mismatches during DNA replication. Genomic instability has been associated with hereditary cancers that is characterised by microsatellite instability (for example hereditary nonpolyposis colon cancer (28)), or chromosomal instability. In

sporadic cancer is less well characterised but believed to not be due to mutations in DNA repair genes or mitotic checkpoint genes. (29)

Carcinogenesis, or the multistep malignant transformation of cells (30), occurs from the homeostatic balanced normal cell, to an aberrantly proliferative one. It requires the loss of systems that protect and control proliferation, joined with the addition of systems that encourage progression of proliferation to form neoplasia.

1.6 Cholangiocarcinoma classification and cells of origin

Cholangiocarcinomas are the second most common form of liver cancer after hepatocellular carcinoma (HCC), and comprise approximately 15% of all primary liver tumours and 3% of gastrointestinal tumours. (31) However unlike HCC, intrahepatic cholangiocarcinoma (iCCA) usually develops from liver that is non-cirrhotic. (32) iCCA is the second most common primary hepatic malignancy. Cholangiocarcinomas arise as lethal malignancies at any point of the biliary tree. They present with features of differentiation of cholangiocytes, the native epithelial cells of the bile ducts.

Cholangiocarcinomas can be classified into three subtypes depending on their anatomical origin. Distal cholangiocarcinomas (dCCA) arise from the common bile duct, specifically the area between the cystic duct origin and the ampulla of Vater. Perihilar cholangiocarcinoma (pCCA) arises from the left and/or right hepatic duct and/or at their junction (33) and finally intrahepatic cholangiocarcinoma (iCCA) arises from the peripheral branches of the biliary tree, proximally to the second confluence of bile ducts. (32)

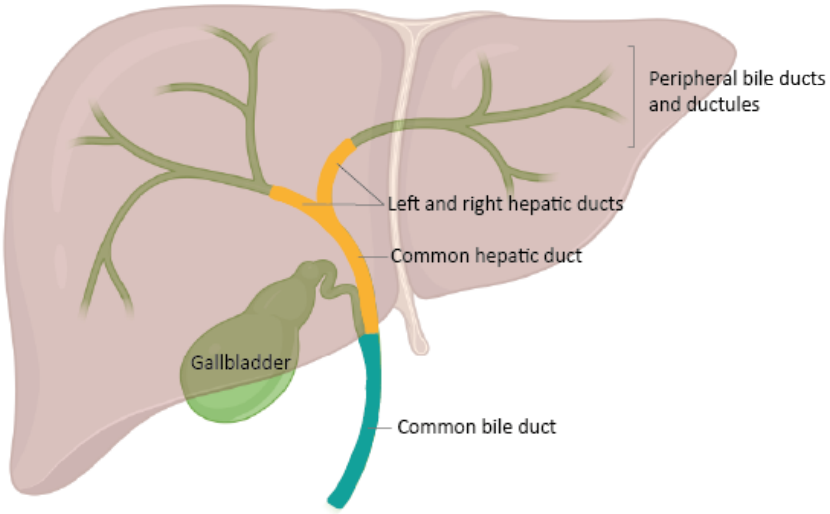

Peripheral bile ducts
and ductules
Left and right hepatic ducts
Common hepatic duct
Gallbladder
Common bile duct

Figure 1

Anatomical site of origin of cholangiocarcinoma. Cholangiocarcinoma is classified as intrahepatic, perihilar, or distal. iCCA presents as a malignancy of the peripheral bile ducts and ductules, illustrated in green. pCCA arises in the right and/or left hepatic duct and at the confluence of these ducts, shown in yellow, and dCCA occurs along the common bile duct illustrated here in blue.

pCCA and dCCA present as either poorly defined nodular sclerosing tumours or less commonly as intraductal papillary tumours (34). The growth patterns, or gross classification of iCCA, allows their classification into three distinct types: mass forming (MF), periductal infiltrating and intraductal growing. (32) MF iCCAs are the most common, reported to form 78% of iCCAs (35). When these are examined using radiology, tumours present as homogenous masses with irregular but well-defined outlines. (32,36) MF tumours are recognisable by their fibrous stroma and/ or coagulative necrosis with some scattered tumour cells. (32) Periductal infiltrating type iCCA spreads along portal tracts causing strictures and/or thickening of the bile ducts and dilation of peripheral bile ducts. These iCCAs are less frequent than MF types and make up 16% of iCCA. They have an elongated or branching morphology and there is some tumour infiltration into the liver parenchyma. (32,34,35) Finally, intraductal iCCA presents with a variety of features including as a polyploid or papillary tumours in a dilated ductal lumen. Intraductal iCCA has a comparatively slow growth and reasonably favourable prognosis. (36)

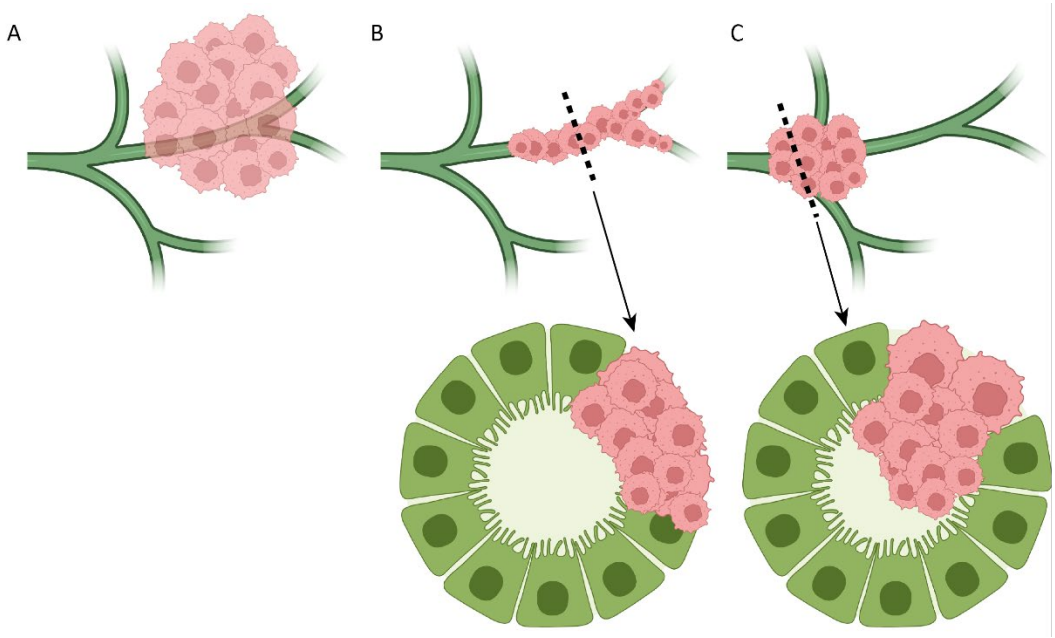

A
B
C

Figure 2

iCCA can be of three gross morphologies. A. Mass forming lesions present as masses in the hepatic parenchyma. B. Periductal infiltrating type grows along the walls of large bile ducts in a longitudinal manner. C. Intraductal growing tumours are polypoid or papillary masses that grow in towards ductal lumen.

The majority of pCCA and dCCA are mucin-producing adenocarcinomas or papillary tumours, whereas iCCAs are more histologically heterogeneous. (37) It's true for all types of cholangiocarcinoma that they can be preceded by Intraductal Papillary Neoplasia of the Bile Duct (IPNB). (33) Histologically, iCCA shows several variants and these are divided into conventional cholangiocarcinoma, cholangiolocelluar carcinoma, and rare variants. (37) Histological determination is important in understanding the consistency of a tumour, and this may be of use when deciding on therapeutic intervention. It is relevant to my project in understanding how my animal models histologically represent human iCCA.

iCCA is typically observed as an invasive adenocarcinoma with tubular structures, acini formation and micropapillary projections. iCCA can be well, moderately or poorly differentiated. (34) iCCA is further diverged into two histological subgroups depending on the size or level of affected duct; either small bile duct type iCCA or large bile duct type iCCA. Small bile duct type and cholangiolocellular carcinoma are derived from interlobular and septal bile ducts. They often form architecture of narrow ductular structures arranged as small monotonous glands or acinar adenocarcinomas with nodular growths invading the liver parenchyma. (38) Additionally, small bile duct type is typically composed of cuboidal or low columnal cells that produce limited or no mucin, lack eosinophilic cytoplasm, and have limited amount of stroma surrounding them. (16,38,39) Choloangiolocelluar carcinoma and small bile duct type iCCAs are thought to originate from bile ductules or as a result of liver inflammation of ducts that occurs in chronic liver diseases such as in cirrhosis. (35) Large bile duct iCCA arises from the large segmental and area bile ducts, or from peribiliary glands. (40) These tumours present with large columnar cells that produce ample mucin and are arranged as a

papillary (35) or large duct. (34) Usually large bile duct type tumours are surrounded by ample fibrous stroma. (16) Rare variants of iCCA include cases, for example, presenting with a mixture of small and large duct types, or with atypical histology such as poorly differentiated tumours, columnar cell tumours that do not produce mucin, or cuboidal cell tumours that do produce mucin. (39)

The cells of origin of iCCA are relevant to my work as I use animal models of iCCA where the progenitor cells are hepatocytes, and this will be discussed later. The cells of origin of iCCA change depending on the type of iCCA that presents e.g. small bile duct and cholangiolocarcinoma iCCAs arise from small intrahepatic ducts and bile ductules. These ducts and ductules are lined with hepatic stem or progenitor cells and cuboidal cholangiocytes. These types of iCCA usually develop on a background of chronic liver disease. (41) Considering large bile duct iCCA, these tumours form from columnar mucous cholangiocytes or peribiliary glands. (16) Despite data in support of the cells of origin above, there is controversy regarding the cellular origins of iCCA, and current evidence is insufficient to determine the definitive origin of iCCA. (41)

1.7 iCCA epidemiology and risk factors

The cancer at the focus of this thesis, intrahepatic cholangiocarcinoma, is the second most common primary hepatic malignancy, and its incidence and mortality is increasing worldwide. (42) This is particularly true of the Western hemisphere, where it is believed to be linked to the increase in obesity, metabolic disease, diabetes and alcohol consumption. (43,44)

The abysmal survival rate of less than 5%, 5 years post prognosis, is in part attributed to the largely asymptomatic presentation of the disease, resulting in late diagnosis when the disease has already begun to metastasise to secondary sites. (41) Despite its classification as a rare cancer, the incidence of cholangiocarcinoma is increasing globally (0.3-6 per 100,000 individuals presenting with cases per year) (45) and the mortality (1-6 deaths per 100,000 per year), increases too. (42) These data do not, however, take into account specific regions of South East Asia such as Northern Thailand, South Korea, and China where the incidence has risen to >6 cases per 100,000 inhabitants. (46) iCCA is known to have a general overall higher incidence in men than women worldwide. (42)

There are several risk factors associated with iCCA, and almost all the risk factors for iCCA share the common characteristic of being associated with bile stasis and chronic inflammation of the biliary epithelium. (47) Interestingly, the risk factors for this cancer also vary significantly with geographic location: infection with the liver flukes *Opisthorchis viverrini* and *Clonorchis sinensis* are the leading risk factor of iCCA in East and Southeast Asia, whereas in the West the risk factors are more generalised. (44,48) High incidence of iCCA correlates with higher infection with liver flukes, especially in the Khon Kaen region of Northeast Thailand. (49) *Opisthorchis viverrini* is endemic in Northeast Thailand where it has a prevalence of 9.4% and is also found in Laos, Vietnam and Cambodia. (50) *Clonorchis sinensis* is found in Korea, China, Taiwan and Vietnam. (51) The pathology of the liver flukes involve completion of a section of their lifecycle in the human liver, specifically after ingestion of the flukes via raw, fermented and/or partially cooked fish, the fluke travel to the distal bile ducts where they develop to a sexually mature stage and where they can exist in humans for 10-25 years. (48)

The presence of the flukes in the distal bile ducts causes occlusion of these ducts and elicits an immune response resulting in periductal fibrosis and scarring. The repeated rounds of injury and repair damage the ducts and this can result in formation of iCCA. The heavier the infection with the fluke, the more thickening of peripheral bile ducts, and this loss of homeostasis between tissue repair, synthesis and degradation of fibrotic tissue leads to fibrosis and can ultimately lead to iCCA. (48,52)

In the West, the risk factors associated with iCCA are less well correlated but primary schlerosing cholangitis is a known risk factor. This disease is characterised by chronic inflammation and peribiliary fibrosis which causes cholestasis and also damages bile ducts which leads to strictures and liver fibrosis. (53) Primary Schlerosing Cholangitis (PSC) is a well-known risk factor for development of iCCA, because of the inflammatory nature of this disease that results in proliferation of progenitor cells. The lifetime incidence of PSC patients with cholangiocarcinoma is reported to be between 5 and 10%. Half of these PSC patients have the diagnosis of cholangiocarcinoma within two years of being diagnosed with PSC. (54) Interestingly, *KRAS* mutations were detected in 30% of bile from PSC patients, and it has been suggested this is an early event that is a contributing factor to the malignant behaviour of cholangiocytes. (55)

Other established risk factors include alcohol consumption, smoking tobacco, obesity, diseases of the biliary tract e.g. choledochal cyst, cholelithiasis, and cirrhosis. (44) Heritable autosomal recessive diseases that can predispose individuals to iCCA include haemochromatosis and Caroli disease. Haemochromatosis affects the hemochromatosis gene (HFE) gene that controls iron absorption, which has been linked to iCCA via a Surveillance, Epidemiology, and End Results (SEER)-Medicare-based study. (56) Caroli disease involves the segmental saccular dilatation of the intrahepatic

bile ducts, which can lead to formation of bile duct stones and the subsequent issues associated with duct blockage. (57) Caroli disease was identified as the strongest risk factor for iCCA in the SEER-Medicare-based study. (56)

1.8 Genetic and epigenetic alterations in iCCA

The genomic profile of iCCA has been investigated in several patient studies over the past decade, and this has allowed the identification of a small number of recurrent genetic alterations. (58) Massive sequencing studies in particular have improved understanding of the causations of iCCA, and these have shown alterations in several areas including kinase signalling; *KRAS, ERBB1, BRAF, PIK3CA, PTEN, SMAD4, FGFR1-3,* cell cycle regulation and DNA damage repair mechanisms; *TP53, CDKN2A, BRCA1, BRCA2,* SWI-SWF complex involvement; *PBRM1, ARID1A, FGFR2* fusions, WNT–CTNNB1 pathway, Hippo signalling; *NF2, SAV1* deletion, epigenetic control of NADPH signalling; *IDH1/2* and de-ubiquitination; *BAP1.* (59–66) The most therapeutically relevant alterations identified in iCCA are associated with epigenetic processes, most notably there are hotspot mutations observed in $IDH1^{R132}$ and $IDH2^{R172}$ that accumulate an oncometabolite, 2-hydroxyglutarate (67) The other most notable mutations are gene fusion events between FGFR2 and several partner genes including *BiCCA* (64,66), *PPHLN1*, (63) *TACC3* and *MGEA5*. (66)

Javle and colleagues, for example, identified driver genes of iCCA using hybrid capture based genomic profiling to examine the mutational landscape of 412 patients (68), Nakamura and colleagues used a combination of whole exome and transcriptome sequencing of 145 patients (69). Lowkey and colleagues used targeted next generation

sequencing (an assay known as the MSK-IMPACT platform) to analyse exons and selected introns of 410 cancer-associated genes in 152 patients. (70) Zou and colleagues analysed with exome sequencing 103 iCCA patients. Zhu *et al* used multiplexed mutational profiling to examine nucleic acids from 200 iCCA patient samples. (71)

Epigenetic alterations in iCCA are among the most important in allowing initiation and progression of the disease. Epigenetic alterations affect tumourigenesis without the need for changes to the DNA sequence. This means that deregulated patterns of methylation, changes to histones and aberrant expression of non-coding RNAs promote incorrect transcription and gene expression. These changes ultimately prevent cells from retaining homeostasis and allows them to grow malignantly.

Interestingly, the genetic heterogeneity of iCCA has been closely related to histological subtypes. (37) For example *IDH1/2* or *FGFR2* mutations have been associated with small bile duct type iCCA, (34,70,72,73) whereas large bile duct type has a higher frequency of *KRAS* and *TP53* mutations, more similar to pCCA or dCCA. (38,67,69,73) Also worth considering is the difference in prevalent mutations between different subtypes of CC; several studies have observed genetic distinction between iCCA, pCCA, and ECC. (60,62,69,74) The genetic changes with changes in anatomical location include iCCA having more *IDH, EPHA2* and *BAP1* mutations, and *FGFR2* fusions compared to extrahepatic tumours that show specific fusions in *PRKACA* and *PRKACB* (41), and *ARID1B*. (69)

Considering genetic predisposition to iCCA, information is limited. (75) Current data is predominantly from Genome Wide Association Studies (GWAS) of patient cohorts diagnosed with PSC (76) who have increased iCCA risk, as they already have a iCCA-

predisposing condition. (77) Whole genome sequencing has been investigated in liver fluke-related iCCA. The two liver flukes most associated with cholangiocarcinoma are *Opisthorchis viverrini* and *Clonorchis sinensis* (62) Overall, tumours in fluke-positive patients have a higher median rate of mutation of 4700 compared to 3143 somatic mutations per tumour in non-fluke related patients. (62) Also observed genetically are the differences between fluke-related and non-fluke related iCCA. A study by Chan-on and colleagues used exome sequencing to reveal distinct mutational profiles in each subtype. For example, they observed that *BAP1* and *IDH1/2* were significantly more mutated in non-fluke related iCCA, than fluke-related iCCA. Interestingly, *TP53, ERBB2,* and *SMAD4* were less commonly mutated in non-fluke related iCCA. (59)

1.8.1 *TP53*

TP53 is a tumour suppressor gene with great importance to the cell in its role as the guardian of the genome. It is often mutated or dysregulated in cancers due to it residing over so many signalling pathways. TP53 as a protein is a sequence-specific transcriptional activator protein and it has many disordered domains and as many as 50 posttranslational modification sites meaning it has many interacting surfaces for other proteins, and residues that can be modified. (78,79) Additionally, TP53 is expressed in several isoforms that can form monomers, dimers or tetramers that can also all have variable functions. The downstream effects of TP53 are vast but these mainly include apoptosis, differentiation, and prevention of cell cycle progression, senescence and DNA repair. The *TP53* gene is activated by a number of stress conditions that mean its levels are elevated and it can act to repair damage being done to a cell. It has such diverse signals and such a protective role in the body that

mutations in the gene are frequently selected for in almost all cancer types. (79) With regard to iCCA, *TP53* is one of the few recurrently mutated genes.

1.8.2 *KRAS* mutations in iCCA

Gain-of-function mutations in *KRAS* form some of the most common mutations in iCCA, and several studies state that *KRAS* mutations are found in around 7-25% of iCCAs (61,68,80–84), with one study stating the recurrence is as high as in 54% of patients. (85) The variation in the numbers of patients presenting with *KRAS*-mutant iCCA is likely due to small sample sizes. One study identified there are two classes of iCCA, the inflammation class and the proliferation class, using integrative genomics. *KRAS* mutations were associated with the proliferation class that has specific copy number changes, oncogene-associated pathway activation, such as of RAS, or MAPK pathway activation, and is correlated with worse survival of the patient. (84) Clinically, the presence of *KRAS* mutations in a Western patient cohort (Memorial Sloan Kettering Cancer Center cohort) was lower than in the China Fudan-iCCA cohort. (86) These data are interesting when considering what treatment should be the mainstay choice of treatment of iCCA depending on geographical location of the patient.

The interaction of *KRAS* and additional mutations has been studied to a small extent in human iCCA, and these data are particularly relevant to this project. One study that combined analysis of genomic (whole exome and targeted exome sequencing) and epigenetic data examined the three most commonly mutated genes in iCCA: *IDH, KRAS* and *TP53*. The study stratified data from 496 patients into groups based on the presence (or absence) of each of these most commonly mutated genes. Each group

was then attributed with a unique mutational and pathway-enrichment profile. *KRAS* group members had the highest interaction count with *TP53*, followed by *ARID1A* and *SMAD4*. These data fed into my decision to concentrate on the interaction between the oncogene *KRAS* and additional loss of function mutations that occur in iCCA. (61)

1.9 Signalling networks in iCCA

Many signalling pathways are involved in the pathogenesis of iCCA, and these include developmental pathways such as Notch, WNT, and transforming growth factor- β (TGF-β) signalling pathways that are significantly activated in iCCA when compared to HCC. Sia *et al* identified molecular subtypes of iCCA that might aid therapeutic intervention depending on the mutations that are present in certain signalling pathways. (87)

Notch signalling is involved in biliary repair, growth, fibrosis, and stem cell niche maintenance. This pathway has been associated with several liver conditions including Alagille syndrome from defective *JAG1* or *NOTCH2* mutations that impair Notch function. (88) Increased Notch activity has been associated with primary liver tumours, and overexpression of Notch receptors occurs in iCCA. (89) Notch signalling is also known to control transdifferentiation of hepatocytes into cholangiocytes during tumour progression and for this reason the overexpression of intracellular NOTCH1 receptor domain in hepatocytes is involved in the development of iCCA mouse models. (90–92)

1.9.1 The Wnt signalling pathway

The Wnt signalling pathway is highly evolutionarily conserved and involves the WNTs; secreted glycoprotein ligands that control many other signalling pathways via β-catenin dependent and independent mechanisms, also known as canonical and non-canonical Wnt pathways, respectively. The WNTs affect many different processes including cell proliferation and migration, cell survival, polarity, and cell fate specification. (93)

Canonical Wnt signalling regulates establishment of the body axis and tissue and organ development in the embryo, and is involved in tissue homeostasis and cell and tissue regeneration in the adult. The protein at the core of canonical Wnt signalling, β-catenin, is a plasma membrane associated protein that interacts with transmembrane protein E-cadherin to connect it to actin filaments and to allow physical signals to be passed from cell to cell. (94,95) β-catenin also has a role as a transcriptional regulator by its interaction with the nuclear DNA binding T cell factor/ lymphoid enhancer-binding factor (TCF/LEF) transcription factor family.

When an activating Wnt signal is not present and therefore β-catenin is not required, it is subject to degradation by the destruction complex, a dynamic assembly of several proteins. (96) During the process of β-catenin destruction, first a single α helix of Axin binds to the serine/threonine kinase glycogen synthase kinase 3 (GSK3) in a groove (97), and adenomatous polyposis coli (APC) binds to a conserved groove on the N-terminal region of Axin called the serine-alanine-methionine-proline (SAMP) repeat. (98). Casein kinase (CK1α) is also part of this complex. β-catenin then binds to Axin via its armadillo repeats (the central region of β-catenin made up of 12 repeats of 3 helices each) and / or to APC by either APC's three 15 amino acid repeats or an alternative region of APC (99); the five N-terminal residues of a region of APC

containing seven 20 amino acid repeats. (100) β-catenin is also targeted for degradation via ubiquitin labelling with β-Transducin repeat containing protein (β-TrCP) (101)

The phosphorylation of β-catenin is facilitated by the binding position of β-catenin to Axin allowing the amino termini of β-catenin close to CK1α and GSK3, and therefore can encounter the active sites of these kinases. CK1α first phosphorylates β-catenin at Ser45 before GSK3 is sequentially permitted to phosphorylate Thr41, Ser37 and Ser33. Once β-catenin is phosphorylated, β-TRCP1 (a component of the SCF ubiquitin ligase complex), recognises it and ubiquitination is permitted to proceed, followed by proteosomal degradation. (96)

In the presence of WNT ligands, β-catenin functions as the molecule at the nexus of this pathway, as illustrated in Figure 3. WNT ligands bind to the G-protein coupled receptor Frizzled causing its activation and subsequent heterodimerisation of Frizzled with LRP5 or LRP6 co-receptors, which allows transduction of an activating signal across the plasma membrane to the Dishevelled proteins. Polymerised, activated Dishevelled proteins inhibit the β-catenin destruction complex. This occurs when the active LRP5/6 receptors are phosphorylated by GSK3, meaning that Axin is sequestered to the plasma membrane by interaction with Dishevelled: this results in the entire destruction complex being localised to the membrane. (102) The sequestration of the destruction complex allows accumulation of β-catenin in the cytoplasm. β-catenin is then able to enter the nucleus and co-activate T-cell factor and lymphoid enhancer factor (TCF/LEF)-mediated transcription of target genes using its conserved armadillo repeat domain. (103) TCF/LEF proteins displace transducin-like Enhancer of split

(TLE)/Groucho complexes and recruit histone-modifying coactivators (104) This then results in many gene transcription that leads to downstream cellular processes.

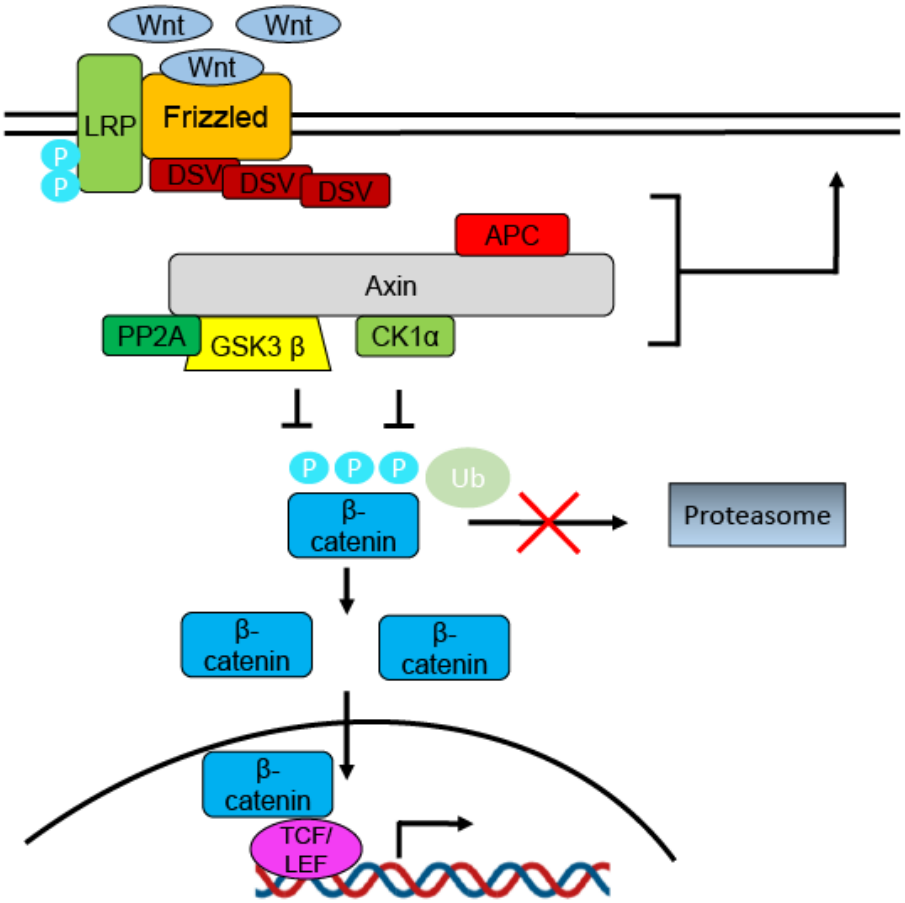

Wnt
Wnt
Wnt
LRP
Frizzled
P
P
DSV
DSV
DSV
APC
Axin
PP2A
GSK3 β
CK1α
P
P
P
Ub
β-
catenin
Proteasome
β-
catenin
β-
catenin
β-
catenin
TCF/
LEF

Figure 3

Canonical Wnt signalling is activated upon the binding of secreted WNT ligands (such as Wnt3a and Wnt1) to the transmembrane Frizzled and LRP co-receptors. LRP is phosphorylated by CK1α and GSK3β, and this leads to the recruitment of Dishevelled (DSV) proteins to the plasma membrane where they polymerise and activate. The Dishevelled proteins inactivate the β-catenin destruction complex consisting of scaffold protein Axin, APC, GSK3β and CK1α. Un-phosphorylated and un-ubiquitinated β-catenin is therefore not targeted for proteosomal degradation.

The WNT- β catenin signalling pathway is activated in most cholangiocarcinomas, and its activation is associated with tumour cell proliferation and survival in both fluke-related (105) and non-fluke related iCCA. (106) The prevalent activation of WNT signalling can be due to several reasons, including the infiltration of macrophages to the stroma releasing WNT ligands, (107,108) due to mutations to canonical WNT signalling, (109) and/or due to DNA methylation. Goeppert *et al* used quantitative mass spectrometry analysis of methyl-CpG immunoprecipitation and whole-genome CpG island arrays to show that there is aberrant methylation of WNT signalling components in iCCA. (110) This study identified that several gene promoters of this pathway including DKK2, WNT3A, and SFRP1/2 in iCCA tissue and cell lines. Similarly, Merino-Azpitarte and colleagues identified in 68 iCCA and 36 pCCA patient samples that the promoter SOX17 is also hypermethylated. (111) With regard to infiltration of macrophages releasing WNT ligands, iCCA is characterised by a highly desmoplastic stroma meaning that the pool of activated macrophages here stimulate β-catenin activation by release of WNT ligands (108) and subsequent iCCA proliferation.

1.9.2 Receptor tyrosine kinase signalling

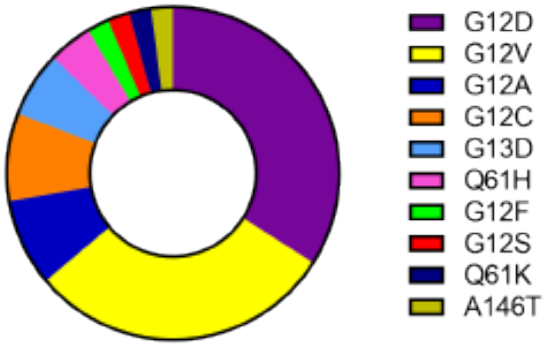
Aberrant activation of receptor tyrosine kinase (RTK) signalling is another molecular feature of iCCA, and other CC subtypes, and KRAS-activating mutations in particular are found in all CCs whereas BRAF mutations are more commonly found in iCCA specifically. (112) RTK signalling via RAS involves the downstream activation of effector pathways. For example: RAF/MAK/ERK and PI3K/AKT signalling.

1.9.3 *KRAS*

KRAS is a member of the small GTPase superfamily and this protein family are linked by their ability to catalyse the hydrolysis of guanosine triphosphate (GTP) to guanidine diphosphate (GDP), which results in the release of a phosphate molecule. The ability of RAS proteins to act as a binary switch makes them powerful signalling molecules. These signals are switched off when GDP-bound and cannot signal their effector proteins, and switched on when GTP-bound. (113)

The cyclical nature of GTP hydrolysis and GDP release is not physiologically relevant of its own fruition, so GTPase activating proteins (GAPs) are required for this process to occur. GAPs increase GTP hydrolysis by 100,000 (114) and the process is also aided by guanine nucleotide exchange factors (GEFs) which catalyse the replacement of GDP with GTP. (113)

KRAS experiences frequent oncogenic substitutions at codon G12, G13, and Q61, and these prevent the catalysis of GTP to GDP by disallowing insertion of the catalytic arginine finger that GAPs use to allow this conversion of biochemical states. (115) In iCCA, several sequencing studies of human tumour tissue carried out sequencing sufficiently deep to establish the mutation occurring when *KRAS* was mutated. I have collated these data in Figure 4 to show that proportionally the most well mutated codon change of *KRAS* in iCCA is G to D at codon 12. Of the 47 tumour samples that were examined, this represents 34% of the *KRAS* mutations. Overall, codon 12 was mutated in 85% of the samples examined, with the remaining mutations occurring at either codon 61 or 146.



Total=47

Figure 4

Proportional representation of oncogenic KRAS *point mutations in iCCA. These data were collated from several genome sequencing studies of patient iCCA tumours. (59,81,83,84) Total n=47.*

1.9.4 RAS effector pathways

One key effector pathway that KRAS activates is the Raf/MEK/ERK axis. After activation, RAS initiates a kinase cascade where kinases are sequentially phosphorylated; RAS phosphorylates the serine threonine kinase RAF, which dimersies and becomes active. Active RAF can then activate MEK by phosphorylation and finally ERK, the terminal effector, is activated and can translocate to the nucleus to interact with transcription factors that control cell proliferation, survival, and differentiation. (116)

Another key effector pathway is the PI3K/AKT signalling pathway that is also activated by KRAS. PI3K/AKT signalling has plays important roles in cell growth, metabolism, proliferation and survival. (117) Phosphatidylinositol-3-kinase (PI3K) is a lipid kinase that forms phosphatidylinositol-3,4,5-trisphosphate (PIP3) from phosphatidylinositol 4,5-bisphosphate (PIP2). PIP3 is an essential secondary messenger for the translocation of AKT to the plasma membrane where it is phosphorylated and activated by phosphoinositide dependent kinases 1 and 2 (PDK1/2). (118)

1.9.5 Hippo signalling

The Hippo pathway regulates organ size during organogenesis by managing the homeostasis of cell proliferation and apoptosis. (119) The upstream activators of Hippo signalling include G-protein coupled receptors (GPCRs), Fat cadherins, and 4.1, ezrin, radixin, moesin (FERM)-domain containing proteins such as NF2, which is encoded by the gene neurofibromin 2 (NF2). (120,121) The Hippo signalling pathway functions through a kinase cascade of the initial activation of Ste20-like serine threonine kinases

MST1 and 2 by an upstream activator protein. MST1/2 then activate LATS 1 and 2 by phosphorylating two highly conserved motifs. Once activated, LATS1 and 2 phosphorylate their transcriptional co-activating protein YAP and TAZ which allows their retention in the cytoplasm so they are degraded by the proteasome and they cannot translocate to the nucleus and activate their associated genes via binding partners. The binding partners of YAP/TAZ are involved in pro-proliferative and anti-apoptotic pathways for example TEAD1-4, SMAD, p73, TBX5, ERBB-4. (122)

In the liver, Hippo signalling is required for control of organ size during development and regeneration, and dysregulation of Hippo signalling for example from loss of NF2 or YAP over-activation results in hepatomegaly, choledochal cysts, and tumourigenesis. (123,124)

In the adult liver, YAP is located in the nucleus of biliary cells and when Hippo signalling is active, it ensures that mature hepatocytes maintain their lineage by maintaining YAP in the cytoplasm and therefore being degraded and not acting as an activator protein. (124) Interestingly, when NF2 is lost and YAP is subsequently translocated to the nucleus of hepatocytes, they dedifferentiate into a progenitor state and can then be re-programmed into ductal lineage cells. This process occurs via NOTCH2 and SOX9 genes being activated by TEAD4 transcription. (124)

1.10 iCCA Molecular pathogenesis

iCCA progresses molecularly via dysregulation of the signalling pathways described above, and also molecular events. (87) It is thought to stem from malignant cholangiocytes and in the case of small duct iCCA, from progenitor cells. Studies have

recently linked stemness to poor prognosis, similar to that observed in HCC. It is thought that iCCA most frequently arises on a background of chronic inflammation, and the pro-inflammatory signals that are present, such as cytokines and growth factors, promote the proliferation of cholangiocytes. (125) The addition of mutations in genetic or epigenetic features causes cells to behave malignantly. (126)

1.11 Interaction with tumour microenvironment

During the progression of tumourigenesis, many changes take place in the protective processes of the body attempting to clear the tumour. The amalgamation of changes in immune cell migration and stroma composition, the remodelling of the extracellular matrix and recruitment of cancer-associated fibroblasts, as well as the altered state of angiogenesis act together to form the tumour microenvironment. iCCA tumours are often highly stromatic, and the stroma is made up of many different cell types. These include cancer-associated endothelial cells, cancer associated fibroblasts (CAFs), and several types of inflammatory cells, including macrophages, neutrophils, natural killer (NK) and T cells. (127,128) The stroma also contains many extracellular matrix proteins like laminin, fibronectin, and several collagens. (129)

Interestingly, the presentation of iCCA tumour stroma are similar to those found around bile ducts of premalignant disease usually seen in primary sclerosing cholangitis or congenital hepatic fibrosis. This is indicative that the stromatic nature of the iCCA tumours is similar to that of the regenerative microenvironment found during ductal repair after damage. The formation of a tumour niche around iCCA tumours links tumour-caused tissue damage to progenitor cell mediated tissue repair. (130)

Cholangiocarcinomas are characterised by a particularly dense stroma that contains many cancer associated fibroblasts. (34) The recruitment of cancer associated fibroblasts in particular is key in the tumourigenesis and metastasis of iCCA, as they encourage extracellular matrix remodelling to become stiffer and therefore affect cellular mechanotransduction. Smooth muscle actin positive CAFs are particularly prominent in iCCA. (127)

1.12 iCCA detection

The only current curative treatment for iCCA is liver resection however, the majority of patients, 70%, are not amenable to surgery due to late stage diagnosis. (131) iCCA is often only detected at an advanced local or metastatic stage due to there being a lack of specific symptoms. (132) Patients report nonspecific symptoms such as abdominal pain, weight loss, nausea, and malaise, but these are already indicative of advanced stage disease. Jaundice is another possible symptom but it is rare as it occurs only when the tumour is located in such a way that the hepatic ducts are obstructed. (133,134) As a result, most patients are diagnosed incidentally or via cross-sectional imaging for another indication. (133) When a patient is suspected of having iCCA, it is suggested that diagnosis would include contrast-enhanced computed tomography (CT) and / or magnetic resonance imaging (MRI), as these investigative techniques allow the examination of local invasion of the cancer, how much of the liver is left intact and what the state of the surrounding organs is. (41,135,136) One such issue that appears is the inability of these techniques to differentiate between iCCA and HCC when applied to small tumours of <2cm in diameter. The distinction between the types of cholangiocarcinoma is also vital to determining effective treatment options. The

potential biomarker carbohydrate antigen 19-9 (CA19-9) has been studied extensively but as it is non-specific for cholangiocarcinoma itself- levels of CA19-9 are altered significantly in cholangitis and cholestasis- let alone determination of different cholangiocarcinoma types, it is inappropriate as a marker for this disease. (137) Presently there is a great need for alternative diagnostic tools for iCCA. One such recent breakthrough in diagnostic tools for iCCA is the use of liquid biopsy. Liquid biopsy comprises several methods that all focus on detection of biomarkers of tumour presence from fluid samples. Fluid (blood, urine, plasma etc.) is taken from the patient in a non-invasive manner, and these samples can then be subjected to a range of methods, including next generation sequencing, to reveal whether there is any tumour-derived genetic material present. (138) Several types of liquid biopsy are currently available for cholangiocarcinoma, including detection of cell-free DNA (cfDNA) and circulating tumour DNA (ctDNA), miRNA and extracellular vesicles and various proteins, cytokines and other metabolites found in serum. (138) Interestingly, a 2016 study by Andersen and Jakobsen used multiplex digital PCR analysis was used to screen for circulating cfDNA of 31 somatic RAS and RAF mutations, including $KRAS^{G12D}$. This screen used multiplex analyses to find mutations and then duplex analyses to identify and quantify these mutations in KRAS, NRAS, BRAF and PIK3CA genes in the plasma. This technique was used on samples from patients with cholangiocarcinoma who had known mutational status, and the mutations found in plasma samples were also present in the tumour.(139) The use of digital multiplex PCR as a diagnostic tool for iCCA would be significant to identify what treatment process would be most successful depending on the mutational profile of the tumour present.

1.13 iCCA Treatment

1.13.1 Surgery

The overall probability of iCCA patients being cured by liver resection is 9.7%, the median overall survival is 40 months and the 5-year survival is 25-40%. (131,133,140) How suitable a patient is to liver resection depends on the location of the tumour in relation to intrahepatic biliary tract and vasculature, and the quality and quantity of the remaining liver parenchyma after tumour removal. Small or peripheral lesions can be removed by atypical or non-anatomical resection, and anatomic hepatectomies are performed on larger tumours. (141) Indeed, because of the late stage diagnosis of iCCA, most patients will require a major hepatectomy where three or more segments are resected, and 50-70% of patients with resectable iCCA undergo hemi- or extended hepatectomy. (133) Despite efforts to ensure patients are suited to surgery, such as inspection for evidence of metastasis that would reduce their likelihood of survival post-surgery, recurrence in the liver after resection still occurs. For this reason, adjuvant therapy is used for liver resection patients, most commonly with the admission of gemcitabine and oxaliplatin, (142) and capecitabine (143).

1.13.2 Transplantation

Liver transplantation for pCCA was originally passed over as an option due to a high rate of recurrence of around 50%. (144,145) However, a study of 216 patients with early stage, unresectable pCCA were treated with neoadjuvant chemoradiotherapy followed by liver transplantation and had a promising 5 year disease-free survival of 65%. (146) For this reason, liver transplantation has been considered as another option

for iCCA treatment, and is thought to be a valuable option for patients with cirrhosis and tumours smaller than 2cm. (147) A small study of iCCA and mixed iCCA-HCC patients showed that the five year survival of patients who had had liver transplants was 65%, and this was further supported by another study that looked back at liver transplant data from previous patients. (147,148) A smaller cohort of patients, where n=6, with large and unresectable iCCA were treated with liver transplantation after treatment with neoadjuvant chemotherapy that caused a long period of disease stability with a 5-year survival of 83%, however recurrence occurred in 50% of patients. (149)

1.13.3 Palliative chemotherapy

When the curative option of surgery is not possible, there are only palliative treatments available. The standard mode of first-line chemotherapy is with cisplatin-gemcitabine which, in the ABC-02 trail of 410 patients when compared to treatment with gemcitabine alone, had a better overall survival. (150) This was also supported by the BT22 study in 2010 examining 83 patients. (151) Gemcitabine is a deoxycytidine analogue that halts DNA polymerisation in a way that prevents exonuclease activity from removing the drug after it has been incorporated into RNA or DNA, thus halting tumour cell growth. Cisplatin forms adducts within the DNA that prevent cell division by DNA distortion. These drugs have complimentary mechanisms and therefore have been used as combination treatments for several other cancers as well as iCCA. (152) Combination treatments of cisplatin-gemcitabine with other chemotherapeutics in a trio have been through phase III randomized clinical trials for advanced stage biliary tract cancer, for example cisplatin, gemcitabine and nab-Paxlitaxel (153) or

FOLFIRINOX (154). Another study showed the efficacy of gemcitabine and S-1 over the use of cisplatin-gemcitabine. (155) Finally of note is the current phase III trial of acelarin (NUC-1031). Acelarin with cisplatin will be compared with gemcitabine and cisplatin combination therapy in a phase III study. (156)

1.13.4 Second line treatment

Following the failure of first-line chemotherapy for treatment of iCCA, the phase III ABC-06 clinical trial of FOLFOX (folinic acid, 5-FU and oxaliplatin) showed that there was a benefit from this second-line chemotherapy treatment. Compared to the control group used in this study, patients treated with FOLFOX had a higher survival at 6 months; 35.5% versus 50.6%, and 12 months; 11.4% versus 25.9%. (157) Thus, FOLFOX is now considered to be a new standard of care as a second-line iCCA treatment.

1.13.5 Targeted therapies

Sequencing studies have been used to examine specific genomic profiles of patient iCCAs to consider whether targeted therapies would be suited to some subgroups. Nepal *et al* used a combination of targeted and whole exome sequencing to identify iCCA-related cancer genes. They concentrated on identifying oncogenic mechanisms and segregating tumours into four subsets that correspond to the potential use of therapeutics: RNA synthesis inhibition in IDH mutants; microtubule modulator for the KRAS mutants; topoisomerase inhibition in the TP53 mutant; and mTOR inhibitors. There are targeted therapies for certain signalling pathways that are aberrantly upregulated in iCCA for examples therapeutics targeting the Notch signalling cascade.

These include gamma-secretase inhibitors which function by blocking proteolytic cleavage of Notch receptors. (158) Also the WNT- β catenin signalling pathway, as described above, is frequently upregulated in iCCA (109) and there are a number of WNT inhibitors that have been shown to successfully reduce and inhibit tumour growth in experimental models of iCCA. For example, the WNT activating ligand Wnt7b is upregulated in human iCCA tumours and also in animal models of iCCA and it was shown that when the macrophages secreting this ligand were depleted in murine models, iCCA growth halted. (108,159) Additionally, the use of WNT- β catenin inhibitors C-59 and ICG-001 have also been shown to reduce tumourigenesis in mouse and rat models (159), and clinical trials are ongoing in the use of WNT- β catenin inhibitors. (109) Targeted therapies that inhibit IDH1 (ivosidenib, IDH305), IDH2 (AG221), and both IDH1/2 (AG881) are currently in testing phase in iCCA patients. (41,160) Finally, a class of inhibitors targeting FGFRs are proving promising in phase II studies. (161–164)

1.14 CRISPR-Cas9 library screening: an indispensable member of the genetic toolkit

CRISPR-Cas9 screening technology has had a significant effect on biological research since it was first developed. The CRISPR (Clustered Regularly Interspace Short Palindromic Repeat/CRISPR-associated nuclease 9) gene editing system is a powerful genome editing tool based on bacterial and archaeaic machinery. CRISPR-mediated gene editing is an efficient and less toxic way to introduce mutations in a controlled manner compared to older techniques. Its use in cancer research has extended from use in animal tumour models to phenotypic screens, to development of cancer gene

therapies. The significant contribution of CRISPR-Cas9 technology to the progression of cancer research is that it allows the examination of functional genomics of targeted gene modification.

The prokaryotic type II CRISPR, and functionally related CRISPR-associated (Cas) genes are a system that was derived from the adaptive immune system of Gram-positive coccus *Streptococcus pyogenes* that acts as a human pathogen. (165) Bacteria exchange DNA via the acquisition of genes from other bacteria or phages, so as to acquire novel traits that affect pathways that will infer environmental advantage such as increased pathogenesis or antibiotic resistance via horizontal gene transfer. However, phage invasion is not always beneficial and thus bacteria have also developed innate and adaptive immune response strategies to interfere with a non-beneficial infection. The constant evolution of phages and bacteria results in a so-called “arms race” where, in the case of *S. pyogenes*, an adaptive immune system is required that can specifically recognise and target an invasive phage after it has already come into contact with it previously by remembering its genome. (166)

The CRISPR array is made up of identical repeats separated by short unique ‘spacer’ sequences that function as the memory part that allows bacteria to recognise intruding phages when they re-infect the bacteria. The Cas component is, in the case of the CRISPR-Cas9 systems used as part of a genetic toolkit, a gene transcribing a nuclease protein that is required to target sequences and perform the subsequent DNA cleavage. The system functions by the *S. pyogenes* Cas9 protein (*sp*Cas9) being targeted to DNA by a single guide RNA (sgRNA) that consist of a 20 nucleotide guide sequence flanked by a scaffold sequence. When the guide sequence binds to the genomic DNA target directly upstream of a specific NGG (any nucleotide-guanine-

guanine motif), a double strand break is introduced approximately 3 base pairs upstream of the protospacer adjacent motif (PAM) sequence. The double strand breaks are then repaired by one of two pathways, either by non-homologous end joining (NHEJ) or homology-directed repair (HDR). NHEJ is an imprecise mechanism that results in insertions and deletions of variable lengths at the site of the double strand break, and HDR, in the presence of a template donor DNA strand, is capable of introducing precise insertions as instructed by the donor template. (165)

1.15 The Origin of CRISPR Library Screening: *in vitro*

Genetic screening using CRISPR technology was first carried out by Wang *et al* in 2014 in the form of a large library screen using 73,000 lentiviral sgRNAs targeting around 7000 human genes in mammalian cells, were used to screen for several parameters. The study importantly showed the possibilities exposed by the use of a loss of function sgRNA screening library to introduce a controlled landscape of mutations. (167)

Likewise, Zhou *et al* also developed a lentiviral library that targeted genes in human cells that allowed knockout and subsequently identification in combination with high-throughput sequencing analysis, the genes required for infection of cells by diphtheria toxins and anthrax. (168)

1.16 The move into living systems: CRISPR-Cas9 *in vivo*

In vivo screening allows examination of the complex microenvironment experienced by tumours cells, their interaction with neighbouring cells, and their 3D organisation, unlike immortalised cancer cell lines used for *in vitro* screening.

Additionally, other key components of the living organism such as the immune system, structural components like the extracellular matrix, and vascularisation interact with cancer cells and this is not explored using *in vitro* screening methods. Although *in vivo* screens can be direct or indirect, only the direct screening method will be discussed here, as it forms the basis of this thesis. Direct *in vivo* screening involves the targeting of cells existing within their natural environment, and techniques exist that successfully allow CRISPR-Cas9 targeting in the lung and brain (169), as well as the liver. (170)

In the case of the liver, naked plasmid DNA can be directly targeted to hepatocytes via hydrodynamic tail vein injection. (171) (172) Weber *et al* were the first to implement CRISPR-Cas9 multiplexed mutagenesis in adult mice by delivery of sgRNAs to the liver which they induced hepatocellular carcinoma (HCC) and iCCA. The screen led to the discovery of cancer genes and the first mouse model of CRISPR-Cas9-induced HCC. (173,174)

1.17 CRISPR-Cas9 library screening as a cancer research tool

The myriad genetic and epigenetic changes found in both tumour suppressor genes and oncogenes that are characteristic of cancer, are also responsible for the difficulty of studying and treating this disease. The ability to model cancer, then, by manipulation of *in vivo* models is paramount to the progression of any study around it.

The CRISPR-*sp*Cas9 system described above provides an easily programmable system that allows targeted gene editing by the direction of the nuclease Cas9 to the desired genomic loci. This occurs via Watson-Crick base pairing of the sgRNA and the genomic

target DNA followed by a protospacer adjacent motif of NGG. The ease and controllability of the sgRNA library system is why it was chosen to identify driver genes of iCCA from a function-based screening basis.

Chapter 2 Materials and Methods

2.1 Generation of CRISPR-*sp*Cas9 screening library

To generate the library, oligonucleotides encoding sgRNAs to target the 91 genes of interest were designed based on the mouse GeCKo V2 library with PCR retrieval sequences as shown in Table 1, as derived from the parallel oligonucleotide retrieval protocol published by the Luo lab. (175) These were then custom synthesised by Twist Biosciences using their silicon chip method. All library sgRNA sequences are listed in Appendix 1.

*Table 1 Oligonucleotide schematic for creation of pooled CRISPR-*sp*Cas9 library*

Forward Primer Binding Sequence	**5' Esp3I cut site**	**Sequence (5' - 3')**	**3' Esp3I cut site**	**Reverse Primer Binding Sequence**
GCACAGTTTCACGATCGCTTTAGC	CGTCTCACACCG	n = 20	GTTTTGAGACG	CTGTCGTGGAATCGCTAAAGCAAG

Oligonucleotides for all target gRNAs and control sequences were amplified from the library pool using the primers in Table 2 on a thermocycler with NEBNext HF PCR Mastermix (New England Biosciences: M0541S), and purified with QIAquick Nucleotide Removal Kit (QIAGEN: 28304). Purified oligos were digested with restriction enzymes to yield overhangs amenable to cloning with Esp31 (New England Biolabs: R0734S) and simultaneously phosphorylated with T4 Polynucleotide Kinase (New England Biolabs: M0201S) at 37°C for 30 minutes and then purified by isopropanol precipitation. This was carried out in triplicate and then samples were pooled together before ligation with T4 ligase (New England Biosciences: M0202S) into the SB-CRISPR plasmid backbone (also called pX330-U6-Chimeric_BB-CBh-hSpCas9, Addgene Plasmid #42230, kindly provided by Dr Roland Rad LMU Munich, who's lab originally generated the plasmid(173)), as shown in Figure 1. Pre ligation, the backbone was digested with

Esp3I, and simultaneously dephosphorylated with FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher Scientific: EF0654). The plasmid products were separated by 2% agarose gel electrophoresis for 45 min at 120V and extracted and purified (Zymoclean Gel Recovery Kit Zymo Research: D4008).

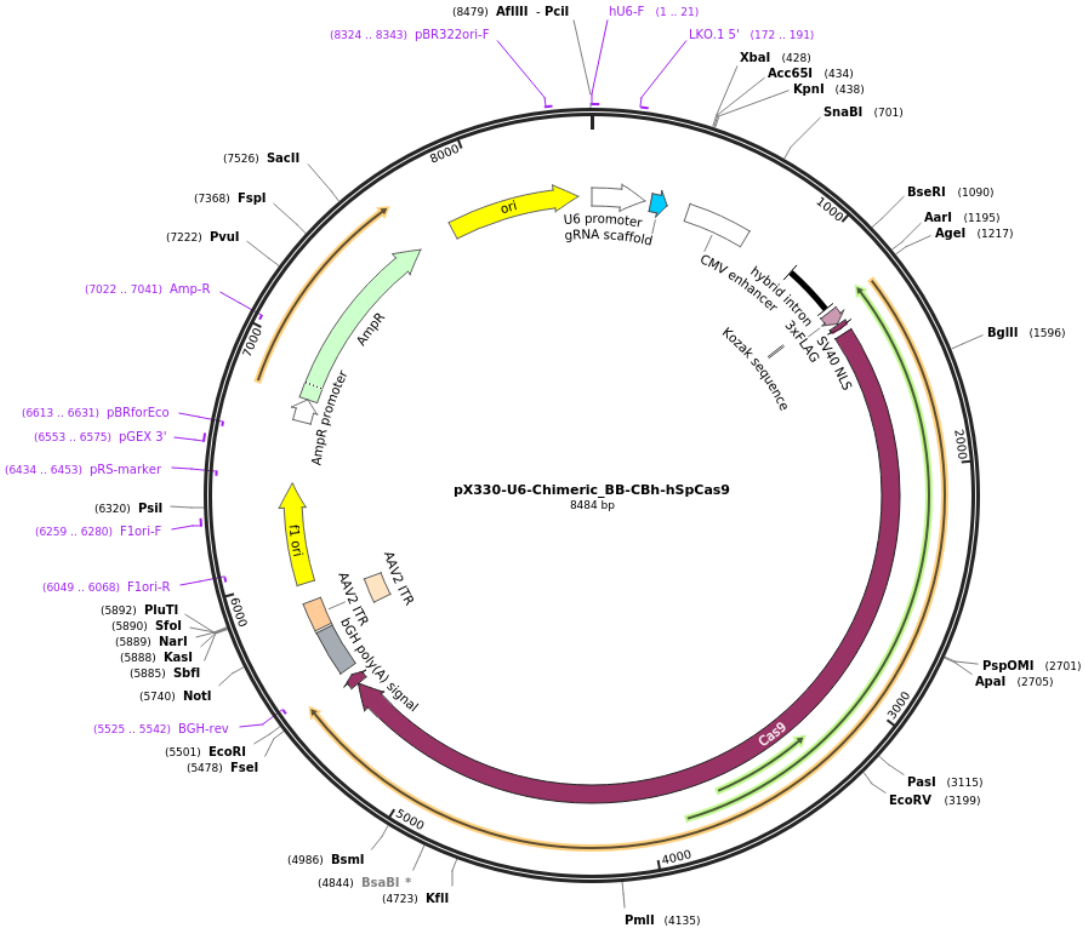

(8479) AflIII - PciI
hU6-F (1 .. 21)
(8324 .. 8343) pBR322ori-F
LKO.1 5' (172 .. 191)
XbaI (428)
Acc65I (434)
KpnI (438)
SnaBI (701)
BseRI (1090)
AarI (1195)
AgeI (1217)
BglII (1596)
PspOMI (2701)
ApaI (2705)
PasI (3115)
EcoRV (3199)
PmlI (4135)
(4723) KflI
(4844) BsaBI *
(4986) BsmI
(5478) FseI
(5501) EcoRI
(5525 .. 5542) BGH-rev
(5740) NotI
(5885) SbfI
(5888) KasI
(5889) NarI
(5890) SfoI
(5892) PluTI
(6049 .. 6068) F1ori-R
(6259 .. 6280) F1ori-F
(6320) PsiI
(6434 .. 6453) pRS-marker
(6553 .. 6575) pGEX 3'
(6613 .. 6631) pBRforEco
(7022 .. 7041) Amp-R
(7222) PvuI
(7368) FspI
(7526) SacII
ori
U6 promoter
gRNA scaffold
CMV enhancer
hybrid intron
3xFLAG
SV40 NLS
Kozak sequence
AmpR
AmpR promoter
f1 ori
AAV2 ITR
AAV2 ITR
bGH poly(A) signal
Cas9
pX330-U6-Chimeric_BB-CBh-hSpCas9
8484 bp
1000
2000
3000
4000
5000
6000
7000
8000

Figure 1

SB-CRISPR plasmid used for the in vivo *experiments described. sgRNAs targeting the selected genes were cloned into the plasmid using the sleeping beauty transposon system.*

Ligation of inserts into SB-CRISPR plasmid backbone via the sleeping beauty inverted repeat transposon system was done in triplicate with T4 ligase (New England Biosciences: M0202S), overnight at 16°C.

*Table 2 Primers for amplifying CRISPR-*sp*Cas9 gRNA library*

Forward Primer	**Reverse Primer**
GCACAGTTTCACGATCGCTTTAGC	CTTGCTTTAGCGATTCCACGACA

Ligation reactions were heat inactivated and then transformed into One Shot Stbl3 Chemically Competent E. Coli as per the manufacturer's instructions (ThermoFisher Scientific: C737303) in 10 technical replicates. The use of triplicate reactions above and technical replicates here was to ensure unbiased, complete library representation. Transformed cells were grown overnight at 37°C on Luria Broth agar in 500cm^2 square plates (ThermoScientific Nunc Square Culture Dish: UY-01929-00). All colonies were then scraped into 2L flasks and cultured for 6 hours in Luria broth with shaking at 250 rpm at 37°C before they were pelleted and the DNA purified (QIAGEN Plasmid Maxi Kit). Amplicon sequencing was carried out to determine representation and equal representation of sgRNA sequences. . The GINI index was used to confirm sgRNA representation in the spCas9-CRISPR library. The GINI index is a measure of perfect equality, represented by 0, to perfect inequality at 1. Originally the GINI index was designed to describe wealth or income inequality in a population, but is applied here to represent whether there is an equal number of reads covering each gRNA sequence – a GINI index of 0 - or whether every read covers one gRNA sequence only, meaning there is only one gRNA sequence present in the whole library. In this case, the library had a GINI coefficient of 0.07, meaning there were no unrepresented gRNAs.

2.2 Cloning of single guide RNA

Single guide RNAs (sgRNA) were cloned into SB-CRISPR plasmid (provided by Dr Roland Rad LMU Munich). The oligonucleotides encoding guide RNAs are listed in Appendix 1 and highlighted. SB-CRISPR plasmid was digested with Esp3I (New England Biolabs: R0734S) or BbsI (New England Biolabs: R3539S) to create overhanging ends suitable for cloning as was done above for the CRISPR-spCas9 library oligos. They were simultaneously dephosphorylated by FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher Scientific: EF0654). The plasmid products were then separated by 2% agarose gel electrophoresis for 45 min at 120V and extracted and purified (Zymoclean Gel Recovery Kit Zymo Research: D4008) sgRNA oligonucleotides were annealed and phosphorylated with T4 Polynucleotide Kinase (New England Biolabs: M0201S) for 30 min at 37°C. The digested backbone and annealed oligo inserts were then ligated overnight at 16°C. Ligation reactions were inactivated by heat and then cloned into One Shot Stbl3 Chemically Competent *E. Coli* as per the manufacturer's instructions (One Shot Stbl3 Chemically Competent *E.coli*; ThermoFisher Scientific: C737303) and grown on LB agar plates overnight at 37°C. Colonies were selected and Sanger sequenced (MRC Human Genetics Unit Technical Services) and correctly ligated plasmids were cultured in liquid LB with ampicillin (1/500) and plasmid DNA extracted as per the manufacturer's instructions. (QIAGEN Plasmid Maxi Kit). Plasmid concentration was determined using a NanoDrop ND-1000 Spectrophotometer (ThermoFisher Scientific).

2.3 *In vivo* Work

2.3.1 Hydrodynamic tail vein injections (HDTVI)

HDTVI and animal work was carried out by myself, Luke Boulter and Konstantinos Gournopanos. Animals were maintained in colonies in 12-hour light–dark cycles and were allowed access to food and water *ad libitum*. Female FVB/N mice were weighed and then injected with physiologic saline solution equal to 10% weight/volume.

Experiments were carried out to induce formation of iCCA in mice using three plasmids. To express human KRAS with a G to D substitution at amino acid 12, the KRAS-IRES-GFP plasmid kindly gifted from Professor Diego Calvisi was used. This plasmid was developed in their lab. (176) This plasmid contains an internal ribosome binding site (IRES) that allows enables the coordinated co-expression of $KRAS^{G12D}$ and GFP, allowing visualisation of where the $KRAS^{G12D}$ protein was expressed by examination for presence of GFP. This system uses CRISPR-based Somatic Oncogene kNock-In for Cancer Modeling (SONIC) technology whereby homology independent repair allows for the integration of $KRAS^{G12D}$ into the genome, and subsequent protein expression. The SB-CRISPR plasmid was also used, as described in section 2.1. Finally, the sleeping beauty 13 (SB13) transposase plasmid. Typically 20µg $KRAS^{G12D}$ plasmid, and *sp*Cas9-CRISPR library or sgRNA plasmids were injected per mouse, and 6µg SB13 transposase plasmid were delivered to the lateral tail vein of mice in 7-10 seconds. When >1 sgRNA was required, plasmids were dissolved to a maximum concentration of 20µg.

2.3.2 Production of K19-CreERT:Pten$^{flox/flox}$:Trp53$^{flox/flox}$:R26RLSL^{tdTomato} mice (KPPTom) Mice

KPPTom mice were generated as stated in Younger and Wilson *et al* (1). Keratin-19-^{CreERT} mice (Jax: 026925) were crossed with animals containing floxed alleles of *Pten* (Jax: 006440) or *Trp53* (Jax: 008462) and a silenced tdTomato reporter targeted to the Rosa26 locus (Jax:007908). All animals in this study are heterozygous for Keratin-19^{CreERT}, homozygous for *Trp53^{flox}* and *Pten^{flox}* alleles, and homozygous for *R26RLS^{LtdTomato}*. Mice received three doses of 4 mg of tamoxifen by oral gavage and followed by 400 mg Thioacetamide in their drinking water to induce liver damage and subsequent tumour formation. All mice were male.

2.3.3 Therapeutic dosing of animal models

KRAS^{G12D}:sgRNA^{Nf2}:sgRNA^{Trp53} mice harbouring tumours were randomized using GraphPad online randomization tool and dosed with either vehicle alone (10% DMSO, 40% PEG300, 5% Tween-80, and 45% saline), 5 mg/kg LGK974, 50 mg/kg Pictilisib, or a combination of both therapeutics. Dosing was carried out starting 7 days following HDTVI.

KPPTom animals were given tamoxifen and Thioacetamide (as detailed above) and at 4 weeks were given either vehicle alone (10% DMSO, 40% PEG300, 5% Tween-80, and 45% saline) or a combination of 5 mg/kg LGK974 and 50 mg/kg pictilisib for 4 weeks. All animals were housed in colonies of five animals.

2.4 Processing of mouse tissue

Mice were euthanised by Schedule 1 method of increasing concentration of carbon dioxide inhalation, and death confirmed by cervical dislocation. Livers were perfused to remove blood with 10mL phosphate–buffered saline through the inferior vena cava, and the hepatic portal vein was cut to allow for drainage. Livers were then excised from the mice and stored in cold PBS until they were photographed. Livers were then dissected and where appropriate tumours were excised. Lobes and tumours were then taken for histology, DNA/ protein extraction and RNA extraction. Tissue for DNA/protein extraction was snap frozen in 2-methylbutane chilled on dry ice, and stored at -80°C. Tissue for RNA was perfused in RNAlater (ThermoFisher Scientific: AM7020) overnight at 4°C before long term storage at -80°C. Tissue liver sections taken for histology were fixed in 4% formaldehyde (VWR Chemicals: 9713.5000) for 24 hours and then transferred to 70% ethanol for 24 hours. Tissue liver sections were then perfused with paraffin (Tissue-Tek VIP 5 Jr.) and mounted in paraffin wax (Tissue-Tek III Paraffin WAX: 4509 and Tissue-Tek Cassette System). 4µmol/L thick slides were sectioned as required onto Thermo Scientific SuperFrost Plus Adhesion slides (10149970).

2.5 Immunohistochemistry

Sections of tissue for immunostaining were dewaxed in 100% xylene for 3 x 3 minutes and then rehydrated via decreasing concentrations of ethanol at 2 x 100%, 75%, 50%, 25% for 2 minutes in each. They were then placed in distilled water for 5 minutes before being subject to antigen retrieval (Table 1) and placed in running tap water for

5 minutes. Samples were mounted in coverplates (Thermo Scientific Shandon Plastic Coverplates) in Sequenza racks (VWR) and washed 3 x 5 minutes in phospho-buffered saline (PBS) and then incubated for 15 minutes with 120µL 3% hydrogen peroxide (H_2O_2) to block endogenous peroxidase activity. Samples were washed 3 x 5 minutes in PBS and then avidin and biotin binding sites were blocked (Avidin/Biotin Blocking Kit, Vector Laboratories: 004303) at room temperature for 15 minutes before protein blocking was carried out with a protein block for 1 hour for pan-species (abcam, ab64226). Primary antibodies were diluted (Antibody diluent, ThermoFisher Scientific: 003118) and incubated at 4°C overnight. Samples were washed 3 x 5 minutes in PBS and then biotinylated secondary antibodies diluted in antibody diluent as above, were added and incubated at room temperature for 1 hour. HRP conjugation to the secondary antibody was carried out using the VECTASTAIN ABC HRP system (Vector Laboratories: PK-4010). Positive staining was detected by DAB detection by incubation of slides for 3-10 minutes with DAB chromogen (DAB Quanto Chromogen and Substrate, ThermoFisher Scientific: 12623957). Samples were washed 2 x 5 minutes PBS then counterstained with Harris haematoxylin for 1.5 minutes followed by rinsing in water until colourless, incubation in lithium carbonate for 5 seconds and then washed in running tap water for 3 minutes. They were then dehydrated through increasing alcohols (25%, 50%, 75% 2 x 100%), 1 minute in each, and cleared for 3 x 45 seconds in xylene. Slides were mounted with DPX (CellPath: SEA-1304-00A).

Table 3 Primary antibodies and conditions used in immunohistochemistry.

Protein	Company, Product code	Dilution	Antigen Retrieval Conditions
αSMA	Sigma, A2547	1/200	Sodium citrate: 15 minutes
COLLAGEN I	SouthernBiotech, 1310-01	1/100	Sodium citrate: 5 minutes
F4/80	Abcam, ab6640	1/400	Proteinase K in 50mL TE buffer, room temperature, 10 minutes
FIBRONECTIN	Abcam, ab2413	1/200	Sodium citrate: 10 minutes
GFP	Abcam, 240	1/200	TE: pressure cooker 4 minutes
Ki67	Abcam, 16667	1/200	TE: pressure cooker 4 minutes
KRT19	DSHB, Troma-III	1/200	TE: pressure cooker 4 minutes
PAK1	Abcam, ab131522	1/200	TE: pressure cooker 4 minutes
panCK	DAKO, Z0622	1/200	Sodium citrate: 5 minutes, followed by 5 minutes cooling and 5 minutes of Proteinase K treatment at 37°C
PCNA	Abcam, ab29	1/2500	Sodium citrate: 10 minutes
P21	Abcam, 188224	1/100	TE: pressure cooker 4 minutes
tdTom (mCherry)	Sicgen, AB0081-200	1/200	Sodium citrate: 10 minutes

2.6 Immunofluorescence

Samples mounted on slides were rehydrated as for immunohistochemistry above. Following antigen retrieval (Table 1) and mounting to a slide rack, slides were incubated in protein block (Abcam: ab64226) for 1 hour at room temperature. Primary antibodies (Table 2) diluted in Antibody Diluent (ThermoFisher Scientific: 003118) and samples were incubated overnight at 4°C. Samples were washed with 3 x PBS and slides incubated with secondary antibodies (Table 3) diluted in Antibody Diluent for 1 hour at room temperature. Samples were washed with 3 x PBS and mounted with DAPI Fluoromount-G (0100-20: SouthernBiotech).

2.7 Immunohistochemical and immunofluorescence image analysis

Immunohistochemically stained slides were scanned with the Hamamatsu NanoZoomer 2.0 Series and NDP Scan software (Hamamatsu). QuPath (177) was used to measure tumour size (area of tissue covered as % of total liver) and tumour number, and all tumour quantification was blinded. QuPath (177) was used to calculate % positive cells.

2.8 DNA and RNA extraction from tissue

50-60mg of snap frozen tissue was used as input for DNA and RNA extraction. DNA extraction was carried out using the DNeasy Blood and Tissue Kit (QIAGEN: 69504) as per the manufacturer's instructions. RNA was extracted by homogenizing the tissue in 500µL TRIzol RNA Isolation Reagent (Invitrogen: 15596026) for 5-15 minutes with a

metal bead in the Qiagen TissueLyser LT (QIAGEN: 69980). Then 100µL of chloroform was added and the sample centrifuged at 12,000 RPM, at 4°C for 15 minutes, then the clear upper fraction was taken into a new Eppendorf and equal volume of isopropanol was added. Then the RNeasy Mini Kit (QIAGEN: 74104) was used as per the manufacturer's instructions from step 5 onwards. RNA was eluted in 10µl of RNase-free water. For downstream sequencing DNA quality and RNA RIN score was found with the Agilent 2100 Bioanalyzer with either DNA 1000, or RNA 6000 chip Nano/Pico assay (performed by Jeffrey Joseph: Technical Services, Institute of Genetics & Cancer, Edinburgh). A minimum RIN number of 8 was used for RNA sequencing.

2.9 RNA sequencing of mouse tumours

RNA was extracted from tumours and from one otherwise genetically and phenotypically wild-type Friend Virus B NIH Jackson (FBV) mouse liver and this sequencing was carried out at the Edinburgh Wellcome Trust Clinical Research Facility as per Younger and Wilson *et al*. (2) Mouse transcriptomes were sequenced by total RNA sequencing with ribonucleotide depletion, but without PolyA enrichment (chosen over polyA enrichment in order to capture non-coding RNA). Ribonucleotide RNA-depleted RNA was DNase treated and purified, then fragmented. Libraries were quantified by fluorometry and assessed for quality and fragment size (Agilent Bioanalyser). Sequencing was performed using the NextSeq 500/550 High-Output v2.5 (150 cycle) Kit on the NextSeq 550 platform (Illumina Inc). Libraries were combined in an equimolar pool based on Qubit and Bioanalyser assay results and run across a single High Output v2.5 Flow Cell. The primary RNA-seq processing, quality control to transcript-level quantitation, was carried out using nf-core/rnaseq v1.4.3dev. (178)

2.10 DNA exome-sequencing of mouse tumours

Exome sequencing was carried out at the Edinburgh Wellcome Trust Clinical Research Facility. Exome sequencing was carried out as per Younger and Wilson *et al*. (1) Libraries were prepared from genomic DNA (gDNA) and was sheared to achieve target DNA fragment sizes of between 150 and 200bp. DNA fragments were processed as adapter-ligated libraries and were purified. Seven hundred fifty nanograms of each prepared gDNA library was hybridized to probes covering the mouse exome and hybridized DNA-probes were amplified to apply unique indexing primers. Library QC: Libraries were quantified by Qubit and sequencing was performed using the NextSeq 500/550 High-Output v2.5 (150 cycle) Kit on the NextSeq 550 platform (Illumina Inc.). Libraries were combined in a single equimolar pool and run on a High-Output v2.5 Flow Cell.

2.11 Proteomic analyses

2.11.1 Reverse phase protein array (RPPA)

RPPA analysis was carried out on samples that were lysed in RPPA lysis buffer (2.5mL Triton X-100, 25mL 0.5 M HEPES pH 7.4, 0.5mL 0.5 EGTA pH 7.5-8.0, 37.5mL 1 M sodium chloride, 0.375mL 1M magnesium chloride, 0.1mL 100mM sodium orthovanadate, 1mL 100mM tetrasodium pyrophosphate, 1mL 1M sodium fluoride, 1 cOmplete mini EDTA-free protease inhibitor tablet (Roche, 11836170001), 1 phosSTOP phosphastase inhibitor table (Roche, 4906845001), 1mL glycerol and 1.9mL distilled water). Lysis was carried out for 10 minutes on ice before being disrupted by the TissueLyser LT for 2 minutes at 50 Hz. Samples were then centrifuged at 1,000 X g for

10 minutes and supernatant taken for RPPA analysis (performed by Kenneth Macleod: HTPU Microarray Services, Cancer Research UK Edinburgh Centre). To identify molecular components of relevant pathways, 120 antibodies were chosen for RPPA analysis, and these are listed in Appendix 2.

2.11.2 Western blotting

Samples were defrosted on ice and 50-100mg of tissue was lysed in RIPA buffer (Thermo Scientific: 89900) with protease and phosphatase inhibitor tablets (PhosSTOP) for 10 minutes on ice before being centrifuged at 1200 X g for 10 minutes and the supernatant taken. Protein was quantified using the Pierce BCA Protein Assay Kit (Thermo Scientific: 23225) as per the manufacturer's instructions. 25µg protein lysate was mixed with 1X NuPAGE Sample Reducing Agent (Invitrogen, NP0009) and 1X NuPAGE LDS Sample Buffer (Invitrogen, NP0007) and loaded onto NuPAGE 4-12% Bis-Tris Gel 10-well gels (Invitrogen, NP0335BOX). Samples were loaded in an equal total volume. The ladder used was Precision Plus Protein Dual Color Standards (BioRad, 1610374). The gels were run in Running Buffer (1X NuPAGE MOPS: Invitrogen: NP0001) at 200 V for 1.25 hours in XCell SureLock Mini-Cell electrophoresis tanks (Invitrogen: EI0001). SDS-PAGE gels were then transferred to PVDF membranes (Cytiva Life Sciences: 10600023), that had been activated in methanol for 1 minute prior. The transfer was carried out in 1X NuPAGE Transfer Buffer (Invitrogen: NP00061) with 0.1% NuPAGE Antioxidant (Invitrogen: NP0005) and 10% methanol. Transfer was carried out for 1.5 hours at 30V using XCell II Blot Module S (Invitrogen: EI9051) placed in 4°C. Once transferred, membranes were blocked in 5% milk (Marvel) in PBS + 0.1% Tween20 (0.1% PBST) at room temperature for 1 hour before they were briefly rinsed

once in 0.1% PBST. Primary antibodies were diluted in 3% BSA in 0.1% PBST and added to the membranes and incubated overnight at 4°C. The following day, membranes were washed 3 x 30 minutes in 0.1% PBST and then incubated with secondary antibodies diluted at 1/20000. Membranes were visualised on ODYSSEY (Li-Cor).

2.12 Quantification and statistical testing

Statistical testing was performed using Graphpad Prism. All experimental groups were analysed for normality with a D'Agostino–Pearson Omnibus test. A Student's t test was used for normally distributed groups that were in groups of 2, or using one-way ANOVA to compare >2 groups. I considered a p value of <0.05 as statistically significant. Data are presented as a mean with SEM for parametric data and mean with SD for nonparametric data. Quantification of protein expression using Western blots was done with Image Studio Lite v5.2 (Li-Cor Biotechnology).

Chapter 3 Results

CRISPR-*sp*Cas9 screening identifies rare driver genes that cooperate with oncogenic $KRAS^{G12D}$

3.1 Introduction

iCCA is considered a rare cancer and this rarity results in the severely limited availability of human iCCA tissue for further analysis. This means that sequencing data on a large enough scale to achieve comprehensive analysis is likewise limited and computationally predicting the mutations that co-occur or are mutually exclusive is difficult. Additionally, previous studies from iCCA have suggested that for the majority of patients, their tumours lack many consistent or targetable mutations between patients. Despite this high level of genetic complexity, a limited number of recurrent mutations between patients have been identified, such as loss of function of mutations in genes including *ARID1A*, *BAP1* and *TP53*, gain of function *KRAS* mutations and neomorphic changes in *IDH1/2*. To overcome the limitations of small sample size in previous studies, work from our lab defined a list of candidate oncogenes by pooling published human exome sequencing data and using IntOGen (179), a tool used to predict driver genes identify novel tumour suppressors that are mutated in iCCA. This study aimed to collate the data contained in five exome sequencing studies, and includes data from a total of 277 patients. While many analyses of iCCA mutations are based on the frequency of a mutation within a patient population, IntOgen is based on more parameters than mutational frequency, which is important in iCCA where high levels of genetic heterogeneity are present and infrequent mutations are likely to modify the effects of mutations in more common oncogenes or tumour suppressors. Our analysis of patient exome sequencing data considered whether a region in the DNA accumulates mutations at a higher frequency than the background mutation rate, functional impact bias, where sites that contain accumulative loss of function mutations identify tumour suppressor genes and identification of where there is

spatial clustering of mutations, which is characteristic of oncogenes. These parameters, in addition to considering the frequency of mutations that were present, provided a list of mutations in 96 driver genes occurring across samples from iCCA patients. This *in silico* screen of patient data was sufficient to identify mutations in multiple genes of interest, but despite its ability to identify mutations that have previously been associated with iCCA (e.g. *ARID1A, BAP1*), those already known to be associated with other cancers and that are present in the COSMIC database, or indeed entirely novel to association with iCCA, this computational approach remained insufficient to determine the interactions between mutations and the mechanisms involved in iCCA growth. For this reason, I developed a CRISPR-spCas9 screening approach whereby the genes that contained mutations in the human patient data and which were defined computationally were utilised to create a library of gRNAs to target mouse genes which was then used to identify functional mutations that can interact to initiate and drive iCCA growth *in vivo* using a murine model.

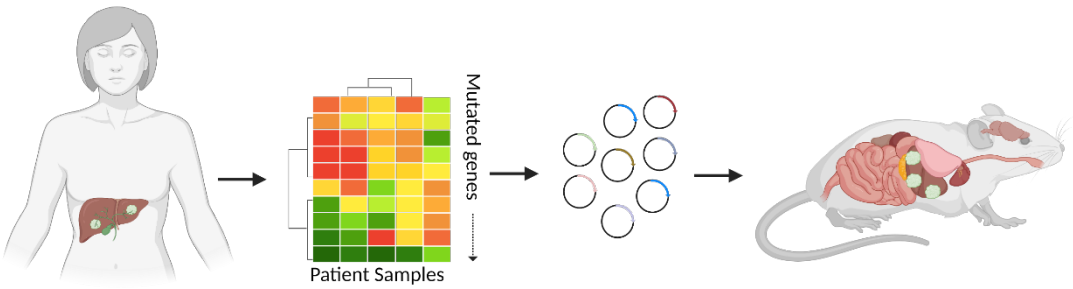

Mutated genes
Patient Samples

Figure 1

Schematic of the process of creating the CRISPR-spCas9 library from patient iCCA data. Exome sequencing data from patients with iCCA tumours were collated and their mutational profiles were examined to reveal novel iCCA driver genes. Mouse homologues to these driver genes were then used as the input of the CRISPR-spCas9 plasmid library screen that allowed formation of iCCA in mice.

3.2 Hypothesis

iCCA presents with tumours containing many driver genes that occur at a low frequency, and these may interact with each other and/or more common mutations to allow progression of tumourigenesis. Using *in vivo* screening of the adult murine liver, I explore the question to establish which combinations of mutations result in iCCA.

3.3 Aims

- Generate a proportionate CRISPR library of plasmids that allow editing of endogenous loci of genes of interest *in vivo* suitable for screening.
- Establish whether any driver genes are capable of forming iCCA originating from hepatocytes.
- Identify whether any of the driver genes found in our patient exome screen synergise with oncogenic $KRAS^{G12D}$ to form iCCA *in vivo*.
- Characterise the tumours formed in the screen as to what mutations they harbour.

3.4 Production of a custom CRISPR-Cas9 library

To introduce mutations into the mouse homologues of the 91 patient iCCA-driver genes found in our *in silico* screen of patient exome sequencing, a library of single guide RNAs (sgRNA) targeting these genes was cloned into the sleeping beauty-CRISPR plasmid (SB-CRISPR). This plasmid was generated and utilised in the liver by Weber *et al* (173). It contains expression cassettes for both sgRNA and spCas9, flanked by sleeping beauty inverted repeats. This means that when the SB-CRISPR plasmid is co-expressed with a SB-transposase vector (SB13) it results in transient expression of CRISPR/*sp*Cas9 and longer term expression of the sgRNA vectors that will be integrated into the genome by the so-called “copy and paste” mechanism employed by the sleeping beauty transposase.

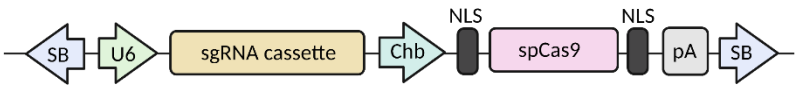

NLS
NLS
SB
U6
sgRNA cassette
Chb
spCas9
pA
SB

Figure 2

Illustration of the component parts of the SB-CRISPR bicistronic expression plasmid used for delivery of components causing multiplexed somatic mutagenesis in murine livers. sgRNA and spCas9 expression cassettes are flanked by sleeping beauty (SB) inverted terminal repeats, allowing the stable integration of these components into the target genome when co-expressed with sleeping beauty transposase (SB13). sgRNA expression is under the control of the U6 promoter, Streptococcus pyogenes Cas9 (spCas9) is driven by a chicken β-actin hybrid intron (Chb) promoter. NLS: nuclear localisation signal, pA: polyadenylation signal.

To identify which human genes there are mouse homologues of, genes were searched for in the Mouse Genome Database and yielded 91 unique homologues. The remaining five human genes did not have a unique mouse homologue and therefore were not included in the *in vivo* screen. These five genes included C2orf76, encoding for a poorly described protein with little known functionality; KRTAP12-1 which encodes a keratin-associated protein that is involved in establishment of hair colour but is not expressed in the liver (180); RBMXL1, a retrogene of the RNA-binding motif protein RBMX, overexpression of which are both involved in acute myeloid leukaemia development (181). SPANXD was also identified, and this gene is a member of the SPANX gene family that are testis-specific genes required for formation of mature spermatozoa and who's loss is associated with testicular embryonal carcinoma (182,183); and finally ZNF181, a member of the KRAB-ZNF protein family that functions to suppress transposable element transcription (184).

For each of the 91 genes to be included in the screening library, I then adapted 3 sgRNAs selected from the mouse GeCKOv2 (Genome-Scale CRISPR Knock-Out) library depository from the Zhang laboratory (185). The GeCKO library contains sgRNA sequences designed specifically to reliably cause gene knock-out by targeting constitutively active exons, and to reduce off target effects. Additionally, a control of 200 non-targeting sgRNAs were also used to ensure that the resulting phenotypes were due to specific *sp*Cas9-induced mutations and not presence of the library and not the action of CRISPR-*sp*Cas9 alone. Therefore a total of 473 sgRNA sequences were prepared as a library.

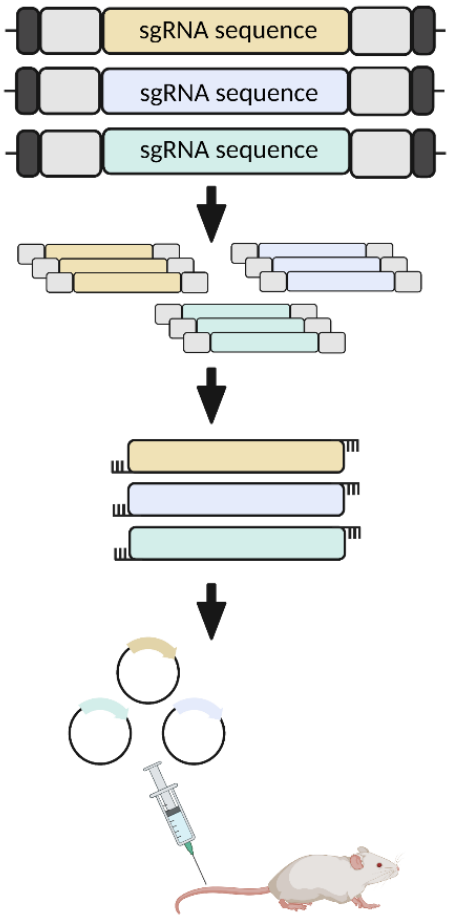
I then cloned the library of sgRNAs into the SB-CRISPR plasmid using an in house optimised protocol, a schematic of which can be seen in Figure 3.

In addition to the library of sgRNAs that allow CRISPR-spCas9-mediated editing of genes which had mutations that occurred at a low frequency in patients, I also wished to examine the cooperative nature of these low frequency mutations with the more common mutation in iCCA, $KRAS^{G12D}$. This mutation results in constitutive activation of KRAS (and thus sustained activation of its downstream pathways) by disallowing GTP hydrolysis and preventing the inactivation of the protein (186). My chosen expression vector for human oncogenic $KRAS^{G12D}$ is tagged with GFP, with an internal ribosome binding site between them, under the control of a pCAGGS promoter. This plasmid was chosen as it allows direct observation of $KRAS^{G12D}$ positive cells by their expression of GFP. Additionally, in the patient exome sequencing data, *KRAS* was the most commonly mutated oncogene. This is intriguing as Ras positive cells are cleared from the liver by way of senescence surveillance, so given the insufficiency of KRAS-expressing cells to form cancer alone, there is likely to be some interaction occurring in these patients between *KRAS* overexpression and the many infrequent loss of function mutations seen in the patient screen. It has been shown that *Trp53* loss can allow the circumvention of Ras-induced senescence, so I hypothesise it is possible other loss of function mutations are sufficient to achieve this.

To deliver these plasmids (SB-CRISPR containing the sgRNA library, SB13 transposon to allow integration and $KRAS^{G12D}$ expression vector) to murine livers, I employed a method called hydrodynamic tail vein injection (HDTVI). This model involves the injection of naked DNA into the lateral tail vein of mice with a volume equal to 20% of the mouse's body weight, and results in preferential uptake in the liver by hepatocytes

when the dose is delivered at a high pressure in <7 seconds. This causes mosaic somatic editing of the hepatocytes, and Weber *et al* (173) first showed that this technique causes somatic multiplex mutagenesis on the background of oncogenic $KRAS^{G12D}$ expression, and that iCCA arises from the cells that uptake the naked DNA.

Figure 3 illustrates the process of the library production.



Libraries are recieved as an oligonucleotide pool. Single guide RNA (sgRNA) sequences are flanked by primer binding sites (black) and restriction enzyme sites (grey).

Oligonucleotides are amplified by PCR.

Restriction digrest produces cloning overhands.

sgRNA oligonucleotides are cloned into the SB-CRISPR plasmid. Amplicon sequencing confirms fair representation in the library.

The sgRNA library is injected into mice with a transposase (SB13) and oncogenic $Kras^{G12D}$.

Figure 3

*Production of a custom CRISPR-*sp*Cas9 sgRNA screening library. Mouse homologues to 91 human genes identified in a patient exome-sequencing screen of iCCA were used to design sgRNAs in triplicate that were obtained as an oligonucleotide pool. All the library sgRNAs contained primer binding sites (black) that were utilised to amplify the sgRNAs before they were restriction digested via their restriction sites (grey). This allowed the creation of cloning overhangs for them to be cloned into the SB-CRISPR plasmid. The plasmid DNA was then cultured in bacteria and the DNA purified. Equal representation of the sgRNAs was confirmed by MiSeq Illumina amplicon sequencing. Finally, the inverted terminal repeats present either side of the gRNAs and spCas9 allowed SB transposase-mediated endogenous editing when these naked DNA plasmids are injected into mice by HDTVI.*

To ensure that the chosen 91 genes are subject to as similar exposure to sgRNAs as possible, MiSeq Illumina amplicon sequencing was performed. These data can be accessed via the NCBI Gene Expression Omnibus (GEO), accession number GSE190770. (1) Following sequencing, analysis using the GINI index was performed, and this determined that the sgRNAs were all represented in the library pool. This resulted in validation that the sgRNA pool was fully representative of all the sgRNAs cloned, and the library had a GINI coefficient of 0.07, meaning that no sgRNA was unrepresented.

3.5 Oncogenic RAS cooperates with novel loss of function mutations to form iCCA

In order to establish whether any of the loss of function mutations interact to drive iCCA tumourigenesis, otherwise genetically wild-type FVB mice were injected with the library of sgRNAs ($iCCA^{Lib}$) or $KRAS^{G12D}$ alone, or with both $KRAS^{G12D}$ and $iCCA^{Lib}$. Mice injected with the $iCCA^{Lib}$ alone (n=14) did not show any tumour formation after 10 weeks, as illustrated in Figure 4A. This indicates that the $iCCA^{Lib}$ loss of function mutations alone were insufficient to cause cancer, Figure 4B shows sample H&E staining of bile ducts from these animals, and they are tumour free. Likewise, animals injected with the $KRAS^{G12D}$ plasmid alone (n=10), 8-weeks post injection, yielded no tumours whatsoever, whereas of those injected with $iCCA^{Lib}$, 8 of 14 animals presented with tumours (Figure 4C). The lack of tumours seen in the $KRAS^{G12D}$ plasmid alone-injected animals is likely due to immune clearance due to Ras oncogene-induced senescence. I stained for GFP to confirm whether there was presence of KRAS in these livers, as shown in Figure 4E, livers injected with $KRAS^{G12D}$ alone were negative for GFP expression, supporting my conclusion that KRAS-expressing cells in these livers have been cleared. Interestingly, there were numerous macroscopic, solid lesions present in

mice injected with iCCA^{Lib} and oncogenic KRAS^{G12D} (Figure 4F), suggesting that at least one loss of function mutation induced by the sgRNA library can co-operate with KRAS^{G12D} and lead to cancer formation. These KRAS-positive tumour cells are shown in Figure 4G.

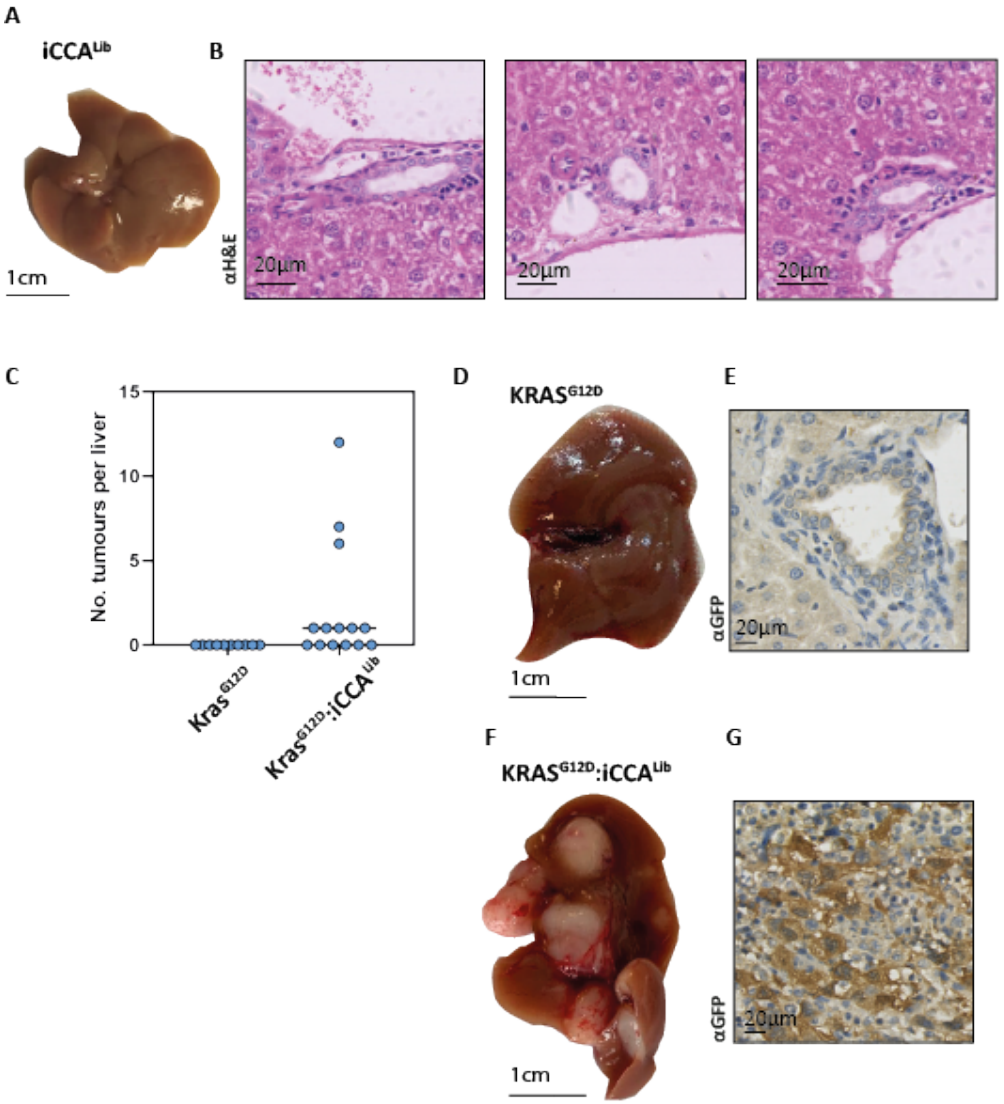

A
iCCA^Ub
1cm
B
αH&E
20μm
20μm
20μm
C
No. tumours per liver
15
10
5
0
Kras G12D
Kras G12D:iCCA Ub
D
KRAS G12D
1cm
E
αGFP
20μm
F
KRAS G12D:iCCA Ub
1cm
G
αGFP
20μm

Figure 4

*CRISPR-*sp*Cas9 screening in murine livers produces Ras positive tumours when co-injected to animals with oncogenic KRAS^{G12D}. A. iCCA^{Lib}-alone injected animals (n=14) formed no tumours in their livers, and when these were stained by H&E, in B., it is clear that the bile ducts are non-tumourigenic. C Quantification of the number of macroscopic tumours seen per mouse liver in the control (n=10) and library (n=14) groups where each dot represents one animal. D. Macroscopic image of a liver injected with KRAS^{G12D} alone. E. GFP stain of ducts from livers of animals injected with KRAS^{G12D} alone. Ducts are GFP negative, as seen by the similarity in colour between the parenchyma and the cuboidal epithelium. F. Macroscopic image of liver injected with KRAS^{G12D} and my custom CRISPR library of sgRNA to induce loss of function mutations (iCCA^{Lib}). G. GFP stain showing RAS-positive tumour cells.*

Two of the $KRAS^{G12D}$:$iCCA^{Lib}$ mice also had metastases to the pancreas, and one to the additional sites of lymph node and diaphragm. In order to exome sequence these tumours, there was insufficient tissue to then also carry out histological analysis. The exome sequencing was informative, and Figure 5 shows macroscopic images of the two pancreas and a list of the mutations they harbour. Interestingly, the animal shown in 5A only contains 2 additional mutations to loss of *Trp53*, whereas that shown in 5B contains an additional 7. Mutations of *Cacna1c* is only seen in this pancreatic metastasis sample, and not in any of the parent liver tumours. However, *Plk2* is mutated in half of the parent liver tumours. *Bcor* is mutated in 20% of the primary tumours, *Fzd2* in 75%, *Cc2d2a* in 40%, *Ccdc66* and *Pik3ca* are only found in this pancreatic lesion, and *Plxb2* and *Ncor1* are found additionally in primary liver tumours from a different animal.

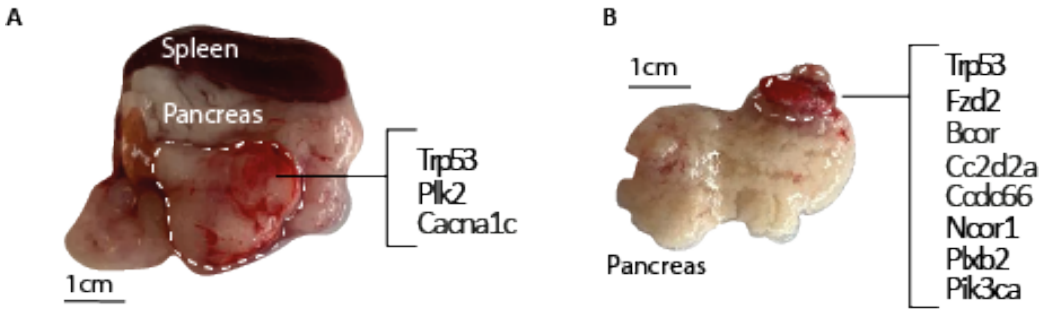

A
Spleen
Pancreas
Trp53
Plk2
Cacna1c
1cm
B
1cm
Pancreas
Trp53
Fzd2
Bcor
Cc2d2a
Ccdc66
Ncor1
Plxdc2
Pik3ca

Figure 5

*Two CRISPR-*sp*Cas9 screen animals yielded metastatic masses in the pancreas.*

Macroscopic images of the pancreas of two $KRAS^{G12D}$*:*$iCCA^{Lib}$ *mice. Exome sequencing revealed mutations present; these are listed alongside.*

3.6 Exome sequencing of tumours revealed which CRISPR-*sp*Cas9 induced mutations were selected for

A total of 14 tumours were excised from five livers and whole exome sequenced, and among these were 2 pancreatic lesions. The pancreatic lesions were chosen with the interest of whether they contain the same mutations as found in the parent liver tumours, or if they harbour different mutations. It was found that of the 91 genes targeted for CRISPR-*sp*Cas9 mediated disruption, certain CRISPR-*sp*Cas9-induced mutations were selected for and we assume at least a subset of these are able to regulate the escape from oncogenic KRAS-induced senescence, as illustrated in Figure 6. On average, 7.5 genes were mutated per tumour, the most frequently mutated gene being *Trp53*, which was mutated in all mice with macroscopic tumours, and indeed only one tumour sample that was exome sequenced did not harbour a *Trp53* mutation. We identified mutations in 30 of the 91 genes that were selected for in these tumours. The full file containing the whole exome sequencing data can be accessed via NCBI Gene Expression Omnibus (GEO) as accession number GSE190770. (1)

Exome sequencing also allowed for the determination that these insertions or deletions (indels) were indeed CRISPR-*sp*Cas9 induced, and not spontaneous. I confirmed that all indels were within 50bp of a gRNA site started or ended at least 3bp upstream of the PAM sequence. The sequences were all examined by hand to confirm that the indels were indeed CRISPR-*sp*Cas9 induced. The introduction of indels was further confirmed by sequencing genomic DNA extracted from the screen tumours compared to a control of wild type FVB livers (not just the FVB reference genome) to verify the mutations.

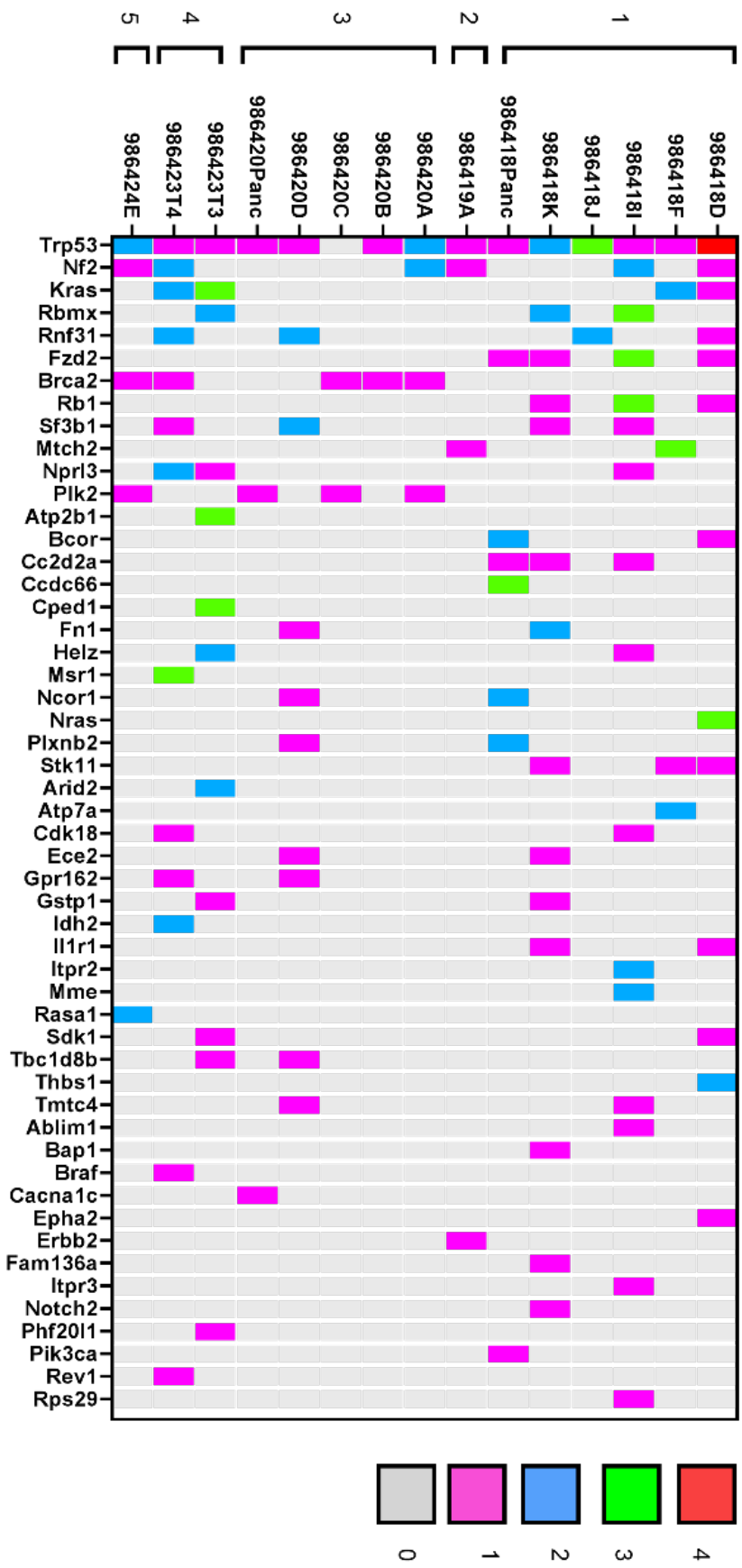

5
4
3
2
1
986424E
986423T4
986423T3
986420Panc
986420D
986420C
986420B
986420A
986419A
986418Panc
986418K
986418J
986418I
986418F
986418D
Trp53
Nf2
Kras
Rbmx
Rnf31
Fzd2
Brca2
Rb1
Sf3b1
Mtch2
Nprl3
Plk2
Atp2b1
Bcor
Cc2d2a
Ccdc66
Cped1
Fn1
Helz
Msr1
Ncor1
Nras
Plxnb2
Stk11
Arid2
Atp7a
Cdk18
Ece2
Gpr162
Gstp1
Idh2
Il1r1
Itpr2
Mme
Rasa1
Sdk1
Tbc1d8b
Thbs1
Tmtc4
Ablim1
Bap1
Braf
Cacna1c
Epha2
Erbb2
Fam136a
Itpr3
Notch2
Phf20l1
Pik3ca
Rev1
Rps29
0
1
2
3
4

Figure 6

*Heat map representation of all mutations present from exome sequencing of CRISPR-*sp*Cas9 screening library tumours. Tumour samples are on the Y axis, and are numbered according to the animal (1-5) they originate from. X axis denotes the gene a mutation is seen in. Colours indicate how many indels are present in the genes; grey: 0, pink: 1, blue: 2, green: 3, red: 4.*

3.7 CRISPR-*sp*Cas9 library screen tumours resemble human iCCA

iCCAs developed in my *in vivo* screen were classified histologically and found to be aggressive adenocarcinoma with poorly differentiated cholangiocellular properties that are representative of human cholangiocarcinoma. They were also examined and confirmed to be iCCA by Dr Tim Kendall, a clinical pathologist at the University of Edinburgh. Dr Kendall also confirmed that the tumours contain identifiable markers of human iCCA. These markers included well differentiated glandular structures surrounded by immune infiltrate or stroma, and regions with less well differentiated tumour that are attributed to sarcomatoid iCCA, as shown in Figure 9A below. Additionally, there are areas of dense immune cells adjacent to and surrounding cancer cells, as seen in 9B, where hepatocytes have been replaced by cancer cells. iCCAs are characterised by significant areas of desmoplastic stroma (127). In tumour tissue driven by expression of $KRAS^{G12D}$ and $iCCA^{Lib,}$ a dense desmoplastic stroma develops adjacent to the cancer cells, indicating that this is not simply epithelial dysplasia, but the formation of a *bona fide* tumour.

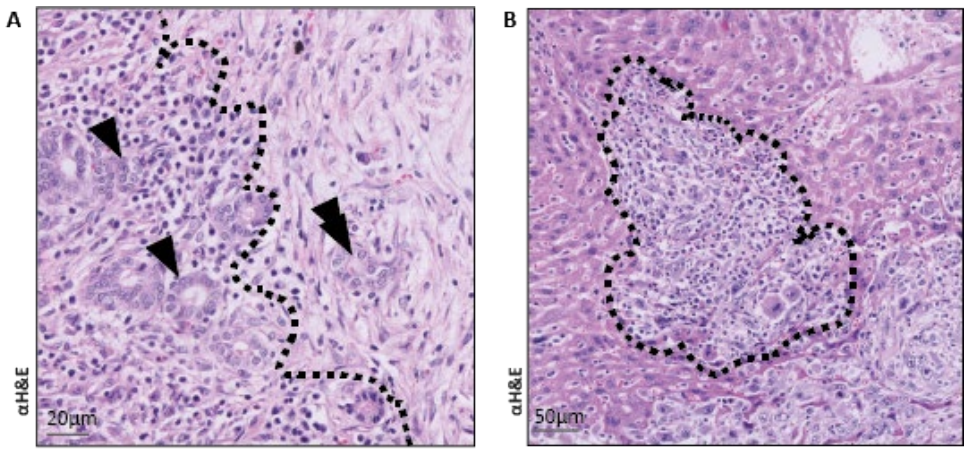

A
αH&E
20μm
B
αH&E
50μm

Figure 9

*Haematoxylin and eosin staining of tumourigenic regions of a murine liver sample from the CRISPR-*sp*Cas9 library screen. A. Single headed arrows illustrate moderately differentiated iCCA that is characterised by irregular columnar shaped cells embedded in immune infiltrate. Cells are moderately pleomorphic and infiltrative. Double arrow illustrates ductular tumour cells surrounded by tumour stroma. The dotted lines outline the regions of immune cells (left) adjacent to stroma (right). B Illustrating an outlined discrete area of immune cells, containing tumourigenic cells in the midst, and less well differentiated tumour tissue in the lower right of the panel.*

As the HDTVI targets hepatocytes and causes subsequent transdifferentiation to malignant cholangiocytes (91) it is important to confirm that the cells expressing GFP are committed to the cholangiocyte lineage. This is observed and confirmed by their ductular morphology and their cuboidal and epithelial cell shape. Therefore the HDTVI delivery has succeeded in allowing oncogenic $KRAS^{G12D}$ to cooperate with the loss of function library mutations to cause iCCA formation from cells of a different lineage – hepatocytic lineage – *in vivo*.

Immunohistochemical staining with an antibody against pan-cytokeratin (panCK) further confirmed this, as can be observed in Figure 10A&B, where there is normal, discrete ductular expression of panCK in murine liver tissue from animals injected with $KRAS^{G12D}$ alone, $KRAS^{G12D}$:$iCCA^{Lib}$ in Figure 10C&D present with irregular structures. There are some cells positive for panCK that are showing non-cholangiocyte-like morphologies in $KRAS^{G12D}$:$iCCA^{Lib}$ tumours, and this further supports that these tumours are discrete from the surrounding liver parenchyma (Figure 10C, surrounding the central ductular structure). panCK positive cells with spindle morphology is characteristic of this rare sarcomatoid type of iCCA. (187,188). Finally, the presence of less well differentiated tumour cells in 10D show the latter progression of these tumours.

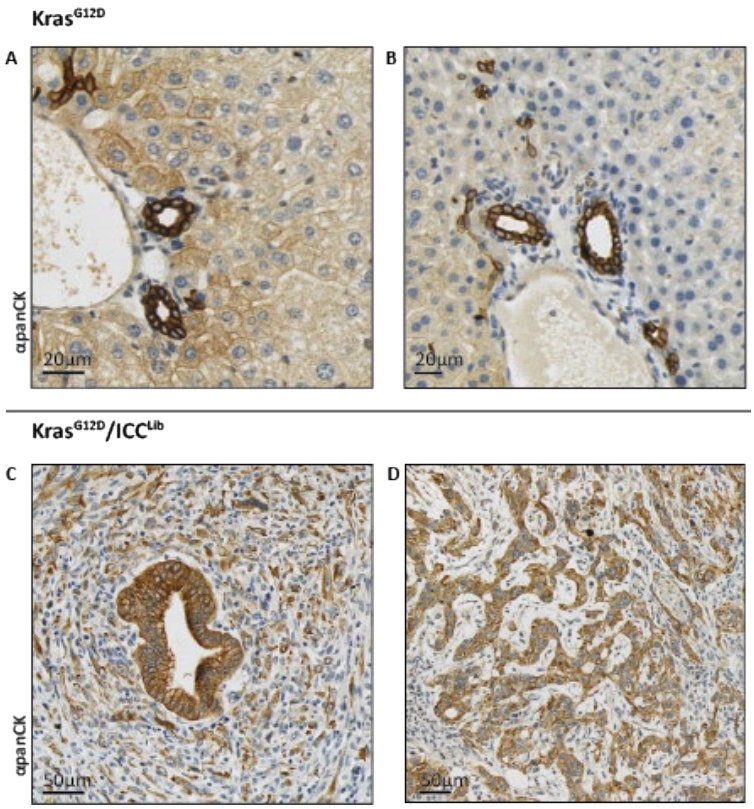

Kras^G12D
A
B
αpanCK
20μm
20μm
Kras^G12D/ICC^Lib
C
D
αpanCK
50μm
50μm

Figure 10

*Tumours from the CRISPR-*sp*Cas9 library screen have ductular morphology and resemble human iCCA. Immunohistochemical stain for pan-cytokeratin in tissue injected with A. oncogenic KRAS^{G12D} and B. KRAS^{G12D}/iCCA^{Lib}.*

To confirm that the tumours have a cholanigocellular phenotype, I stained also for keratin 19, which is used to show biliary identity and to differentiate between HCC and iCCA. Figure 11 illustrates the difference in wild-type FVB murine liver stained with CK19 (Figure 11A) to $KRAS^{G12D}$ (11B) and $KRAS^{G12D}$/$iCCA^{Lib}$ (11C, D and E). The expression of CK19 in $KRAS^{G12D}$ highly resembles WT expression: there is clear and discrete staining of the biliary compartment. In comparison to this, the tumour cells seen in 11C are highlighted by the CK19 staining as iCCA. Areas of iCCA are surrounded by Keratin 19 positive cells in 11D showing that even areas that have lost their biliary morphology in these tumours are still expressing markers of biliary lineage. Interestingly, we also observe in 11E some spindle-like cells that are positive for keratin 19 indicative of their cholanigocellular origin, but not at all representing the usual cuboidal epithelial morphology seen in wild type tissue, or the usual iCCA tumour morphology.

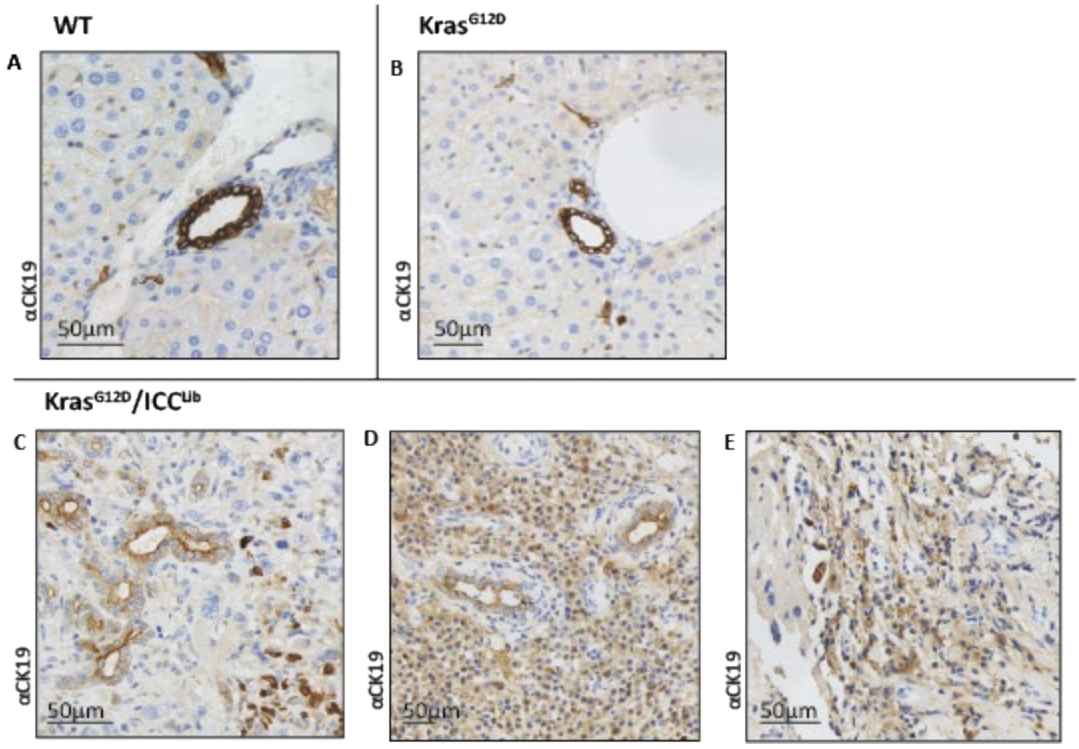

WT
A
αCK19
50μm
Kras^G12D
B
αCK19
50μm
Kras^G12D/ICC^Lib
C
αCK19
50μm
D
αCK19
50μm
E
αCK19
50μm

Figure 11

Cytokeratin 19 staining illustrates tumours originate from the biliary compartment. A. Otherwise wild-type (WT) FVB mouse liver tissue from mice euthanised at 8 weeks expresses CK19 in the biliary compartment in a restricted manner. B. Animals injected by HDTVI with $KRAS^{G12D}$ express CK19 in a pattern representative of the WT animals. C, D and E: $KRAS^{G12D}$/$iCCA^{Lib}$ HDTVI mice express CK19 in ducts and also in tumourous tissue.

Another key aspect of human iCCA is the nature of the tumour stroma consisting of many α-smooth muscle actin positive myofibroblasts: cancer associated fibroblasts, CAFs. (127,189) I investigated this in more detail by staining tumours from the library screen immunohistologically for the cancer associated fibroblast marker α-smooth muscle actin (α-SMA). Figure 12A shows the lack of cells expressing this marker in the tissue from HDTVI of $KRAS^{G12D}$ alone, unsurprising as these livers contained no tumours. When compared to the great number of CAFs present in 12B, a sample image from the liver of an animal injected with $KRAS^{G12D}$:$iCCA^{Lib}$ that is tumourous, the difference is stark. Interestingly, in patients, those with resected tumours presenting with high levels of α-SMA had significantly lower survival and larger tumours than those with lower levels of α-SMA. (190) The presence of such an array of α-SMA positive CAFs is therefore indicative of the severity of the tumours observed in my $KRAS^{G12D}$:$iCCA^{Lib}$ animals.

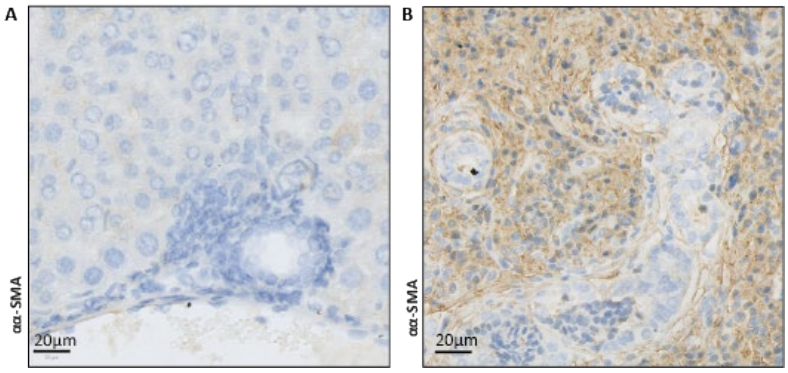

A
αα-SMA
20μm
B
αα-SMA
20μm

Figure 12

Cancer associated fibroblasts, as shown by their positivity for α-smooth muscle actin (α-SMA). CAFs are not present in HDTVI with A. oncogenic KRAS^{G12D} alone, where no tumours were present, however they surround the cancerous ducts in B, injected with oncogenic KRAS^{G12D} and iCCA^{Lib}.

3.8 An extensive tumour microenvironment supports $KRAS^{G12D}$:$iCCA^{Lib}$ tumours

Given the strong presence of stroma in this model, I proceeded to examine the stromal composition in more detail by staining for pan-fibronectin and collagen 1. Pan-fibronectin is also secreted by CAFs, while the alignment of fibronectin by CAFs is associated with allowing directional migration of cancer cells (191). It has been established as a marker for epithelial to mesenchymal transition (EMT). EMT involves epithelial cells losing their apical-basal polarity and cell to cell contacts and gaining mesenchymal characteristics including the ability to migrate and more cell to extracellular matrix connections. (192) Collagen 1 is an important component of the iCCA stroma and is a known biomarker for fibrosis in cholangiocarcinoma, and it is also excreted by CAFs. Figure 13 shows positivity surrounding ductular structures, highlighted in dashed lines, for pan-fibronectin in 13A and collagen 1 in 13C. The images here are sample images of the stromal components observed in all $KRAS^{G12D}$:$iCCA^{Lib}$ tumours. Figure 13B shows pan-Fib staining of a liver from an otherwise wild-type, non-cancerous FVB mouse culled at 10 weeks (WT). Fibronectin is expressed weakly in the cytoplasm of the biliary components of the tumour, and in the surrounding stroma, and also strongly in the ECM of the cancerous ductal cells. As fibronectin is usually expressed in the ECM, the presence of fibronectin in the cytoplasm is surprising, and illustrates how dysregulated the structural components of these tumours are. The extensive positivity seen for collagen 1 surrounding the particularly ductal components of the tumours is further indicative of the dense stroma seen in this model, especially when compared to the WT liver in 13D.

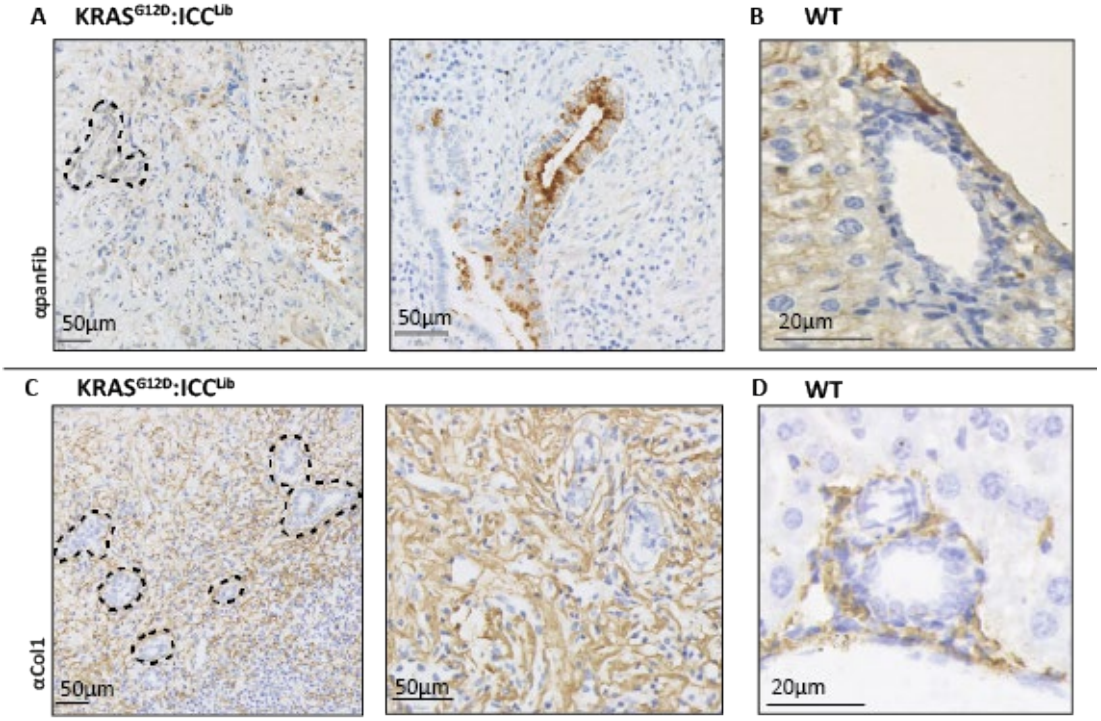

A KRAS^{G12D}:ICC^{Lib}
B WT
αpanFib
50μm
50μm
20μm
C KRAS^{G12D}:ICC^{Lib}
D WT
αCol1
50μm
50μm
20μm

Figure 13

Immunohistochemical staining of acellular components pan-fibronectin and Collagen 1. A. $KRAS^{G12D}$:$iCCA^{Lib}$ animals and B. an otherwise wild-type FBV mouse liver stained with pan-fibronectin (panFib). C. $KRAS^{G12D}$:$iCCA^{Lib}$ animals and D. an otherwise wild-type FBV mouse liver stained with Collagen 1.

The tumour microenvironment consists of non-malignant cells, for example the recruited immune cells, as well as the acellular compartment which is the extracellular matrix. Further to establishing the presence of many CAFs, I characterised other components that support the tumours excised from $KRAS^{G12D}$:$iCCA^{Lib}$ mice, as it has been suggested that an abundance of immune cells contributes to the desmoplastic nature of iCCA. (189) I stained for cell types involved in the immune response- macrophages expressing F4/80 and foxP3 positive regulatory T-cells, which can be observed in Figure 14. The infiltration of macrophages and regulatory T-cells observed is indicative of the human-like iCCA desmoplastic microenvironment in these tumours.

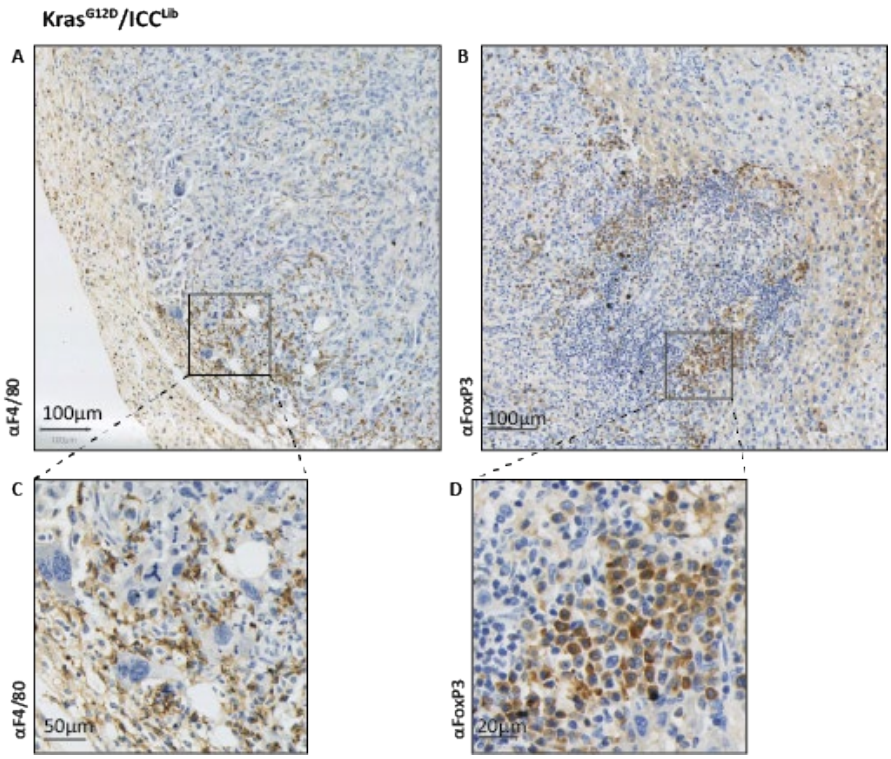

Kras^G12D/ICC^Lib
A
B
αF4/80
αFoxP3
100µm
100µm
C
D
αF4/80
αFoxP3
50µm
20µm

Figure 14

Macrophages and regulatory T-cells form part of the tumour microenvironment of animals injected with $KRAS^{G12D}$ and $iCCA^{Lib}$, as shown by the positive expression of F4/80 and FoxP3, respectively. A. F4/80 staining showing presence of macrophages in the context of a larger tumour area, B. showing presence of regulatory T-cells in the larger tumour area, with C and D showing a more detailed cell morphology.

3.9 $KRAS^{G12D}$ interacts with loss of *Trp53* to form a reliable model of iCCA

Following the characterisation of the screen tumours histologically and by using exome sequencing to examine the mutations present, I noted the presence of *Trp53* mutations in all but 1 of the tumours exome sequenced. Again this highlighted the significance of this gene as an additive mutation on the background of Ras by allowing Ras mutant cells to circumvent senescence surveillance, a known function of *Trp53*-loss. Additionally, in the screen of patient data, these two mutations are the most common and are highly co-occurring.

To investigate further the co-occurrence of gene interactions between mutations in iCCA, we looked back at the patient dataset and found that of the canonical driver mutations (mutations that have been previously associated with iCCA), except for *KRAS* and *TRP53*, the level of mutual exclusivity between canonical driver genes is high and consistent. This means that the mutations we see recurring across studies do not depend on each other for iCCA tumourigenesis. In addition to this, when this analysis was expanded to include COSMIC and novel mutations, we observed discrete patterns of co-occurrence and mutual exclusivity. This is indicative that canonical driver genes such as *KRAS* and the large group of infrequently mutated genes are interacting in a novel manner.

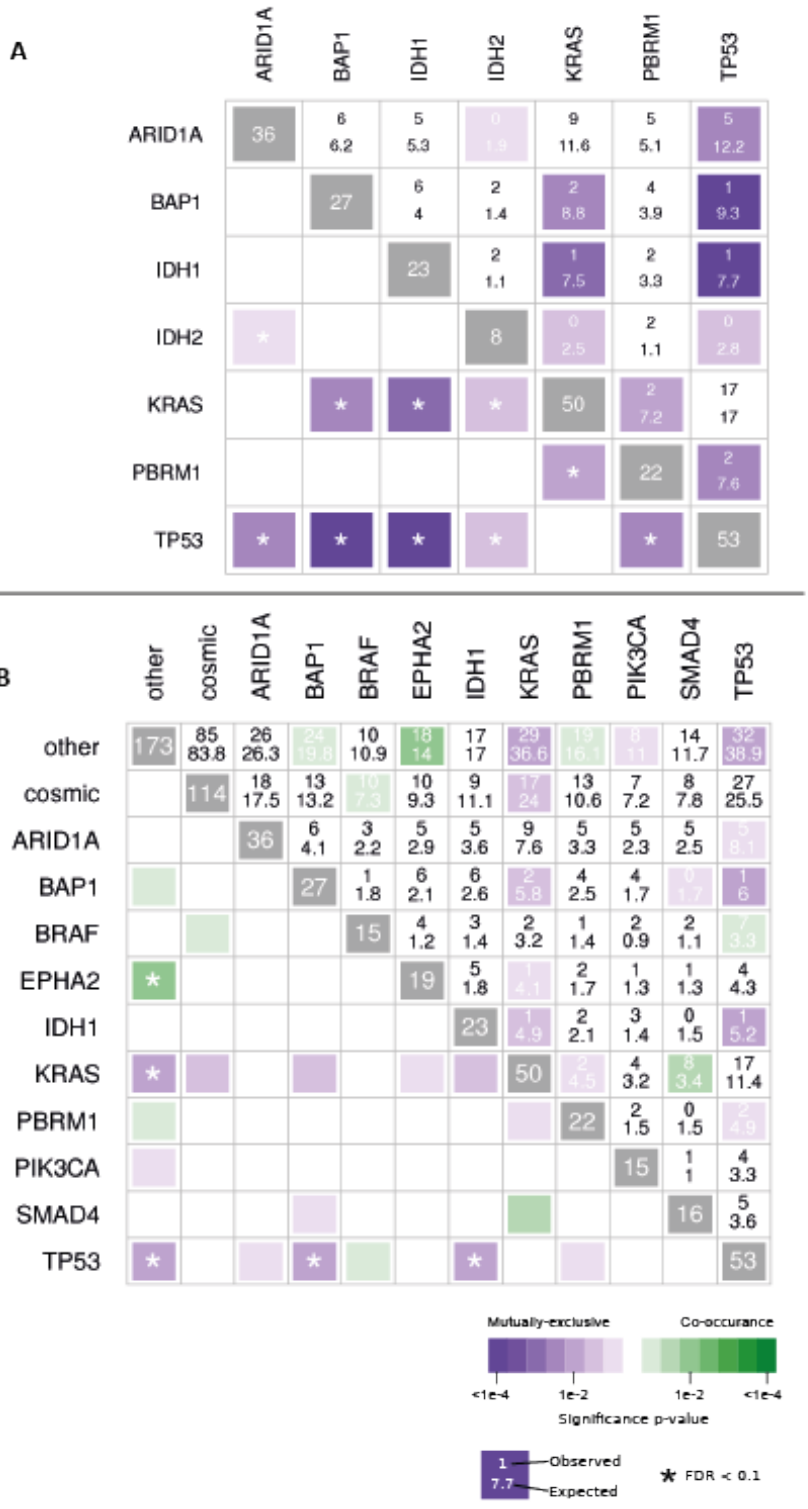

A
ARID1A
BAP1
IDH1
IDH2
KRAS
PBRM1
TP53
B
other
cosmic
Mutually-exclusive
Co-occurance
<1e-4
1e-2
1e-2
<1e-4
Significance p-value
Observed
Expected
FDR < 0.1

Figure 15

Co-occurrence and mutual exclusivity analysis of iCCA patient exome sequencing data. A shows there is a high level of mutational exclusivity in iCCA between canonical driver mutations (mutations that have been associated with iCCA in previous studies, and are well-known driver genes of iCCA). B indicates the co-occurrence and mutual exclusivity of COSMIC drivers that co-occur with ≥15 times in the human sample set and all novel drivers are combined in other.

Given the observation of mutations in *KRAS* and *TP53* co-occurrence in the patient data, I wished to address which other mutations may act in an additive manner on the background of oncogenic $KRAS^{G12D}$ expression and *Trp53* deletion in mice. I generated a murine model whereby only sgRNAs against *Trp53* were used to introduce CRISPR-spCas9 mediated mutation into animals, along with the plasmid expressing $KRAS^{G12D}$. By comparison of the transcriptomes of my CRISPR-spCas9 screen tumours to this model it was possible to identify with more reliability which mutations may be functioning in an additive manner to the existing mutational profile.

These mice yielded discrete, macroscopic tumours in as little as 31 days, with a median survival of 40 days, and these tumours were GFP positive, indicating the presence of my $KRAS^{G12D}$ plasmid, and also have ductular iCCA morphology as shown by the positivity for CK19 staining, observed in Figure 16.

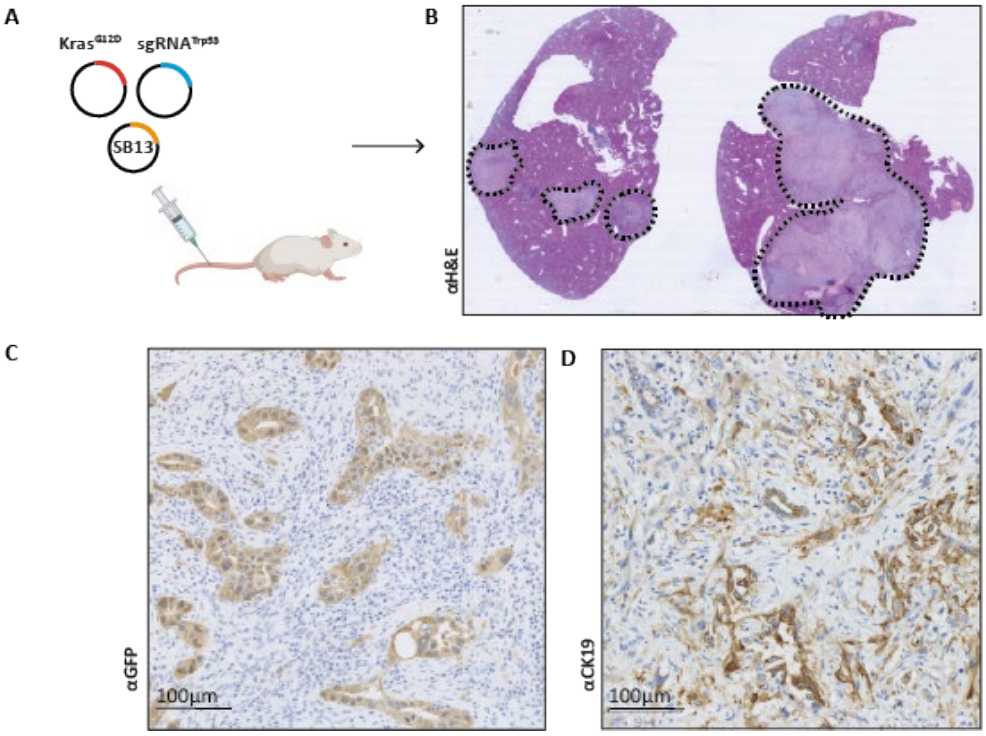

A
Kras^G12D
sgRNA^Trp53
SB13
B
αH&E
C
αGFP
100μm
D
αCK19
100μm

Figure 16

Establishment of an iCCA model with oncogenic KRAS^{G12D} and loss of Trp53. A. Schematic of the model. Naked DNA plasmids (SB-CRISPR -sgRNA^{Trp53}, KRAS^{G12D}, SB13 transposon) are injected by HDTVI into 5 week old FVB female mice. Around 40 days post injection the mice develop macroscopic liver cancers and are euthanised at end point. B H&E staining of a sample liver shows these discrete macroscopic tumours that are outlined. C. GFP stained KRAS-positive tumours where cancerous ductal morphology can be observed. D. these tumours are also CK19 positive.

3.10 CRISPR-*sp*Cas9 screening identifies rare drivers of iCCA

Having established this model with a defined genotype that produces reliably reproducible iCCA in mice, I then used it as a known genotype to which I could compare the CRISPR-*sp*Cas9 library screen tumours with. Having defined the mutations present in the library screen tumours with exome sequencing I wished to examine the transcriptome using RNA-sequencing.

The library tumours clustered significantly on principal component analysis (PCA), as shown in blue in Figure 17, and indeed they clustered differently to wild type bile ducts (purple) that clustered close to each other. PCA allows a dataset that exists in multiple dimensions to be flattened and simplified into a 2D representation, which is effective in teasing out variation between datasets. Additionally, screen tumours were transcriptionally similar to tumours from the established model of overexpression of $KRAS^{G12D}$ and deletion of *Trp53*.

RNA-sequencing analysis most importantly identified a subset of tumours that were transcriptionally distinct from the others and that were transcriptionally different to tumours which were driven only by $KRAS^{G12D}$ and *Trp53*-loss, and these all contained mutations in *Nf2* and *Trp53*. As can be observed in Figure 17, four of the five screen tumours cluster separately from the main cluster of screen tumours, as shown in green. This is interesting as it is indicative that the additional deletion of Nf2 is capable of functionally cooperating with the established model of oncogenic $KRAS^{G12D}$ and *Trp53* deletion, and this has an effect on the iCCA phenotype.

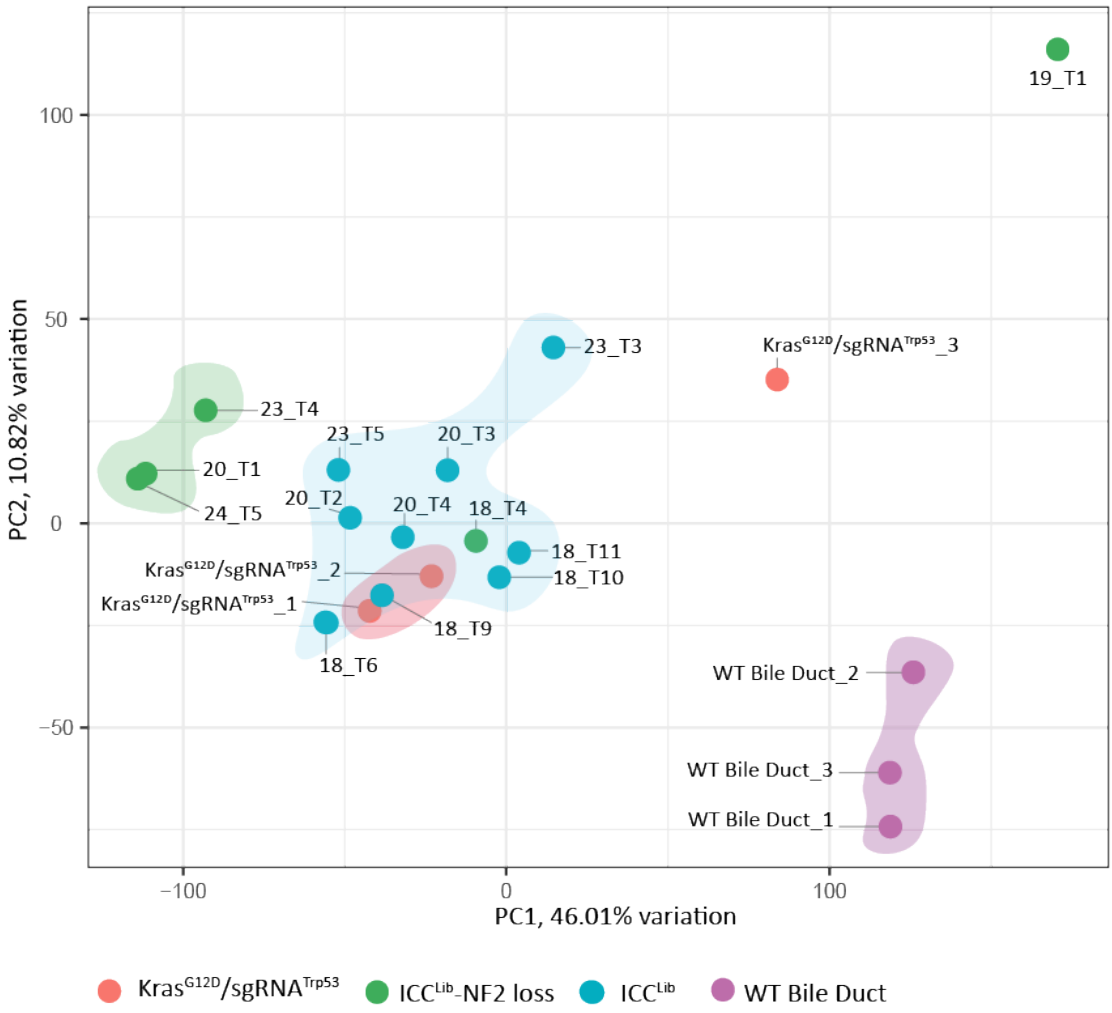

19_T1
100
50
0
−50
PC2, 10.82% variation
23_T3
Kras^G12D/sgRNA^Trp53_3
23_T4
23_T5
20_T3
20_T1
20_T2
20_T4
18_T4
24_T5
18_T11
18_T10
Kras^G12D/sgRNA^Trp53_2
Kras^G12D/sgRNA^Trp53_1
18_T9
18_T6
WT Bile Duct_2
WT Bile Duct_3
WT Bile Duct_1
−100
0
100
PC1, 46.01% variation
Kras^G12D/sgRNA^Trp53
ICC^Lib-NF2 loss
ICC^Lib
WT Bile Duct

Figure 17

*RNA-sequencing identified a subset of CRISPR-*sp*Cas9 screen tumours that all harbour mutations in Neurofibromin 2. Principal component analysis illustrates the clustering of screen tumours (iCCA^{Lib}) compared with KRAS^{G12D}:sgRNA^{Trp53} tumours and wild type FBV (non-cancerous, and non-modified) isolated bile ducts. Tight clustering of 3 sample tumours harbouring loss of Nf2 are illustrated by the green points and clustering by the green shadow (iCCA^{Lib}-Nf2 loss), and another point is less tightly clustered. The wild type bile ducts cluster together far away from all the other tumourous samples and are shown in purple. The iCCA^{Lib} samples that do not harbour Nf2 mutations are shown in blue and cluster closely together, as illustrated by the light blue shadow. 2 of the KRAS^{G12D}:sgRNA^{Trp53} tumours, shown in red, cluster very closely together and the other somewhat outlies.*

3.11 Discussion

Somatic multiplex mutagenesis has been previously used to demonstrate that editing of hepatocytes leads to the generation of iCCA in HDTVI of mice, but with a limited library of only 10 sgRNAs (173). This work acted as a foundation to my research whereby I was able to use the SB-CRISPR plasmid and transposase (SB13) system to endogenously edit a large library of 91 genes that we identified in our lab as functional loss of function driver mutations of iCCA in humans. Considering the rarity of non-fluke associated iCCA and the subsequent difficulty of acquiring sufficient tissue samples for large scale sequencing analysis, this study provided insight into how the genetic changes in this disease interact to drive tumour initiation. To date, a number of CRISPR screens in human cancer cells have been used but these fail to recapitulate the selective pressure found in tissues and initiated by the immune system as it attempts to clear the tumour. By developing a mouse model of iCCA in which we can investigate the genetic complexity of this cancer, we can more closely represent human disease where an immune and tumour microenvironment co-evolve with the cancer. Furthermore these types of model allow us to understand the interaction of the tumour with the surrounding tissues, including acellular structures such as the ECM that would not be possible in a screen of cultured cells.

My screen identified which mutations were capable of cooperating with oncogenic $KRAS^{G12D}$ to form tumours, this is indicative of which are capable of allowing Ras positive cells to be retained in the liver for long enough to become tumourigenic. Nf2 was identified to act in an additive manner to cause tumourigenesis to *Trp53* deletion and overexpression of oncogenic KRAS. The presence of additional mutations despite *Trp53* loss and KRAS overexpression being sufficient themselves to cause

tumourigenesis raises the possibility that other loss of function mutations could be capable of standing in for *Trp53* loss: Ras mutant cells may be able to escape senescence surveillance through the mutation of other genes than *Trp53*. In fact, this may provide an explanation as to the great number of infrequently mutated genes we observe in iCCA as a cancer: the deletion of many genes are capable of allowing Ras-driven tumourigenesis. To continue this work I therefore chose to examine the cancer-forming ability of several of the hits from this screen: *Cdkn2a, Brca2, Epha2, Ncor1, Nf2, Plk2, Plxb2* and *Rnf31*.

Another dimension to consider in this CRISPR-*sp*Cas9 library screen is the choice of $KRAS^{G12D}$ as the gain of function mutation on which to base the rest of the screen. The use of other oncogenes or fusion mutations as the background of these experiments would also give great insight into the cooperative nature of loss of function mutations with alternative background drivers. For example, several studies have observed the presence of missense mutations in the isocitrate dehydrogenase 1 and 2 genes (IDH1/2) (67,73,193), and this would therefore be an interesting choice of gain of function mutations to look at instead of $KRAS^{G12D}$ as done here.

Chapter 4 Results

Validation of CRISPR-*sp*Cas9 screen-identified genes

4.1 Introduction

Following the identification of potential driver genes in the CRISPR-*sp*Cas9 screen, I validated these *in vivo* using the HDTVI and SB transposon as for the pooled library screen. The difference being instead of injecting with an entire library of sgRNAs, I injected these animals with a single sgRNA against the gene(s) of interest. This was performed to confirm whether these mutations are sufficient to cause tumourigenesis *in vivo* by cooperating with over-expression of human $KRAS^{G12D}$ and/or *Trp53* loss.

The liver has evolved a protective mechanism whereby cells expressing activated oncogenes, such as $KRAS^{G12D}$, produce secreted molecules such as proteases, interleukins, inflammatory cytokines, chemokines, MMPs and growth factors, the so called senescence associated secretome or SASP phenotype. (194) These signals are detected by neighbouring cells that signal to induce an immune response, ultimately resulting in the apoptosis and clearance of the mutant cells in question from the liver. (195) This normally protective process can be overcome if the oncogenic cells also contain additional loss of function mutations, though which mutations are sufficient to overcome this is not known. In this chapter I present the experiments carried out to establish whether any of my genes of interest are capable of circumventing this senescence associated clearance system and cause tumourigenesis. Firstly the relevant background literature, and justifications for my chosen genes of interest are introduced, then I examine the results of the knockout of each gene *in vivo*, in turn.

4.1.1 Senescence Surveillance in the liver

Senescence is a cellular phenomenon first identified in 1961 by Hayflick and Moorhead in human diploid fibroblast cultures whereby proliferation gradually declined until the cells were no longer undergoing population doubling or growth at all. This occurred despite ample space and nourishment for the cells. (196) Further study has shown that this type of senescence is a result of telomere shortening gradually happening after each population doubling, known as replicative senescence, and is protective against DNA damage by limiting the capacity of cells to undergo long-term uncontrolled division. There is another type of senescence that occurs prior to telomere shortening that can be triggered by a number of stressors for example oncogene expression or tumour suppressor inactivation (as well as other stressors such as exposure to DNA damaging reagents and disruption of chromatin). (2)

The process of stable cell cycle arrest known as cellular senescence that is entered by cells upon the activation of oncogenes, provides a key protective mechanism against tumourigenesis. In addition to suppressing tumour cell development by inducing states of cell cycle arrest, senescence surveillance also has a role in inducing a response against pre-malignant cells by activating cells immune response. In the liver, cells undertake oncogene-induced senescence and this causes the targeted cells to produce cytokines, chemokines and growth factors, thus targeting them for clearance by immune cells that are involved in the CD4+ T-cell mediated adaptive immune response. This was shown by Kang *et al* who delivered oncogenic $Nras^{G12V}$ to hepatocytes by HDTVI and allowed stable integration of transposable elements that allowed oncogene expression. They showed that these elements were totally cleared from the murine liver 60 days post injection, whereas $Nras^{G12V/D38A}$ cells, that are an

inactive Nras effector loop mutant, were not cleared. (3) This mutant Nras is incapable of downstream signalling and therefore these cells cannot induce senescence, despite also having a mutation that would normally render Nras constitutively active. Additionally, immune cells are found to be gathered around senescent $Nras^{G12V}$ hepatocytes which is concurrent with the idea that immune cells regulate the removal of senescent hepatocytes from the liver.

4.2 Chosen Genes of interest

Led by the CRISPR-spCas9 screen (Chapter 3), several genes were selected to ascertain whether they were sufficient to cause iCCA in vivo in isolation from the rest of the library.

4.2.1 *Epha2* and *KRAS* interact in *in vitro* systems

Eph receptors interact with their corresponding Eph receptor-interacting, ephrin, ligands to form a signalling network that is highly conserved across many organisms. (197) Epha2 is a gene encoding for a protein from the Eph receptor family, an important class of receptor tyrosine kinases that constitute a family of 14 Eph receptors and 8 related ephrin ligands (5). Eph receptor-ephrin signalling is involved with cell to cell adhesion, or repulsion thus giving them a role in axon guidance and angiogenesis (6). Eph receptors and their ligands are usually expressed on adjacent cell populations meaning that they control cell movement by differing expression of said receptors or ligands. In humans, Epha2 and its corresponding ligand ephrin A1 can signal in both a forward (ligand-held cell to receptor-held cell), and in reverse where

the receptor can signal back to the ligand expressing cell. This receptor-ligand complex controls many cellular processes in embryonic development, angiogenesis and tumourigenesis.

In 2016, Porazinski *et al* showed that *Epha2* is required for epithelial tissues that express oncogenic Ras to be detected by neighbouring cells and to be extruded from epithelial sheets of tissue. They showed that this process is dependent on signalling via E-cadherin cell to cell interactions and cytoskeletal signalling, and that Epha2-ephrin A signalling is responsible for the detection and extrusion of oncogenic Ras-expressing cells. (7) Further to this, in 2021 Hill *et al* showed that a similar process happens in the adult pancreas where $Kras^{G12D}$ oncogene-expressing cells are cleared from different pancreatic compartments in an Epha2-dependent manner. (8)

These data were of particular interest to me given the requirement of *Epha2* to aid the occlusion of oncogenic Ras cells. For this reason I considered whether loss of *Epha2* would allow tumourigenesis to progress in my model because the lack of *Epha2* would allow oncogenic Ras cells to be retained for a sufficient time in the liver. Additionally, *EPHA2* loss has been previously described to occur in iCCA. One study of 30 iCCA patients found that iCCA was frequently mutated, and could particularly promote lymphatic metastasis of iCCA. (198) Interestingly, *EPHA2* overexpression has also been observed in iCCA, where aberrant activation of *EPHA2*'s downstream signalling pathways (such as mammalian target of rapamycin complex 1 (mTORC1) and extracellular signal-regulated kinase (ERK) pathways). (199)

Due to these previous reports of *EPHA2* dysregulation in iCCA, and its presence in our own human exome sequencing screen, I chose this gene to examine in more detail.

4.2.2 *BRCA2*

One of the two major breast cancer susceptibility genes, *BRCA2* loss has also been implicated in the progression of iCCA. In one study investigating the frequency of *BRCA2* mutations in biliary tract cancers, *Brca2* was mutated in 2.6% of iCCA patients examined. (9) Although rare, *BRCA2*-related iCCA has been identified as a subgroup of iCCA. *BRCA2* functions as a core mediator of homologous recombination where it recruits the recombinase RAD51 to sites of double stranded DNA breaks. (10) *BRCA2*-mutant cells are deficient in homologous recombination which leads to genomic instability and increased likelihood of malignant transformation. It has been reported that *BRCA2* somatic biallelic inactivation infers sensitivity to poly ADP ribose polymerase (PARP) inhibition, which illustrates the potential of *BRCA2* as a gene that could be targeted therapeutically. (200)

4.2.3 *CDKN2A*

The *CDKN2A* (cyclin dependent kinase inhibitor 2A) locus encodes the cyclin kinase inhibitors p16ink4a, and also the alternate reading frame product p14arf. (201) Under normal conditions, CDKN2A acts as a cell-cycle inhibitor protein where it causes cell cycle arrest in G1 phase, whereas mutant Cdkn2a is unable to form complexes with cyclin-dependent kinases and cyclin D, therefore allowing the cell to undergo mitosis. *CDKN2A* contains mutations leading to loss of function in many types of human cancer,

and in this aspect has been likened to Trp53. Several studies have identified mutations in *CDKN2A* in intrahepatic cholangiocarcinoma. (68,70,202–204)

4.2.4 *PLK2*

Polo-like kinase 2 is a gene encoding for the serine/threonine protein kinase of the same name. The polo-like kinases are a protein family that act by binding and phosphorylating proteins at a specific motif using their POLO box domains. (12) Plk2 is involved in synaptic plasticity and memory. Plk2 signals via the small GTPases Ras and Rap, and it does this by eliminating the Ras activator RasGRF1 and leads to phosphorylation-dependent ubiquitin and proteosomal degradation of the Rap inhibitor SPAR. (13) PLK2 phosphorylation also stimulates the Ras inhibitor SynGAP and the Rap activator PDZGEF. *PLK2* is a *TP53* target gene, and thus is a tumour suppressor gene, therefore I considered the possibility that *Plk2* could substitute for loss of *Trp53* in the presence of oncogenic $KRAS^{G12D}$.

4.2.5 *RNF31*

Ring finger protein 31 (RNF31) encodes for an E3 ubiquitin-protein ligase that is responsible for targeting proteins for proteosomal degradation by labelling with poly-ubiquitin chains. RNF31 plays a key role in NF-KB activation, a pathway that has key roles in cell survival and suppression of cell death by upregulating anti-apoptotic molecules that inhibit caspases. (205) In order to become activated, polyubiquitination of IKKγ is required to allow the release of the NF-κB complex, where it can then enter

the nucleus and induce target gene expression. This process is carried out by the ubiquitin chain assembly complex (LUBAC), of which RNF31 is the main component. (206) It has been shown that RNF31 cleavage reduces its ability to activate the protective NF-κB signalling and that cleavage site mutation inhibits NF-κB-mediated apoptosis. (207) Additionally, RNF31 plays a role in the control and induction of inflammatory responses, which was shown in human hepatocytes where knockdown of RNF31 resulted in a reduced inflammatory response. (208) The loss of *Rnf31* in my CRISPR-*sp*Cas9 screen was particularly interesting given that RNF31 overexpression is observed in breast cancer, and has been attributed the role of stabilising estrogen receptor α, which results in cancer progression as it transcribes pro-proliferative genes. (209)

4.2.6 *NCOR1*

NCOR1 was first identified as a corepressor of thyroid hormone and retinoic acid receptors that can control nuclear receptor transcriptional expression in the absence of their ligands. (210) Further studies showed that *NCOR1* interacts with other transcription factors and acts as a transcriptional regulator in ways aside from via nuclear receptors. It has a wide role in tissue-specific gene expression regulation during differentiation, cell homeostasis and metabolism and differentiation. These processes include, for example, adipocyte regulation (211), muscle physiology (212), and control of immune tolerance regulation by dendritic cells. (213) With respect to its role in disease, macrophage *NCOR1* has been shown to mediate atherosclerosis repression (214), and is an important regulator of transcription in naïve and effector CD4+ T cells (215). Overexpression of *NCOR1* mRNA has been shown to co-regulate

estrogen receptor α signalling, and thus has been attributed as an independent prognostic for breast cancer (216). Indeed it has been implicated also as an independent predictor of tamoxifen resistance in breast cancer patients (217), and decreased *NCOR1* expression in a mouse model correlated with accumulated tamoxifen resistance (218). Its increased expression and changes to localisation have been altered in prostate and bladder cancer. (219,220)

4.2.7 *PLXNB2*

PLXNB2 is a semaphorin-binding surface receptor involved in cell to cell contacts by way of controlling cell contractility, and interacting with the cytoskeleton to control cell tension and adhesion. (221) Widely expressed in the nervous system, *PLXNB2* signals via Rho family small GTPases to control cell migration and morphogenesis. *PLXNB2* also regulates GTPases RAC and CDC42 in macrophages. (222) The interaction with these signalling pathways is interesting in the context of cancer formation given the commonality of signalling pathway dysregulation. *PLXNB2* also has a role in axon guidance, and is involved in the organisation of the actomyosin network during embryonic stem and neuroprogenitor cell development. (223) *Plxb2* has also been shown to have a role in the control of differentiation and motility of neurons. (224) The reduction of *PLXNB2* in breast cancer has been associated with worse prognosis, as well as with local recurrence. (225) Conversely to its loss being associated with poor breast cancer prognosis, *PLXNB2* is consistently upregulated in all human gliomas, and the mechanism is thought to be a result of interaction of *PLXNB2* with actin-based cytoskeletal dynamics and migration via interaction with semaphorin SEMA4C. (226)

The diversity of the roles of *PLXNB2,* and its role in small GTPase signalling were intriguing as to how its loss may be influencing tumour formation in my screen.

4.2.8 NF2 is at the nexus of many key signalling pathways

NF2 is a member of a cytoskeletal-associated FERM (4.1 protein/Ezrin/Radixin/Moesin) domain protein family, encoded by the neurofibromatosis type 2 gene (*NF2*). (120) The loss of the *NF2* gene causes its namesake disease neurofibromatosis 2 which is a disease resulting in multiple benign tumours forming in the nervous system including characteristic formation of bilateral vestibular schwannomas that arise from Schwann cells making up the myelin sheath surrounding neurons, menangiomas that arise from arachnoid cap cells within the meninges and ependymomas arising from ependymal cells that line ventricles in the brain and central spinal canal. (227) In our exome sequencing analysis based on iCCA patients, *NF2* was mutated in 2 patients, and these patients also contained loss of function mutations in *TRP53*, but did not contain mutations in *KRAS*.

NF2 is at the nexus of many signalling pathways including those involved in development and in other cellular processes such as differentiation and proliferation. NF2 negatively regulates the EGFR-Ras-ERK pathway, the Hippo signalling pathway, TGFβ signalling, receptor tyrosine kinases, Notch and Hedgehog signalling.

Hippo signalling regulates organ size in mammals using a kinase cascade of core proteins. NF2 is a classical activating protein of Hippo signalling by its interaction with Kibra forming a scaffolding complex that localises the kinases MST and LATS to the cell membrane. These kinases then phosphorylate the co-activator proteins YAP/TAZ to

the nucleus where they bind to the TEAD transcription factors. This leads to their retention in the cytoplasm and are able to be degraded by the proteosomal system.

NF2 has also been implicated in WNT/ β catenin signalling, which is key for regulation of developmental embryonic processes. NF2 has been shown, in pancreatic ductal adenocarcinoma cell lines and animal models, to be associated with inhibition of growth. This occurred by NF2 causing the inhibition of WNT/ β-catenin signalling, which has been shown to occur by its interaction with LRP6 where it blocks phosphorylation. (228)

4.3 Hypothesis

I hypothesise some of these chosen genes that are outcomes of the CRISPR-Cas9 screen can cause tumourigenesis *in vivo*, and when injected in combination with oncogenic $KRAS^{G12D}$ and *Trp53* loss, they accelerate tumourigenesis.

4.4 Aims

- Validate CRISPR-*sp*Cas9 screen *in vivo* by testing the tumour forming ability of selected genes: *Brca2, Cdkn2a, Epha2, Nf2, Ncor1, Plk2, Plxb2* and *Rnf31.*
- Find which of these genes are capable of causing tumour formation *in vivo* in combination with oncogenic $KRAS^{G12D}$, and not as a part of the CRISPR-*sp*Cas9 pooled library.

4.5 *In vivo* validation of CRISPR-*sp*Cas9 screen-informed driver genes

To test which of the genes described above initiate cancer when they are deleted alongside overexpression of oncogenic $KRAS^{G12D}$, and in the absence or presence of *Trp53* deletion, I deleted these genes *in vivo*. I assessed whether these genes act in an additive nature to the tumourigenesis of $KRAS^{G12D}$:$sgRNA^{Trp53}$, or if their loss is sufficient to cause cancer formation with overexpression of $KRAS^{G12D}$, and thus act in a compensatory manner for the loss of *Trp53*.

4.6 *Epha2, Plk2, Plxb2,* and *Rnf31* do not substitute for *Trp53* loss

4.6.1 *Plk2* and *Rnf31*

The CRISPR-*sp*Cas9 library screen presented loss of function mutations in *Plk2* in four samples, including in a pancreatic metastasis. Additionally, there were a total of seven mutations in *Rnf31*. Motivated by the commonality of these mutations in the library screen, and by their interesting molecular biology, I deleted *Plk2* or *RNf31* in groups of 5 week-old FVB mice in the presence or absence of *Trp53* loss, and in the presence of oncogenic $KRAS^{G12D}$.

Although macroscopic tumours formed in $KRAS^{G12D}$:$sgRNA^{Trp53/Plk2}$ animals, as observed in Figure 1A, when their survival was compared to that of $KRAS^{G12D}$:$sgRNA^{Trp53}$ animals, there was no statistically significant difference between the survival of animals containing additional mutations (Figure 1B). This indicates that the additional loss of either *Plk2* did not further cooperate with the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model to accelerate tumourigenesis. Additionally, when *Plk2* was deleted without the additional deletion of

Trp53, no tumours formed at all (Figure 1B), and animals were euthanised at a chosen end point of 58 days. These data strongly imply that the loss of *Plk2* alone is insufficient to allow cancer formation with the expression of oncogenic $KRAS^{G12D}$, however greater number of mice would be required to confirm this. The tumours that formed in this model were discrete and histologically representative of $KRAS^{G12D}$:$sgRNA^{Trp53}$ tumours (Figure 1C-E). I also quantified the tumour burden of these animals, and it did not appear to be significantly different compared to that of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model. Likewise the tumour number did not appear to be significantly different. However, as mentioned above the number of mice included in these experiments would need to be increased in order to perform statistically significant comparisons. These data also support that the loss of *Plk2* is not additive to the acceleration of tumourigenesis of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model.

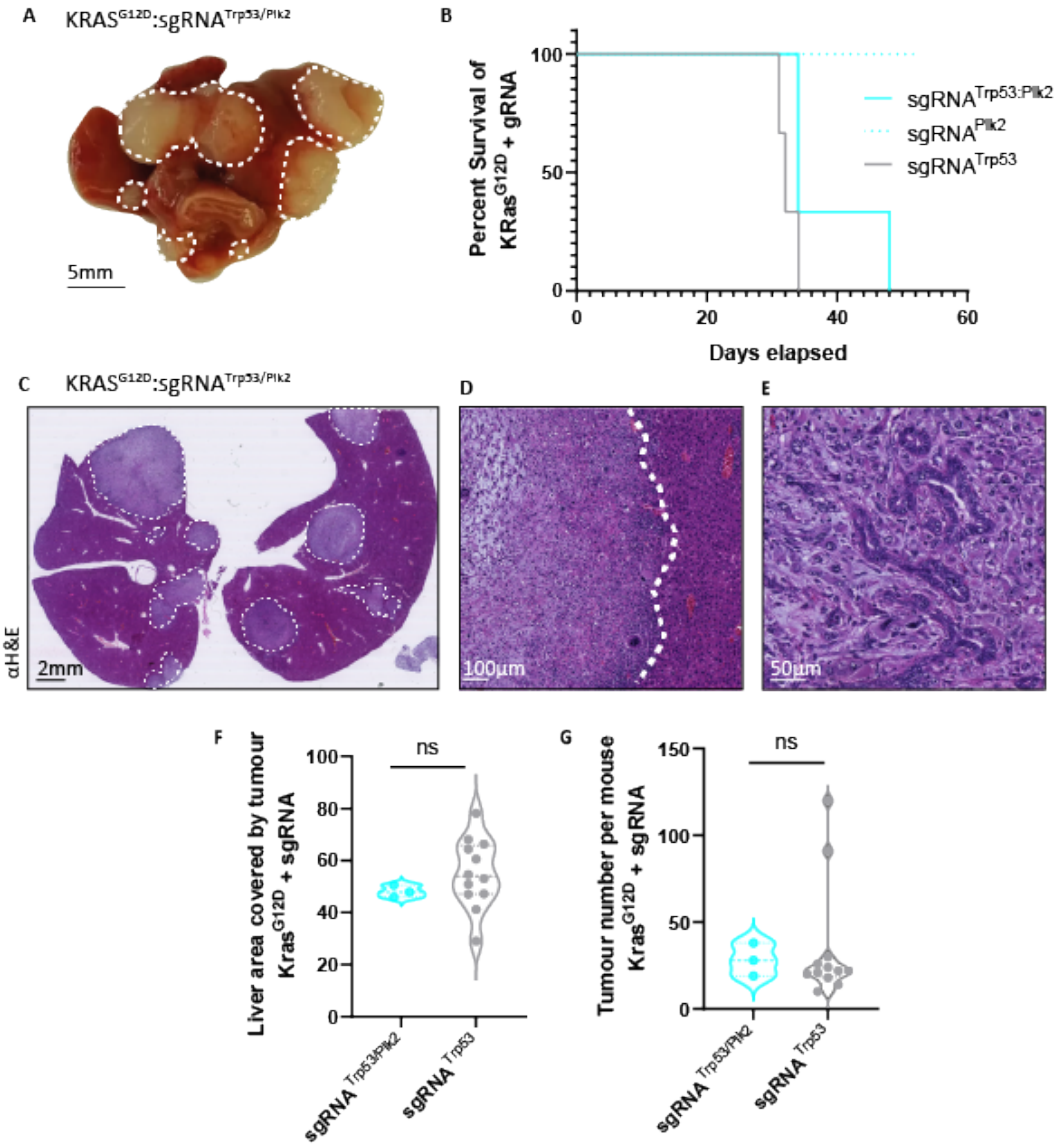

A
KRAS^{G12D}:sgRNA$^{Trp53/Plk2}$
5mm
B
Percent Survival of KRas^{G12D} + gRNA
100
50
0
0
20
40
60
Days elapsed
sgRNA$^{Trp53:Plk2}$
sgRNA^{Plk2}
sgRNA^{Trp53}
C
KRAS^{G12D}:sgRNA$^{Trp53/Plk2}$
αH&E
2mm
D
100μm
E
50μm
F
Liver area covered by tumour Kras^{G12D} + sgRNA
100
80
60
40
20
0
ns
sgRNA $^{Trp53/Plk2}$
sgRNA ^{Trp53}
G
Tumour number per mouse Kras^{G12D} + sgRNA
150
100
50
0
ns
sgRNA $^{Trp53/Plk2}$
sgRNA ^{Trp53}

Figure 1

Loss of Plk2 *and* Trp53 *on a background of oncogenic* $KRAS^{G12D}$ *(n=3) expression results in formation of tumours. Control group of* Trp53 *loss alone on a background of oncogenic* $KRAS^{G12D}$ *(n=12). A. Macroscopic* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Plk2}$ *murine liver with tumours outlined by dashed white lines. B. Kaplan-Meier survival curve of animals injected with* $Kras^{G12D}$ *and sgRNAs against* Plk2 *and/or* Trp53*. C. H&E stain of whole liver with tumours indicated with white dashed lines. D. Showing the point of meeting between tumour tissue (left) and liver parenchyma (right). E. iCCA tumour in this model. F. Quantification of tumour burden, measured as percentage of whole liver covered by tumour mass and G. Tumour number in this model, both compared to* $KRAS^{G12D}$*:*$sgRNA^{Trp53}$ *animals.*

Similarly to the KRAS^{G12D}:sgRNA$^{Trp53/Plk2}$ animals, loss of *Rnf31* and *Trp53* and expression of oncogenic KRAS^{G12D} resulted in formation of macroscopic iCCA tumours (Figure 2A). However, 1/3 of these animals did not require euthanasia and instead were culled at the chosen end point of 58 days alongside the KRAS^{G12D}:sgRNA^{Rnf31} animals, as shown in the Kaplan-Meier curve in Figure 2B. The tumours were histologically similar to the KRAS^{G12D}:sgRNA^{Trp53} model (Figure 2C-E). Interestingly, despite the tumour burden not being significantly different to that of the KRAS^{G12D}:sgRNA^{Trp53} model (Figure 2F), the tumour number was fewer. This shows these animals harboured larger singular tumours that had perfused further throughout the liver, but that were possibly less dangerous to the animals than the tumours present in the KRAS^{G12D}:sgRNA^{Trp53} model.

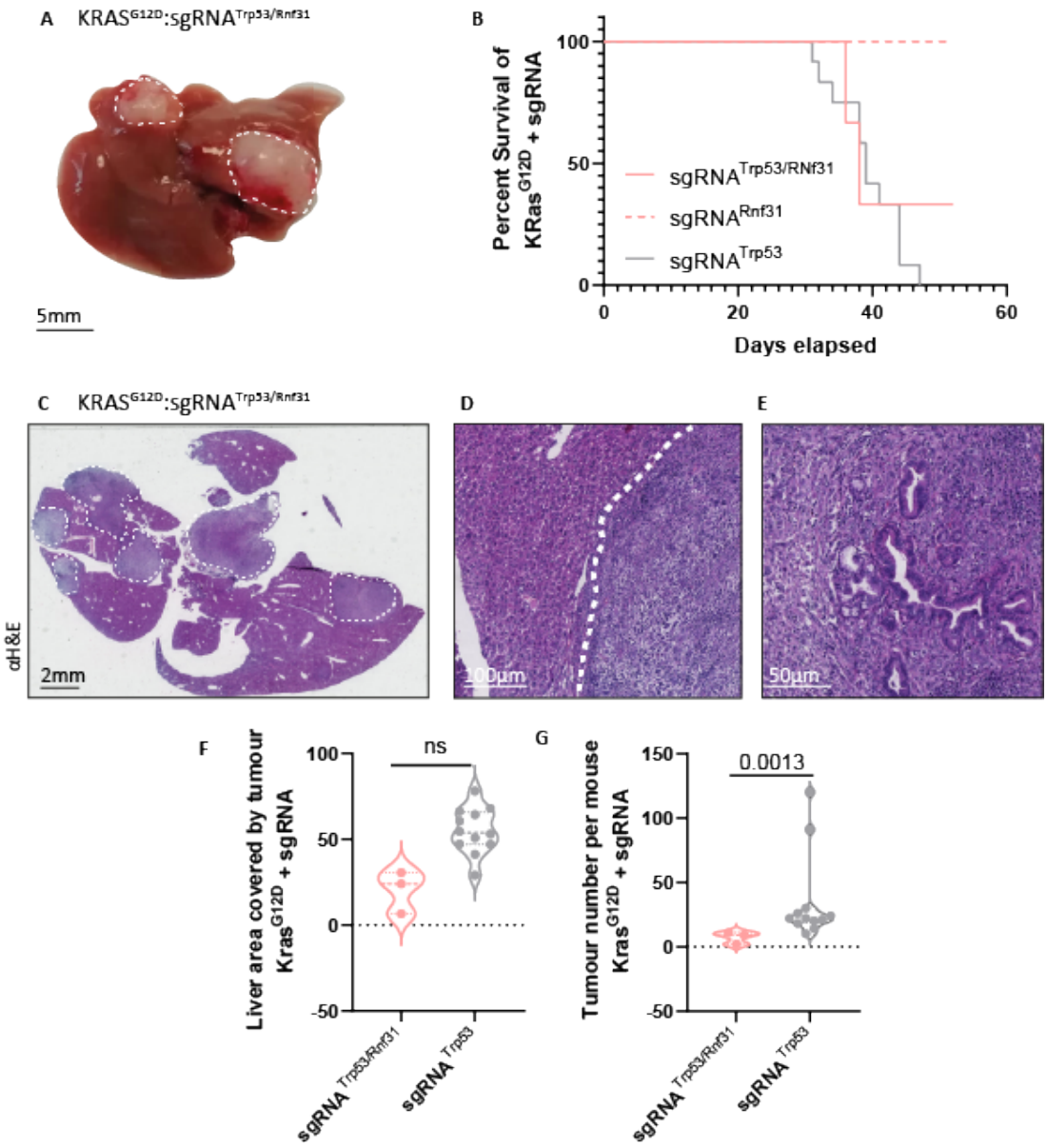

A
KRAS^{G12D}:sgRNA$^{Trp53/Rnf31}$
5mm
B
Percent Survival of Kras^{G12D} + sgRNA
100
50
0
sgRNA$^{Trp53/RNf31}$
sgRNA^{Rnf31}
sgRNA^{Trp53}
0
20
40
60
Days elapsed
C
KRAS^{G12D}:sgRNA$^{Trp53/Rnf31}$
αH&E
2mm
D
100μm
E
50μm
F
Liver area covered by tumour Kras^{G12D} + sgRNA
ns
100
50
0
-50
sgRNA$^{Trp53/Rnf31}$
sgRNA^{Trp53}
G
Tumour number per mouse Kras^{G12D} + sgRNA
0.0013
150
100
50
0
-50
sgRNA$^{Trp53/Rnf31}$
sgRNA^{Trp53}

Figure 2

Loss of Rnf31 *and* Trp53 *on a background of oncogenic KRAS^{G12D} (n=3) expression results in formation of tumours. Control group of* Trp53 *loss alone on a background of oncogenic KRAS^{G12D} (n=12). A. Macroscopic KRAS^{G12D}:sgRNA$^{Trp53/Rnf31}$ murine liver with tumours outlined by dashed white lines. B. Kaplan-Meier survival curve of animals injected with Kras^{G12D} and sgRNAs against* Rnf31 *and/or* Trp53*. C. H&E stain of whole liver with tumours indicated with white dashed lines. D. Showing the point of meeting between tumour tissue (left) and liver parenchyma (right). E. iCCA tumour in this model. F. Quantification of tumour burden, measured as percentage of whole liver covered by tumour mass and G. Tumour number in this model, both compared to KRAS^{G12D}:sgRNA^{Trp53} animals.*

4.6.2 *Epha2* loss allows retention of $KRAS^{G12D}$ positive cells *in vivo*

There is an established cooperation between loss of *Epha2* and overexpression of oncogenic RAS^{G12D} *in vitro*. Additionally, in the computational analysis of iCCA patient exome sequencing data, *EPHA2* was the most recurrently mutated of the "novel" gene group. Given the role of this gene in permitting RAS positive cells to be retained in the liver for sufficient time for them to become tumourigenic: for them to avoid hepatic senescent surveillance, I examined the early stages of tumour formation. To do this, I deleted *Epha2* with the overexpression of oncogenic $KRAS^{G12D}$ *in vivo* and took animals over a course of time points commencing from two days post HDTVI of cancer-inducing plasmids. In order to demonstrate that loss of function mutations are required for the retention and subsequent tumourigenesis of RAS positive cells, I used a control whereby a sgRNA with a scrambled guide sequence ($sgRNA^{Scrm}$) was used in place of *Epha2.* Using QuPath Positive Cell Detection, I quantified the number of KRAS positive cells over the time course, for all fixed and stained tissue. As can be observed in sample images shown in Figure 3A, the $KRAS^{G12D}$:$sgRNA^{Scrm}$ animals, despite the peak of KRAS positive cells at day seven, these animals succeed in clearing RAS cells from their livers and the number of positive cells drops to 0 at days 14 and 21. The $KRAS^{G12D}$:$sgRNA^{Trp53}$ animals, however, have a higher number of KRAS positive cells at day 7, and succeed in retention of KRAS positive cells over days 14 and 21. Histologically, the number of larger (considered as >5 cells clustered together; counted by hand; 10 field replicates of 100μm^2 counted per stained and fixed piece of tissue) KRAS positive lesions in $KRAS^{G12D}$:$sgRNA^{Epha2}$ animals is greater at day 2, as shown in Figure B and C. However, as the time course progresses, the number of KRAS positive cells decreases and histologically represents the $KRAS^{G12D}$:$sgRNA^{Scrm}$ control animals.

As these animals were euthanised at these early time points, and tumourigenesis was interrupted, it is indeterminate whether the deletion of *Epha2* is simply slowing the process of clearance, or whether it is overcoming the mechanism of senescence surveillance completely, and whether tumours would form if tumourigenesis would be left to complete term. Given the absence of any KRAS positive lesions in the latter 21 day time points, I predict that the process of senescence surveillance has been slowed, but not prevented, by deletion of *Epha2* in this model.

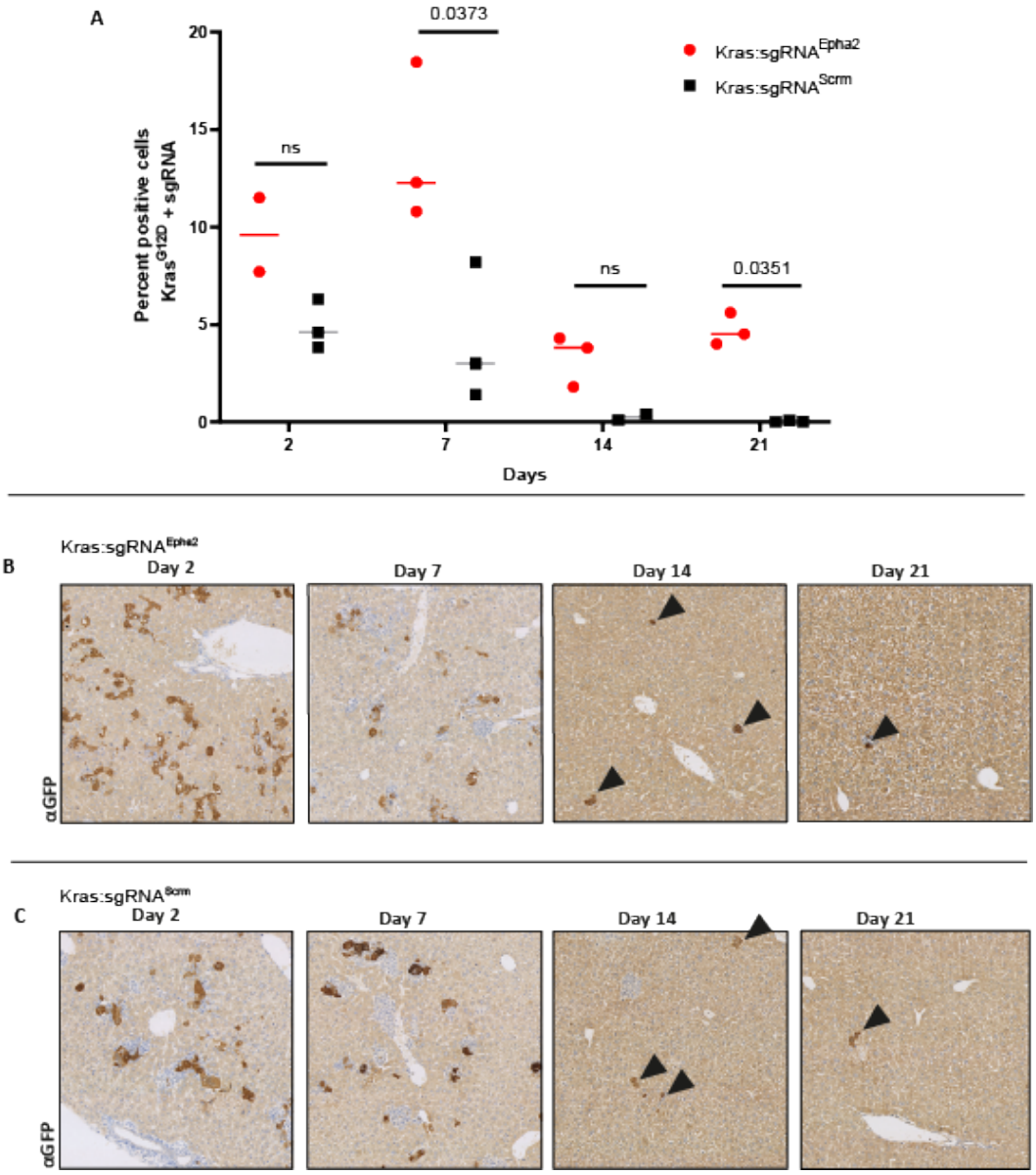

A
0.0373
Kras:sgRNA^Epha2
Kras:sgRNA^Scrm
Percent positive cells
Kras^G12D + sgRNA
ns
ns
0.0351
20
15
10
5
0
2
7
14
21
Days
B
Kras:sgRNA^Epha2
Day 2
Day 7
Day 14
Day 21
αGFP
C
Kras:sgRNA^Scrm
Day 2
Day 7
Day 14
Day 21
αGFP

Figure 4

Loss of Epha2 *on a background of oncogenic KRAS^{G12D} expression results in retention of RAS positive cells from the liver. (KRAS^{G12D}:sgRNA^{Epha2} (red) experimental numbers: Day 2 n=2, Day 7 n=3, Day 14 n=3, Day 21 n=3). Control group of loss of Scrambled (Scrm) non-targeting sgRNA on background of oncogenic KRAS^{G12D} expression. (KRAS^{G12D}:sgRNA^{Scrm} (grey) experimental numbers: Day 2 n=3, Day 7 n=3, Day 14 n=2, Day 21 n=3) A. Percentage positive cells over time course over 2, 7, 14 and 21 days for KRAS^{G12D}:sgRNA^{Epha2} (red) and KRAS^{G12D}:sgRNA^{Scrm} (grey). B GFP staining identifies KRAS^{G12D} positive cells being cleared from the liver over time in KRAS^{G12D}:sgRNA^{Epha2} animals and C. in KRAS^{G12D}:sgRNA^{Scrm} animals.*

4.6.3 *Brca2* loss decelerates tumour formation

Brca2 was lost in a total of five tumour samples, from three different animals in the CRISPR-*sp*Cas9 screen and this, in addition to its established role in cancer, motivated me to select it to examine the result of its deletion in more detail.

Deletion of *Brca2* alongside deletion of *Trp53* and overexpression of $KRAS^{G12D}$ resulted in very limited tumour formation. Animals were euthanised at 58 days and they presented with only 0-2 medium sized lesions per animal, as illustrated in Figure 5A. Two animals did not present with tumours at all, and when compared to the survival of either $KRAS^{G12D}$:$sgRNA^{Trp53}$ or $KRAS^{G12D}$:$sgRNA^{Trp53/Scrm}$ animals, survived significantly longer than either control group (Figure 5B). I introduced the $KRAS^{G12D}$:$sgRNA^{Trp53/Scrm}$ controls to this experiment to ensure that the presence of an additional sgRNA (the non-targeting $sgRNA^{Scrm}$), does not affect the tumourigenic nature of these models. The survival of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ and $KRAS^{G12D}$:$sgRNA^{Trp53/Scrm}$ animals did not alter significantly, and they histologically represent the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model. In the case of *Brca2* loss, the tumours that did form represented iCCA (Figure 5C). When I quantified the tumour burden and tumour number, they are significantly reduced compared to the $KRAS^{G12D}$:$sgRNA^{Trp53/Scrm}$ animals. The absence of any tumours in these models could be attributed to *Brca2* deletion sensitising the cells to clearance from the liver by the immune system. The loss of *Brca2* decelerates tumour formation, and indicates that this mutation has a protective effect on the animals harbouring oncogenic $KRAS^{G12D}$ and *Trp53* deletion.

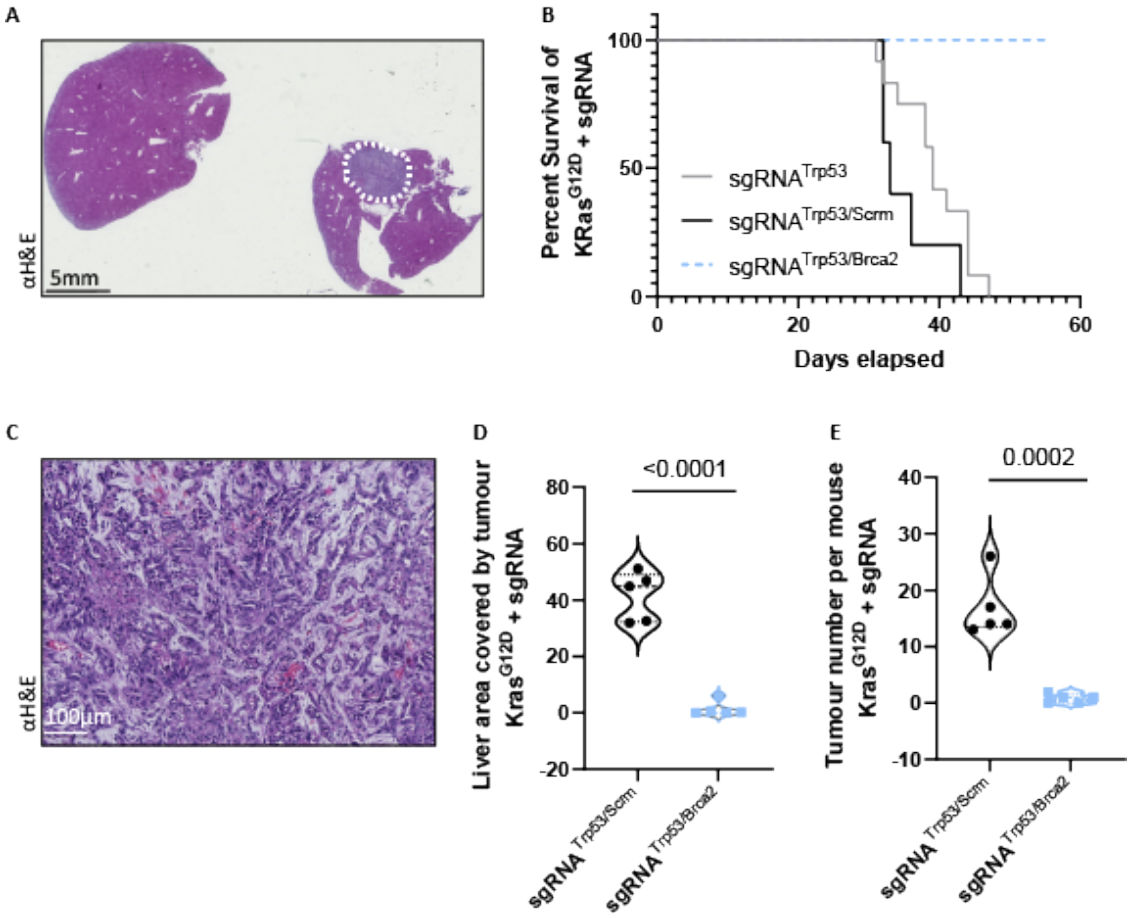

A
αH&E
5mm
B
Percent Survival of KRas^G12D + sgRNA
100
50
0
0
20
40
60
Days elapsed
sgRNA^Trp53
sgRNA^Trp53/Scrm
sgRNA^Trp53/Brca2
C
αH&E
100µm
D
Liver area covered by tumour Kras^G12D + sgRNA
<0.0001
80
60
40
20
0
-20
sgRNA Trp53/Scrm
sgRNA Trp53/Brca2
E
Tumour number per mouse Kras^G12D + sgRNA
0.0002
40
30
20
10
0
-10
sgRNA Trp53/Scrm
sgRNA Trp53/Brca2

Figure 5

Loss of Brca2 *and* Trp53 *on a background of oncogenic KRAS^{G12D} (n=5) expression results in deceleration of tumour formation. Control group of loss of Scrambled (Scrm) non-targeting sgRNA and Trp53 on a background of oncogenic KRAS^{G12D} (n=5). A. Macroscopic KRAS^{G12D}:sgRNA$^{Trp53/Brca2}$ murine liver with a singular tumour outlined by dashed white lines. B. Kaplan-Meier survival curve of animals injected with plasmids expressing KRAS^{G12D} and sgRNAs to induce mutations in* Brca2 *and/or* Trp53, *and compared animals a non-targeting Scrambled sgRNA. C. H&E stain of tumour detail: tumour represents iCCA. D. Quantification of tumour burden, measured as percentage of whole liver covered by tumour mass and E. Tumour number in this model, both compared to KRAS^{G12D}:sgRNA^{Trp53} animals.*

4.6.4 Loss of *Ncor1*

Ncor1, the nuclear co-repressor, was mutated in a total of 2 samples in the CRISPR-*sp*Cas9 library, one of which included a pancreatic metastasis in the CRISPR-*sp*Cas9 screen. Given the known roles of *NCOR1* breast, prostate and bladder cancer (217,219,220), I decided to examine this in more detail. As with the genes already described in this chapter, I deleted *Ncor1* in animals with the additional deletion of *Trp53* and a non-targeting scrambled sgRNA, and with the overexpression of oncogenic $KRAS^{G12D}$. As illustrated in Figure 6A, the survival of animals harbouring loss of *Ncor1* was not significantly altered compared to animals with loss of *Trp53* alone (grey), or to those with *Trp53* and a non-targeting scrambled sgRNA (black).

To confirm that the expression of oncogenic $KRAS^{G12D}$ is required for tumourigenesis to occur, I also injected animals with sgRNA against *Trp53* and *Scrambled* without KRAS (grey dashed line in 6A). These animals harboured no tumours at all, and were euthanised at the set end point of 58 days post HDTVI. This confirms that the presence of the oncogene is required for tumours to form.

The similar nature of the animals that did have oncogenic KRASG12D expression, was reflected in the tumour burden and tumour number in Figure 6B&C, whereby neither burden or number are significantly different between these models. (Whereas $sgRNA^{Trp53/Scrm}$ animals present with no tumours at all.)

These data indicate several things. Firstly that the presence of another sgRNA in the HDTVI has no significant effect on tumourigenesis. Secondly, these data show that the presence of oncogenic $KRAS^{G12D}$ is a requirement for formation of tumours in this model, and that loss of function mutations alone are insufficient to cause tumour

formation. This is in line with what is known about senescence surveillance in the liver whereby oncogenic RAS is a requirement for cancer formation. Finally, the loss of *Ncor1* does not significantly affect the survival or tumour formation on the background of $KRAS^{G12D}$ overexpression and *Trp53* loss.

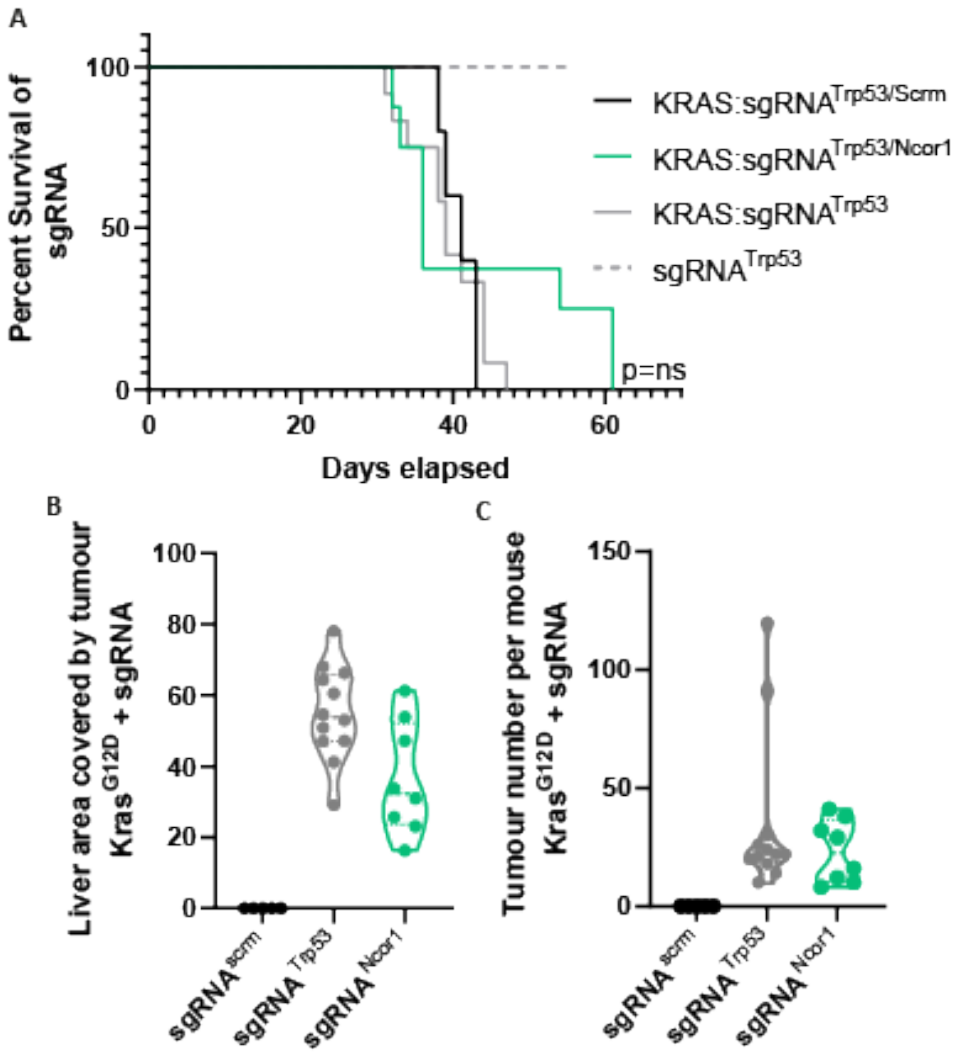

A
Percent Survival of sgRNA
100
50
0
0
20
40
60
Days elapsed
KRAS:sgRNA Trp53/Scrm
KRAS:sgRNA Trp53/Ncor1
KRAS:sgRNA Trp53
sgRNA Trp53
p=ns
B
Liver area covered by tumour Kras G12D + sgRNA
100
80
60
40
20
0
sgRNA scrm
sgRNA Trp53
sgRNA Ncor1
C
Tumour number per mouse Kras G12D + sgRNA
150
100
50
0
sgRNA scrm
sgRNA Trp53
sgRNA Ncor1

Figure 6

Loss of Ncor1 *and* Trp53 *on a background of oncogenic* $Kras^{G12D}$ *(n=8) expression results in little change to survival compared to loss of* Trp53 *(n=12), or with* Trp53 *loss and a non-targeting Scrambled (Scrm) sgRNA control (n=5), as shown in Kaplan-Meier in A, where p value (p) of survival between* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Ncor1}$ *and either control is non-significant (ns). B. Illustrating tumour burden and C., tumour number.*

4.7 *Ncor1* loss results in metastases

Despite there being no significant difference in the presentation of tumourigenesis in the livers of $KRAS^{G12D}$:$sgRNA^{Trp53/Ncor1}$ animals compared to the $KRAS^{G12D}$-containing controls, there was however tumours in various sites other than the liver, which are summarised in Table 1. Most notably were consistent macroscopic tumours to the pancreas and lungs in all $KRAS^{G12D}$:$sgRNA^{Trp53/Ncor1}$ animals. Other macroscopic tumours were present in the intestines, ovaries and diaphragm of one animal. 62.5% of the mice had macroscopically visible tumours in the common bile duct (CBD), and 75% had tumours, or KRAS positive lesions, to the spleen. I considered the tendency of iCCA to metastasise to the lymph nodes in humans, these were examined by eye but there were no visible tumours.

	CBD	LUNG	PANCREAS	SPLEEN	INTESTINE	OVARY	DIAPHRAGM
1077	✓	✓	✓	✓	☒	☒	✓
1078	☒	✓	✓	✓	☒	☒	☒
1076	✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓✓
1608	✓✓	✓✓	✓✓	✓			
1610	✓✓	✓✓	✓✓	✓			
1620	☒	✓✓	✓✓	☒			
1609	☒	✓✓	✓	✓			
1611	✓✓	✓✓	✓✓	☒			

Table 1.

Tumour formation in several organs other than the liver in $KRAS^{G12D}$:$sgRNA^{Trp53/Ncor1}$ animals. Singular tick (✓) indicates the presence of microscopically visualised lesions of >5 KRAS positive cells (visualised via GFP staining). Double tick (✓✓) indicates the presence of solid, macroscopic tumours of >500uM. Cross (☒) indicates the absence of tumours in these tissues, and blank spaces are from organs that were not collected for these animals. Mouse numbers are listed in the first column. Common bile duct: CBD.

I was particularly interested in the presence of lung metastases in the $KRAS^{G12D}$:$sgRNA^{Trp53/Ncor1}$ mice because the presence of tumour cells here means they travelled from the intraperitoneal cavity where the liver sits to the thoracic cavity, through the diaphragm. To visualise the tumour cells, I stained tissues for GFP to confirm their KRAS positive status. Given the reliable delivery of plasmids to the liver specifically by HDTVI, and the KRAS positive nature of these cells, I consider them to be metastases from the primary site of liver tumours. Figure 7 shows the presence of macroscopic KRAS positive metastatic tumours in A, KRAS positive lesions in B and detail of the nature of the tumours present in C and D.

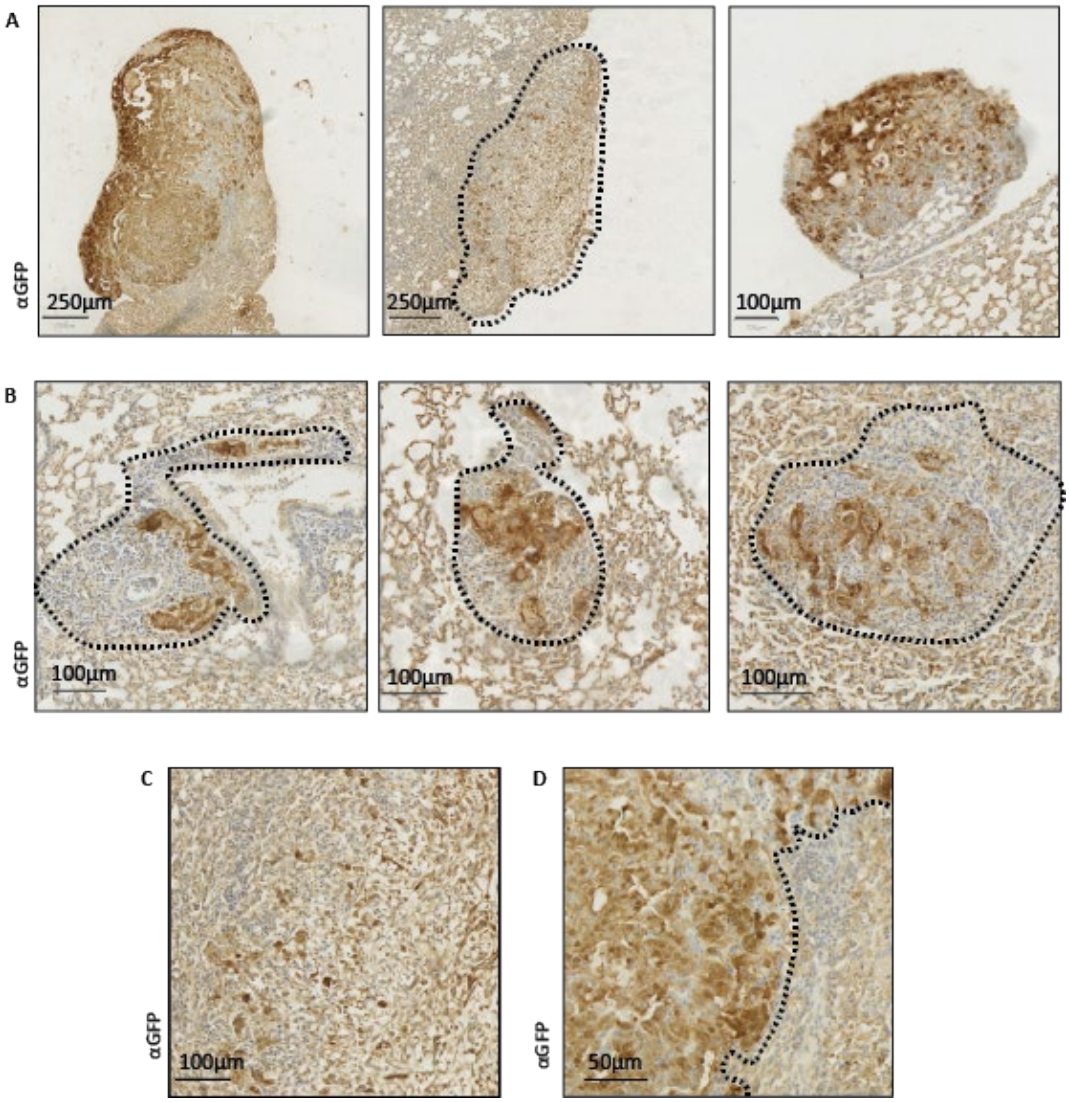

A
αGFP
250µm
250µm
100µm
B
αGFP
100µm
100µm
100µm
C
αGFP
100µm
D
αGFP
50µm

Figure 7

$KRAS^{G12D}$:$sgRNA^{Trp53/Ncor1}$ mice present consistently with metastases to the lungs. GFP-stained lung tissue from $KRAS^{G12D}$:$sgRNA^{Trp53/Ncor1}$ mice. A. Macroscopic tumours, B. tumourigenic lesions. C. Tumours are stromatic in nature and D. indicating the clear difference between KRAS positive cells and non-tumourous lung tissue.

To examine further the presence of the lung lesions in my *Ncor1* loss model, I also stained two control models, $KRAS^{G12D}$:$sgRNA^{Trp53}$ and $KRAS^{G12D}$:$sgRNA^{Trp53/Scrm}$ to examine whether they contained KRAS positive lesions. Additionally, due to my lack of expertise in lung histology, I also stained otherwise wild type FBV lung from mice that were euthanised at 10 weeks old (FVB:WT) to ensure that what I was observing was not artefact of my antibody. There are lesions present in both control groups: those with loss of Trp53 alone, and with the additional loss of a non-targeting scrambled sgRNA, on the background of oncogenic $KRAS^{G12D}$, in the lungs and pancreas, and in one animal additionally in the spleen. Table 2 summaries these data. Interestingly there are no tumours at all present in the common bile duct (CBD), unlike when *Ncor1* is lost and tumours form in 5/8 animals. There is no metastasis to the ovaries, intestine or diaphragm.

		CBD	LUNG	PANCREAS	SPLEEN	INTESTINE	OVARY	DIAPHRAGM
KRAS/ TRP53	1680	☒	✓	☒	☒	☒	☒	☒
	1677	☒	✓	☒	☒	☒	☒	☒
	1679	☒	☒	☒	☒	☒	☒	☒
KRAS/ TRP53/SCRM	1602	☒	✓✓	✓	✓	☒	☒	☒
	1600	☒	✓	✓	☒	☒	☒	☒
	1603	☒	☒	☒	☒	☒	☒	☒
	1601	☒	✓	☒	☒	☒	☒	☒
	1621	☒	✓	☒	☒	☒	☒	☒

Table 2

Metastasis occurs in KRAS^{G12D}:sgRNA^{Trp53} and KRAS^{G12D}:sgRNA$^{Trp53/Scrm}$ mice. Singular tick (✓) indicates the presence of microscopically visualised lesions of >5 KRAS positive cells (visualised via GFP staining). Double tick (✓✓) indicates the presence of solid, macroscopic tumours of >500uM. Cross (☒) indicates the absence of tumours in these tissues, and blank spaces are from organs that were not collected for these animals. Mouse numbers are listed in the first column. Common bile duct: CBD.

I stained these tissues with GFP to show the presence of KRAS positive cells. Figure 8A shows the FVB lung tissue stained for GFP, illustrating how non-cancerous lung tissue appears. Figure 8B illustrates the only macroscopic tumour that was present in either of these control models, and Figure 8C shows the presence of KRAS positive lesions in the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model. Together these data show that the HDTVI of $KRAS^{G12D}$: and sgRNAs against *Trp53* and either *Ncor1* or *Scrambled* sgRNA cause metastases of tumours past the diaphragm into the lungs.

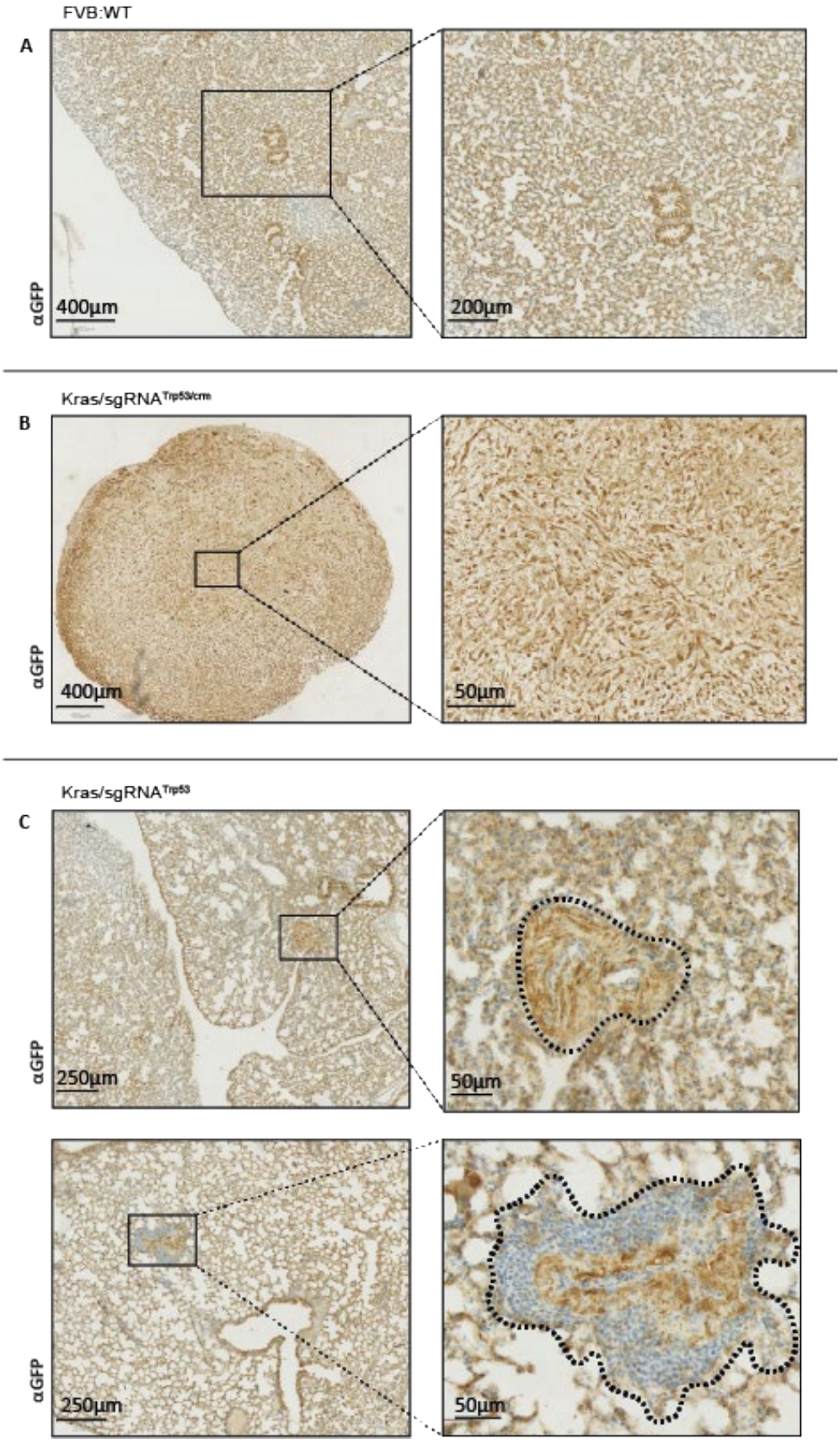

FVB:WT
A
αGFP
400μm
200μm
Kras/sgRNA Trp53lcrm
B
αGFP
400μm
50μm
Kras/sgRNA Trp53
C
αGFP
250μm
50μm
αGFP
250μm
50μm

Figure 8

KRAS positive cells metastasise to the lungs of KRAS^{G12D}:sgRNA^{Trp53} and KRAS^{G12D}:sgRNA$^{Trp5/Scrm}$ mice. A. Otherwise wild-type FVB mouse lung stained for GFP to act as a negative control and to observe any non-specific antibody binding. B. Macroscopic KRAS^{G12D}:sgRNA$^{Trp5/Scrm}$ lung tumour and C. two examples of the small KRAS positive lesions observed in KRAS^{G12D}:sgRNA^{Trp53} mice.

4.7.1 *Cdkn2a* loss doesn't significantly alter KRAS^{G12D}:sgRNA^{Trp53} mice

The established tumour suppressor gene CDKN2A and its transcribed protein p16^{INK4a} has a key role in the cell cycle and senescence via the regulation of cyclin dependent kinase and cyclin D complexes. I deleted *Cdkn2a* in mice as the other genes described in this chapter and, as illustrated in Figure 9, found that there appeared to be no significant difference between their survival and that of the KRAS^{G12D}:sgRNA^{Trp53} control animals. Likewise the tumour number and tumour burden were not significantly altered compared to either control group of pre-existing mutations. However, due to the small number of animals used here, a greater number should be used to increase the statistical significance. It was surprising that the loss of *Cdkn2a* did not alter tumourigenesis in the liver in any way considering its role in various cancer types.

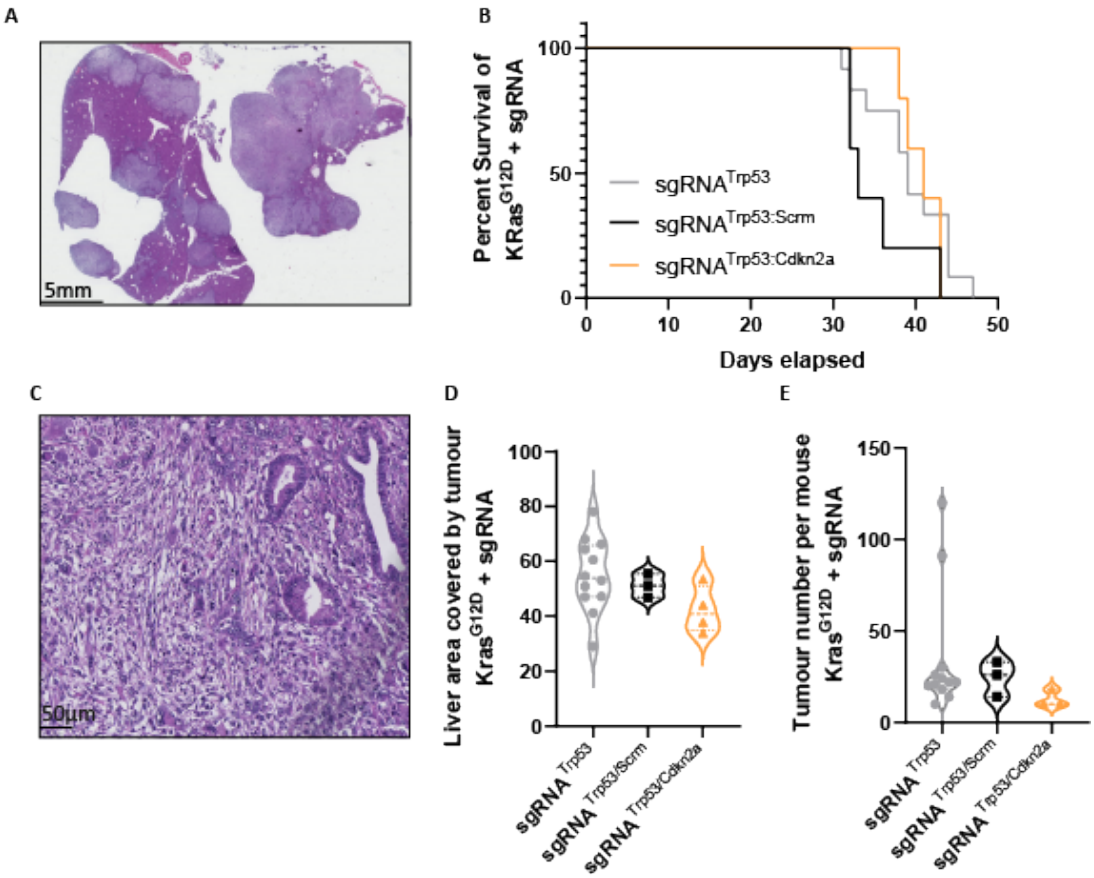

A
5mm
B
Percent Survival of KRas^G12D + sgRNA
100
50
0
sgRNA^Trp53
sgRNA^Trp53:Scrm
sgRNA^Trp53:Cdkn2a
0
10
20
30
40
50
Days elapsed
C
50µm
D
Liver area covered by tumour KRas^G12D + sgRNA
100
80
60
40
20
0
sgRNA Trp53
sgRNA Trp53/Scrm
sgRNA Trp53/Cdkn2a
E
Tumour number per mouse KRas^G12D + sgRNA
150
100
50
0
sgRNA Trp53
sgRNA Trp53/Scrm
sgRNA Trp53/Cdkn2a

Figure 9

Loss of Cdkn2a *and* Trp53 *interact with oncogenic* $KRAS^{G12D}$ *to form tumours* in vivo*. A. Macroscopic H&E stained image. Liver parenchyma is purple, and tumours are lighter in colour. B. Kaplan-Meier showing survival of* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Cdkn2a}$ *in orange (n=4),* $KRAS^{G12D}$*:*$sgRNA^{Trp53}$ *in grey (n=12), and* $KRAS^{G12D}$*:*$sgRNA^{Trp5/Scrm}$ *in black (n=3). All these animals were taken at end point. C. H&E detail illustrating iCCA tumour. D. Tumour burden presented as percentage of total liver area that is covered by tumour. E. Tumour number per animal.*

4.7.1 *Plxb2* cooperates with, but does not substitute for, $KRAS^{G12D}$

Plxb2 was chosen to be examined by single guide HDTVI because of its prevalence in the exome sequencing of the CRISPR-*sp*Cas9 library where it was mutated in two samples, one of which was a metastasis to the pancreas, and because it has not been associated before with iCCA. I was interested in identifying novel driver genes of iCCA, as well as re-identifying canonical drivers such as *Brca2*. As shown in Figure 10 below, loss of *Plxb2* in combination with expression of oncogenic $KRAS^{G12D}$ and loss of *Trp53* results in the formation of many tumours that are representative of iCCA. The survival of animals harbouring loss of *Plxb2* and *Trp53* was significantly reduced compared to the loss of *Trp53* alone, as was the tumour burden. The tumour number was greater, however. This is interesting as the survival is very similar to that of $KRAS^{G12D}$:$sgRNA^{Trp53}$, and overall tumour burden is lower, *Plxb2* loss does not accelerate tumourigenesis but the presence of many more, smaller, tumours show it is more successful at initiating tumour formation, if not progressing it. The tumours were GFP positive, indicating that the tumour cells are expressing KRAS.

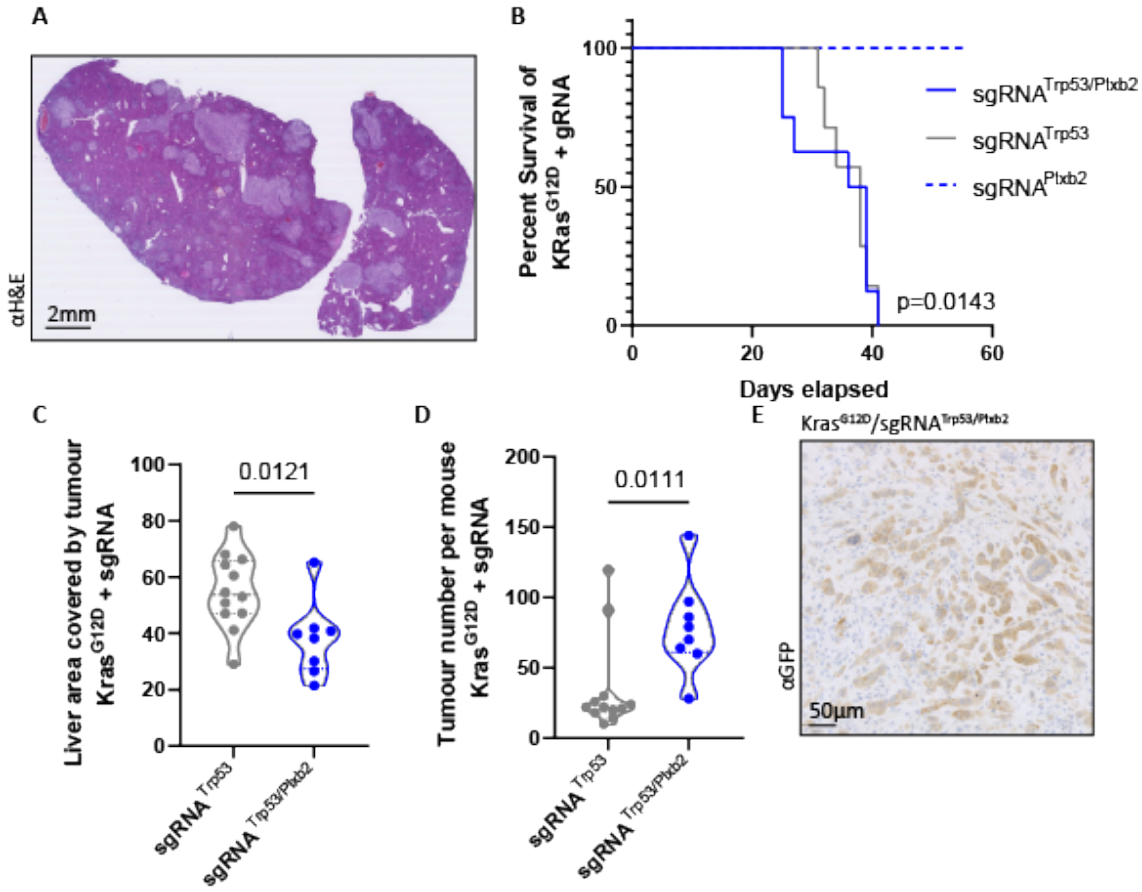

A
αH&E
2mm
B
Percent Survival of KRas^G12D + gRNA
100
50
0
0
20
40
60
Days elapsed
sgRNA^Trp53/Plxb2
sgRNA^Trp53
sgRNA^Plxb2
p=0.0143
C
Liver area covered by tumour Kras^G12D + sgRNA
100
80
60
40
20
0
0.0121
sgRNA Trp53
sgRNA Trp53/Plxb2
D
Tumour number per mouse Kras^G12D + sgRNA
200
150
100
50
0
0.0111
sgRNA Trp53
sgRNA Trp53/Plxb2
E
Kras^G12D/sgRNA^Trp53/Plxb2
αGFP
50μm

Figure 10

Plxb2 interacts with KRAS^{G12D} to form tumours in vivo. *A. Macroscopic H&E stained image. Liver parenchyma is purple, and tumours are lighter in colour. B. Kaplan-Meier showing survival of KRAS^{G12D}:sgRNA$^{Trp53/Plxb2}$ in blue and KRAS^{G12D}:sgRNA^{Trp53} in grey, all these animals were taken at end point (p value shown between these two data), and KRAS^{G12D}:sgRNA^{Plxb2} in dashed blue lines, taken at 58 days. C. Tumour burden presented as percentage of total liver area that is covered by tumour. D. Tumour number per animal. E. GFP stain indicating Kras positive cells in a tumour portion of KRAS^{G12D}:sgRNA$^{Trp53/Plxb2}$.*

4.7.2 *Nf2* loss accelerates tumourigenesis

Led again by the CRISPR-*sp*Cas9 screen data, the exome sequencing revealed that *Nf2* was the second most mutated gene in the screen after *Trp53*. Additionally, the RNA sequencing showed that tumours harbouring *Nf2* loss clustered closely on PCA, indicating their transcriptional similarity. Given the significance of *Nf2* as a driver gene of iCCA, as with the other single guide RNAs above, I used HDTVI to introduce sgRNAs that caused deletion of *Nf2*, and these were co-injected with $KRAS^{G12D}$, and another group was injected sgRNAs to cause the additional deletion of *Trp53*.

The Kaplan-Meier survival curve in Figure 11A illustrates that the deletion of *Trp53* and *Nf2* alone with no oncogenic $KRAS^{G12D}$ yielded no tumours. However, *Nf2* loss in combination with oncogenic $KRAS^{G12D}$ did form large discrete tumours. This shows that Nf2 is capable of overcoming senescence that $KRAS^{G12D}$ induces in the liver. However, the combined $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ mice formed highly aggressive and invasive cancer significantly sooner than the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model mice, with a median survival of 14 days compared to 39 days of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model mice. The $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ livers contained tumours that covered a smaller area of the liver overall (Figure 10B), however there were a significantly larger number of tumours (Figure 10C). These data suggest that mutations in *Trp53* and *Nf2* suppressor genes may synergize, and *Nf2* and *Trp53* are not functionally redundant because together they increase the retention of $KRAS^{G12D}$ cells in the liver, and subsequently allow them to be retained and more cancers to form. The accelerated mortality the $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ mice is possibly a result of the large number of tumours forming, rather than due to their size.

I examined the tumours with *Nf2* deletion histologically and those containing *Nf2* deletion and $KRAS^{G12D}$ expression were solid and representative of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model tumours (Figure 11D-F). However, the co-deleted $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ tumours were particularly diffuse and invasive.

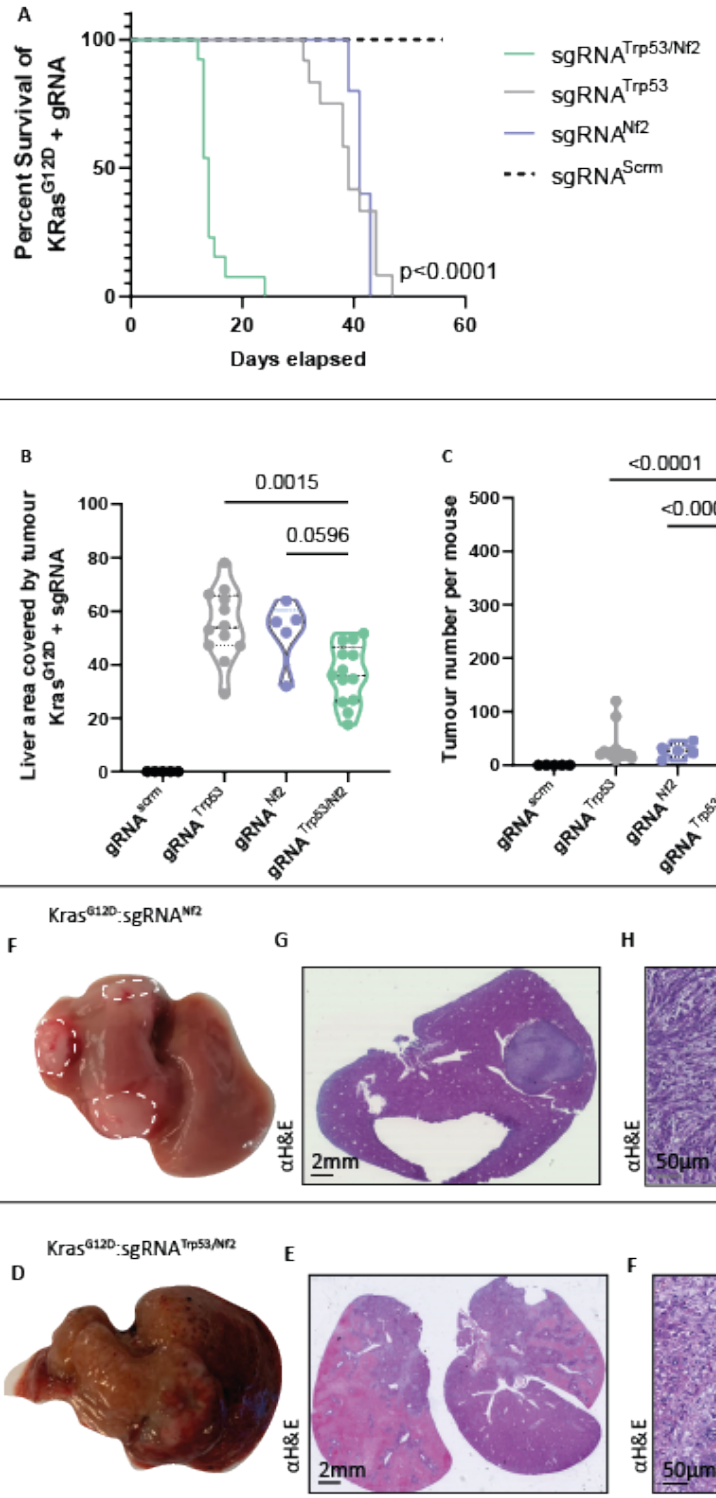

A
Percent Survival of KRas^{G12D} + gRNA
100
50
0
0
20
40
60
Days elapsed
p<0.0001
sgRNA$^{Trp53/Nf2}$
sgRNA^{Trp53}
sgRNA^{Nf2}
sgRNA^{Scrm}
B
Liver area covered by tumour KRas^{G12D} + sgRNA
100
80
60
40
20
0
0.0015
0.0596
gRNA ^{scrm}
gRNA ^{Trp53}
gRNA ^{Nf2}
gRNA $^{Trp53/Nf2}$
C
Tumour number per mouse
500
400
300
200
100
0
<0.0001
<0.0001
gRNA ^{scrm}
gRNA ^{Trp53}
gRNA ^{Nf2}
gRNA $^{Trp53/Nf2}$
Kras^{G12D}:sgRNA^{Nf2}
F
G
H
αH&E
2mm
αH&E
50µm
Kras^{G12D}:sgRNA$^{Trp53/Nf2}$
D
E
F
αH&E
2mm
αH&E
50µm

Figure 11

Nf2 *cooperates with oncogenic KRAS^{G12D} to form tumours* in vivo*. A. Kaplan-Meier indicates survival of KRAS^{G12D} expressing animals with either loss of* Trp53 *(n=12),* Nf2 *(n=5), both* Trp53 *and* Nf2 *(n=13) or a Scrambled (Scrm) non-targeting sgRNA (n=5). B. Macroscopic image of KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ liver, C. and D. H&E stain of this liver with D. showing tumour cells. E. Macroscopic KRAS^{G12D}:sgRNA^{Nf2} liver, F. H&E stain of this liver with G. showing tumour cells.*

The $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ Model Recapitulates a Rare form of iCCA

Histopathologically, the *Nf2* loss model expresses changes in vasculature and proliferative ability that were examined by staining. These data established that this model is closely representative of a rare sarcomatoid type of human disease, spindle cell cholangiocarcinoma. Figure 12 illustrates this in more detail where the macroscopic image of an H&E stained liver is shown in A, and details of the sarcomatoid nature are highlighted as "S" in Figure 12B-D. The tumour cells in this model were highly invasive and CK19 and GFP staining confirmed the presence of these cells migrating throughout the liver, as illustrated in Figure 12E&F.

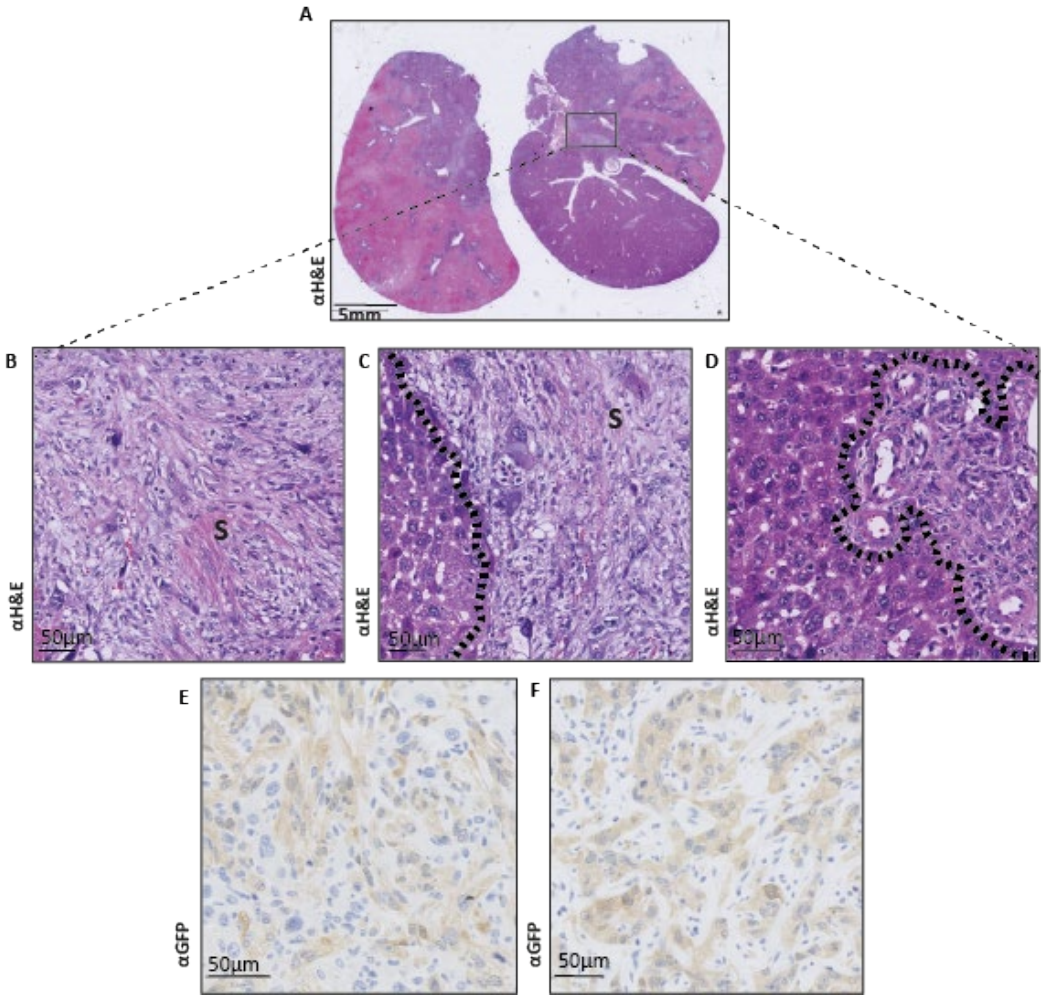

A
αH&E
5mm
B
αH&E
S
50µm
C
αH&E
S
50µm
D
αH&E
50µm
E
αGFP
50µm
F
αGFP
50µm

Figure 12

The KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ model recapitulates a rare form of iCCA, sarcomatoid iCCA. A. H&E staining of this model. A-D detail on tumour formation with sarcomatoid aspects denoted as "S". GFP staining confirming presence of KRAS positive cells in this model.

To establish whether the sarcomatoid phenotype is a result of early stage cancer forming events, a time course was used, from day 2, post-HDTVI, to day 14. Figure 13 shows the number of $KRAS^{G12D}$ positive cells, as indicated by their GFP-positivity, and there was no significant difference in the co-deletion livers and *Trp53* deletion. These data further support the earlier observation that the tumour burden of the $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ model at end point is not significantly different from the single deletion of *Trp53*.

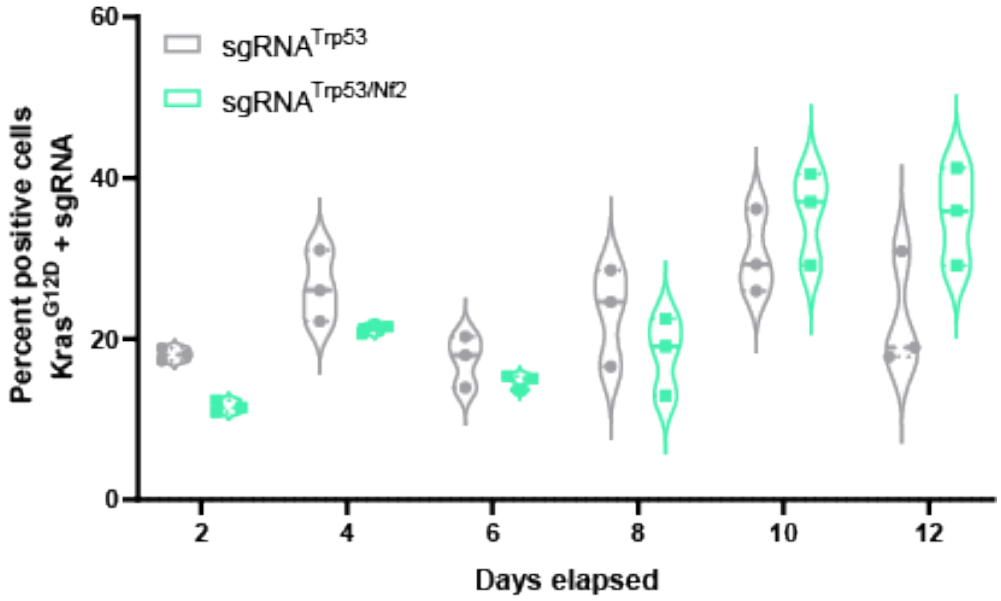

60
40
20
0
sgRNA^Trp53
sgRNA^Trp53/Nf2
Percent positive cells
Kras^G12D + sgRNA
2
4
6
8
10
12
Days elapsed

Figure 13

Early time course of KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ and KRAS^{G12D}:sgRNA^{Trp53} animals culled at number of days elapsed since the HDTVI of tumour-inducing plasmids. P=non-significant between any of the KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ and KRAS^{G12D}:sgRNA^{Trp53} animals.

4.8 Discussion

CRISPR-*sp*Cas9 library screening using a hydrodynamic tail vein injection model of murine intrahepatic cholangiocarcinoma identified nine genes of interest. I tested whether these genes were capable of driving iCCA tumourigenesis in combination with oncogenic $Kras^{G12D}$. Using functional genetics, I tested the outcome of my CRISPR-*sp*Cas9 library screen to identify loss of function mutations in driver genes of iCCA. I identified *Nf2* loss as a potent driver of tumourigenesis.

To address the possibility that there are a subset of rare driver genes of iCCA that are causative of metastasis, I would propose a screen design based again on patient data. Primary iCCAs from patients with confirmed or unconfirmed metastases would be isolated as single cells and sequenced to identify *de novo* mutations. Filtering of these mutations would allow rationalisation as to whether there are any metastatic drivers present. These drivers would then be collated into a CRISPR-spCas9 metastatic library containing bar codes allowing for ease of sequencing. The metastatic library would be injected into mice via HDTVI and their tissues harvested post euthanasia. The presence of lung metastases in the control model containing loss of *Trp53* and guides against a scrambled sgRNA poses the question as to whether this is a side effect of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model itself, or due to the presence of more sgRNAs. Given the CRISPR-*sp*Cas9 library screen tumours had metastasis to the pancreas in only 2 animals, is indicative that it is not due to the sheer volume of gRNAs injected that determines whether metastasis occurs.

Loss of *Rnf31*, although insufficient to cause tumour formation when deleted in addition to oncogenic $KRAS^{G12D}$ alone, did present with fewer and larger, (and

potentially less lethal) tumours when it was deleted with the additional deletion of *Trp53*. This was shown by one of the animals not reaching an end point and instead being euthanised at 58 days. Although there is a minor chance this may be an artefact of the HDTVI (the possibility that this particular injection was not delivered in <7-10 seconds, for example, and thus affecting the permeability of hepatocyte uptake of the plasmids. However this was always recorded and taken into account when assessing animals later, and this animal did receive the correct dosage in <10 seconds), it is interesting as the loss of *Rnf31* in this model may have a protective role when $KRAS^{G12D}$ is expressed, and *Trp53* is lost. *Rnf31* is an atypical ubiquitin ligase that forms the linear ubiquitin chain assembly complex (LUBAC) in combination with the ubiquitin-associated proteins RBCK1 and SHARPIN and promotes linear ubiquitination of an adaptor of the IκB kinases and subsequent activation of NF-κB signalling. (229) It has also been shown that conditional deletion of the ubiquitin-binding domain of RNF31 in B-cells leads to the impairment or lack of canonical NF-κB and ERK signalling. (230)

The role of *Rnf31* in activating these pathways is important when considering the changes in this model when it is deleted. The over-activation of KRAS, as occurs in my $KRAS^{G12D}$:$sgRNA^{Rnf31}$ model, leads to downstream activation of the MAPK/MEK/ERK signalling axis, and the over-expression of ERK in a mouse model of iCCA with KRAS activation and Trp53 loss has been previously established (231). Therefore the absence of *Rnf31* may be acting to reduce ERK signalling, which leads to the decrease in the tumourigenic nature of this KRAS-driven model. Another possibility is the absence of *Rnf31* is changing the immune cell landscape of this model. It has been shown that *Rnf31* loss exposes pancreatic ductal adenocarcinoma cells to be killed by CD8+ T-cells, as the deletion of *Rnf31* in mouse tumours caused infiltration of this cell type. (232)

Additionally this study also showed that *Rnf31* protects tumour cells from TNF-mediated caspase 8 cleavage and therefore from removal by apoptosis. These processes provide possible explanations as to how *Rnf31* loss may mitigate the tumourigenic nature of oncogenic KRAS^{G12D} and *Tpr53* deletion.

The absence of tumours in animals that were subject to HDTVI of KRAS^{G12D}:sgRNA$^{Trp53/Brca2}$ is fascinating, considering that the hypothesis considered that additional loss of function mutations would result in increased tumour burden, and the associated decrease in survival. Whereas in these data I found that the opposite occurred. Indeed, when *Brca2* was lost, only one animal presented with a single macroscopic tumour, and the rest resembled non-cancerous animals. Seeing as the other two permutations in this model, oncogenic KRAS^{G12D} expression and loss of *Trp53*, cause iCCA in mice, I consider that the loss of *Brca2* in addition to *Trp53* loss causes cells themselves to be non-viable, meaning that either the non-cancerous, non-mutated cells are able to out-compete the mutated cells, or the double loss of function mutant cells are destroyed by the immune system.

On the presence of metastatic lesions when *Ncor1* is lost on the background of the KRAS^{G12D}:sgRNA^{Trp53} model, I consider these data require further investigation. The possibility that the presence of at least some metastatic lesions are present as a result of the HDTVI model itself would need to be assessed in more detail. I propose this to be done using HDTVI of a scrambled non-targeting sgRNA on the background of oncogenic KRAS^{G12D}. Examination and staining of additional tissue to the liver following this HDTVI would give insight into whether it is the sgRNA regardless of its CRISPR-*sp*Cas9-inducing ability, or the presence of oncogenic KRAS^{G12D} that is causative of this phenotype. Additionally, an experiment involving injection of the KRAS^{G12D} plasmid

alone, and subsequent lung analysis, will show whether these KRAS positive cells are reaching the lung, be that via the HDTVI itself or movement of these cells from their primary site in the liver before they are cleared by senescence surveillance.

I chose to examine further the severely accelerated phenotype identified in $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals.

Chapter 5 Results

Characterisation of *Nf2* loss model

5.1 Introduction

The CRISPR-*sp*Cas9 library screen (described in chapter 3) lead to the identification of *Nf2* as a gene of interest in that its deletion is potentially tumourigenic, so I wished to explore this further by examining the detailed mechanisms by which the loss of *Nf2* leads to iCCA formation. To do this, I examined the signalling pathway changes that occur when I delete *Nf2* alone, or with the additional deletion of *Trp53* with overexpression of oncogenic $KRAS^{G12D}$. Having defined the changes in signalling pathway components following *Nf2*-loss, I then further examined the function of these signals by targeting deregulated signalling pathways with small molecule therapeutics. The genes and pathways relevant to this chapter will be introduced below and then I will go on to describe the signals are required for aggressive cancer growth following Nf2-loss.

5.1.1 NF2 as a signalling molecule

Nf2, and its protein product, NF2, are well known to be involved in the Hippo signalling pathway where it acts to negatively regulate pathway activation via Mst1/2 phosphorylation. NF2 also sits at the nexus of many other signalling pathways including PI3K and Wnt/ β-catenin signalling. NF2 has been established as a regulator of Wnt/ β-catenin signalling in several studies. It does this by prevention of LRP6 phosphorylation, which is a process required for the transduction of Wnt signalling. (228) In further support of the importance of NF2 in the Wnt/ β-catenin signalling pathway, aberrant β-catenin activity was found in breast cancer cells lacking NF2, and found that β-catenin and NF2 interact directly. (233) NF2 has also been shown to

inhibit phosphatidylinositol 3-kinase (PI3K), a key kinase in the PI3K/AKT signalling pathway. (234)

5.1.2 Interactions of Hippo and Wnt/ β-catenin signalling

The Hippo pathway (as described in section 1.9.5), is a regulator of Wnt/ β-catenin signalling, which is perhaps not surprising considering they regulate similar biological processes. (235) These pathways interact in several places, that I will cover here, and are illustrated in Figure 5.1.2 The main effector proteins in Hippo signalling, YAP/TAZ, physically interact with both β-catenin and Dishevelled protein (DVL) to regulate phosphorylation state and subsequently control the levels of β-catenin. (236) NF2 also controls the phosphorylation state of LRP5/6, the transmembrane receptor that is interacting with Frizzled to transfer Wnt-ligand signals across the extracellular matrix. TAZ knockdown significantly increased DVL phosphorylation, and subsequently increased nuclear levels of β-catenin. It has also been shown that deletion of MST1/2 led to β-catenin activation in HCC. (235) Additionally, there has been a link established that TAZ acts as a downstream component of Wnt signalling: in the absence of Wnt signalling activation, the β-catenin destruction complex also causes the degradation of TAZ. This is due to the behaviour of the E3-ubiquitin ligase βTrCP where, when lost, βTrCP results in the increased activation of TAZ. This occurs because β-catenin and GSK3β activity are required for the interaction of TAZ and βTrCP. (237)

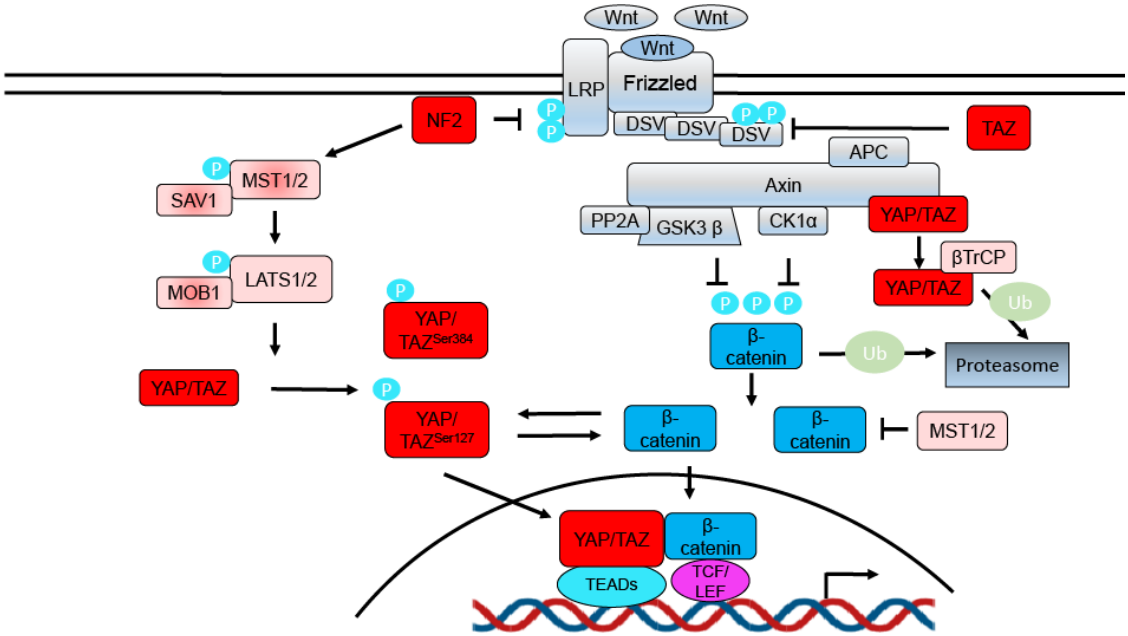

Wnt
Wnt
Wnt
Frizzled
LRP
DSV
DSV
DSV
NF2
TAZ
APC
Axin
PP2A
GSK3 β
CK1α
YAP/TAZ
βTrCP
YAP/TAZ
Ub
Proteasome
β-catenin
Ub
MST1/2
SAV1
MOB1
LATS1/2
YAP/TAZ
YAP/TAZSer384
YAP/TAZSer127
β-catenin
β-catenin
MST1/2
YAP/TAZ
β-catenin
TEADs
TCF/LEF

Figure 5.1.2

Hippo and Wnt signalling interact in several ways. Hippo signalling components are shown in red, Wnt signalling components in blue and the proteasome in grey. Phosphorylation events are turquoise "P". Ubiquitination events are green "Ub". Cytoplasmic YAP/TAZ plays a role in the β-catenin destruction complex, the E3-ubiquitin ligase βTrCP causes the degradation of YAP/TAZ and also β-catenin. YAP and β-catenin can interact themselves, and this results in the cooperation of these proteins and subsequent gene transcription.

5.2 Hypothesis

Nf2 loss cooperating with oncogenic $KRAS^{G12D}$ expression and *Trp53* deletion will result in an identifiable, *Nf2*-dependant signalling pathway signature that can then be targeted using therapeutic intervention. Additionally, the mechanism by which deletion of *Nf2* causes such prolific tumourigenesis can be identified using molecular methods.

5.3 Aims

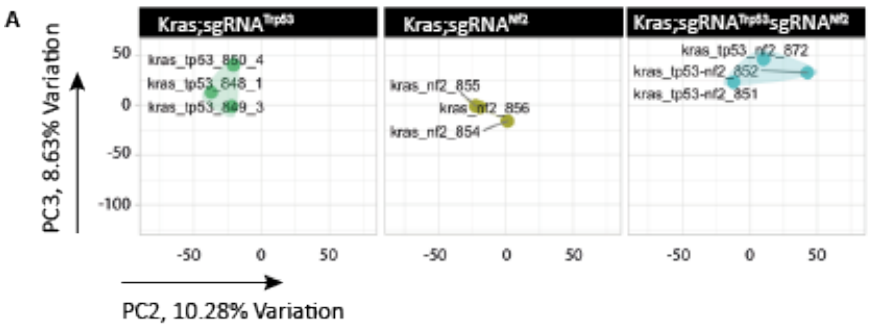
- Establish how- via which molecular signalling pathways- the loss of *Nf2* cooperates with oncogenic $KRAS^{G12D}$ expression and *Trp53* deletion to accelerate iCCA tumourigenesis.
- Identify changes in signalling pathways in this model with the intention of targeting these to reduce tumour formation with select therapeutic agents.

5.4 *Nf2* loss causes formation of aggressive tumours expressing a Wnt/AKT signature

The sarcomatoid KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ tumours (Chapter 4, Figure 11) were examined using bulk RNA sequencing to define the changes in transcripts compared to when either *Trp53* or *Nf2* alone are lost. Based on principal component analysis (PCA), tumours harbouring loss of both *Trp53* and *Nf2* on the background of Kras^{G12D} were transcriptionally distinct from cancers with either of these mutations alone, as observed in Figure 1A. Protein annotation through evolutionary relationship (PANTHER) analysis was used to examine the up and down-regulated genes in the KRAS^{G12D}:sgRNA^{Trp53} or KRAS^{G12D}:sgRNA^{Nf2} vs the KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ groups. PANTHER is a tool that uses different input analyses including gene function, ontology pathway analysis and statistics to examine bulk RNA-seq. (238)

The lack of protocadherins resulting from *Nf2* loss was examined in more detail by re-examination of the CRISPR-*sp*Cas9 library screen results (Chapter 3). Upon closer inspection of the differential expression of transcribed genes in the library tumours harbouring *Nf2* deletion compared to the KRAS^{G12D}:sgRNA^{Trp53} model, I found that there were downregulation of some cadherins and protocadherins. These include clustered protocadherins *Pcdhb7*, *Pcdhb4*, *Pcdhb16* and *Pcdhb19*, the cadherins *Cdh6* and *Cdh17*, and the atypical cadherin *Fat3*. These data further supported that the loss of protocadherins could result in phenocopy of the KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ model. Given that between KRAS^{G12D}:sgRNA^{Trp53} and KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ tumours, both have altered expression of atypical cadherins I hypothesised that this was responsible for the tumourigenic nature of *Nf2* deletion in these models. Protocadherins are calcium-dependent signalling molecules that comprise the largest group of the

cadherin family of cell to cell adhesion proteins. The relevance of their being clustered or non-clustered is dependent on their genomic locus. The clustered protocadherins are present on a single genomic locus and this organisation is highly conserved, compared to the non-clustered protocadherins that are found throughout the genome. (239)


A
Kras;sgRNA^Trp53
Kras;sgRNA^Nf2
Kras;sgRNA^Trp53sgRNA^Nf2
kras_tp53_850_4
kras_tp53_848_1
kras_tp53_849_3
kras_nf2_855
kras_nf2_856
kras_nf2_854
kras_tp53_nf2_872
kras_tp53-nf2_852
kras_tp53-nf2_851
PC3, 8.63% Variation
PC2, 10.28% Variation

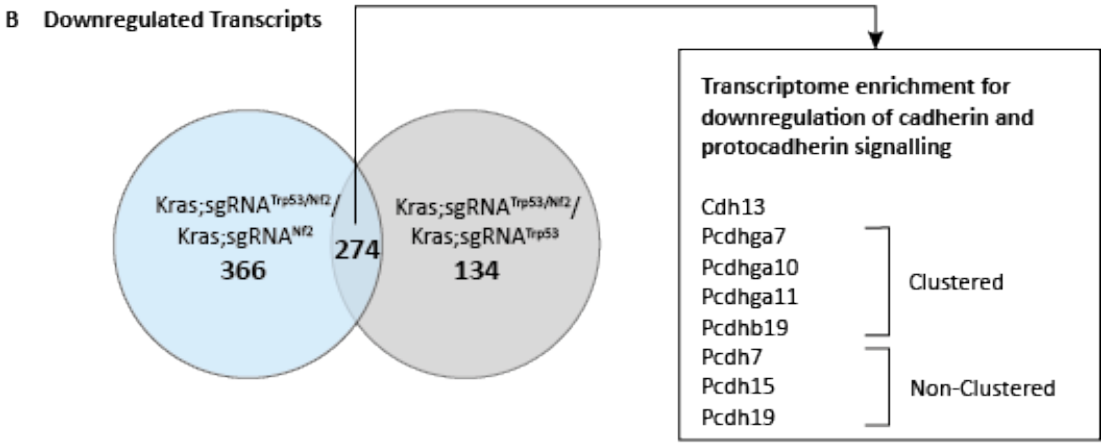

B Downregulated Transcripts
Kras;sgRNA^Trp53/Nf2 / Kras;sgRNA^Nf2
366
274
Kras;sgRNA^Trp53/Nf2 / Kras;sgRNA^Trp53
134
Transcriptome enrichment for downregulation of cadherin and protocadherin signalling
Cdh13
Pcdhga7
Pcdhga10
Pcdhga11
Pcdhb19
Clustered
Pcdh7
Pcdh15
Pcdh19
Non-Clustered

Figure 1

RNA-sequencing confirmed the transcriptional differentiation between KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ mice vs both KRAS^{G12D}:sgRNA^{Trp53} and KRAS^{G12D}:sgRNA^{Nf2}. A. PCA illustrates the close clustering of animals belonging to the different model groups. B. Transcripts that were downregulated in KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ compared to both KRAS^{G12D}:sgRNA^{Trp53} and KRAS^{G12D}:sgRNA^{Nf2}. Numbers represent downregulated transcripts. Transcriptome enrichment for cadherin 13 and several protocadherins were identified, these are listed in B. The status of gene clustering or non-clustering are also indicated.

To examine whether the protocadherins were themselves capable of causing iCCA *in vivo*, that was similar histologically and pathologically to my sarcomatoid iCCA KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ model, I carried out an experiment involving different groups of protocadherins. The protocadherins were split into three mini libraries: clustered, non-clustered and total protocadherins, as listed in Figure 1B, and these injected via HDTVI into 5 week old FVB female mice along with SB13 and my plasmid expressing Kras^{G12D}, as shown by the schematic in Figure 2A. The Kaplan-Meier curve in Figure 2B shows the survival of these animals that were euthanised at end point, or at 58 days post injection. The clustered and non-clustered groups of animals yielded 3/5 and 2/5 presenting with iCCA tumours, respectively, whereas the total protocadherins were entirely clear of any cancer. To establish that these tumours were of Kras origin and to determine whether any of the livers that were negative for tumours weren't harbouring any Ras positive cells, I stained the tissue for GFP. The clustered protocadherin library yielded KRAS negative bile ducts and KRAS positive tumours representing iCCA, and dense KRAS positive tumour stroma as shown in the panels in Figure 2C. The non-clustered protocadherin library yielded Kras negative ducts in the animals whose livers were not cancerous, and iCCA positive for KRAS in the tumourous animals, Figure 2D illustrates this. Finally, HDTVI of the total protocadherin library resulted in the formation of no tumours in any of these animals, as shown by the lack of tumour and any KRAS positive cells in Figure 2D. The tumours formed in these experiments did not yield tumours that histologically represent sarcomatoid iCCA as observed in my KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ model, rather they closely resemble the solid, discrete tumours observed in the KRAS^{G12D}:sgRNA^{Trp53} mice. These data indicate that the disruption of protocadherins are not directly responsible for the sarcomatoid tumours in the KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ model, as the deletion of various combinations

of protocadherins do not phenocopy *Nf2* and *Trp53* loss. However, it is indicative that individual or some combination of clustered and non-clustered protocadherins do substitute for *Trp53* loss.

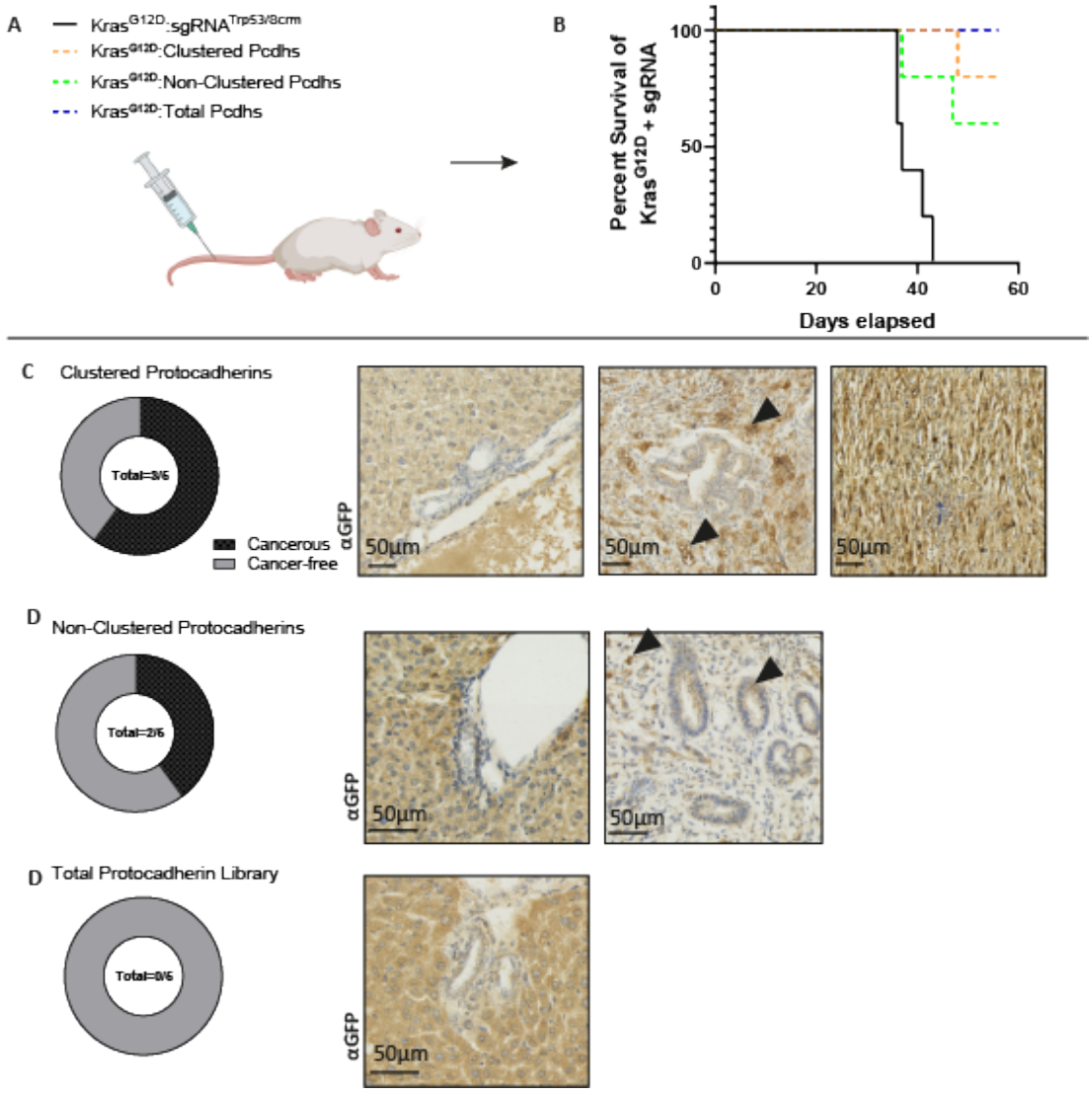

A
KrasG12D:sgRNATrp53/8crm
KrasG12D:Clustered Pcdhs
KrasG12D:Non-Clustered Pcdhs
KrasG12D:Total Pcdhs
B
Percent Survival of KrasG12D + sgRNA
100
50
0
0
20
40
60
Days elapsed
C
Clustered Protocadherins
Total=3/6
Cancerous
Cancer-free
αGFP
50μm
50μm
50μm
D
Non-Clustered Protocadherins
Total=2/6
αGFP
50μm
50μm
D
Total Protocadherin Library
Total=0/6
αGFP
50μm

Figure 2

Loss of the protocadherins did not result in a phenocopy of KRAS^{G12D} expression and deletion of Trp53 *and* Nf2*. A. Three mini libraries of clustered (Pcdhga7, Pcdhga10, Pcdhga11 and Pcdhb19), non-clustered (Pcdh7, Pcdh15 and Pcdh19) and total protocadherins (Pcdhs) were delivered to mice via HDTVI. B. Kaplan-Meier curve of survival where the animals were euthanised at end point, or at 58 days. C. Of the KRAS^{G12D}:Clustered Pcdhs, 3/5 animals presented with iCCA tumours. GFP staining to identify Ras positive cells from left to right illustrates a normal bile duct negative of Ras positive cells, Ras positive cancer representing iCCA, and dense Ras-positive stroma. D. Showing non-clustered Pcdhs where 2/5 animals presented with iCCA-like tumours and GFP staining illustrating a normal bile duct and iCCA tumour. E. HDTVI with KRAS^{G12D}:Total Pcdhs resulted in no tumour formation in any animals, as indicated by the GFP stain of normal bile duct to show absence of Ras positive cells. Arrows indicate Kras positive cells.*

In addition to the downregulation of these protocadherin transcripts, I also observed changes in components associated with Wnt signalling. These included upregulation of *Wnt9b*, and suppression of Wnt signalling inhibitors *Nkd2* and *Sfrp2*. These data suggest that changes in Wnt ligand levels and changes in the expression of negative Wnt ligand regulators is responsible for the mediation of iCCA carcinogenesis.

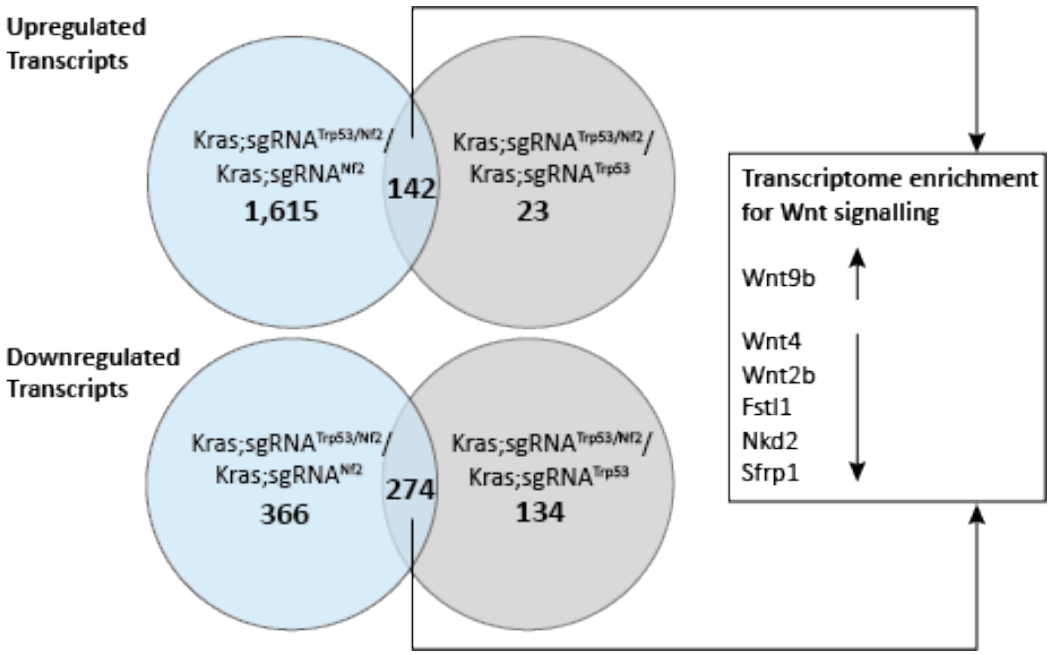

Upregulated Transcripts
Kras;sgRNA Trp53/Nf2 / Kras;sgRNA Nf2
1,615
142
Kras;sgRNA Trp53/Nf2 / Kras;sgRNA Trp53
23
Downregulated Transcripts
Kras;sgRNA Trp53/Nf2 / Kras;sgRNA Nf2
366
274
Kras;sgRNA Trp53/Nf2 / Kras;sgRNA Trp53
134
Transcriptome enrichment for Wnt signalling
Wnt9b
Wnt4
Wnt2b
Fstl1
Nkd2
Sfrp1

Figure 3

Transcripts of Wnt signalling components are up and down regulated in $Kras^{G12D}$*:*$sgRNA^{Nf2}$ *animals compared to either deletion of* Trp53 *or* Nf2 *alone. The upregulated transcript of Wnt9b is shown at the top. Numbers represent the number of changed transcripts. Downregulated transcripts are shown in the lower portion.*

5.5 Wnt/β-catenin and PI3K signalling are activated in the sarcomatoid, dual loss model of *Trp53* and *Nf2*

Considering there are altered transcripts in Wnt signalling, I screened this model for signalling pathway components that are changed. To do this I carried out a reverse phase protein array (RPPA). As illustrated in Figure 4, an RPPA is a multiplexed platform that allows the detection of 120 chosen molecular components of important and relevant signalling pathways (these are listed in Appendix 2). The RPPA allowed molecular signalling-relevant changes to be identified with the intention of identifying those that could be targeted therapeutically.

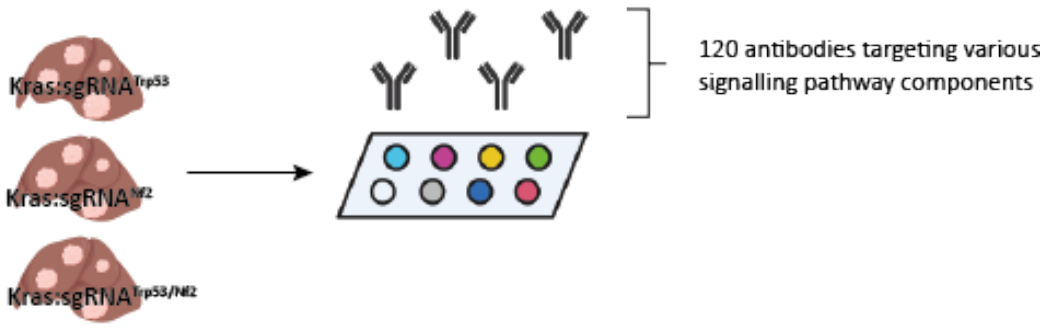

Kras:sgRNA^Trp53
Kras:sgRNA^Nf2
Kras:sgRNA^Trp53/Nf2
120 antibodies targeting various
signalling pathway components

Figure 4

Schematic of Reverse Phase Protein Array (RPPA) analysis carried out. Tumours harboured in the three models listed were excised and protein purified. The samples were then exposed to a collection of 120 antibodies targeting relevant and interesting signalling pathway components.

I examined the $KRAS^{G12D}$ expressing tumours that harboured mutations in either *Trp53*, *Nf2* or both of these genes. I found that there was an increased signature of Wnt and PI3K signalling when both *Nf2* and *Trp53* are lost in the presence of oncogenic $KRAS^{G12D}$. The changes in molecular components included increased inhibitory phosphorylation of GSK3β, which is involved in the control of β-catenin phosphorylation that targets its degradation, when *Nf2* was lost either alone or in addition to *Trp53*, as shown in Figure 5A. Additionally, I found activation of β-catenin signalling in the model containing loss of both *Trp53* and *Nf2* compared to either loss alone (Figure 5B). These data demonstrate that canonical Wnt signalling is altered to be highly activated in my model where Nf2 is lost in addition to Trp53. It also confirms that the transcriptional changes we observe translate into changes in signalling pathways. Importantly, elevated levels of phosphorylated AKT at Serine 473, as shown in Figure 5C, indicates that AKT/PI3K signalling is also elevated, and considering the cellular importance of this pathway . I confirmed the increase of active, non-phosphorylated β-catenin and active, phosphorylated-AKT via Western blot of protein lysates, as shown in Figure 5D and E.

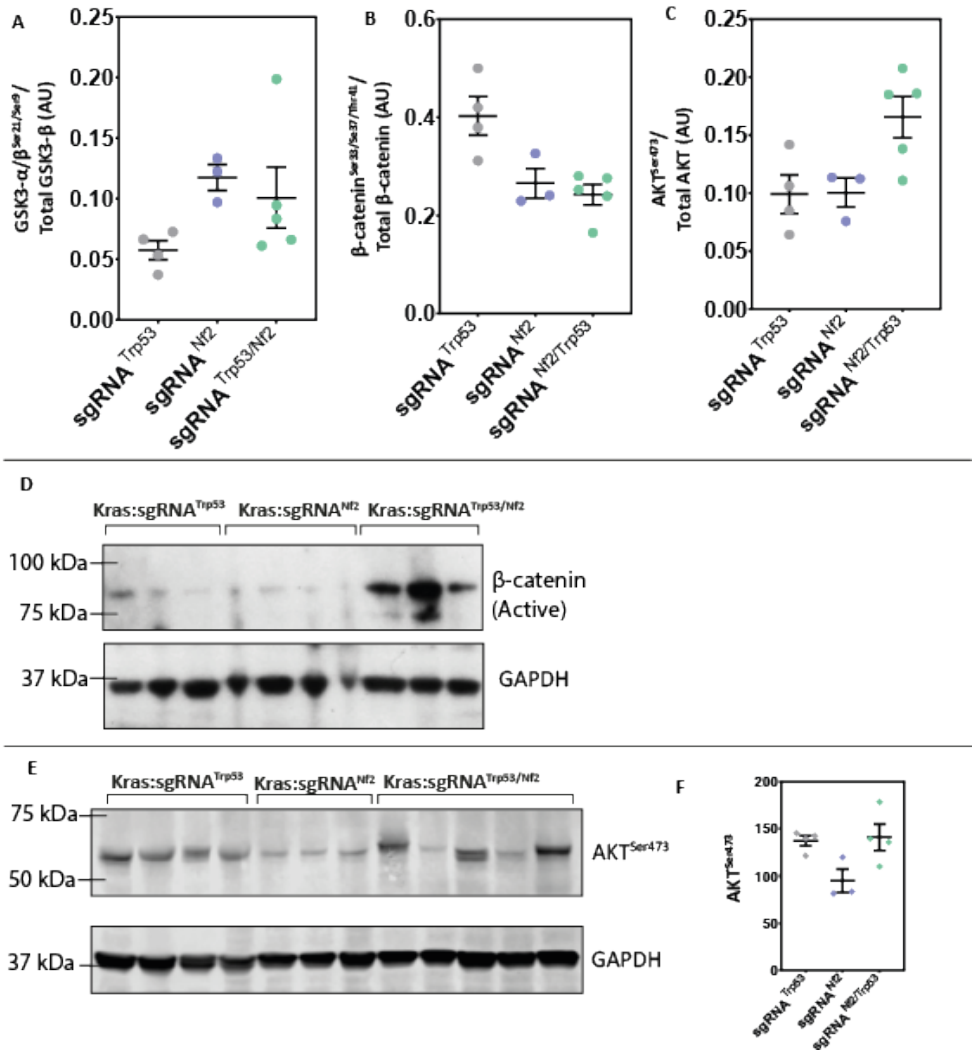

A
B
C
D
E
F
sgRNA Trp53
sgRNA Nf2
sgRNA Trp53/Nf2
sgRNA Nf2/Trp53
Kras:sgRNA Trp53
Kras:sgRNA Nf2
Kras:sgRNA Trp53/Nf2
100 kDa
75 kDa
37 kDa
50 kDa
β-catenin (Active)
GAPDH
AKT Ser473

Figure 5

Wnt and AKT active signalling pathway components are upregulated in KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals compared to either Kras^{G12D}:sgRNA^{Trp53} or KRAS^{G12D}:sgRNA^{Nf2} animals. A. Reverse Phase Protein Analysis (RPPA) shows increase in phosphorylated glycogen synthase kinase 3 α and β, B. increase in active (non-phosphorylated) β-catenin, C shows increase in P-AKT all in KRAS^{G12D}:sgRNA^{Nf2} animals compared to deletion of either gene alone in addition to KRAS^{G12D}. D. Western blotting for active, non-phosphorylated β-catenin and E. for phosphorylated AKT. F. RPPA analysis for phosphorylated AKT alone.

To examine the protein levels in my tumours histologically, I stained for active β-catenin, as shown in Figure 6, and although all the tumours express β-catenin, it is greatly increased in the animals lacking both *Trp53* and *Nf2*, compared to either gene deleted alone. The staining here expresses β-catenin as membranous, which is interesting considering active β-catenin would be nuclear.

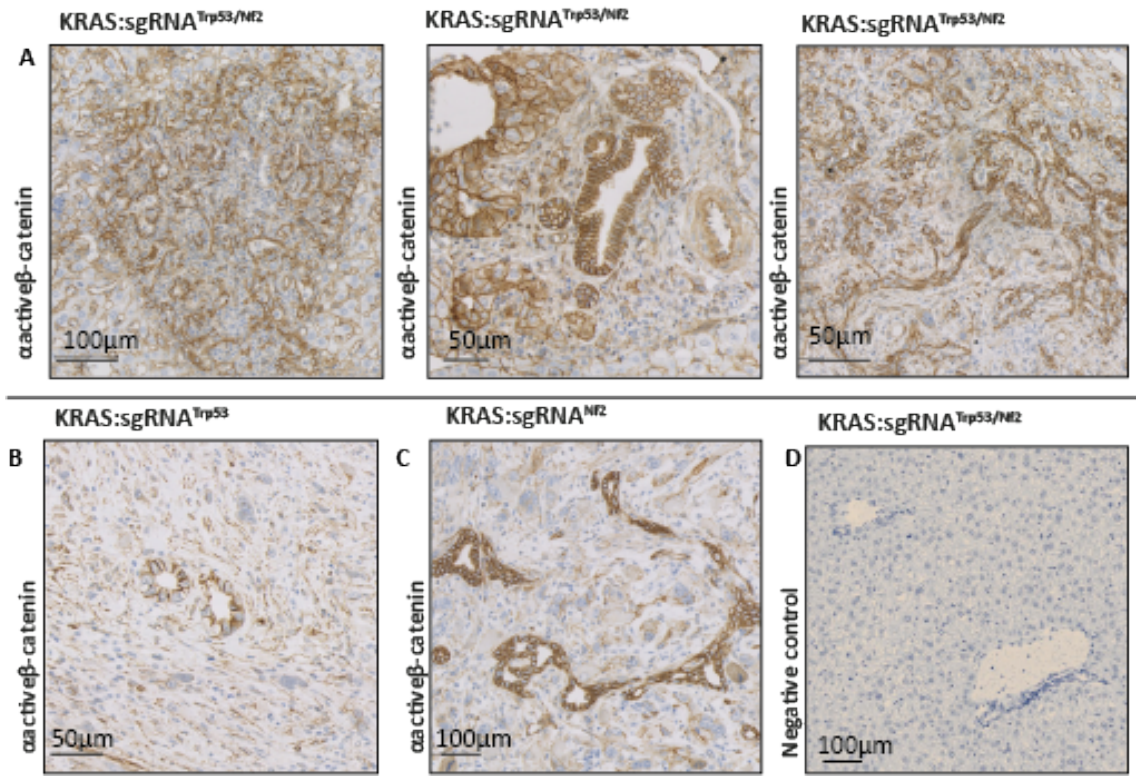

A
KRAS:sgRNA Trp53/Nf2
αactiveβ-catenin
100µm
KRAS:sgRNA Trp53/Nf2
αactiveβ-catenin
50µm
KRAS:sgRNA Trp53/Nf2
αactiveβ-catenin
50µm
B
KRAS:sgRNA Trp53
αactiveβ-catenin
50µm
C
KRAS:sgRNA Nf2
αactiveβ-catenin
100µm
D
KRAS:sgRNA Trp53/Nf2
Negative control
100µm

Figure 6

Active (non-phosphorylated) β-catenin is increased in KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals compared to those with loss of Nf2 or Trp53 alone. A. KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ tumours contain more positive staining for β-catenin than B. KRAS^{G12D}:sgRNA^{Trp53} or C. KRAS^{G12D}:sgRNA^{Nf2}. D. Negative, secondary only control for β-catenin.

I also stained for phosphorylated-AKT, as shown in Figure 7, and the increase in positive staining from $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals in Figure 7A, compared to deletion of either gene alone, either *Trp53* in Figure 7B, or *Nf2* in Figure 7C, can be clearly seen.

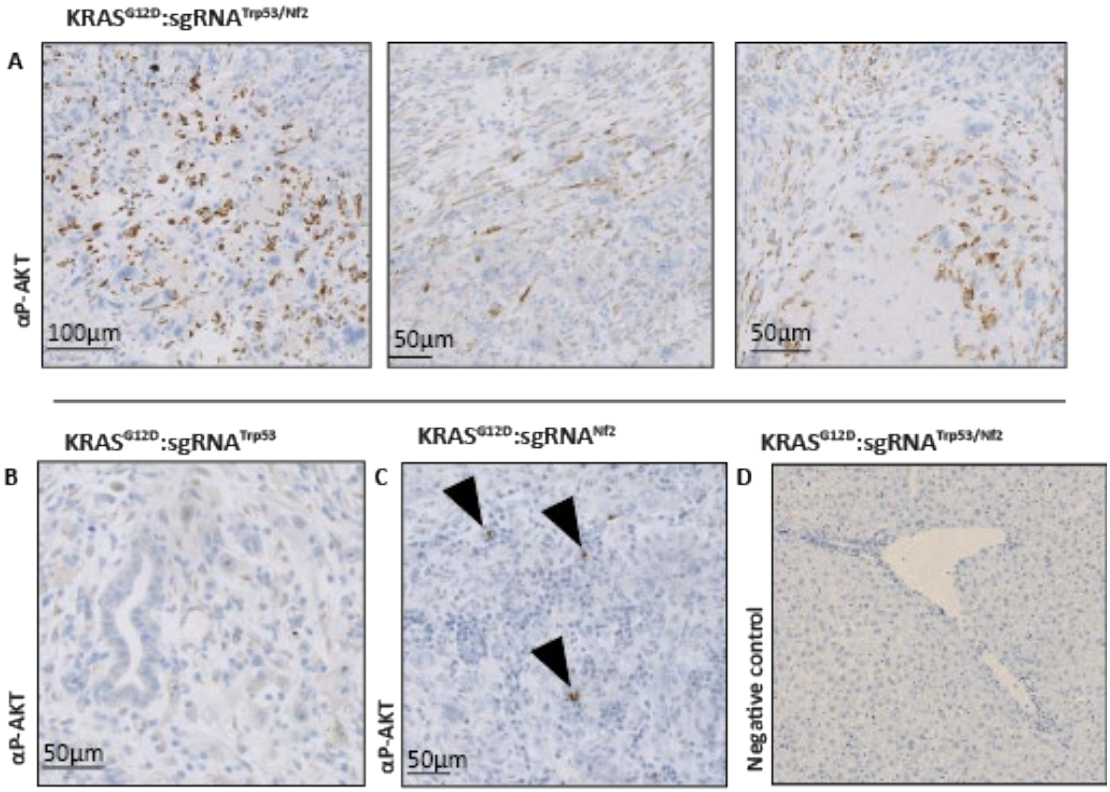

KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$
A
αP-AKT
100μm
50μm
50μm
KRAS^{G12D}:sgRNA^{Trp53}
B
αP-AKT
50μm
KRAS^{G12D}:sgRNA^{Nf2}
C
αP-AKT
50μm
KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$
D
Negative control

Figure 7

Phosphorylated AKT is increased in KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals compared to those with loss of Nf2 or Trp53 alone. A. KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ tumours contain more positive staining for P-AKT than B. KRAS^{G12D}:sgRNA^{Trp53} or C. Kras^{G12D}:sgRNA^{Nf2}. Arrows indicate positive cells. D. Negative control.

The combined results from my RNA-seq and RPPA, suggest that both the Wnt/β-catenin and PI3K/AKT pathways are hyper activated when there is dual loss of both *Trp53* and *Nf2* on a background of $KRAS^{G12D}$ compared to the loss of either gene alone.

5.6 Treatment with therapeutic inhibitors reduced tumourigenesis

To examine this relationship directly *in vivo*, mice baring mutations in both *Trp53* and *Nf2* on the background of $KRAS^{G12D}$ were treated with a Porcupine inhibitor LGK974 and a PI3K inhibitor, Pictilisib. LGK974 works by preventing the palmitoylation of Wnt ligands which leads of reduced Wnt ligand secretion and Pictilisib prevents the conversion of PIP2 into PIP3 meaning that AKT phosphorylation, and therefore activation, is decreased. Mice were induced to form sarcomatoid $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ iCCA tumours by HDTVI before they were randomised 7 days post-injection and then treated daily with LGK974, Pictilisib, a combination of both or a vehicle control, as shown in the schematic in Figure 8A. Animals were euthanised at end point, and their survival was significantly improved when treated with the either, or both therapeutic than with the vehicle control, as shown in Figure 8B. To confirm these cancers were iCCA of Ras origin I stained them histologically for GFP and treatment with each of LGK974, Pictilisib, combination or vehicle all presented with KRAS-positive iCCA, as can be observed in Figure 8C-F, respectively. There was no detectable toxicity in any of the treated animals, even those treated with the combination of both drug treatments.

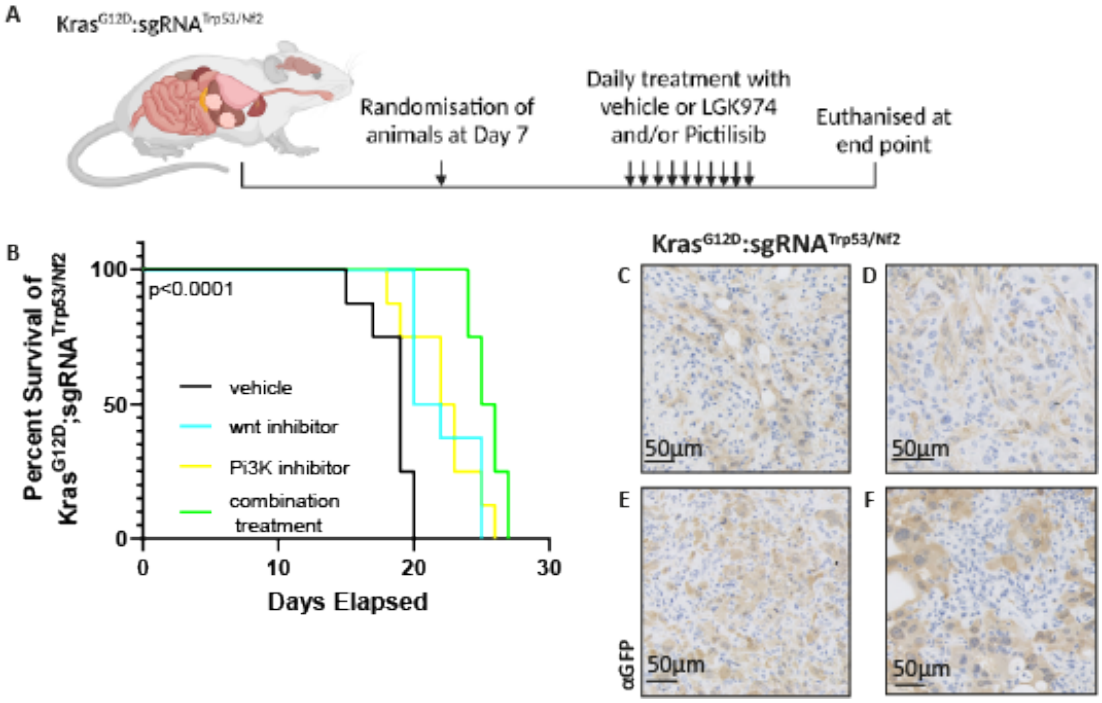

A
Kras^{G12D}:sgRNA$^{Trp53/Nf2}$
Randomisation of animals at Day 7
Daily treatment with vehicle or LGK974 and/or Pictilisib
Euthanised at end point
B
Percent Survival of Kras^{G12D};sgRNA$^{Trp53/Nf2}$
p<0.0001
vehicle
wnt inhibitor
Pi3K inhibitor
combination treatment
Days Elapsed
Kras^{G12D}:sgRNA$^{Trp53/Nf2}$
C
D
E
F
αGFP
50μm

Figure 8

Treatment of KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals with Wnt pathway inhibitor LGK974 or AKT pathway inhibitor Pictilisib. A. Schematic of treatment, KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals harbouring tumours were randomised at 7 days post HDTVI and then given daily oral treatment with either therapeutic, a combination of both, or a vehicle control. The animals were euthanised at end point and their survival is shown in the Kaplan-Meier in B. Positivity for GFP indicates Ras positive cells are present in the tumours from treatment with C. LGK974, D. Pictilisib, E. combination treatment with both therapeutics and F. vehicle control. P value between combination treatment and vehicle control is <0.0001.

The comparison of survival between each therapeutic, or the combination of both, to the vehicle control is statistically significant in all cases, as can be observed in Figure 9A, B and C. The combination treatment of both therapeutics is more effective at improving survival than either therapeutic alone (Figure 9D and 9E). Finally, the comparison of the survival of animals on each therapeutic to each other did not yield a significant difference (Figure 9F), indicating that the most successful treatment is indeed the combination of both Wnt/ β-catenin and PI3K/AKT signalling.

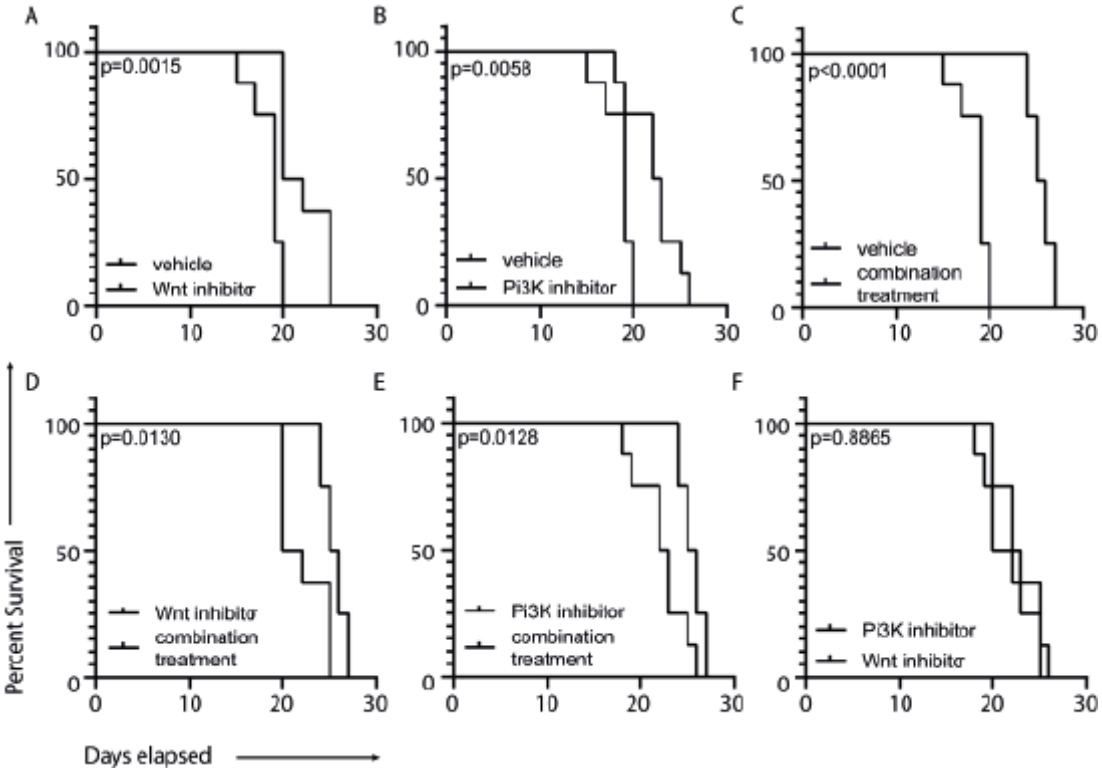

A
B
C
D
E
F
p=0.0015
p=0.0058
p<0.0001
p=0.0130
p=0.0128
p=0.8865
vehicle
Wnt inhibitor
Pi3K inhibitor
combination treatment
Percent Survival
Days elapsed
0
10
20
30
50
100

Figure 9

Isolated data from the treatment of KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals with Wnt pathway inhibitor LGK974 or AKT pathway inhibitor Pictilisib, a combination of both or a vehicle control.

When I examined the tumour burden in the therapeutically treated animals, there was a larger tumour burden in the combination treated animals, as these cancers had the time to spread further throughout the liver. Tumour burden was measured as a percentage of liver covered by tumour over the total liver area, using QuPath and measuring by hand. It is interesting that despite the greater tumour burden, these animals still survived longer, which is indicative of the therapeutic treatment reducing the lethality, rather than purely the volume of the tumours present.

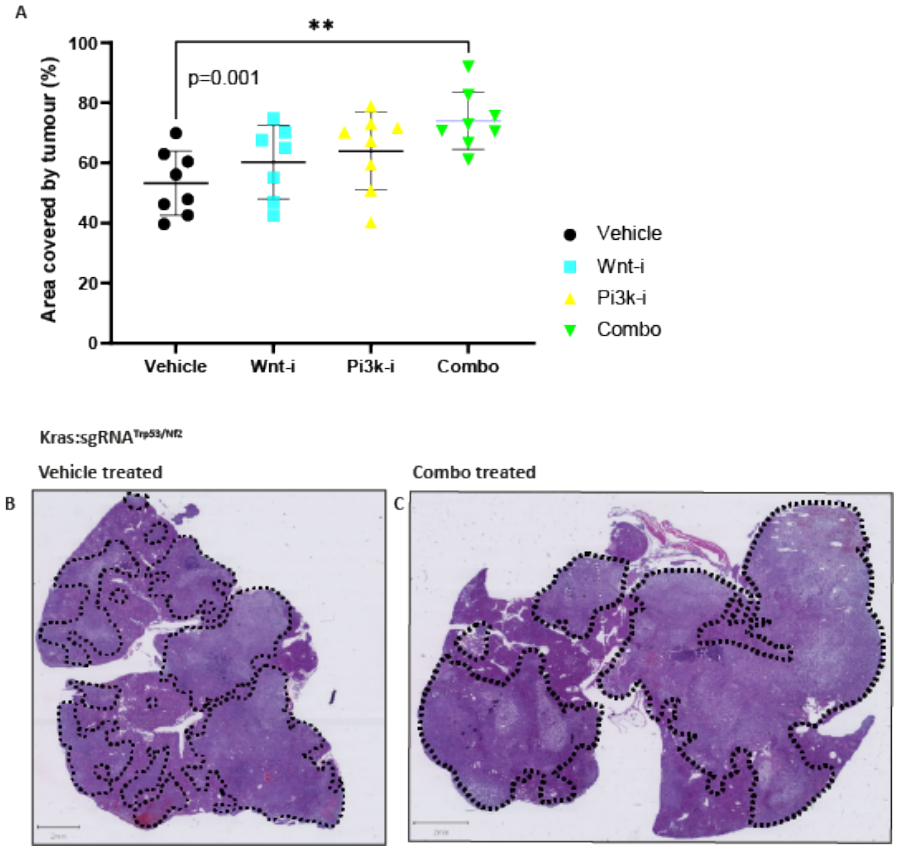

A
**
p=0.001
Area covered by tumour (%)
100
80
60
40
20
0
Vehicle
Wnt-i
Pi3k-i
Combo
Vehicle
Wnt-i
Pi3k-i
Combo
Kras:sgRNA Trp53/Nf2
Vehicle treated
Combo treated
B
C

Figure 10

A. Tumour burden presented as a proportion of the liver covered by tumour. B. H&E staining of $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals that were either treated with vehicle control in B. or with both Pictilisib and LGK974 in C. The tumours are outlined in black.

5.7 Apoptosis is increased when *Nf2* is deleted

These data show that increase in both Wnt/β-catenin and PI3K/AKT signalling stimulate the development of sarcomatoid iCCA, but the mechanism by which they do this was still unclear. Therefore I stained my tissue for the apoptosis marker cleaved caspase 3, and found that apoptosis is suppressed in $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals.

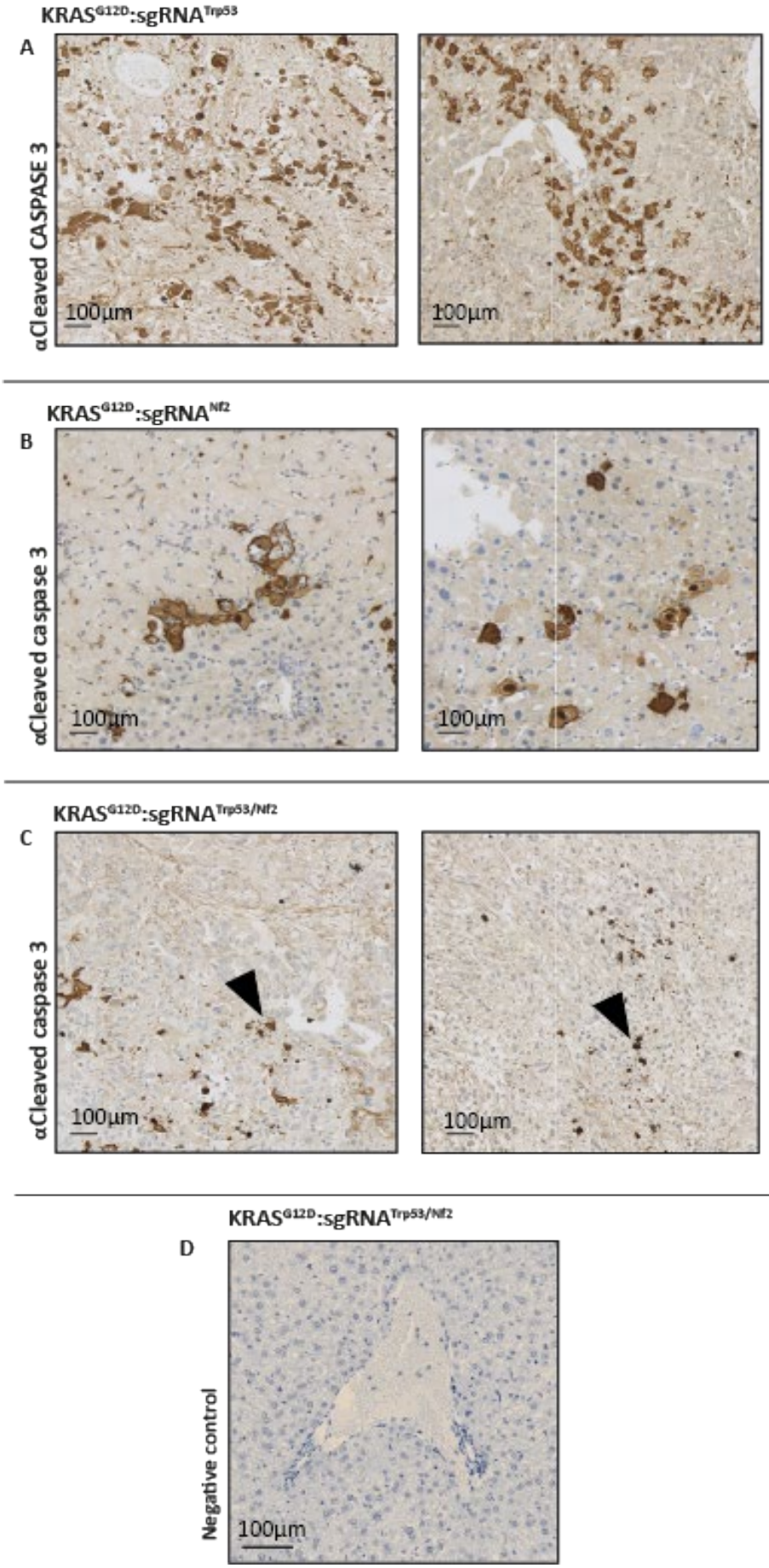

KRAS^{G12D}:sgRNA^{Trp53}
A
αCleaved CASPASE 3
100µm
100µm
KRAS^{G12D}:sgRNA^{Nf2}
B
αCleaved caspase 3
100µm
100µm
KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$
C
αCleaved caspase 3
100µm
100µm
KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$
D
Negative control
100µm

Figure 11

Immunohistochemical staining for cleaved caspase 3. A. $KRAS^{G12D}$:$sgRNA^{Trp53}$ animals express more cleaved caspase 3 than either B. $KRAS^{G12D}$:$sgRNA^{Nf2}$ or C. $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals. D. Negative control.

To further examine the suppression of apoptosis in this model, I looked at other markers for apoptosis in my RPPA data. The apoptotic transcription factor FOXO3A, as observed in Figure 12A, is decreased in animals with the deletion of *Nf2*, as is the pro-apoptotic enzyme CASPASE-3 (Figure 12B) (which is also supportive of the immunohistochemical staining for CASPASE3 above). The inhibitor of apoptosis BCL-2 is increased (Figure 12C) when there is *Nf2* loss, also supporting the decrease of apoptosis observed when *Nf2* is deleted. The decrease of phosphorylated BAD at Ser112 observed here (Figure 12D) is interesting as phosphorylation at this site and another, Ser136, leads to the dissociation of BAD from heterodimerisation with BCL-2 family members and BCL-X_L, which is a pro-apoptotic association. (240) However levels of inactive, anti-apoptotic phosphorylated BAD are decreased when Nf2 is lost in my model. Likewise, as observed in Figure 12E, phosphorylated BIM is also decreased when *Nf2* is lost, but as a pro-apoptotic member of the Bcl-2 family this is also in line with my *Nf2* loss models having decreased apoptosis.

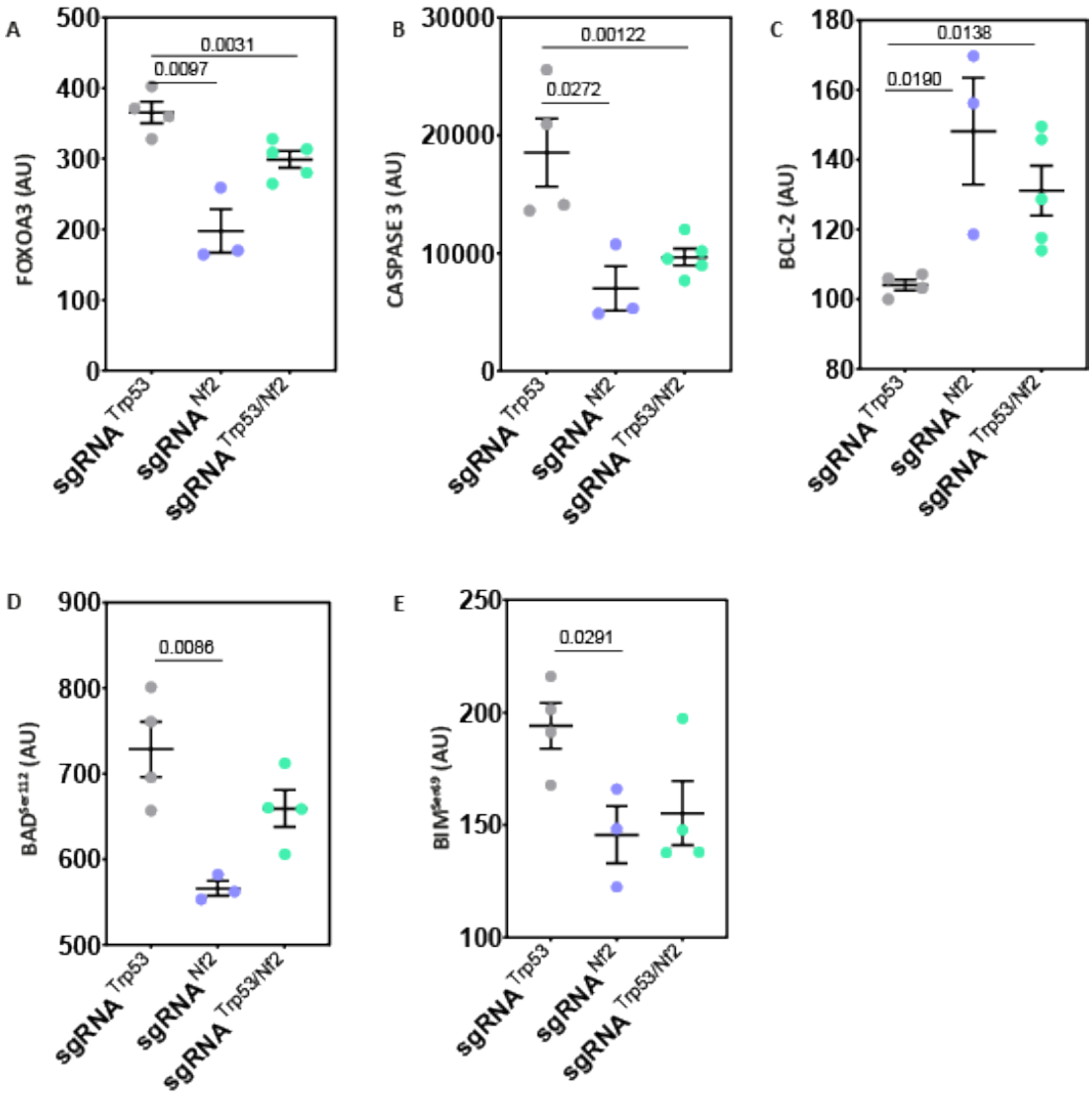

A
FOXOA3 (AU)
500
400
300
200
100
0
0.0031
0.0097
sgRNA Trp53
sgRNA Nf2
sgRNA Trp53/Nf2
B
CASPASE 3 (AU)
30000
20000
10000
0
0.00122
0.0272
C
BCL-2 (AU)
180
160
140
120
100
80
0.0138
0.0190
D
BAD^{Ser112} (AU)
900
800
700
600
500
0.0086
E
BIM^{EL} (AU)
250
200
150
100
0.0291

Figure 12

Apoptotic signalling pathway components are changed in the RPPA. Changes are measured in absorbance units (AU) and are from signalling components: A. FOXO3A, B. CASPASE-3, C. BCL-2, D. phosphorylated BAD at Ser112 and E. phosphorylated BIM at Ser49.

To link the loss of apoptosis in cells lacking *Nf2* with the molecular mechanism by which this process may occur, I investigated the NF2-interacting protein p21-activated kinase 1 (PAK1). Given evidence that NF2 can interact directly in an inhibitory manner with p21-activated kinase 1 (PAK1) (241), I assessed protein lysates for PAK1 levels with the expectation that loss of *Nf2* would lead to increased PAK1 levels. As can be observed in Figure 14A, when I blotted for PAK1, it is increased in $KRAS^{G12D}$:$sgRNA^{Nf2}$ animals compared to either $KRAS^{G12D}$:$sgRNA^{Trp53}$ or $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals. This became more evident when I quantified the triplicate of Western blots, as shown in Figure 14B.

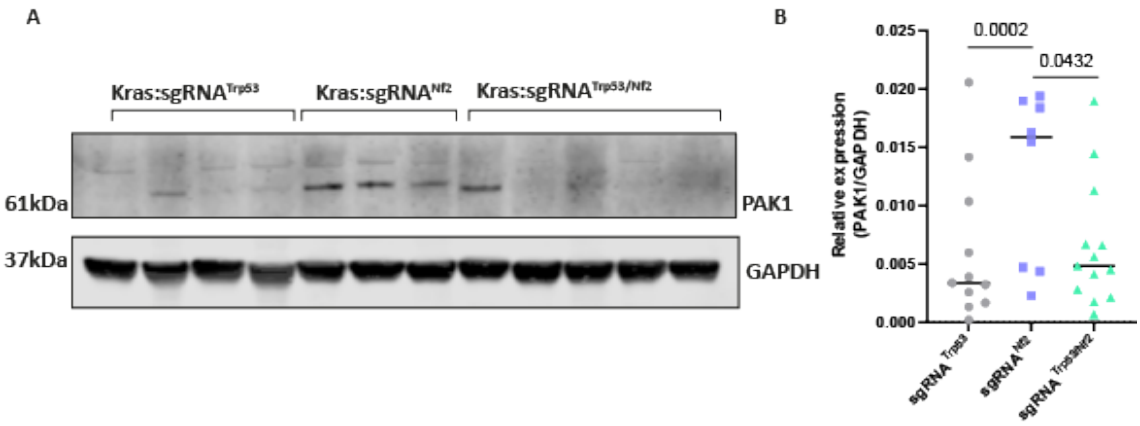

A
Kras:sgRNA^Trp53
Kras:sgRNA^Nf2
Kras:sgRNA^Trp53/Nf2
61kDa
PAK1
37kDa
GAPDH
B
Relative expression (PAK1/GAPDH)
0.025
0.020
0.015
0.010
0.005
0.000
0.0002
0.0432
sgRNA Trp53
sgRNA Nf2
sgRNA Trp53Nf2

Figure 14

PAK1 is increased in protein lysates of livers from animals lacking Nf2, compared to those lacking Trp53. A. Western blotting for PAK1 in KRAS^{G12D}:sgRNA^{Trp53} KRAS:sgRNA^{Nf2} and KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ animals. B. Quantification of PAK1 expression as relative expression to GAPDH, of three Western blots.

I considered the increase in PAK1 in the $KRAS^{G12D}$:$sgRNA^{Nf2}$ model could be a result of changes in upstream Rac/cdc42 signalling, as PAKs are established effector proteins of Rac/cdc42 signalling and they interact specifically and directly with these proteins. I tested protein lysates for the presence of cdc42, as presented in Figure 15A, but it is clear that the presence of cdc42 is not significantly increased when this Western blot was quantified in Figure 15B.

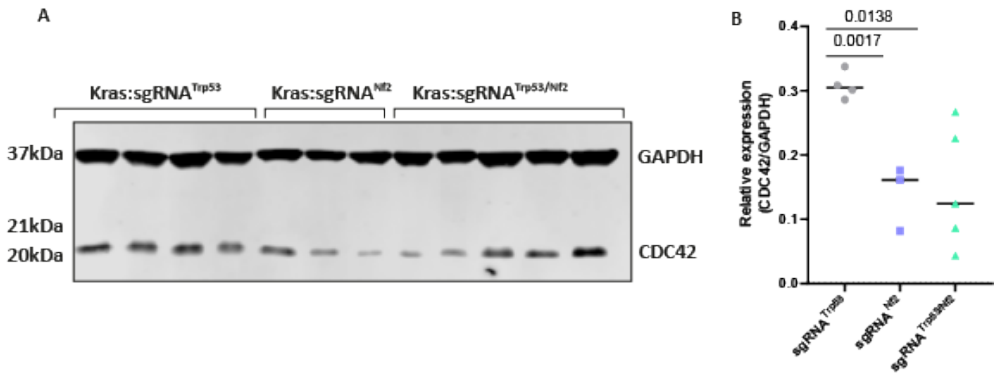

A
Kras:sgRNA^Trp53
Kras:sgRNA^Nf2
Kras:sgRNA^Trp53/Nf2
37kDa
21kDa
20kDa
GAPDH
CDC42
B
0.0138
0.0017
Relative expression (CDC42/GAPDH)
0.4
0.3
0.2
0.1
0.0
sgRNA Trp53
sgRNA Nf2
sgRNA Trp53Nf2

Figure 15

Cdc42 is increased in protein lysates of livers from animals lacking Trp53, compared to those lacking Nf2. A. Western blotting for cdc42 in $KRAS^{G12D}$*:*$sgRNA^{Trp53}$ *,* $KRAS^{G12D}$*:*$sgRNA^{Nf2}$ *and* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Nf2}$ *animals. B. Quantification of cdc42 expression as relative expression to GAPDH.*

5.8 Changes in cellular senescence when *Nf2* is lost

Loss of the cell cycle inhibitory protein p21 has been previously associated with loss of *Nf2*, and is a known marker of cellular senescence. (242) Given the requirement of additional loss of function mutations to permit Ras positive cells to escape senescence surveillance in the liver, I wished to explore the relationship between the concurrent loss of NF2 and state of cellular senescence. I examined the RPPA levels of P21 (Figure 13A) and these data showed that P21 is reduced in the animal models with deletion of *Nf2*. I observe that P21 is reduced also in the protein lysates of my samples. The Western blots I carried out in triplicate and quantified as relative P21 expression compared to GAPDH, is shown in Figure 13B, and 13C shows an exemplar blot. Interestingly, as P21 is a direct target of TRP53 I would have expected the loss of Trp53 to result in the suppression of P21 via Trp53-dependent pathways, however as P21 is also regulated independently of Trp53, this may provide an explanation for the presence of P21 in the $KRAS^{G12D}$:$sgRNA^{Trp53}$ model, and its low levels in the other two models.

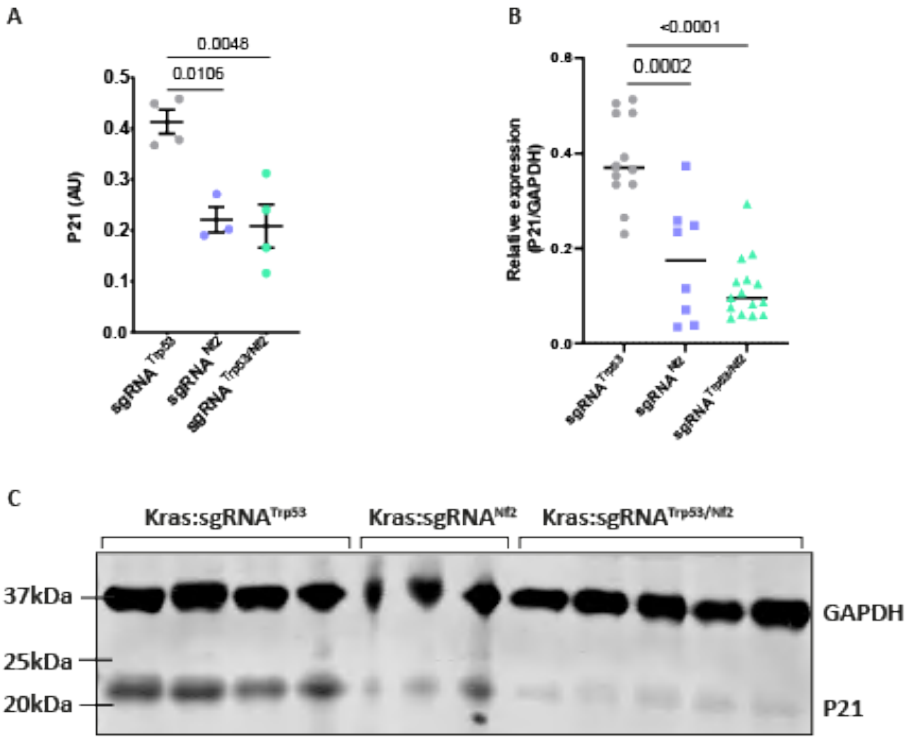

A
0.0048
0.0106
P21 (AU)
0.5
0.4
0.3
0.2
0.1
0.0
sgRNA Trp53
sgRNA Nf2
sgRNA Trp53/Nf2
B
<0.0001
0.0002
Relative expression (P21/GAPDH)
0.8
0.4
0.2
0.0
sgRNA Trp53
sgRNA Nf2
sgRNA Trp53/Nf2
C
Kras:sgRNA Trp53
Kras:sgRNA Nf2
Kras:sgRNA Trp53/Nf2
37kDa
25kDa
20kDa
GAPDH
P21

Figure 13

P21 levels are decreased when Nf2 is deleted in vivo. A. RPPA, measured in absorbance units (AU), B. Quantification of Western blots carried out in triplicate on all three models. C. exemplar Western blot of protein lysates from $KRAS^{G12D}$*:*$sgRNA^{Trp53}$*,* $KRAS^{G12D}$*:*$sgRNA^{Nf2}$ *and* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Nf2}$ *animals.*

Nf2 deletion cooperates with $KRAS^{G12D}$ to cause iCCA formation *in vivo*, showing that it is capable of evading senescence surveillance in the liver. To establish whether there were any changes in senescence during the early stages of tumour formation in when *Nf2* is lost, I examined the levels of P21 in the early time course of $KRAS^{G12D}$:$sgRNA^{Trp53}$ vs $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals. These data are shown in Figure 14A: there is an overall trend of decrease in P21 positive cells in $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals from maximum levels at 4 days post HDTVI of cancer forming plasmids, to the end point when mice were euthanised. There is a less discernible trend in the $KRAS^{G12D}$:$sgRNA^{Trp53}$ mice, but they do appear to retain consistent levels of P21 over days 4-12, and then these increase dramatically once the animals are euthanised at their end point. The immunohistochemical staining for P21 in the two models also reflects these trends where Figure 14B shows the increase in P21 levels in $KRAS^{G12D}$:$sgRNA^{Trp53}$ animals, and 14C shows the $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ mice.

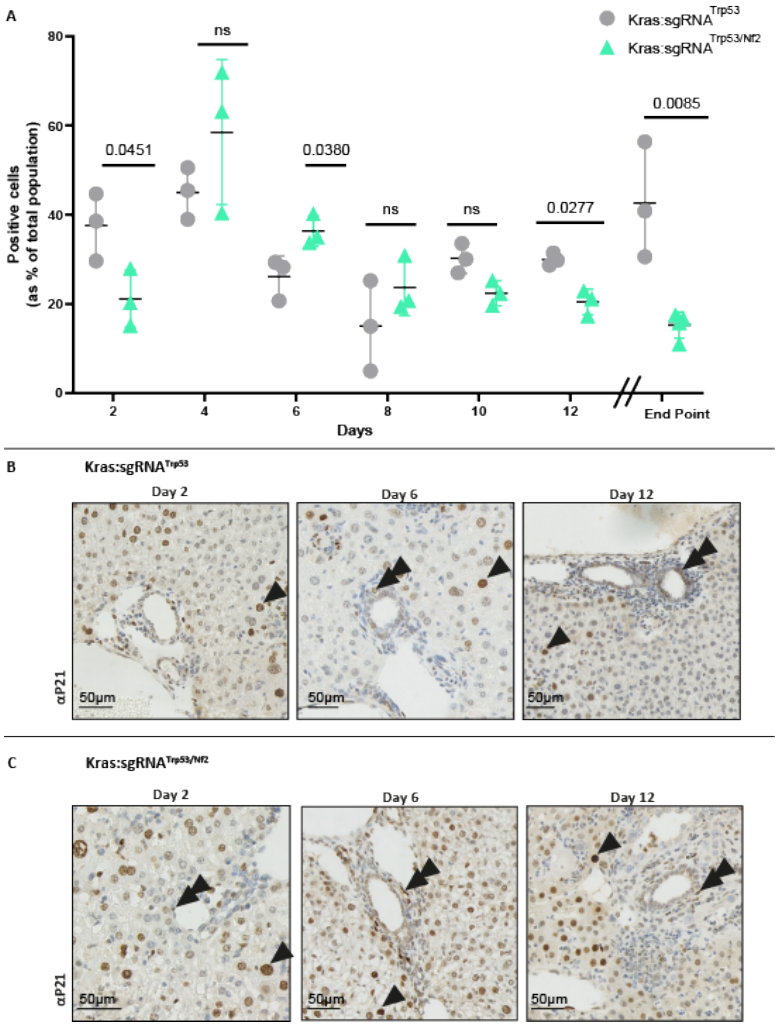

A
Kras:sgRNA Trp53
Kras:sgRNA Trp53/Nf2
Positive cells
(as % of total population)
80
60
40
20
0
0.0451
ns
0.0380
ns
ns
0.0277
0.0085
2
4
6
8
10
12
End Point
Days
B
Kras:sgRNA Trp53
Day 2
Day 6
Day 12
αP21
50μm
C
Kras:sgRNA Trp53/Nf2
Day 2
Day 6
Day 12
αP21
50μm

Figure 14

P21 positive cells, indicative of their senescent state, are measured in $KRAS^{G12D}$*:*$sgRNA^{Trp53}$ *and* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Nf2}$ *animals taken over a time course of early cancer formation, the days they were taken post HDTVI are indicated, A. B. Immunohistochemical staining for P21 on* $KRAS^{G12D}$*:*$sgRNA^{Trp53}$ *tissue, positive cells are indicated by arrows, and double arrows indicate P21 positive cholangiocytes. Likewise C. is* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Nf2}$ *tissue stained for P21.*

I also compared the amount of P21 expressed in end point tumours of the $KRAS^{G12D}$:$sgRNA^{Trp53}$ and $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ models to tissue containing mutations in $KRAS^{G12D}$:$sgRNA^{Nf2}$ and also to animals that were not injected with oncogenic Kras and only have deletion of *Trp53* and *Nf2*, that show no tumour formation at all, and were culled at a chosen end point of 58 days. I determined the number of positive cells of each of the aforementioned models, Figure 15A, where $sgRNA^{Trp53/Nf2}$ animals had the fewest number of cells expressing P21, the low number of senescent cells present is indicative of the lack of tumourous cells in this model. This is shadowed in Figure 15B where I stained this tissue immunohistochemically for P21 and there were a severely limited number of positive cells present. $KRAS^{G12D}$:$sgRNA^{Nf2}$ animals in 15C presented with a large number of P21 positive cells, as did $KRAS^{G12D}$:$sgRNA^{Trp53}$ and $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ animals, all in both the tumour and peri-tumour areas, as illustrated by Figures 15D and E where the black dotted lines indicate tumour and peri-tumour area.

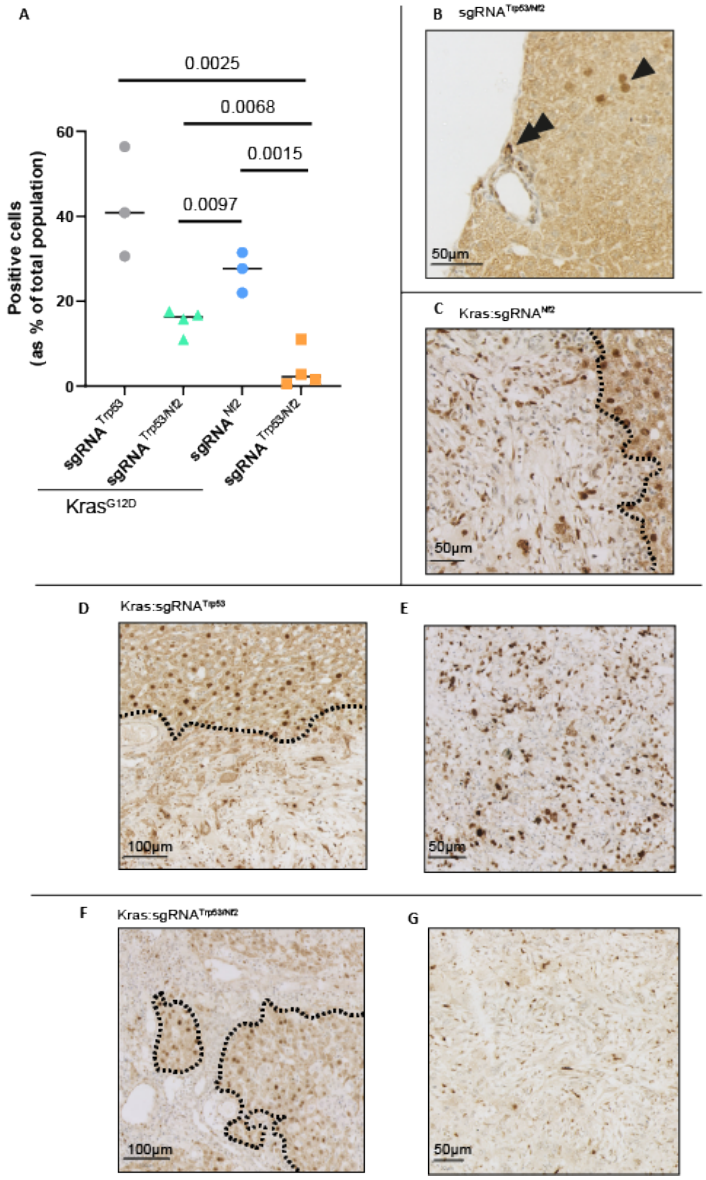

A
0.0025
0.0068
0.0015
0.0097
Positive cells
(as % of total population)
60
40
20
0
sgRNA Trp53
sgRNA Trp53/Nf2
sgRNA Nf2
sgRNA Trp53/Nf2
KrasG12D
B sgRNATrp53/Nf2
50μm
C Kras:sgRNANf2
50μm
D Kras:sgRNATrp53
100μm
E
50μm
F Kras:sgRNATrp53/Nf2
100μm
G
50μm

Figure 15

Expression of P21 in end point tissue. A. Percentage positive cell count of $KRAS^{G12D}$*:*$sgRNA^{Trp53}$*,* $KRAS^{G12D}$*:*$sgRNA^{Nf2}$*,* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Nf2}$ *and* $sgRNA^{Trp53/Nf2}$ *animals. B. Immunohistochemical staining for P21 in* $sgRNA^{Trp53/Nf2}$ *model, C.* $KRAS^{G12D}$*:*$sgRNA^{Nf2}$*, D. and E.* $KRAS^{G12D}$*:*$sgRNA^{Trp53}$*, F. and G.* $KRAS^{G12D}$*:*$sgRNA^{Trp53/Nf2}$*. Tumour mass is indicated by black dotted lines. Arrow indicates P21 positive hepatocytes and double arrow indicates P21 positive cholangiocytes.*

5.9 Treatment of non-RAS related model with therapeutics also greatly reduced tumourigenesis

To assess whether Wnt/ β-catenin and PI3K/AKT signalling are concurrently activated in a non-Ras driven model of iCCA, I used an existing mouse model whereby we delete *Trp53* and *Pten* in biliary epithelial cells specifically using a Keratin19-Cre^{ERT} knock in mouse line. When these animals are administered with the hepatotoxin Thioacetamide, the combination of mutations present and Thioacetamide cause formation of well differentiated cholanigocellular lesions around the portal tracts in the liver. Additionally, cells that recombine *Trp53* and *Pten* are labelled with tdTomato and can be visualised.

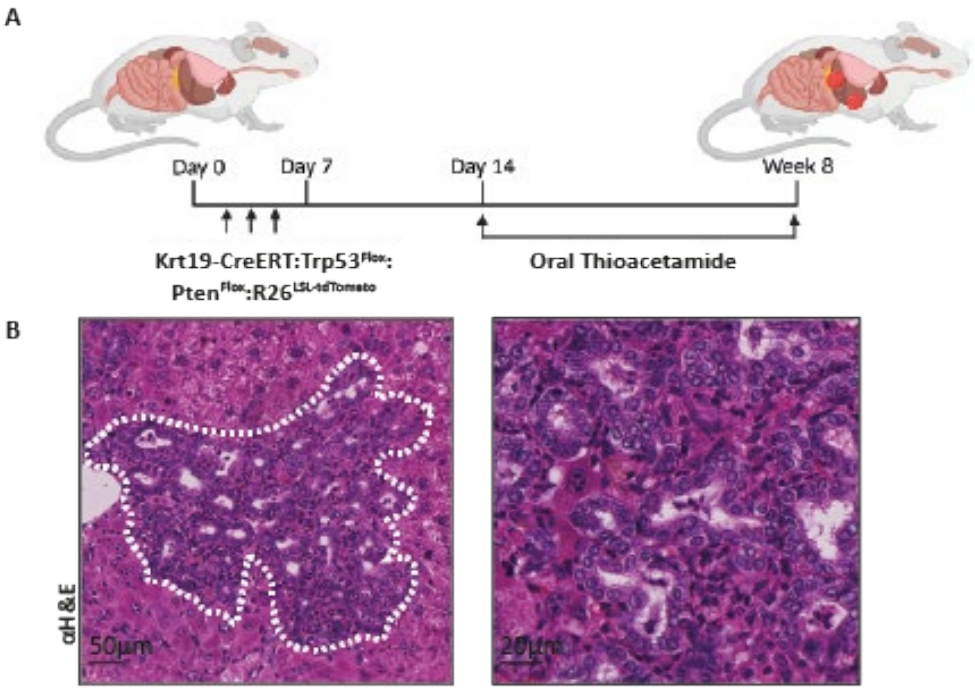

A
Day 0
Day 7
Day 14
Week 8
Krt19-CreERT:Trp53Flox:
PtenFlox:R26LSL-tdTomato
Oral Thioacetamide
B
αH&E
50μm
20μm

Figure 17

A non-Ras driven model of iCCA. A. Schematic of K19-CreERT:Trp53^{Flox}:Pten^{Flox}:R26^{tdTomato} mice are treated with the hepatotoxin Thioacetamide, leading to formation of RFP positive tumours after 8 weeks of treatment. B. H&E immunostaining of this tissue. White dashed line indicates tumour mass.

To assess whether the presence of a drug-treatable Wnt/β-catenin and PI3K/AKT signature is present in a non-Ras driven model of iCCA, CreERT:Trp53^{Flox}:Pten^{Flox}:R26^{tdTomato} mice were treated with LGK974 and Pictilisib, and another group treated with a vehicle control four weeks after the initial administration of Thioacetamide, when tumours are beginning to form. As shown in Figure 18, the number of tdTomato positive cells significantly decreases compared to vehicle treated animals, and only 33% of the cohort developed iCCA when treated with the combination of therapeutics compared to 80% of those treated with vehicle. When these livers were stained for β-catenin and phosphorylated-AKT, the tumours were strongly positive for these signalling components.

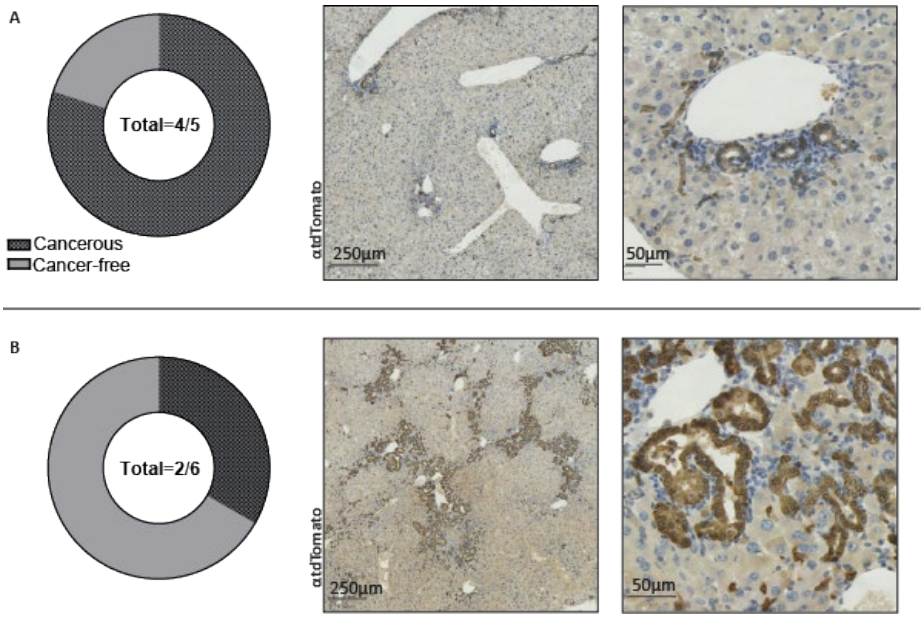

A
Total=4/5
Cancerous
Cancer-free
αtdTomato
250μm
50μm
B
Total=2/6
αtdTomato
250μm
50μm

Figure 18

CreERT:Trp53^{Flox}:Pten^{Flox}:R26^{tdTomato} mice treated with A., vehicle control, compared to animals treated with a combination of both LGK974 and Pictilisib, B. Representative of which animals present with cancerous lesions, and those that are cancer free, and immunohistochemical staining for tdTomato-positive cancer cells.

To ensure that the Nf2 loss model was not altering Hippo signalling, a pathway that it has been shown to partake in, I stained $KRAS^{G12D}$:$sgRNA^{Trp53/Nf2}$ tissue for YAP/TAZ. As a key component of Hippo signalling, nuclear YAP/TAZ is representative of active Hippo signalling. As shown in Figure 19, the tumourous ductal cells are not positive for YAP/TAZ, and the cells that do contain nuclear YAP/TAZ are adjacent to the cancerous cells. For this reason, I concluded that in this model the loss of *Nf2* is not affecting Hippo signalling.

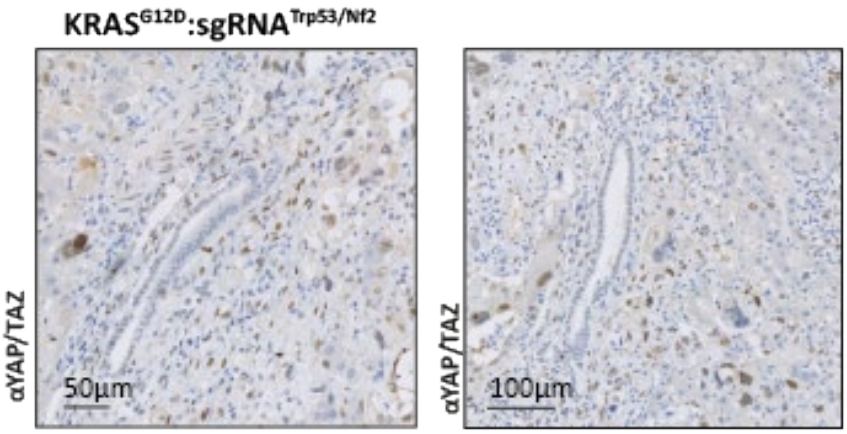

KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$
αYAP/TAZ
50μm
αYAP/TAZ
100μm

Figure 19

YAP/TAZ staining on KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ HDTVI tissue shows that cancerous cells with ductal morphology do not express nuclear YAP/TAZ.

5.10 Discussion

NF2 sits at the nexus of many signalling pathways and its perhaps most established role is in exerting control over the Hippo signalling pathway, however when I stained for YAP/TAZ in KRAS:sgRNA$^{Trp53/Nf2}$ tissue, I observed no nuclear YAP/TAZ in the nuclei of tumour cells. Surrounding hepatocytes were YAP/TAZ positive but not the tumours themselves. This then raised the question as to by which mechanism *Nf2* loss causes the sarcomatoid iCCA phenotype I see in my model. Having established that apoptosis is reduced in cells with *Nf2* deletion, I examined the levels of PAK1 as this is a known interacting protein of NF2, and indeed PAK1 levels were increased when *Nf2* was lost, but not when *Trp53* was lost also.

Considering the presence of P21 positive cells in my KRAS^{G12D}:sgRNA$^{Trp53/Nf2}$ model, this was interesting as it provides evidence of an unusual relationship between *Trp53* loss and increased p21, as Trp53 directly regulates p21 expression. Therefore these data are indicative of *Trp53*-independent p21 activation.

The establishment of an aggressive, sarcomatoid phenotype model of iCCA with accelerated mortality that lacks mutations in *Nf2* and *Trp53*, and expresses human oncogenic KRAS^{G12D} was key in the latter discovery of which pathways were dysregulated. The loss of Nf2 ultimately led me to find that there was dysregulation in both the Wnt/ β-catenin and PI3K/AKT signalling pathways, which could then be targeted therapeutically in order to reduce tumourigenesis and extend survival. I consider that the Nf2 loss model was a key stepping-stone to the discovery of the upregulation of Wnt/ β-catenin and PI3K/AKT signalling. I would suggest that these data could feed into the treatment of patients with iCCA, but under specific

circumstances: assessment of whether they were experiencing upregulation of Wnt/ β-catenin and PI3K/AKT signalling.

Chapter 6 Concluding Remarks and Future Perspectives

This work aimed to identify which rare loss of function driver mutations interact with oncogenic $KRAS^{G12D}$ mutations in order to cause iCCA formation. As iCCA is a particularly complex disease at the genetic, molecular and cellular levels, there is a limited number of effective therapies for treatment. In iCCA, a small number of defined mutations are involved in formation of this disease, but they are insufficient targets for effective therapies because patients do not necessarily possess these mutations. This highlights the need for more effective, widespread therapeutic options for iCCA patients with various mutational landscapes.

To address the greatly heterogeneous genetic nature of iCCA, I firstly utilised patient exome sequencing data to generate a *sp*Cas9 CRISPR library of rare driver mutations to identify which of these are favoured in the formation of iCCA *in vivo* alongside one of the more commonly mutated oncogenes in iCCA, $KRAS^{G12D}$. Given the presence of *Trp53* loss in all of the resultant screen tumours, this led to the identification of *Nf2* deletion as a potent driver gene in iCCA tumourigenesis in addition to KRAS and *Trp53* mutations.

The model of aggressive iCCA harbouring loss of function mutations in *Trp53* and *Nf2*, and expression of oncogenic KRAS was used to address the underlying molecular mechanisms of iCCA. This model histologically closely represents a rare but lethal form of iCCA

I identified concurrent upregulation of Wnt/β-catenin and PI3K/AKT signalling in this model, and established that this was targetable with inhibitors of these pathways. The discovery of changes in these pathways in another, non-Ras driven model of iCCA

further supports that the genetic complexity present is responsible for changes in iCCA that result in its aggressive nature. I propose a mechanism of action for how *Nf2* deletion may contribute to the tumourigenesis of iCCA via Wnt/ β-catenin and PI3K/AKT signalling.

Therefore this study has allowed the translation of the human genetic diversity observed in iCCA into changes in specific signalling cascades that change despite the underlying genetics of the model in question. These data illustrate promising potential wide-reaching therapeutic opportunities for iCCA.

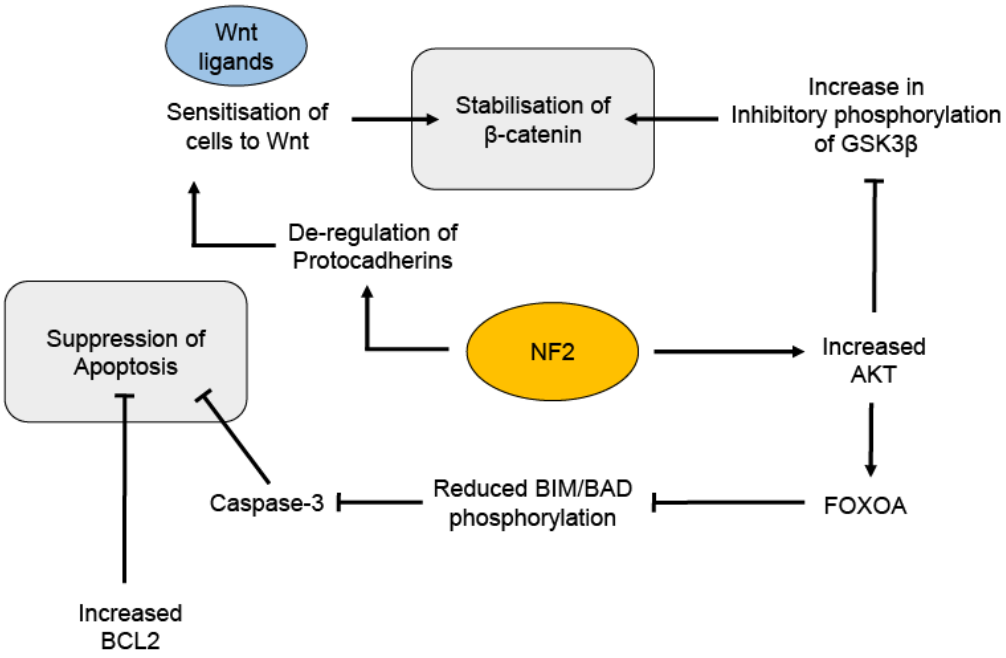

Wnt ligands
Sensitisation of cells to Wnt
Stabilisation of β-catenin
Increase in Inhibitory phosphorylation of GSK3β
De-regulation of Protocadherins
Suppression of Apoptosis
NF2
Increased AKT
Caspase-3
Reduced BIM/BAD phosphorylation
FOXOA
Increased BCL2

Figure 1

Proposed mechanism of action of Nf2 *deletion on accelerating sarcomatoid iCCA tumourigenesis.* Nf2 *deletion leads to increased* AKT *signalling and stabilisation of β-catenin, which results in increased signalling through these two pathways.*

Chapter 7 Appendix

Appendix 1

Single guide RNA sequences used in this project.

gRNAs included within library pool		
Gene Name	Target ID	gRNA sequence 5'-3'
Abl1	Abl1_1	CCCTTCATACCGCAGCGAGA
Abl1	Abl1_2	AAGGGAGGGTGTACCACTAC
Abl1	Abl1_3	CGAGAGCCGCTTCAACACTC
Ablim1	Ablim1_1	GTCATAGATCGCCTTGACCT
Ablim1	Ablim1_2	GAAAGTATCTATTCTAGACC
Ablim1	Ablim1_3	CTTACAGTCGGGATGCCACA
Acvr1b	Acvr1b_1	GTGTCTACCATAACCGCCAG
Acvr1b	Acvr1b_2	GTCTACGACCTCTCCACGTC
Acvr1b	Acvr1b_3	AGGGCCGGTTCGGGGAAGTA
Akap9	Akap9_1	GATGCCGTCTCGCTGTTTAA
Akap9	Akap9_2	GCCGCCATTAAACAGCGAGA
Akap9	Akap9_3	GCGCACCTGCTTCAGCCCTT
Apbb1	Apbb1_1	CGAGACGGATTCCGATCTAC
Apbb1	Apbb1_2	GGAATCCGTCTCGAAAGCGT
Apbb1	Apbb1_3	GAAGGACAATGTCGCCTCCT
Aqp7	Aqp7_1	ACTCCCCAGGCATTCGTGAC
Aqp7	Aqp7_2	GGCATCCTTGTTACCGTCCT
Aqp7	Aqp7_3	CATTCGGCCTAGTGCACAAT
Arhgap8	Arhgap8_1	GGCCCTCACCGCACTACTTC
Arhgap8	Arhgap8_2	TCCTCCAGGGGATGACCGCC
Arhgap8	Arhgap8_3	TATAGTCGTTCTCCACATGC
Arid1a	Arid1a_1	GCCCTGCTGGCCATACGCAC
Arid1a	Arid1a_2	GGTCCCTGTTGTTGCGAGTA
Arid1a	Arid1a_3	TAACCCATACTCGCAACAAC
Arid2	Arid2_1	TACCAGCAACACAGCGTGTC
Arid2	Arid2_2	ATAGCGAAGTCCACTTCATT
Arid2	Arid2_3	CACTTTACTGCTCGCTAATG
Atf7ip	Atf7ip_1	CCGTACGTACTTCTCACAGG
Atf7ip	Atf7ip_2	TCCAAAACGTTTGGTTAACC
Atf7ip	Atf7ip_3	GGCAGTGCTCACCGAGCTGC
Atp2b1	Atp2b1_1	TTACGTGCATTATCCCCTTC
Atp2b1	Atp2b1_2	AATTGCAGCCATAGTATCAT
Atp2b1	Atp2b1_3	CCTGCTCAATTCGACTCTGC
Atp7a	Atp7a_1	AACACGGTATTGGTTAAGAC
Atp7a	Atp7a_2	CAAGGTGTTCAGCGCATTAA
Atp7a	Atp7a_3	GCAGCCGAAGTACCTCAAAT
Bap1	Bap1_1	TCAAATGGATCGAAGAGCGC
Bap1	Bap1_2	CCGCCGCAAGGTTTCTACGT
Bap1	Bap1_3	GCCCACGCTGAGCCGAATGA
Bcor	Bcor_1	GATGAGGCCGAATCCAACGA

Bcor	Bcor_2	CTTACGGGCTGCATACGAGC
Bcor	Bcor_3	CGACGCTTCAAAAGCCAGGC
Braf	Braf_1	TCATAGGTACCCGCAAGATG
Braf	Braf_2	GAGGCCCTATTGGACAAATT
Braf	Braf_3	GTTGTTCCGTGATGCATCTG
Brca2	Brca2_1	TAGGACCGATAAGCCTCAAT
Brca2	Brca2_2	AAAGCTCCTCAAACCAATTG
Brca2	Brca2_3	GGACTAGCAACATCTACCAC
Cacna1c	Cacna1c_1	CAACCGCCACACGGCCGCCC
Cacna1c	Cacna1c_2	CTCCTCGCAGGTTCCAACTA
Cacna1c	Cacna1c_3	GTTGGAGTCGTCTTCCGGAA
Camkk1	Camkk1_1	ACAGGCCGTCCTCCCCCCAG
Camkk1	Camkk1_2	TCTCCTGATACACACGCTCC
Camkk1	Camkk1_3	TCTGTCTTCACTTTCGTTGT
Cc2d2a	Cc2d2a_1	ACTCACTCCGGGGCCCCCGC
Cc2d2a	Cc2d2a_2	AGGCAGAAAGCGCATTGTTG
Cc2d2a	Cc2d2a_3	CAACAATGCGCTTTCTGCCT
Ccdc66	Ccdc66_1	TGAATACTGCTTATCCGTAC
Ccdc66	Ccdc66_2	CACTACCGGTACTTGTCCAG
Ccdc66	Ccdc66_3	TGCTCATGTTCACTCGTGAC
Cdk18	Cdk18_1	GCCCTGCACTGCTATTCGAG
Cdk18	Cdk18_2	GAACCAGCGCCGATTCTCCA
Cdk18	Cdk18_3	TCCCAGTCAGATATCGGCTT
Cdkn2a	Cdkn2a_1	GTGCGATATTTGCGTTCCGC
Cdkn2a	Cdkn2a_2	CCCAACGCCCCGAACTCTTT
Cdkn2a	Cdkn2a_3	GGGGTACGACCGAAAGAGTT
Chd4	Chd4_1	GCTTCGGATAAATCCTCGTC
Chd4	Chd4_2	TCGGACCCTCACCAACTACA
Chd4	Chd4_3	TCCTGACTTACCTAACAAAT
Cped1	Cped1_1	TAGTCGACACATGTGCTCCC
Cped1	Cped1_2	TGGTCTAGGTCCGGAGCTGC
Cped1	Cped1_3	GTCGACTAAGCCTACATCAA
Cpt1a	Cpt1a_1	GACTATGCGCTACTCGCTGA
Cpt1a	Cpt1a_2	CACTCACGATGTTCTTCGTC
Cpt1a	Cpt1a_3	GGGTTACTCACGTGGTGTCT
Dbndd2	Dbndd2_1	CATAGGTAGCATCTCGTCTA
Dbndd2	Dbndd2_2	GTTTATTGATCTTGCGGATC
Dbndd2	Dbndd2_3	GATGTGGTCCTGTCAGACGT
Dchs1	Dchs1_1	TAGGTAGGGCTATCAAATCG
Dchs1	Dchs1_2	TGTGTGTACCACCCGGACCC
Dchs1	Dchs1_3	TAGAGCGCGCAGTGTATAGA
Ece2	Ece2_1	GAACGTCGCGCTGCACGAGT
Ece2	Ece2_2	CTCGTGCAGCGCGACGTTCA
Ece2	Ece2_3	GACCATGGATGTTCGAGCAC
Ehhadh	Ehhadh_1	AGCTGCGTTCCTCTTGCACC
Ehhadh	Ehhadh_2	CTATCGGATTGCCAATGCAA
Ehhadh	Ehhadh_3	GCCATCCAAGGCGTGGCTCT
Elf3	Elf3_1	CGACTTCTCCCGCTGCGACA

Elf3	Elf3_2	CTAGCGAGCGGCCCCAGTTC
Elf3	Elf3_3	CAAAGGGGCCCGAGTCGCCT
Epha2	Epha2_1	CAACGTGGTATCCGGCGACC
Epha2	Epha2_2	TTCGCTGTCGAAGCACGCAA
Epha2	Epha2_3	CTGCTGACCGTGATCTCGTC
Erbb2	Erbb2_1	CTTACAGGCCCGGGAACGAT
Erbb2	Erbb2_2	GCTAGACAACCGAGACCCTT
Erbb2	Erbb2_3	ACATGGACACCAATCGTTCC
Erbb3	Erbb3_1	CCCCTTGCAGACTTCGTGAC
Erbb3	Erbb3_2	CGGGGAACCCAGGTCTACGA
Erbb3	Erbb3_3	GCCCTTACCTAACCTCCGAG
Esrra	Esrra_1	TGTACTTCTGCCGTCCGCCG
Esrra	Esrra_2	GTACGTCCTGCTGAAAGCTC
Esrra	Esrra_3	TGCGACACCAGAGCGTTCAC
Etv6	Etv6_1	GCATGGCGTGCTCTTCCGGT
Etv6	Etv6_2	GTGAACATGAAGCGGAGTCG
Etv6	Etv6_3	CTGACAGGGGGTACGTTTCC
Fam136a	Fam136a_1	GCGCCGCTACCTGCATCTTC
Fam136a	Fam136a_2	CGAGCGCTGCCATGCGCCTC
Fam136a	Fam136a_3	CATGGCAGCGCTCGATGCAT
Fam174b	Fam174b_1	CCACGATCACCGTCGCCTTC
Fam174b	Fam174b_2	TCGCAGAGCCGAGTCCGCTT
Fam174b	Fam174b_3	GCCATTTCCACACGTTCCGC
Fh1	Fh1_1	AATTGGGCGAACTCACACGC
Fh1	Fh1_2	CGTGTAGAGTTCGACACCTT
Fh1	Fh1_3	AAAATCCAAAGAGTTTGCGC
Fn1	Fn1_1	CCAGGTCTCCCCCACGACGT
Fn1	Fn1_2	GACCTACCTAGGCAACGCCC
Fn1	Fn1_3	CCTACAAGATTGGCGACAAG
Fzd2	Fzd2_1	ATCGACGGCGACCTGCTGAG
Fzd2	Fzd2_2	GGCGCGGTGGTGAGTAGCGC
Fzd2	Fzd2_3	TGCGAGCATTTCCCGCGTCA
Gpr162	Gpr162_1	AGTCCAACGGCTCGGCTATC
Gpr162	Gpr162_2	TCTCCGGATAGCCGAGCCGT
Gpr162	Gpr162_3	GTGTGTGGCGATCATGTCCG
Gstp1	Gstp1_1	GTCAGCCAGCAGCATTCGCA
Gstp1	Gstp1_2	ATTGTCTACTTCCCAGTTCG
Gstp1	Gstp1_3	TCACATAGTTGGTGTAGATG
Heatr3	Heatr3_1	CACTACCAGACTTAGCCCGT
Heatr3	Heatr3_2	GCCGGCGCACAGCATCCCGA
Heatr3	Heatr3_3	ATTGACCTGGCTGTCTCCGT
Helz	Helz_1	GCAACGAGTAATGTCCTGTC
Helz	Helz_2	CTTCTACGTGACATCCAGAC
Helz	Helz_3	GCTGATGAAGATTGTAGGCA
Hrct1	Hrct1_1	GAAACCGAGTCCGACAAGCC
Hrct1	Hrct1_2	GTGGTGGTGAACGCCCACAC
Hrct1	Hrct1_3	CGCCGGCCATGGAAGAGTCG
Idh1	Idh1_1	TGGGCCTGTAAGAATTACGA

Idh1	Idh1_2	GGCCCAAGCTATGAAGTCCG
Idh1	Idh1_3	AGTCTTCAATTGACTTATCC
Idh2	Idh2_1	ATGTTCCGGATCGTTCCGTT
Idh2	Idh2_2	AGAGCCCTAACGGAACGATC
Idh2	Idh2_3	TCGAGCTGGCACGTTCAAGT
Il1r1	Il1r1_1	AGTCCCGGTCCGCTGATATG
Il1r1	Il1r1_2	GCCGTATGTCCTATACGTTC
Il1r1	Il1r1_3	TTGCTTCCCCCGGAACGTAT
Itpr2	Itpr2_1	GGACATCGTGTCCCTGTACG
Itpr2	Itpr2_2	TGACCGAGCCCTCCGCGTAC
Itpr2	Itpr2_3	GTTCCCCTGTTTCGCCTGCT
Itpr3	Itpr3_1	AGCATCCAGCGTCACCCGCA
Itpr3	Itpr3_2	GCTGGATGCTACGGGCAATG
Itpr3	Itpr3_3	GGTCTTACCTCGGAACTTCT
Kras	Kras_1	TGAGTATGACCCTACGATAG
Kras	Kras_2	AGCAGCGTTACCTCTATCGT
Kras	Kras_3	TAGAACAGTAGACACGAAAC
Mme	Mme_1	TGCTCGACTGATTCAGAATA
Mme	Mme_2	GGCCAGTAGCATCAGATAAC
Mme	Mme_3	TCCCAGTTATCTGATGCTAC
Msr1	Msr1_1	GCGTTCCGTGTCTATAAGGT
Msr1	Msr1_2	TGAACGTGCGTCAAATTTCA
Msr1	Msr1_3	TTCCTTGATTTCGTCAGTCC
Mtch2	Mtch2_1	CTTTGACCGAGTTATCAAAG
Mtch2	Mtch2_2	GGCACCTACCATAGCAAAAG
Mtch2	Mtch2_3	AAAATATTTCGTCCTATTGT
Ncor1	Ncor1_1	GGATATGAACAGTTTCACTC
Ncor1	Ncor1_2	AATGTATTAGGCCTCAAGAA
Ncor1	Ncor1_3	TATGCCTTACCTGCTGGTGT
Nf2	Nf2_1	GATCCGCACCGTGAATGTCT
Nf2	Nf2_2	GAAGCTCATGCGAGAAGCGA
Nf2	Nf2_3	CTTGGCGTCATATGCTGTCC
Notch2	Notch2_1	GTCGTCGATATTCCGCTCAC
Notch2	Notch2_2	GCCTCCGTTGACGCAGGGCG
Notch2	Notch2_3	TACGAGTGCACCTGCCAAGT
Nprl3	Nprl3_1	AGCGGCGCTCCTCATGCTGC
Nprl3	Nprl3_2	CTGCCAGTACCTCACTCGAG
Nprl3	Nprl3_3	AAGGTTCTCGGATGTTATTC
Nras	Nras_1	CTAGCATACCTGTCGGGTCT
Nras	Nras_2	GACCTCAGCCAAGACCCGAC
Nras	Nras_3	TTGCAGATATTAACCTCTAC
Pbrm1	Pbrm1_1	CAAACTCATTTCTTGTTCGA
Pbrm1	Pbrm1_2	GTTGTAGCCACAAATCCATC
Pbrm1	Pbrm1_3	ACATCACTTACCACTTTGGA
Phf20l1	Phf20l1_1	CACGATCCAGCTGGGTCGAT
Phf20l1	Phf20l1_2	GATATGATGAGTGGATTTAC
Phf20l1	Phf20l1_3	TGTTGTCCTATCGACCCAGC
Pik3ca	Pik3ca_1	GCGCACTATTTATGACCCAG

Pik3ca	Pik3ca_2	TCACCATGCCGTCATACTCC
Pik3ca	Pik3ca_3	CAGAAGTCCAAGACTTTCGA
Pik3r1	Pik3r1_1	AGATGCGTCTCGTACCAAAA
Pik3r1	Pik3r1_2	GGAGTACACCCGTACTTCCC
Pik3r1	Pik3r1_3	GTGAATTACCTGCTGGTATT
Plch2	Plch2_1	TAGCCCGTCGCCAGCGTACC
Plch2	Plch2_2	CCGTAGTCCCCCAAGCCAAA
Plch2	Plch2_3	GATATTGGTCCCTGGTACGC
Plk2	Plk2_1	CAGAAGTCCGATACTACCTC
Plk2	Plk2_2	AAAACTGCACGACATGCTTA
Plk2	Plk2_3	ACGAACAAGAAATCTTGCAC
Plxnb2	Plxnb2_1	CTCAGATGGCCGGATCCTTA
Plxnb2	Plxnb2_2	TACTCTGCAGAAGTGCCGTC
Plxnb2	Plxnb2_3	GGAGTCACGACACTGAGCGC
Rasa1	Rasa1_1	ACACGCCTTCTATCTTCTAC
Rasa1	Rasa1_2	CTATAGCAGAAGAACGCCTC
Rasa1	Rasa1_3	CGTGACTGTAATAACCTATT
Rb1	Rb1_1	AGAAATCGATACCAGTACCA
Rb1	Rb1_2	TGACATAGCATTATCAACCT
Rb1	Rb1_3	TTGGGAGAAAGTTTCATCCG
Rbmx	Rbmx_1	CCTTAAGCCCCTGTGTCCCG
Rbmx	Rbmx_2	CTCTTGACTTATTCGTTTCT
Rbmx	Rbmx_3	CGAGAAACGAATAAGTCAAG
Rev1	Rev1_1	ACGGCGTCTGCAATATTTAG
Rev1	Rev1_2	TGAAATGAAGTTGCAGTCCG
Rev1	Rev1_3	TTGGTCGTAGGGCGGGTACA
Rexo4	Rexo4_1	TGATACTTACTCAGTGGGCG
Rexo4	Rexo4_2	CCCTAGCTTGACAAAAGCCT
Rexo4	Rexo4_3	GTGTCGATCGTGAACCAGTA
Rnf31	Rnf31_1	CATACAACCGTAGTACATCC
Rnf31	Rnf31_2	AGGGTGGCCGGGATGTACTA
Rnf31	Rnf31_3	TAACCCCGTCTTTCGCAGCA
Rnf43	Rnf43_1	GAGACGCTTACCCCGGCGGG
Rnf43	Rnf43_2	CAGGGGCGAGGAGCTCGTCG
Rnf43	Rnf43_3	TTTCCACAGGCCCGAATGGC
Rps29	Rps29_1	CTCACCAAGAGCGGGAACCC
Rps29	Rps29_2	GAAGGACATAGGCTTCATTA
Rps29	Rps29_3	TGGGTCACCAGCAGCTCTAC
Sdk1	Sdk1_1	AGAACTCACGTGCTACGCTC
Sdk1	Sdk1_2	TCCTGTTCCGCACCACGCAG
Sdk1	Sdk1_3	CATCACTCGCTGTTGTAGCA
Setd2	Setd2_1	GATCTCTTTCGGACCGACAT
Setd2	Setd2_2	GTCGGTCCGAAAGAGATCGA
Setd2	Setd2_3	ACTGCATTCGCTTAATATCC
Sf3b1	Sf3b1_1	GTCCTCCAAAGATTGCCGAT
Sf3b1	Sf3b1_2	GCCTGGATATCATGCCCCCG
Sf3b1	Sf3b1_3	TATATCATTAAGCAACGCCA
Smad4	Smad4_1	GCCAAGTAATCGCGCATCAA

Smad4	Smad4_2	TCCGTTGATGCGCGATTACT
Smad4	Smad4_3	ACAACCCGCTCATAGTGATA
Smu1	Smu1_1	TCCAGATGAATATAGCGCTC
Smu1	Smu1_2	ATTGATAGAGCTTCGTGAAT
Smu1	Smu1_3	CTTGCCTCGAAACAAATCGA
Snrpn	Snrpn_1	GGAACTCCACCTCCACCTGT
Snrpn	Snrpn_2	TGGGGAATAGGTACACCTGC
Snrpn	Snrpn_3	TTCACAGGTCATGACCCCAC
Stk11	Stk11_1	GCGCCCTACGTATATGGTGA
Stk11	Stk11_2	TGTACAGCACGTCCACAAGC
Stk11	Stk11_3	ATTCCAGGCCGTCAATCAGC
Tbc1d8b	Tbc1d8b_1	CTAGGGACTAATCGCTGAAG
Tbc1d8b	Tbc1d8b_2	GTCTGGTGTAGGATGCGAAA
Tbc1d8b	Tbc1d8b_3	CCGTCCTCTCCAGTAACTGC
Tgfbr1	Tgfbr1_1	ATCTATTCAAGTAATCGAAA
Tgfbr1	Tgfbr1_2	CTTCTAGAGAAGAGCGTTCA
Tgfbr1	Tgfbr1_3	AGTGATGGATCCTCTTCATT
Thbs1	Thbs1_1	AAGGGGCCCCGGTCGCCGAC
Thbs1	Thbs1_2	CCCTTCACCAGTCGGCGACC
Thbs1	Thbs1_3	GCTATCCGCACCAACTACAT
Tmtc1	Tmtc1_1	GTAGAATGACATTGACCGCG
Tmtc1	Tmtc1_2	TTATACTCACCGCCTCAGTG
Tmtc1	Tmtc1_3	CCATGTCGCCGAGAGCTATG
Tmtc4	Tmtc4_1	GGCCTTTACCGGTGTACTGT
Tmtc4	Tmtc4_2	AGGATTAACTACTACCTGTC
Tmtc4	Tmtc4_3	CCCCAAGGGGCGTGTCTGAC
Trp53	Trp53_1	AGTGAAGCCCTCCGAGTGTC
Trp53	Trp53_2	AACAGATCGTCCATGCAGTG
Trp53	Trp53_3	TGAGGGCTTACCATCACCAT
Ttc8	Ttc8_1	CGATCTATGCACGCAGATGC
Ttc8	Ttc8_2	CGACCTATCACTAGCTCATC
Ttc8	Ttc8_3	CGGCCAGGTACCTGATCATA
Wdr62	Wdr62_1	ATGACATGGTCCTCAATGTT
Wdr62	Wdr62_2	GTCACTATGTACTTCCCATC
Wdr62	Wdr62_3	ATGTGCTTACCCTGCTAAGT
Zfp317	Zfp317_1	CCATAAATAGGTTATCAGGT
Zfp317	Zfp317_2	TTATAGCAATCTAAGTTCAC
Zfp317	Zfp317_3	GAACTACATCTAAAGTCAAA
Zfp36l2	Zfp36l2_1	CGATATCGACTTCTTGTGCA
Zfp36l2	Zfp36l2_2	CGCGTGCGCCAAACGCTCGC
Zfp36l2	Zfp36l2_3	GCCGCCGCTCGTCCGCGTTG

Appendix 2

Antibody list for RPPA analysis

4E-BP1 P Ser65	LKB1
Akt	MEK 1/2 P Ser217/221
Akt P Ser473	MEK1/2
AMPK alpha	MEK1/2 P Ser217/221
AMPK alpha P Thr172	MEK6 [EP558Y]
ATM/ATR Substrate P Ser/Thr	Met
Aurora A/B/C P Thr288/Thr232/Thr198	Met P Tyr1234
Bad P Ser112	Met P Tyr1349
Bcl-2	MMP21 [EP1277Y]
Bcl-x	MNK1 (MKNK) P Thr197,Thr202
beta-Catenin	MNK1 (MKNK) P Thr197,Thr202
beta-Catenin P Ser33,Ser37,Thr41	MSK1 P Ser376
beta-Tubulin	mTOR (7C10)
Bid	mTOR P Ser2448
Bim P Ser69	p21 CIP/WAF1
c-Abl	p21 CIP/WAF1 p Thr145
c-Abl P Y245	p44/42 MAPK (ERK1/2)
c-Abl P Y412 (247C7)	p44/42 MAPK (ERK1/2) P Thr202/Thr185,Tyr204/Tyr187
Caspase 3	p53
CDK1 (cdc2)	p53 P Ser15
Chk1 P Ser345	PARP
Cleaved Notch 1 (val1744)	PI3 Kinase p110-alpha
c-Myc	PKC (pan) P Ser660 (beta-2)
c-Myc P Thr58,Ser62	PKC-alpha
CrkL	PLC-gamma1
CrkL P Tyr207	PLC-gamma1 P Tyr783
Cyclin D1	Profilin (C56B8)
Cyclin D1 P Thr286	Prohibitin
E-Cadherin	PTEN
EGFR P Tyr1173	PTEN P Ser380,Thr382,Thr383
EGFR P Y1068	Puma
EGFR P Y992	PYK2 [EP206Y]
EphA2 (D4A2)	PYK2 P Y402
EphA2 P Ser897	Raf P Ser259
ErbB-1/EGFR	Rap1
ErbB-2/Her2/EGFR P Tyr1248/Tyr1173	Ras
Ezrin P T567/Radixin T564/Moesin T588	Rb
FLT3 P Tyr591 P Tyr591	Rb P Ser807,Ser811
FOXO1 (C29H4)	Rb P Ser780
FOXO3a (75D8)	Rb P Ser780
FRA1 (R20)	RhoA (67BC)
Gab1	Rock1 (C8F7)
Grb2	SAPK/JNK (JNK2)
Grb2 P Y237	SHP2 P Tyr542
GSK-3-alpha/beta P Ser21/Ser9	Slug (C19G7
GSK-3-beta	Smad1/5 P Ser463/Ser465
Histone H2A.X P Ser139	Smad2/3 P Ser465/Ser423,Ser467/Ser425
HSP27 (HSPB1)	Src
HSP27 (HSPB1) P Ser78	Src (family) P Tyr416
IGF-1R beta	Stat1 P Ser727
IGF-1R beta P Tyr1162,Tyr1163	Stat3 P Ser727
IkB-alpha	Stat3 P Tyr705

IKK alpha/beta P Ser176/Ser177	Stat5
ILK1 (4G9)	Stat5 P Tyr694
Integrin Beta 1 [EP1041Y]	Tau
IRS-1	Tau Phospho/non Phos ser 305
IRS-1 P S636/639	TGF beta (56E4)
JAK1	VEGFR P Tyr1175
JAK1 P Tyr1022,Thr1023	XIAP
Ki-67 (Annexin II, p36)	YAP1 [EP1674Y]

Chapter 8 Bibliography

1. Younger NT, Wilson ML, Martinez Lyons A, Jarman EJ, Meynert AM, Grimes GR, et al. In vivo modeling of patient genetic heterogeneity identifies new ways to target cholangiocarcinoma. Cancer Res. 2022 Apr 15;82(8):1548–1559.
2. Rouiller C, editor. The Liver: Morphology, Biochemistry, Physiology. revised. Academic Press; 2013.
3. Boyer JL. Bile formation and secretion. Compr Physiol. 2013 Jul;3(3):1035–1078.
4. Rouiller CH. The Liver: Morphology, Biochemistry, Physiology [Internet]. London: Academic Press Inc. Ltd.; 1964 [cited 2021 Aug 30]. Available from: https://books.google.co.uk/books?hl=en&lr=&id=HSngBAAAQBAJ&oi=fnd&pg=PP1&dq=liver+physiology+function&ots=6XJ5R0ogdy&sig=yww3HwO6S54867c2eE-iop3OhZk&redir_esc=y#v=onepage&q&f=false
5. Lorente S, Hautefeuille M, Sanchez-Cedillo A. The liver, a functionalized vascular structure. Sci Rep. 2020 Oct 1;10(1):16194.
6. Seo W, Jeong W-I. Hepatic non-parenchymal cells: Master regulators of alcoholic liver disease? World J Gastroenterol. 2016 Jan 28;22(4):1348–1356.
7. Gao B, Jeong W-I, Tian Z. Liver: An organ with predominant innate immunity. Hepatology. 2008 Feb;47(2):729–736.
8. Vekemans K, Braet F. Structural and functional aspects of the liver and liver sinusoidal cells in relation to colon carcinoma metastasis. World J Gastroenterol. 2005 Sep 7;11(33):5095–5102.
9. Racanelli V, Rehermann B. The liver as an immunological organ. Hepatology. 2006 Feb;43(2 Suppl 1):S54–62.
10. Higashi T, Friedman SL, Hoshida Y. Hepatic stellate cells as key target in liver fibrosis. Adv Drug Deliv Rev. 2017 Nov 1;121:27–42.
11. Dranoff JA, Wells RG. Portal fibroblasts: Underappreciated mediators of biliary fibrosis. Hepatology. 2010 Apr;51(4):1438–1444.
12. Tabibian JH, Masyuk AI, Masyuk TV, O'Hara SP, LaRusso NF. Physiology of cholangiocytes. Compr Physiol. 2013 Jan;3(1):541–565.
13. Han Y, Glaser S, Meng F, Francis H, Marzioni M, McDaniel K, et al. Recent advances in the morphological and functional heterogeneity of the biliary epithelium. Exp Biol Med. 2013 May;238(5):549–565.
14. Doctor RB, Fouassier L. Emerging roles of the actin cytoskeleton in cholangiocyte function and disease. Semin Liver Dis. 2002 Aug;22(3):263–276.
15. Dave HD, Al Obaidi NM. Physiology, Biliary. StatPearls. Treasure Island (FL): StatPearls Publishing; 2018.

16. Komuta M, Govaere O, Vandecaveye V, Akiba J, Van Steenbergen W, Verslype C, et al. Histological diversity in cholangiocellular carcinoma reflects the different cholangiocyte phenotypes. Hepatology. 2012 Jun;55(6):1876–1888.

17. Cairns J. Mutation selection and the natural history of cancer. Nature. 1975 May 15;255(5505):197–200.

18. Marshall CJ. Tumor suppressor genes. Cell. 1991 Jan 25;64(2):313–326.

19. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021 May;71(3):209–249.

20. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000 Jan 7;100(1):57–70.

21. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4;144(5):646–674.

22. Cooper GM. Oncogenes. books.google.com; 1995.

23. Baxter E, Windloch K, Gannon F, Lee JS. Epigenetic regulation in cancer progression. Cell Biosci. 2014 Aug 19;4:45.

24. McDuff FKE, Turner SD. Jailbreak: oncogene-induced senescence and its evasion. Cell Signal. 2011 Jan;23(1):6–13.

25. Knudson AG. Antioncogenes and human cancer. Proc Natl Acad Sci USA. 1993 Dec 1;90(23):10914–10921.

26. Wang L-H, Wu C-F, Rajasekaran N, Shin YK. Loss of tumor suppressor gene function in human cancer: an overview. Cell Physiol Biochem. 2018 Dec 12;51(6):2647–2693.

27. Levine AJ. Tumor suppressor genes. Bioessays. 1990 Feb;12(2):60–66.

28. Preston BD, Albertson TM, Herr AJ. DNA replication fidelity and cancer. Semin Cancer Biol. 2010 Oct 1;20(5):281–293.

29. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability--an evolving hallmark of cancer. Nat Rev Mol Cell Biol. 2010 Mar;11(3):220–228.

30. Yuspa SH. Overview of carcinogenesis: past, present and future. Carcinogenesis. 2000 Mar;21(3):341–344.

31. Khan AS, Dageforde LA. Cholangiocarcinoma. Surg Clin North Am. 2019 Apr;99(2):315–335.

32. Lee DH, Lee JM. Primary malignant tumours in the non-cirrhotic liver. Eur J Radiol. 2017 Oct;95:349–361.

33. Deoliveira ML, Schulick RD, Nimura Y, Rosen C, Gores G, Neuhaus P, et al. New staging system and a registry for perihilar cholangiocarcinoma. Hepatology. 2011 Apr;53(4):1363–1371.

34. Nakanuma Y, Kakuda Y. Pathologic classification of cholangiocarcinoma: New concepts. Best Pract Res Clin Gastroenterol. 2015 Apr;29(2):277–293.

35. Nakanuma Y, Sato Y, Harada K, Sasaki M, Xu J, Ikeda H. Pathological classification of intrahepatic cholangiocarcinoma based on a new concept. World J Hepatol. 2010 Dec 27;2(12):419–427.

36. Chung YE, Kim M-J, Park YN, Choi J-Y, Pyo JY, Kim YC, et al. Varying appearances of cholangiocarcinoma: radiologic-pathologic correlation. Radiographics. 2009 Jun;29(3):683–700.

37. Kendall T, Verheij J, Gaudio E, Evert M, Guido M, Goeppert B, et al. Anatomical, histomorphological and molecular classification of cholangiocarcinoma. Liver Int. 2019 May;39 Suppl 1:7–18.

38. Liau J-Y, Tsai J-H, Yuan R-H, Chang C-N, Lee H-J, Jeng Y-M. Morphological subclassification of intrahepatic cholangiocarcinoma: etiological, clinicopathological, and molecular features. Mod Pathol. 2014 Aug;27(8):1163–1173.

39. Hayashi A, Misumi K, Shibahara J, Arita J, Sakamoto Y, Hasegawa K, et al. Distinct clinicopathologic and genetic features of 2 histologic subtypes of intrahepatic cholangiocarcinoma. Am J Surg Pathol. 2016 Aug;40(8):1021–1030.

40. Igarashi S, Sato Y, Ren XS, Harada K, Sasaki M, Nakanuma Y. Participation of peribiliary glands in biliary tract pathophysiologies. World J Hepatol. 2013 Aug 27;5(8):425–432.

41. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol. 2020 Sep;17(9):557–588.

42. Bertuccio P, Malvezzi M, Carioli G, Hashim D, Boffetta P, El-Serag HB, et al. Global trends in mortality from intrahepatic and extrahepatic cholangiocarcinoma. J Hepatol. 2019 Jul;71(1):104–114.

43. Rizvi S, Gores GJ. Molecular pathogenesis of cholangiocarcinoma. Dig Dis. 2014 Jul 14;32(5):564–569.

44. Clements O, Eliahoo J, Kim JU, Taylor-Robinson SD, Khan SA. Risk factors for intrahepatic and extrahepatic cholangiocarcinoma: A systematic review and meta-analysis. J Hepatol. 2020 Jan;72(1):95–103.

45. Banales JM, Cardinale V, Carpino G, Marzioni M, Andersen JB, Invernizzi P, 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 Apr 20;13(5):261–280.

46. Thunyaharn N, Promthet S, Wiangnon S, Suwanrungruang K, Kamsa-ard S. Survival of Cholangiocarcinoma Patients in Northeastern Thailand after Supportive Treatment. Asian Pac J Cancer Prev. 2013 Nov 30;14(11):7029–7032.

47. Khan SA, Tavolari S, Brandi G. Cholangiocarcinoma: Epidemiology and risk factors. Liver Int. 2019 May;39 Suppl 1:19–31.

48. Sithithaworn P, Yongvanit P, Duenngai K, Kiatsopit N, Pairojkul C. Roles of liver fluke infection as risk factor for cholangiocarcinoma. J Hepatobiliary Pancreat Sci. 2014 May;21(5):301–308.

49. Kamsa-Ard S, Luvira V, Suwanrungruang K, Kamsa-Ard S, Luvira V, Santong C, et al. Cholangiocarcinoma Trends, Incidence, and Relative Survival in Khon Kaen, Thailand From 1989 Through 2013: A Population-Based Cancer Registry Study. J Epidemiol. 2019 May 5;29(5):197–204.

50. Jongsuksuntigul P, Imsomboon T. Opisthorchiasis control in Thailand. Acta Trop. 2003 Nov;88(3):229–232.

51. Sithithaworn P, Andrews RH, Nguyen VD, Wongsaroj T, Sinuon M, Odermatt P, et al. The current status of opisthorchiasis and clonorchiasis in the Mekong Basin. Parasitol Int. 2012 Mar;61(1):10–16.

52. Kaewpitoon N, Kaewpitoon S-J, Pengsaa P, Sripa B. Opisthorchis viverrini: the carcinogenic human liver fluke. World J Gastroenterol. 2008 Feb 7;14(5):666–674.

53. Saffioti F, Mavroeidis VK. Review of incidence and outcomes of treatment of cholangiocarcinoma in patients with primary sclerosing cholangitis. World J Gastrointest Oncol. 2021 Oct 15;13(10):1336–1366.

54. Gupta A, Dixon E. Epidemiology and risk factors: intrahepatic cholangiocarcinoma. Hepatobiliary Surg Nutr. 2017 Apr;6(2):101–104.

55. Kubicka S, Kühnel F, Flemming P, Hain B, Kezmic N, Rudolph KL, et al. K-ras mutations in the bile of patients with primary sclerosing cholangitis. Gut. 2001 Mar;48(3):403–408.

56. Petrick JL, Yang B, Altekruse SF, Van Dyke AL, Koshiol J, Graubard BI, et al. Risk factors for intrahepatic and extrahepatic cholangiocarcinoma in the United States: A population-based study in SEER-Medicare. PLoS One. 2017 Oct 19;12(10):e0186643.

57. Jang MH, Lee YJ, Kim H. Intrahepatic cholangiocarcinoma arising in Caroli's disease. Clin Mol Hepatol. 2014 Dec;20(4):402–405.

58. DiPeri TP, Javle MM, Meric-Bernstam F. Next generation sequencing for biliary tract cancers. Expert Rev Gastroenterol Hepatol. 2021 May;15(5):471–474.

59. Chan-On W, Nairismägi M-L, Ong CK, Lim WK, Dima S, Pairojkul C, et al. Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers. Nat Genet. 2013 Dec;45(12):1474–1478.

60. Farshidfar F, Zheng S, Gingras M-C, Newton Y, Shih J, Robertson AG, et al. Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. Cell Rep. 2017 Jun 27;19(13):2878–2880.

61. Nepal C, O'Rourke CJ, Oliveira DVNP, Taranta A, Shema S, Gautam P, et al. Genomic perturbations reveal distinct regulatory networks in intrahepatic cholangiocarcinoma. Hepatology. 2018 Sep;68(3):949–963.

62. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, et al. Whole-Genome and Epigenomic Landscapes of Etiologically Distinct Subtypes of Cholangiocarcinoma. Cancer Discov. 2017 Oct;7(10):1116–1135.

63. Sia D, Losic B, Moeini A, Cabellos L, Hao K, Revill K, et al. Massive parallel sequencing uncovers actionable FGFR2-PPHLN1 fusion and ARAF mutations in intrahepatic cholangiocarcinoma. Nat Commun. 2015 Jan 22;6:6087.

64. Wu Y-M, Su F, Kalyana-Sundaram S, Khazanov N, Ateeq B, Cao X, et al. Identification of targetable FGFR gene fusions in diverse cancers. Cancer Discov. 2013 Jun;3(6):636–647.

65. Ross JS, Wang K, Gay L, Al-Rohil R, Rand JV, Jones DM, et al. New routes to targeted therapy of intrahepatic cholangiocarcinomas revealed by next-generation sequencing. Oncologist. 2014 Mar;19(3):235–242.

66. Borad MJ, Champion MD, Egan JB, Liang WS, Fonseca R, Bryce AH, et al. Integrated genomic characterization reveals novel, therapeutically relevant drug targets in FGFR and EGFR pathways in sporadic intrahepatic cholangiocarcinoma. PLoS Genet. 2014 Feb 13;10(2):e1004135.

67. Wang P, Dong Q, Zhang C, Kuan PF, Liu Y, Jeck WR, et al. Mutations in isocitrate dehydrogenase 1 and 2 occur frequently in intrahepatic cholangiocarcinomas and share hypermethylation targets with glioblastomas. Oncogene. 2013 Jun 20;32(25):3091–3100.

68. Javle M, Bekaii-Saab T, Jain A, Wang Y, Kelley RK, Wang K, et al. Biliary cancer: Utility of next-generation sequencing for clinical management. Cancer. 2016 Dec 15;122(24):3838–3847.

69. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, et al. Genomic spectra of biliary tract cancer. Nat Genet. 2015 Sep;47(9):1003–1010.

70. Lowery MA, Ptashkin R, Jordan E, Berger MF, Zehir A, Capanu M, et al. Comprehensive molecular profiling of intrahepatic and extrahepatic cholangiocarcinomas: potential targets for intervention. Clin Cancer Res. 2018 Sep 1;24(17):4154–4161.

71. Zhu AX, Borger DR, Kim Y, Cosgrove D, Ejaz A, Alexandrescu S, et al. Genomic profiling of intrahepatic cholangiocarcinoma: refining prognosis and identifying therapeutic targets. Ann Surg Oncol. 2014 Nov;21(12):3827–3834.

72. Arai Y, Totoki Y, Hosoda F, Shirota T, Hama N, Nakamura H, et al. Fibroblast growth factor receptor 2 tyrosine kinase fusions define a unique molecular subtype of cholangiocarcinoma. Hepatology. 2014 Apr;59(4):1427–1434.

73. Kipp BR, Voss JS, Kerr SE, Barr Fritcher EG, Graham RP, Zhang L, et al. Isocitrate dehydrogenase 1 and 2 mutations in cholangiocarcinoma. Hum Pathol. 2012 Oct;43(10):1552–1558.

74. Andersen JB, Thorgeirsson SS. Genetic profiling of intrahepatic cholangiocarcinoma. Curr Opin Gastroenterol. 2012 May;28(3):266–272.

75. Braconi C, Roessler S, Kruk B, Lammert F, Krawczyk M, Andersen JB. Molecular perturbations in cholangiocarcinoma: Is it time for precision medicine? Liver Int. 2019 May;39 Suppl 1:32–42.

76. Alberts R, de Vries EMG, Goode EC, Jiang X, Sampaziotis F, Rombouts K, et al. Genetic association analysis identifies variants associated with disease progression in primary sclerosing cholangitis. Gut. 2018 Aug;67(8):1517–1524.

77. Melum E, Franke A, Schramm C, Weismüller TJ, Gotthardt DN, Offner FA, et al. Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. Nat Genet. 2011 Jan;43(1):17–19.

78. Meek DW, Anderson CW. Posttranslational modification of p53: cooperative integrators of function. Cold Spring Harb Perspect Biol. 2009 Dec;1(6):a000950.

79. Oren M. Decision making by p53: life, death and cancer. Cell Death Differ. 2003 Apr;10(4):431–442.

80. Robertson S, Hyder O, Dodson R, Nayar SK, Poling J, Beierl K, et al. The frequency of KRAS and BRAF mutations in intrahepatic cholangiocarcinomas and their correlation with clinical outcome. Hum Pathol. 2013 Dec;44(12):2768–2773.

81. Zou S, Li J, Zhou H, Frech C, Jiang X, Chu JSC, et al. Mutational landscape of intrahepatic cholangiocarcinoma. Nat Commun. 2014 Dec 15;5:5696.

82. Ong CK, Subimerb C, Pairojkul C, Wongkham S, Cutcutache I, Yu W, et al. Exome sequencing of liver fluke-associated cholangiocarcinoma. Nat Genet. 2012 May 6;44(6):690–693.

83. Farshidfar F, Zheng S, Gingras M-C, Newton Y, Shih J, Robertson AG, et al. Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles. Cell Rep. 2017 Mar 14;18(11):2780–2794.

84. Sia D, Hoshida Y, Villanueva A, Roayaie S, Ferrer J, Tabak B, et al. Integrative molecular analysis of intrahepatic cholangiocarcinoma reveals 2 classes that have different outcomes. Gastroenterology. 2013 Apr;144(4):829–840.

85. Tannapfel A, Benicke M, Katalinic A, Uhlmann D, Köckerling F, Hauss J, et al. Frequency of p16(INK4A) alterations and K-ras mutations in intrahepatic cholangiocarcinoma of the liver. Gut. 2000 Nov;47(5):721–727.

86. Dong L, Lu D, Chen R, Lin Y, Zhu H, Zhang Z, et al. Proteogenomic characterization identifies clinically relevant subgroups of intrahepatic cholangiocarcinoma. Cancer Cell. 2021 Dec 28;

87. Sia D, Tovar V, Moeini A, Llovet JM. Intrahepatic cholangiocarcinoma: pathogenesis and rationale for molecular therapies. Oncogene. 2013 Oct 10;32(41):4861–4870.

88. Turnpenny PD, Ellard S. Alagille syndrome: pathogenesis, diagnosis and management. Eur J Hum Genet. 2012 Mar;20(3):251–257.

89. Geisler F, Strazzabosco M. Emerging roles of Notch signaling in liver disease. Hepatology. 2015 Jan;61(1):382–392.

90. Sekiya S, Suzuki A. Intrahepatic cholangiocarcinoma' ' can arise from Notch-mediated conversion of hepatocytes. J Clin Invest. 2012 Nov;122(11):3914–3918.

91. Fan B, Malato Y, Calvisi DF, Naqvi S, Razumilava N, Ribback S, et al. Cholangiocarcinomas can originate from hepatocytes in mice. J Clin Invest. 2012 Aug;122(8):2911–2915.

92. Wang J, Dong M, Xu Z, Song X, Zhang S, Qiao Y, et al. Notch2 controls hepatocyte-derived cholangiocarcinoma formation in mice. Oncogene. 2018 Jun;37(24):3229–3242.

93. Anastas JN, Moon RT. WNT signalling pathways as therapeutic targets in cancer. Nat Rev Cancer. 2013 Jan;13(1):11–26.

94. Niessen CM, Gottardi CJ. Molecular components of the adherens junction. Biochim Biophys Acta. 2008 Mar;1778(3):562–571.

95. Valenta T, Hausmann G, Basler K. The many faces and functions of β-catenin. EMBO J. 2012 Jun 13;31(12):2714–2736.

96. Kimelman D, Xu W. beta-catenin destruction complex: insights and questions from a structural perspective. Oncogene. 2006 Dec 4;25(57):7482–7491.

97. Dajani R, Fraser E, Roe SM, Yeo M, Good VM, Thompson V, et al. Structural basis for recruitment of glycogen synthase kinase 3beta to the axin-APC scaffold complex. EMBO J. 2003 Feb 3;22(3):494–501.

98. Spink KE, Polakis P, Weis WI. Structural basis of the Axin-adenomatous polyposis coli interaction. EMBO J. 2000 May 15;19(10):2270–2279.

99. Ha N-C, Tonozuka T, Stamos JL, Choi H-J, Weis WI. Mechanism of phosphorylation-dependent binding of APC to beta-catenin and its role in beta-catenin degradation. Mol Cell. 2004 Aug 27;15(4):511–521.

100. Liu J, Xing Y, Hinds TR, Zheng J, Xu W. The third 20 amino acid repeat is the tightest binding site of APC for beta-catenin. J Mol Biol. 2006 Jun 30;360(1):133–144.

101. Latres E, Chiaur DS, Pagano M. The human F box protein beta-Trcp associates with the Cul1/Skp1 complex and regulates the stability of beta-catenin. Oncogene. 1999 Jan 28;18(4):849–854.

102. Cong F, Schweizer L, Varmus H. Wnt signals across the plasma membrane to activate the beta-catenin pathway by forming oligomers containing its receptors, Frizzled and LRP. Development. 2004 Oct;131(20):5103–5115.

103. Xu W, Kimelman D. Mechanistic insights from structural studies of beta-catenin and its binding partners. J Cell Sci. 2007 Oct 1;120(Pt 19):3337–3344.

104. Lien W-H, Fuchs E. Wnt some lose some: transcriptional governance of stem cells by Wnt/β-catenin signaling. Genes Dev. 2014 Jul 15;28(14):1517–1532.

105. Yothaisong S, Thanee M, Namwat N, Yongvanit P, Boonmars T, Puapairoj A, et al. Opisthorchis viverrini infection activates the PI3K/ AKT/PTEN and Wnt/β-catenin signaling pathways in a Cholangiocarcinogenesis model. Asian Pac J Cancer Prev. 2014;15(23):10463–10468.

106. Tokumoto N, Ikeda S, Ishizaki Y, Kurihara T, Ozaki S, Iseki M, et al. Immunohistochemical and mutational analyses of Wnt signaling components and target genes in intrahepatic cholangiocarcinomas. Int J Oncol. 2005 Oct 1;

107. Boulter L, Guest RV, Kendall TJ, Wilson DH, Wojtacha D, Robson AJ, et al. WNT signaling drives cholangiocarcinoma growth and can be pharmacologically inhibited. J Clin Invest. 2015 Mar 2;

108. Loilome W, Bungkanjana P, Techasen A, Namwat N, Yongvanit P, Puapairoj A, et al. Activated macrophages promote Wnt/β-catenin signaling in cholangiocarcinoma cells. Tumour Biol. 2014 Jun;35(6):5357–5367.

109. Perugorria MJ, Olaizola P, Labiano I, Esparza-Baquer A, Marzioni M, Marin JJG, et al. Wnt-β-catenin signalling in liver development, health and disease. Nat Rev Gastroenterol Hepatol. 2019 Feb;16(2):121–136.

110. Goeppert B, Konermann C, Schmidt CR, Bogatyrova O, Geiselhart L, Ernst C, et al. Global alterations of DNA methylation in cholangiocarcinoma target the Wnt signaling pathway. Hepatology. 2014 Feb;59(2):544–554.

111. Merino-Azpitarte M, Lozano E, Perugorria MJ, Esparza-Baquer A, Erice O, Santos-Laso Á, et al. SOX17 regulates cholangiocyte differentiation and acts as a tumor suppressor in cholangiocarcinoma. J Hepatol. 2017 Jul;67(1):72–83.

112. Moeini A, Sia D, Bardeesy N, Mazzaferro V, Llovet JM. Molecular pathogenesis and targeted therapies for intrahepatic cholangiocarcinoma. Clin Cancer Res. 2016 Jan 15;22(2):291–300.

113. Rajalingam K, Schreck R, Rapp UR, Albert S. Ras oncogenes and their downstream targets. Biochim Biophys Acta. 2007 Aug;1773(8):1177–1195.

114. Gideon P, John J, Frech M, Lautwein A, Clark R, Scheffler JE, et al. Mutational and kinetic analyses of the GTPase-activating protein (GAP)-p21 interaction: the C-terminal domain of GAP is not sufficient for full activity. Mol Cell Biol. 1992 May;12(5):2050–2056.

115. Scheffzek K, Ahmadian MR, Kabsch W, Wiesmüller L, Lautwein A, Schmitz F, et al. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. Science. 1997 Jul 18;277(5324):333–338.

116. Cox AD, Der CJ. Ras history: The saga continues. Small GTPases. 2010;1(1):2–27.

117. Hemmings BA, Restuccia DF. PI3K-PKB/Akt pathway. Cold Spring Harb Perspect Biol. 2012 Sep 1;4(9):a011189.

118. Osaki M, Oshimura M, Ito H. PI3K-Akt pathway: its functions and alterations in human cancer. Apoptosis. 2004 Nov;9(6):667–676.

119. Kim W, Jho E-H. The history and regulatory mechanism of the Hippo pathway. BMB Rep. 2018 Mar;51(3):106–118.

120. Evans DGR. Neurofibromatosis type 2 (NF2): a clinical and molecular review. Orphanet J Rare Dis. 2009 Jun 19;4:16.

121. Bao Y, Hata Y, Ikeda M, Withanage K. Mammalian Hippo pathway: from development to cancer and beyond. J Biochem. 2011 Apr;149(4):361–379.

122. Kim M-K, Jang J-W, Bae S-C. DNA binding partners of YAP/TAZ. BMB Rep. 2018 Mar;51(3):126–133.

123. Benhamouche S, Curto M, Saotome I, Gladden AB, Liu C-H, Giovannini M, et al. Nf2/Merlin controls progenitor homeostasis and tumorigenesis in the liver. Genes Dev. 2010 Aug 15;24(16):1718–1730.

124. Yimlamai D, Christodoulou C, Galli GG, Yanger K, Pepe-Mooney B, Gurung B, et al. Hippo pathway activity influences liver cell fate. Cell. 2014 Jun 5;157(6):1324–1338.

125. Woo HG, Lee J-H, Yoon J-H, Kim CY, Lee H-S, Jang JJ, et al. Identification of a cholangiocarcinoma-like gene expression trait in hepatocellular carcinoma. Cancer Res. 2010 Apr 15;70(8):3034–3041.

126. Bridgewater J, Galle PR, Khan SA, Llovet JM, Park J-W, Patel T, et al. Guidelines for the diagnosis and management of intrahepatic cholangiocarcinoma. J Hepatol. 2014 Jun;60(6):1268–1289.

127. Sirica AE, Gores GJ. Desmoplastic stroma and cholangiocarcinoma: clinical implications and therapeutic targeting. Hepatology. 2014 Jun;59(6):2397–2402.

128. Tamma R, Annese T, Ruggieri S, Brunetti O, Longo V, Cascardi E, et al. Inflammatory cells infiltrate and angiogenesis in locally advanced and metastatic cholangiocarcinoma. Eur J Clin Invest. 2019 May;49(5):e13087.

129. Szendröi M, Lapis K. Distribution of fibronectin and laminin in human liver tumors. J Cancer Res Clin Oncol. 1985;109(1):60–64.

130. Lorenzini S, Bird TG, Boulter L, Bellamy C, Samuel K, Aucott R, et al. Characterisation of a stereotypical cellular and extracellular adult liver progenitor cell niche in rodents and diseased human liver. Gut. 2010 May;59(5):645–654.

131. Spolverato G, Vitale A, Cucchetti A, Popescu I, Marques HP, Aldrighetti L, et al. Can hepatic resection provide a long-term cure for patients with intrahepatic cholangiocarcinoma? Cancer. 2015 Nov 15;121(22):3998–4006.

132. Forner A, Vidili G, Rengo M, Bujanda L, Ponz-Sarvisé M, Lamarca A. Clinical presentation, diagnosis and staging of cholangiocarcinoma. Liver Int. 2019 May;39 Suppl 1:98–107.

133. Mazzaferro V, Gorgen A, Roayaie S, Droz Dit Busset M, Sapisochin G. Liver resection and transplantation for intrahepatic cholangiocarcinoma. J Hepatol. 2020 Feb;72(2):364–377.

134. Ejaz A, Cloyd JM, Pawlik TM. Advances in the Diagnosis and Treatment of Patients with Intrahepatic Cholangiocarcinoma. Ann Surg Oncol. 2020 Feb;27(2):552–560.

135. Iavarone M, Piscaglia F, Vavassori S, Galassi M, Sangiovanni A, Venerandi L, et al. Contrast enhanced CT-scan to diagnose intrahepatic cholangiocarcinoma in patients with cirrhosis. J Hepatol. 2013 Jun;58(6):1188–1193.

136. Kim SH, Lee CH, Kim BH, Kim WB, Yeom SK, Kim KA, et al. Typical and atypical imaging findings of intrahepatic cholangiocarcinoma using gadolinium ethoxybenzyl diethylenetriamine pentaacetic acid-enhanced magnetic resonance imaging. J Comput Assist Tomogr. 2012 Dec;36(6):704–709.

137. Sinakos E, Saenger AK, Keach J, Kim WR, Lindor KD. Many patients with primary sclerosing cholangitis and increased serum levels of carbohydrate antigen 19-9 do not have cholangiocarcinoma. Clin Gastroenterol Hepatol. 2011 May 1;9(5):434–9.e1.

138. Rompianesi G, Di Martino M, Gordon-Weeks A, Montalti R, Troisi R. Liquid biopsy in cholangiocarcinoma: Current status and future perspectives. World J Gastrointest Oncol. 2021 May 15;13(5):332–350.

139. Andersen RF, Jakobsen A. Screening for circulating RAS/RAF mutations by multiplex digital PCR. Clin Chim Acta. 2016 May 12;458:138–143.

140. de Jong MC, Nathan H, Sotiropoulos GC, Paul A, Alexandrescu S, Marques H, et al. Intrahepatic cholangiocarcinoma: an international multi-institutional analysis

of prognostic factors and lymph node assessment. J Clin Oncol. 2011 Aug 10;29(23):3140–3145.

141. Si A, Li J, Yang Z, Xia Y, Yang T, Lei Z, et al. Impact of Anatomical Versus Non-anatomical Liver Resection on Short- and Long-Term Outcomes for Patients with Intrahepatic Cholangiocarcinoma. Ann Surg Oncol. 2019 Jun;26(6):1841–1850.

142. Edeline J, Benabdelghani M, Bertaut A, Watelet J, Hammel P, Joly J-P, et al. Gemcitabine and Oxaliplatin Chemotherapy or Surveillance in Resected Biliary Tract Cancer (PRODIGE 12-ACCORD 18-UNICANCER GI): A Randomized Phase III Study. J Clin Oncol. 2019 Mar 10;37(8):658–667.

143. Primrose JN, Fox RP, Palmer DH, Malik HZ, Prasad R, Mirza D, et al. Capecitabine compared with observation in resected biliary tract cancer (BILCAP): a randomised, controlled, multicentre, phase 3 study. Lancet Oncol. 2019 May;20(5):663–673.

144. Meyer CG, Penn I, James L. Liver transplantation for cholangiocarcinoma: results in 207 patients. Transplantation. 2000 Apr 27;69(8):1633–1637.

145. Robles R, Figueras J, Turrión VS, Margarit C, Moya A, Varo E, et al. Spanish experience in liver transplantation for hilar and peripheral cholangiocarcinoma. Ann Surg. 2004 Feb 1;239(2):265–271.

146. Darwish Murad S, Kim WR, Harnois DM, Douglas DD, Burton J, Kulik LM, et al. Efficacy of neoadjuvant chemoradiation, followed by liver transplantation, for perihilar cholangiocarcinoma at 12 US centers. Gastroenterology. 2012 Jul;143(1):88–98.e3; quiz e14.

147. Sapisochin G, Facciuto M, Rubbia-Brandt L, Marti J, Mehta N, Yao FY, et al. Liver transplantation for “very early” intrahepatic cholangiocarcinoma: International retrospective study supporting a prospective assessment. Hepatology. 2016 Oct;64(4):1178–1188.

148. Sapisochin G, Fidelman N, Roberts JP, Yao FY. Mixed hepatocellular cholangiocarcinoma and intrahepatic cholangiocarcinoma in patients undergoing transplantation for hepatocellular carcinoma. Liver Transpl. 2011 Aug;17(8):934–942.

149. Lunsford KE, Javle M, Heyne K, Shroff RT, Abdel-Wahab R, Gupta N, et al. Liver transplantation for locally advanced intrahepatic cholangiocarcinoma treated with neoadjuvant therapy: a prospective case-series. Lancet Gastroenterol Hepatol. 2018 May;3(5):337–348.

150. Weigt J, Malfertheiner P. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. Expert Rev Gastroenterol Hepatol. 2010 Aug;4(4):395–397.

151. Okusaka T, Nakachi K, Fukutomi A, Mizuno N, Ohkawa S, Funakoshi A, et al. Gemcitabine alone or in combination with cisplatin in patients with biliary tract

cancer: a comparative multicentre study in Japan. Br J Cancer. 2010 Aug 10;103(4):469–474.

152. Bergman AM, Ruiz van Haperen VW, Veerman G, Kuiper CM, Peters GJ. Synergistic interaction between cisplatin and gemcitabine in vitro. Clin Cancer Res. 1996 Mar;2(3):521–530.

153. Shroff RT, Javle MM, Xiao L, Kaseb AO, Varadhachary GR, Wolff RA, et al. Gemcitabine, Cisplatin, and nab-Paclitaxel for the Treatment of Advanced Biliary Tract Cancers: A Phase 2 Clinical Trial. JAMA Oncol. 2019 Jun 1;5(6):824–830.

154. Phelip JM, Desrame J, Edeline J, Barbier E, Terrebonne E, Michel P, et al. Modified FOLFIRINOX versus CISGEM chemotherapy for patients with advanced biliary tract cancer (PRODIGE 38 AMEBICA): A randomized phase II study. J Clin Oncol. 2021 Oct 18;JCO2100679.

155. Mizusawa J, Morizane C, Okusaka T, Katayama H, Ishii H, Fukuda H, et al. Randomized Phase III study of gemcitabine plus S-1 versus gemcitabine plus cisplatin in advanced biliary tract cancer: Japan Clinical Oncology Group Study (JCOG1113, FUGA-BT). Jpn J Clin Oncol. 2016 Apr;46(4):385–388.

156. McNamara MG, Bridgewater J, Palmer DH, Faluyi O, Wasan H, Patel A, et al. A Phase Ib Study of NUC-1031 in Combination with Cisplatin for the First-Line Treatment of Patients with Advanced Biliary Tract Cancer (ABC-08). Oncologist. 2021 Apr;26(4):e669–e678.

157. Lamarca A, Palmer DH, Wasan HS, Ross PJ, Ma YT, Arora A, et al. ABC-06 | A randomised phase III, multi-centre, open-label study of active symptom control (ASC) alone or ASC with oxaliplatin / 5-FU chemotherapy (ASC+mFOLFOX) for patients (pts) with locally advanced / metastatic biliary tract cancers (ABC) previously-treated with cisplatin/gemcitabine (CisGem) chemotherapy. JCO. 2019 May 20;37(15_suppl):4003–4003.

158. Cigliano A, Wang J, Chen X, Calvisi DF. Role of the Notch signaling in cholangiocarcinoma. Expert Opin Ther Targets. 2017 May;21(5):471–483.

159. Boulter L, Guest RV, Kendall TJ, Wilson DH, Wojtacha D, Robson AJ, et al. WNT signalling drives cholangiocarcinoma growth and can be pharmacologically inhibited. The Journal of Clinical Investigation2. 2014;125(3):1269–1285.

160. Popovici-Muller J, Lemieux RM, Artin E, Saunders JO, Salituro FG, Travins J, et al. Discovery of AG-120 (Ivosidenib): A First-in-Class Mutant IDH1 Inhibitor for the Treatment of IDH1 Mutant Cancers. ACS Med Chem Lett. 2018 Apr 12;9(4):300–305.

161. Efficacy of TAS-120, an Irreversible Fibroblast Growth Factor Receptor Inhibitor (FGFRi), in Patients With Cholangiocarcinoma and FGFR Pathway Alte... | OncologyPRO [Internet]. [cited 2022 Jan 9]. Available from: https://oncologypro.esmo.org/meeting-resources/esmo-asia-2018-congress/Efficacy-of-TAS-120-an-Irreversible-Fibroblast-Growth-Factor-

Receptor-Inhibitor-FGFRi-in-Patients-With-Cholangiocarcinoma-and-FGFR-Pathway-Alterations-Previously-Treated-With-Chemotherapy-and-Other-FGFRi-s

162. Javle M, Lowery M, Shroff RT, Weiss KH, Springfeld C, Borad MJ, et al. Phase II Study of BGJ398 in Patients With FGFR-Altered Advanced Cholangiocarcinoma. J Clin Oncol. 2018 Jan 20;36(3):276–282.

163. Javle MM, Shroff RT, Zhu A, Sadeghi S, Choo S, Borad MJ, et al. A phase 2 study of BGJ398 in patients (pts) with advanced or metastatic FGFR-altered cholangiocarcinoma (CCA) who failed or are intolerant to platinum-based chemotherapy. J Clin Oncol. 2016 Feb;34(4_suppl):335–335.

164. Mazzaferro V, Shaib W, Rimassa L, Harris W, Personeni N, El-Rayes B, et al. PD-019 ARQ 087, an oral pan- fibroblast growth factor receptor (FGFR) inhibitor, in patients (pts) with advanced and/or metastatic intrahepatic cholangiocarcinoma (iCCA). Ann Oncol. 2016 Jun;27:ii109.

165. Le Rhun A, Escalera-Maurer A, Bratovič M, Charpentier E. CRISPR-Cas in Streptococcus pyogenes. RNA Biol. 2019 Apr;16(4):380–389.

166. Qi X, Zhang J, Zhao Y, Chen T, Xiang Y, Hui J, et al. The applications of CRISPR screen in functional genomics. Brief Funct Genomics. 2017 Jan;16(1):34–37.

167. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. Science. 2014 Jan 3;343(6166):80–84.

168. Zhou Y, Zhu S, Cai C, Yuan P, Li C, Huang Y, et al. High-throughput screening of a CRISPR/Cas9 library for functional genomics in human cells. Nature. 2014 May 22;509(7501):487–491.

169. Chow RD, Guzman CD, Wang G, Schmidt F, Youngblood MW, Ye L, et al. AAV-mediated direct in vivo CRISPR screen identifies functional suppressors in glioblastoma. Nat Neurosci. 2017 Oct;20(10):1329–1341.

170. Serebrenik YV, Shalem O. CRISPR mutagenesis screening of mice. Nat Cell Biol. 2018 Nov;20(11):1235–1237.

171. Zhang G, Budker V, Wolff JA. High levels of foreign gene expression in hepatocytes after tail vein injections of naked plasmid DNA. Hum Gene Ther. 1999 Jul 1;10(10):1735–1737.

172. Sebestyén MG, Budker VG, Budker T, Subbotin VM, Zhang G, Monahan SD, et al. Mechanism of plasmid delivery by hydrodynamic tail vein injection. I. Hepatocyte uptake of various molecules. J Gene Med. 2006 Jul;8(7):852–873.

173. Weber J, Öllinger R, Friedrich M, Ehmer U, Barenboim M, Steiger K, et al. CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. Proc Natl Acad Sci USA. 2015 Nov 10;112(45):13982–13987.

174. Xue W, Chen S, Yin H, Tammela T, Papagiannakopoulos T, Joshi NS, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature. 2014 Oct 16;514(7522):380–384.

175. Read A, Gao S, Batchelor E, Luo J. Flexible CRISPR library construction using parallel oligonucleotide retrieval. Nucleic Acids Res. 2017 Jun 20;45(11):e101.

176. Mou H, Ozata DM, Smith JL, Sheel A, Kwan S-Y, Hough S, et al. CRISPR-SONIC: targeted somatic oncogene knock-in enables rapid in vivo cancer modeling. Genome Med. 2019 Apr 16;11(1):21.

177. QuPath | Quantitative Pathology & Bioimage Analysis [Internet]. [cited 2022 May 10]. Available from: https://qupath.github.io/

178. Ewels P, Hammarén R, Peltzer A, Moreno D, Rfenouil, Garcia M, et al. nf-core/rnaseq: nf-core/rnaseq version 1.3. Zenodo. 2019;

179. IntOGen - Cancer Mutations Browser [Internet]. [cited 2022 Apr 18]. Available from: https://www.intogen.org/search

180. Shibuya K, Obayashi I, Asakawa S, Minoshima S, Kudoh J, Shimizu N. A cluster of 21 keratin-associated protein genes within introns of another gene on human chromosome 21q22.3. Genomics. 2004 Apr;83(4):679–693.

181. Prieto C, Nguyen DTT, Liu Z, Wheat J, Perez A, Gourkanti S, et al. Transcriptional control of CBX5 by the RNA binding proteins RBMX and RBMXL1 maintains chromatin state in myeloid leukemia. Nat Cancer. 2021 Jul 5;2:741–757.

182. Zendman AJW, Zschocke J, van Kraats AA, de Wit NJW, Kurpisz M, Weidle UH, et al. The human SPANX multigene family: genomic organization, alignment and expression in male germ cells and tumor cell lines. Gene. 2003 May;309(2):125–133.

183. Cheung H-H, Yang Y, Lee T-L, Rennert O, Chan W-Y. Hypermethylation of genes in testicular embryonal carcinomas. Br J Cancer. 2016 Jan 19;114(2):230–236.

184. Yang P, Wang Y, Macfarlan TS. The Role of KRAB-ZFPs in Transposable Element Repression and Mammalian Evolution. Trends Genet. 2017 Nov;33(11):871–881.

185. Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR screening. Nat Methods. 2014 Aug;11(8):783–784.

186. Vatansever S, Erman B, Gümüş ZH. Oncogenic G12D mutation alters local conformations and dynamics of K-Ras. Sci Rep. 2019 Aug 13;9(1):11730.

187. Nakajima T, Tajima Y, Sugano I, Nagao K, Kondo Y, Wada K. Intrahepatic cholangiocarcinoma with sarcomatous change. Clinicopathologic and immunohistochemical evaluation of seven cases. Cancer. 1993 Sep 15;72(6):1872–1877.

188. Aishima S, Kuroda Y, Asayama Y, Taguchi K, Nishihara Y, Taketomi A, et al. Prognostic impact of cholangiocellular and sarcomatous components in combined hepatocellular and cholangiocarcinoma. Hum Pathol. 2006 Mar;37(3):283–291.

189. Gentilini A, Pastore M, Marra F, Raggi C. The Role of Stroma in Cholangiocarcinoma: The Intriguing Interplay between Fibroblastic Component, Immune Cell Subsets and Tumor Epithelium. Int J Mol Sci. 2018 Sep 22;19(10).

190. Chuaysri C, Thuwajit P, Paupairoj A, Chau-In S, Suthiphongchai T, Thuwajit C. Alpha-smooth muscle actin-positive fibroblasts promote biliary cell proliferation and correlate with poor survival in cholangiocarcinoma. Oncol Rep. 2009 Apr;21(4):957–969.

191. Erdogan B, Ao M, White LM, Means AL, Brewer BM, Yang L, et al. Cancer-associated fibroblasts promote directional cancer cell migration by aligning fibronectin. J Cell Biol. 2017 Nov 6;216(11):3799–3816.

192. Brabletz T, Kalluri R, Nieto MA, Weinberg RA. EMT in cancer. Nat Rev Cancer. 2018 Feb;18(2):128–134.

193. Borger DR, Tanabe KK, Fan KC, Lopez HU, Fantin VR, Straley KS, et al. Frequent mutation of isocitrate dehydrogenase (IDH)1 and IDH2 in cholangiocarcinoma identified through broad-based tumor genotyping. Oncologist. 2012;17(1):72–79.

194. Coppé J-P, Patil CK, Rodier F, Sun Y, Muñoz DP, Goldstein J, et al. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. PLoS Biol. 2008 Dec 2;6(12):2853–2868.

195. Kang T-W, Yevsa T, Woller N, Hoenicke L, Wuestefeld T, Dauch D, et al. Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. Nature. 2011 Nov 9;479(7374):547–551.

196. Hayflick L, Moorhead PS. The serial cultivation of human diploid cell strains. Exp Cell Res. 1961 Dec;25:585–621.

197. Drescher U. Eph family functions from an evolutionary perspective. Curr Opin Genet Dev. 2002 Aug;12(4):397–402.

198. Sheng Y, Wei J, Zhang Y, Gao X, Wang Z, Yang J, et al. Mutated EPHA2 is a target for combating lymphatic metastasis in intrahepatic cholangiocarcinoma. Int J Cancer. 2019 May 15;144(10):2440–2452.

199. Cui X-D, Lee M-J, Kim J-H, Hao P-P, Liu L, Yu G-R, et al. Activation of mammalian target of rapamycin complex 1 (mTORC1) and Raf/Pyk2 by growth factor-mediated Eph receptor 2 (EphA2) is required for cholangiocarcinoma growth and metastasis. Hepatology. 2013 Jun;57(6):2248–2260.

200. Bryant HE, Schultz N, Thomas HD, Parker KM, Flower D, Lopez E, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature. 2005 Apr 14;434(7035):913–917.

201. Foulkes WD, Flanders TY, Pollock PM, Hayward NK. The CDKN2A (p16) gene and human cancer. Mol Med. 1997 Jan;3(1):5–20.

202. Joseph NM, Tsokos CG, Umetsu SE, Shain AH, Kelley RK, Onodera C, et al. Genomic profiling of combined hepatocellular-cholangiocarcinoma reveals similar genetics to hepatocellular carcinoma. J Pathol. 2019 Jun;248(2):164–178.

203. Jang S, Chun S-M, Hong S-M, Sung CO, Park H, Kang HJ, et al. High throughput molecular profiling reveals differential mutation patterns in intrahepatic cholangiocarcinomas arising in chronic advanced liver diseases. Mod Pathol. 2014 May;27(5):731–739.

204. Churi CR, Shroff R, Wang Y, Rashid A, Kang HC, Weatherly J, et al. Mutation profiling in cholangiocarcinoma: prognostic and therapeutic implications. PLoS One. 2014 Dec 23;9(12):e115383.

205. Oeckinghaus A, Ghosh S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harb Perspect Biol. 2009 Oct;1(4):a000034.

206. Wu C-J, Conze DB, Li T, Srinivasula SM, Ashwell JD. Sensing of Lys 63-linked polyubiquitination by NEMO is a key event in NF-kappaB activation [corrected]. Nat Cell Biol. 2006 Apr;8(4):398–406.

207. Joo D, Tang Y, Blonska M, Jin J, Zhao X, Lin X. Regulation of Linear Ubiquitin Chain Assembly Complex by Caspase-Mediated Cleavage of RNF31. Mol Cell Biol. 2016 Dec 15;36(24):3010–3018.

208. Li S, Zheng X, Hu Y, You K, Wang J. RNF31 mediated ubiquitination of A20 aggravates inflammation and hepatocyte apoptosis through the TLR4/MyD88/NF-κB signaling pathway. Chem Biol Interact. 2021 Oct 1;348:109623.

209. Zhu J, Zhao C, Kharman-Biz A, Zhuang T, Jonsson P, Liang N, et al. The atypical ubiquitin ligase RNF31 stabilizes estrogen receptor α and modulates estrogen-stimulated breast cancer cell proliferation. Oncogene. 2014 Aug 21;33(34):4340–4351.

210. Hörlein AJ, Näär AM, Heinzel T, Torchia J, Gloss B, Kurokawa R, et al. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature. 1995 Oct 5;377(6548):397–404.

211. Li P, Fan W, Xu J, Lu M, Yamamoto H, Auwerx J, et al. Adipocyte NCoR knockout decreases PPARγ phosphorylation and enhances PPARγ activity and insulin sensitivity. Cell. 2011 Nov 11;147(4):815–826.

212. Yamamoto H, Williams EG, Mouchiroud L, Cantó C, Fan W, Downes M, et al. NCoR1 is a conserved physiological modulator of muscle mass and oxidative function. Cell. 2011 Nov 11;147(4):827–839.

213. Ahad A, Stevanin M, Smita S, Mishra GP, Gupta D, Waszak S, et al. Ncor1: putting the brakes on the dendritic cell immune tolerance. iScience. 2019 Sep 27;19:996–1011.

214. Oppi S, Nusser-Stein S, Blyszczuk P, Wang X, Jomard A, Marzolla V, et al. Macrophage NCOR1 protects from atherosclerosis by repressing a pro-atherogenic PPARγ signature. Eur Heart J. 2020 Mar 1;41(9):995–1005.

215. Hainberger D, Stolz V, Zhu C, Schuster M, Müller L, Hamminger P, et al. NCOR1 orchestrates transcriptional landscapes and effector functions of CD4+ T cells. Front Immunol. 2020 Apr 3;11:579.

216. Zhang Z, Yamashita H, Toyama T, Sugiura H, Ando Y, Mita K, et al. NCOR1 mRNA is an independent prognostic factor for breast cancer. Cancer Lett. 2006 Jun 8;237(1):123–129.

217. Girault I, Lerebours F, Amarir S, Tozlu S, Tubiana-Hulin M, Lidereau R, et al. Expression Analysis of Estrogen Receptor α Coregulators in Breast Carcinoma | Clinical Cancer Research | American Association for Cancer Research. Clin Cancer Res. 2003 Apr 1;

218. Lavinsky RM, Jepsen K, Heinzel T, Torchia J, Mullen TM, Schiff R, et al. Diverse signaling pathways modulate nuclear receptor recruitment of N-CoR and SMRT complexes. Proc Natl Acad Sci USA. 1998 Mar 17;95(6):2920–2925.

219. Khanim FL, Gommersall LM, Wood VHJ, Smith KL, Montalvo L, O'Neill LP, et al. Altered SMRT levels disrupt vitamin D3 receptor signalling in prostate cancer cells. Oncogene. 2004 Sep 2;23(40):6712–6725.

220. Abedin SA, Thorne JL, Battaglia S, Maguire O, Hornung LB, Doherty AP, et al. Elevated NCOR1 disrupts a network of dietary-sensing nuclear receptors in bladder cancer cells. Carcinogenesis. 2009 Mar;30(3):449–456.

221. Ohta K, Mizutani A, Kawakami A, Murakami Y, Kasuya Y, Takagi S, et al. Plexin: a novel neuronal cell surface molecule that mediates cell adhesion via a homophilic binding mechanism in the presence of calcium ions. Neuron. 1995 Jun;14(6):1189–1199.

222. Roney KE, O'Connor BP, Wen H, Holl EK, Guthrie EH, Davis BK, et al. Plexin-B2 negatively regulates macrophage motility, Rac, and Cdc42 activation. PLoS One. 2011 Sep 23;6(9):e24795.

223. Junqueira Alves C, Dariolli R, Haydak J, Kang S, Hannah T, Wiener RJ, et al. Plexin-B2 orchestrates collective stem cell dynamics via actomyosin contractility, cytoskeletal tension and adhesion. Nat Commun. 2021 Oct 14;12(1):6019.

224. Van Battum E, Heitz-Marchaland C, Zagar Y, Fouquet S, Kuner R, Chédotal A. Plexin-B2 controls the timing of differentiation and the motility of cerebellar granule neurons. Elife. 2021 Jun 8;10.

225. Malik MFA, Ye L, Jiang WG. Reduced expression of semaphorin 4D and plexin-B in breast cancer is associated with poorer prognosis and the potential linkage with oestrogen receptor. Oncol Rep. 2015 Aug;34(2):1049–1057.

226. Le AP, Huang Y, Pingle SC, Kesari S, Wang H, Yong RL, et al. Plexin-B2 promotes invasive growth of malignant glioma. Oncotarget. 2015 Mar 30;6(9):7293–7304.

227. Zhou L, Hanemann CO. Merlin, a multi-suppressor from cell membrane to the nucleus. FEBS Lett. 2012 May 21;586(10):1403–1408.

228. Kim M, Kim S, Lee SH, Kim W, Sohn MJ, Kim HS, et al. Merlin inhibits Wnt/β-catenin signaling by blocking LRP6 phosphorylation. Cell Death Differ. 2016 Oct;23(10):1638–1647.

229. Ikeda F, Deribe YL, Skånland SS, Stieglitz B, Grabbe C, Franz-Wachtel M, et al. SHARPIN forms a linear ubiquitin ligase complex regulating NF-κB activity and apoptosis. Nature. 2011 Mar 31;471(7340):637–641.

230. Sasaki Y, Sano S, Nakahara M, Murata S, Kometani K, Aiba Y, et al. Defective immune responses in mice lacking LUBAC-mediated linear ubiquitination in B cells. EMBO J. 2013 Sep 11;32(18):2463–2476.

231. O'Dell MR, Huang JL, Whitney-Miller CL, Deshpande V, Rothberg P, Grose V, et al. Kras(G12D) and p53 mutation cause primary intrahepatic cholangiocarcinoma. Cancer Res. 2012 Mar 15;72(6):1557–1567.

232. Frey N, Tortola L, Egli D, Janjuha S, Rothgangl T, Marquart KF, et al. Loss of Rnf31 and Vps4b sensitizes pancreatic cancer to T cell-mediated killing. Nat Commun. 2022 Apr 4;13(1):1804.

233. Morrow KA, Das S, Meng E, Menezes ME, Bailey SK, Metge BJ, et al. Loss of tumor suppressor Merlin results in aberrant activation of Wnt/β-catenin signaling in cancer. Oncotarget. 2016 Apr 5;7(14):17991–18005.

234. Rong R, Tang X, Gutmann DH, Ye K. Neurofibromatosis 2 (NF2) tumor suppressor merlin inhibits phosphatidylinositol 3-kinase through binding to PIKE-L. Proc Natl Acad Sci USA. 2004 Dec 28;101(52):18200–18205.

235. Li N, Lu N, Xie C. The Hippo and Wnt signalling pathways: crosstalk during neoplastic progression in gastrointestinal tissue. FEBS J. 2019 Oct;286(19):3745–3756.

236. Varelas X, Miller BW, Sopko R, Song S, Gregorieff A, Fellouse FA, et al. The Hippo pathway regulates Wnt/beta-catenin signaling. Dev Cell. 2010 Apr 20;18(4):579–591.

237. Azzolin L, Zanconato F, Bresolin S, Forcato M, Basso G, Bicciato S, et al. Role of TAZ as mediator of Wnt signaling. Cell. 2012 Dec 21;151(7):1443–1456.

238. www.pantherdb.org/ [Internet]. [cited 2022 Apr 23]. Available from: http://www.pantherdb.org/

239. Kim S-Y, Yasuda S, Tanaka H, Yamagata K, Kim H. Non-clustered protocadherin. Cell Adh Migr. 2011 Apr;5(2):97–105.

240. Fang X, Yu S, Eder A, Mao M, Bast RC, Boyd D, et al. Regulation of BAD phosphorylation at serine 112 by the Ras-mitogen-activated protein kinase pathway. Oncogene. 1999 Nov 18;18(48):6635–6640.

241. Kissil JL, Wilker EW, Johnson KC, Eckman MS, Yaffe MB, Jacks T. Merlin, the product of the Nf2 tumor suppressor gene, is an inhibitor of the p21-activated kinase, Pak1. Mol Cell. 2003 Oct;12(4):841–849.

242. Wu H, Chen Y, Wang ZY, Li W, Li JQ, Zhang L, et al. Involvement of p21 (waf1) in merlin deficient sporadic vestibular schwannomas. Neuroscience. 2010 Sep 29;170(1):149–155.

In Vivo Modeling of Patient Genetic Heterogeneity Identifies New Ways to Target Cholangiocarcinoma

Nicholas T. Younger[1], Mollie L. Wilson[1], Anabel Martinez Lyons[1], Edward J. Jarman[1], Alison M. Meynert[1], Graeme R. Grimes[1], Konstantinos Gournopanos[1], Scott H. Waddell[1], Peter A. Tennant[1], David H. Wilson[1], Rachel V. Guest[2], Stephen J. Wigmore[2], Juan Carlos Acosta[3], Timothy J. Kendall[4], Martin S. Taylor[1], Duncan Sproul[1], Pleasantine Mill[1], and Luke Boulter[1]

ABSTRACT

Intrahepatic cholangiocarcinoma (ICC) is an aggressive malignancy of the bile ducts within the liver characterized by high levels of genetic heterogeneity. In the context of such genetic variability, determining which oncogenic mutations drive ICC growth has been difficult, and developing modes of patient stratification and targeted therapies remains challenging. Here we model the interactions between rare mutations with more common driver genes and combine *in silico* analysis of patient data with highly multiplexed *in vivo* CRISPR-spCas9 screens to perform a functional *in vivo* study into the role genetic heterogeneity plays in driving ICC. Novel tumor suppressors were uncovered, which, when lost, cooperate with the RAS oncoprotein to drive ICC growth. Focusing on a set of driver mutations that interact with KRAS to initiate aggressive, sarcomatoid-type ICC revealed that tumor growth relies on Wnt and PI3K signaling. Pharmacologic coinhibition of Wnt and PI3K *in vivo* impeded ICC growth regardless of mutational profile. Therefore, Wnt and PI3K activity should be considered as a signature by which patients can be stratified for treatment independent of tumor genotype, and inhibitors of these pathways should be levied to treat ICC.

Significance: This work shows that, despite significant genetic heterogeneity, intrahepatic cholangiocarcinoma relies on a limited number of signaling pathways to grow, suggesting common therapeutic vulnerabilities across patients.

Introduction

Intrahepatic cholangiocarcinomas (ICC) are epithelial tumors of the bile duct comprised of malignant ducts surrounded by an extensive stroma (1). ICC driven by infection with the liver fluke, *Opisthorchis viverrini* is endemic in South East Asia and although historically seen as a rare malignancy in the West, sporadic, nonfluke-associated disease has increased in incidence in the UK, Europe, and the United States over the last four decades. Currently, surgical resection is the only curative option for patients diagnosed with this cancer; however, of the ∼30% of patients who have disease that is amenable to surgery, 70% of those patients relapse following resection (2). In patients where surgery is not an option, the standard of care is palliative chemotherapy, which extends life by approximately 3 to 6 months (3). Early studies using either patient ICC samples (4) or mouse models (5) demonstrated that oncogenic mutations in *Kras* (typically $Kras^{G12D}$) and loss-of-function mutations in *Trp53* cooperate to initiate tumor formation. Recent genomic data, however, has challenged whether mutations in this oncogene and tumor suppressor pair often co-occur in human ICC (6, 7). Instead, these sequencing data suggest that alternate or less-frequent mutations cooperate with more dominant oncogenes (such as mutant *Kras*) to promote tumorigenesis. Deep sequencing of ICC has uncovered that a high level of genetic heterogeneity exists within patient ICC samples (8, 9). Although a recurring set of mutations in canonical genes has been identified (1), many infrequent mutations have also been detected. The functional contribution of these infrequent *de novo* changes to affect disease progression and modulate therapeutic resistance or susceptibility remains unclear.

To identify and prioritize gain-of-function or loss-of-function mutations in a patient dataset of ICC, we use a computational pipeline, IntOGen (10), to generate a high-confidence list of candidate driver genes, of which, 64 have not previously been assigned as being cancer drivers. To recapitulate the clonal competition observed in human tumorigenesis, we developed an *in vivo* CRISPR-SpCas9 system, which simultaneously screens the candidate gene set against either $KRAS^{G12D}$ or $NRAS^{G12V}$ oncogenes. This identified a subset of genes in which human ICC-derived mutations genetically interact with RAS to initiate and accelerate ICC formation. Among these, we found that loss-of-function of Neurofibromin 2 (*Nf2*) interacts with mutant RAS to initiate tumor formation independent of *Trp53* status, highlighting again that as seen in patient data, RAS mutant cells do not strictly rely on *Trp53* loss to initiate ICC. Loss of *Nf2* results in the formation of aggressive and poorly differentiated sarcomatoid-type ICC. These tumors are driven by dysregulation of Wnt–PI3K signaling, highlighting a novel therapeutic avenue that could be used to target ICC growth.

[1]MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, United Kingdom. [2]Clinical Surgery, University of Edinburgh, Royal Infirmary of Edinburgh, Edinburgh, United Kingdom. [3]Cancer Research UK Edinburgh Centre, Institute of Genetics and Cancer, Crewe Road South, Edinburgh, United Kingdom. [4]Centre for Inflammation Research, University of Edinburgh, Edinburgh, United Kingdom.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

N.T. Younger and M.L. Wilson contributed equally to this article.

Current address for N.T. Younger and D.H. Wilson: Centre for Inflammation Research, University of Edinburgh, Edinburgh, United Kingdom.

Corresponding Author: Luke Boulter, MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh EH4 2XU, UK. E-mail: luke.boulter@ed.ac.uk

Cancer Res 2022;82:1548–59

doi: 10.1158/0008-5472.CAN-21-2556

Materials and Methods

Identification and processing of genomic data from patient datasets

Alignment and preprocessing of publicly available data

Exome-seq FASTQ files from Chan-on and colleagues (8) were downloaded from the European Nucleotide Archive with accession no. PRJEB4445. Exome-seq FASTQ files from Sia and colleagues (11) were downloaded from the Gene Expression Omnibus Database with accession GSE63420. The Cancer Genome Atlas (TCGA) BAM files were downloaded from the Genomic Data Commons after receiving access to individual patient BAM files. Input FASTQ files were aligned to the Hg19 reference genome. Duplicates were marked and base quality score recalibration was carried out and local indel realignment was performed. Ensemble variant calling was performed. TCGA data were input as reads which were aligned to the Hg38 reference so were re-mapped using the Hg38toHg19. IntOgen was run on each cohort individually, and then all combined. IntOgen outputs were used to build a network of known and inferred functional interactions and was clustered into modules by connection density and each module was annotated with pathway enrichments.

Evolutionary dependency (co-occurrence and mutual exclusivity) was scored using SELECT (version 1.6) with default parameters. The weighted mutual information (wMI) *P* value was used for color coding significance and FDR < 0.1 used as multitesting corrected threshold of significance. For the aggregate analysis of genes ("cosmic" and "other" groups), individual tumors were binary encoded as 1 if they contained a putative driver mutation in any gene of the corresponding gene list, and 0 otherwise.

Further experimental details about our computational approach can be can be found in Supplementary Materials and Methods.

Design and preparation of sgRNA plasmids for *in vivo* editing

Generation of pooled sgRNA screening library

Oligonucleotides encoding sgRNAs targeting the set of predicted drivers were designed using spacer sequences from the mouse GeCKo V2 library (Supplementary Table S1; ref. 12). Library-specific PCR retrieval arms (13) and our schematic for library preparation is described in Supplementary Table S1. Complete sgRNA oligos for all target genes and control sequences were custom synthesized by Twist Biosciences. Purified amplicons were then digested with Esp3I, phosphorylated and ligated into the SB-CRISPR plasmid backbone.

Generation of single gRNAs

Single gRNAs (Supplementary Table S2; Supplementary Fig. S3A) were cloned into SB-CRISPR plasmids kindly provided by Professor Dr. Roland Rad (LMU Munich). SB-CRIPSR was digested with Esp3I or Bbsl.

Animal work

All animal work was performed under the UK Home Office project license held by Dr. Luke Boulter (PFD31D3D4). Animals were maintained in colonies in 12-hour light–dark cycles and were allowed access to food and water *ad libitum*.

Hydrodynamic tail vein injection

Female, FVB/N mice from Charles River, and were used at 4 to 6 weeks of age. For the hydrodynamic tail vein injection, animals were injected with a physiologic saline solution (10% w/v) containing plasmids into the lateral tail vein. The typical injection contained 6 μg of PGK-SB13, 20 μg of CAG-Kras^{G12D} or CAG-Nras^{G12V}, and 20 μg of SB-CRISPR gRNA plasmid. In models that relied on a combination of gRNA plasmids, plasmids were dissolved to a maximum concentration of 20 μg and were mixed such that they were balanced pools of each gRNA. For screening studies, the gRNA library was injected at 20 μg and we determined gRNA representation by Sanger sequencing prior to injection and plotted the GINI index for each library (Supplementary Fig. S3B).

Keratin-19-CreER^T;pten$^{flox/flox}$;trp53$^{flox/flox}$;R26R^{LSLtdTomato} mice (KPPTom)

Keratin-19-CreER^T mice (Jax: 026925) were crossed with animals containing floxed alleles of *Pten* (Jax: 006440) or *Trp53* (Jax: 008462) and a silenced tdTomato reporter targeted to the Rosa26 locus (Jax: 007908). All animals in this study are heterozygous for Keratin-19CreER^T, homozygous for *Trp53^{flox}* and *Pten^{flox}* alleles, and homozygous for *R26R^{LSLtdTomato}*. Mice received three doses of 4 mg of tamoxifen by oral gavage and followed by 400 mg thioacetamide in their drinking water. All mice were male.

Therapeutic dosing of animal models

Animals baring *KRas^{G12D};gRNA^{Nf2};gRNA^{Trp53}* tumors were randomized using GraphPad online randomization tool and dosed with either vehicle alone (10% DMSO, 40% PEG300, 5% Tween-80, and 45% saline), 5 mg/kg LGK974, 50 mg/kg pictilisib, or a combination of the two daily starting 7 days following hydrodynamic injection. KPPTom animals were given tamoxifen and thioacetamide (as detailed above) and at 4 weeks were given either vehicle alone (10% DMSO, 40% PEG300, 5% Tween-80, and 45% saline) or a combination of 5 mg/kg LGK974 and 50 mg/kg pictilisib for 4 weeks. All animals were housed in colonies of five animals.

Isolation of RNA and DNA

Both DNA and RNA extraction used 50 to 100 mg of snap frozen tissue. DNA was extracted from tissue using the DNeasy Blood and Tissue Kit (Qiagen) as per the manufacturer's instructions. RNA was extracted using TRIzol RNA Isolation Reagent (Invitrogen), precipitated with chloroform, and cleaned up using the RNeasy Mini Kit (Qiagen). For sequencing, applications DNA and RNA quality (RIN score) was quantified using the Agilent 2100 Bioanalyzer. A minimum RIN threshold of 8 was used for RNA sequencing (RNA-seq).

RNA-seq

Libraries were prepared from total-RNA following rRNA depletion. rRNA-depleted RNA was then DNase treated and purified prior to fragmentation Libraries were quantified by fluorometry and assessed for quality and fragment size using the Agilent Bioanalyser Sequencing was performed using the NextSeq 500/550 High-Output v2.5 (150 cycle) Kit on the NextSeq 550 platform (Illumina Inc). Libraries were combined in an equimolar pool based on Qubit and Bioanalyser assay results and run across a single High Output v2.5 Flow Cell.

RNA-seq data processing and analysis

The primary RNA-seq processing, quality control to transcript-level quantitation, was carried out using nf-core/rnaseq v1.4.3dev (https://github.com/ameynert/rnaseq; ref. 14).

DNA exome sequencing

Libraries were prepared from genomic DNA (gDNA) and was sheared to achieve target DNA fragment sizes of between 150 and 200bp. DNA fragments were processed as adapter-ligated libraries and were purified. Seven hundred fifty nanograms of each prepared gDNA

library was hybridized to probes covering the mouse exome and hybridized DNA-probes were amplified to apply unique indexing primers. Library QC: Libraries were quantified by Qubit and sequencing was performed using the NextSeq 500/550 High-Output v2.5 (150 cycle) Kit on the NextSeq 550 platform (Illumina Inc.). Libraries were combined in a single equimolar pool and run on a High-Output v2.5 Flow Cell.

CRISPR/Cas9-editing validation and structural variant calling

DNA sequences from exome-sequencing of tumors arising from the RAS^{G12}-library screens and were aligned to the FVB mouse reference genome; subsequently, indels within 50bp upstream or downstream of sgRNA target sites were called. To determine if indels were likely due to SpCas9 activity, the interval of each indel was observed using Integrative Genomics Viewer (15), and overlain with sgRNA library binding sites. Further information on identifying the outcome of CRISPR editing is detailed in the Supplementary Materials and Methods.

Data curation and deposition

All RNA and exome-sequencing data pertaining to this manuscript is deposited on the NCBI Gene Expression Omnibus (GEO) as accession number GSE190770.

Histology and IHC

Livers were perfused with phosphate-buffered saline and dissected into 10% neutral-buffered formalin. Fixed tissue was processed in wax blocks and sectioned 4 μmol/L thick. Sections for immunostaining were dewaxed in xylene and rehydrated. Following antigen retrieval and sections were incubated with primary and secondary antibodies as detailed in Supplementary Table S3. Histologic assessment was undertaken by a consultant liver histopathologist working at the national liver transplant center (TJK) with experience in the comparative pathology of animal models of primary liver cancer.

FUnGI immunostaining and clearing

We adapted a previously published protocol for FUnGI staining and imaging (16). Eight millimeters of cores were are taken from liver and cancer tissue and sectioned at 200 μm intervals using a Krumdeick Tissue Slicer and fixed in formalin. Tissues were blocked and incubated overnight with primary and secondary antibodies in a series of permeabilizing washes before being cleared in FUnGI solution, which contains high levels of fructose. Clarified tissues are incubated with DAPI and mounted in FUNGI for confocal analysis.

Quantification of tumor burden

Histologic sections containing tumors were scanned using a Nanozoomer slide scanner with a 40× objective lens. Files were then imported into QuPath (https://qupath.github.io) and tumor tissue was manually annotated. Tumor burden represents the area of tissue occupied by tumor and number is the number of discrete tumors in the tissue. All tumor analysis was blinded.

Reverse phase protein arrays

Snap frozen, dissected tumor tissue was provided to the Human Tumor Profiling Unit (HTPU) at the Cancer Research UK Edinburgh Centre. The target proteins analyzed by RPPA are listed in Supplementary Table S4. Reverse phase protein array (RPPA) analysis was carried out using established protocols for nitrocellulose-based arrays (17). Slide images were acquired using an InnoScan 710-IR scanner (Innopsys) with laser power and gain settings optimized for highest readout without saturation of the fluorescence signal. The relative fluorescence intensity of each array feature was quantified using Mapix software (Innopsys).

Statistical analysis

All experimental groups were analyzed for normality using a D'Agostino–Pearson Omnibus test. Groups that were normally distributed were compared with either a two-tailed Student *t* test (for analysis of two groups) or using one-way ANOVA to compare multiple groups, with a *post hoc* correction for multiple testing. Nonparametric data were analyzed using a Wilcoxon–Mann–Whitney *U* test when comparing two groups or a Kruskal–Wallis test when comparing multiple nonparametric data. Throughout, $P < 0.05$ was considered significant. Data are represented as mean with SEM for parametric data or median with SD for nonparametric data.

All figures were laid out with Adobe Illustrator and graphics were created with BioRender.com.

Ethics approval and consent to participate

All human data included in this manuscript comes from previously published, consented studies. Animal work performed here is approved by the UK Home Office license provided to L. Boulter.

Availability of supporting data

All -omics data are deposited in publicly accessible databases, and all other datasets used and analyzed during the current study are available within the manuscript and its additional files.

Results

Identifying candidate causative mutations that drive ICC growth

ICC contains a range of infrequently mutated genes without known function. Identifying a consensus group of driver mutations in ICC using exome and genome sequencing has been challenging, due in a large part to tissue availability. Nonetheless, a number of studies have demonstrated recurrent ICC mutations including neomorphic alterations in *IDH1* and *IDH2*, loss-of-function mutations in *PBRM1*, *BAP1*, *TP53*, *ARID1A*, and gain-of-function mutations in *KRAS*. Despite their identification, the presence of these mutations in a tumor is not a strong predictor of therapeutic outcome (18, 19) and for approximately 30% of patients with ICC a driver mutation cannot be identified (20).

To determine whether all patient tumors contain potential driver mutations, we used a computational pipeline based around the driver prediction tool IntOGen (10). This method utilizes a combination of functional impact bias (OncodriveFM), spatial clustering (OncodriveCLUST), and corrected frequency (MutSigCV) to define whether particular genomic regions have a mutational rate beyond that which is expected, have a bias towards clustered mutations or those that are likely to impact functional domains, such as those that are regulatory or catalytic (summarized in Supplementary Fig. S1). Having filtered out hypermutated samples (Supplementary Table S1), we used this pipeline to analyze the variants identified in the genomes of 277 sporadic and fluke-associated ICCs from four distinct studies (7, 11, 21, 22), summary information on the aggregated cohort can be found in Supplementary Figs. S2A to S2H and Supplementary Table S2. Following processing, 55% of samples ($N = 152$) carried ≥2 predicted drivers whereas 18% ($N = 50$) carried none (**Fig. 1A**; Supplementary Table S3). Of these predicted driver mutations, approximately one third of mutations were already known to occur in ICC or were present in genes in the COSMIC database (**Fig. 1B**). The

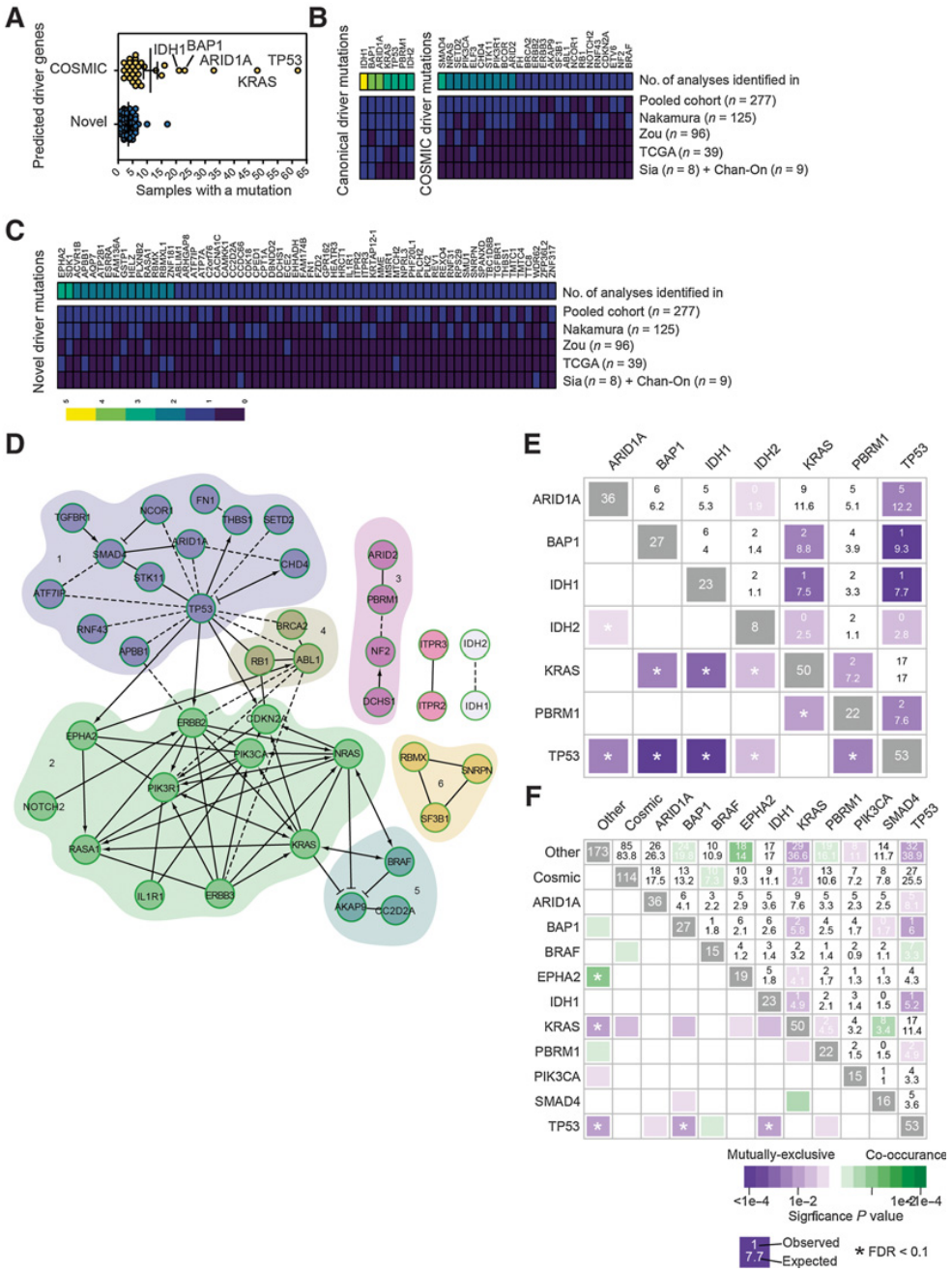


Figure 1.

In silico screening identifies novel drivers of ICC. **A,** The number of samples with mutations in driver genes identified following analysis with IntOGen. Samples clustered into those that had been previously identified in ICC or are present in the COSMIC database (yellow points), or those mutated genes that have not previously been assigned as being cancer drivers (blue points). **B,** Frequency of samples containing known ICC or COSMIC mutation in each of the individual cohorts collated in this study and in the pooled datasets. **C,** Frequency of samples containing a predicted novel oncogenic mutation identified by IntOGen. In both **B** and **C,** heatmaps represent the frequency at which each mutation is found within each study and the aggregated frequency between all studies. Top bar represents the number of times a mutation was identified between studies. **D,** Pathway interaction analysis of putative ICC driver mutations identified by IntOGen based on known functional (solid lines) and predicted physical (dotted lines) interactions. Numbers 1 to 6 represent distinct gene relationship modules based on predicted or known genetic interactions. **E,** Co-occurrence and mutual exclusivity analysis demonstrates that there is a high level of mutational exclusivity between canonical driver mutations in patients with ICC. **F,** Co-occurrence and mutual exclusivity of COSMIC drivers that co-occur ≥15 times in the human sample set and combined "other," which includes all novel drivers. $N = 277$ patient exomes or genomes with matched, noncancerous tissue.

remaining two thirds of mutations were novel and occurred in fewer than 8% of ICC cases. Indeed, the majority of novel mutations were only found in three to four patients, corresponding to ∼1.5% of the patient cohort (**Fig. 1C**; Supplementary Tables S4 and S5). To explore whether these low-frequency predicted drivers are involved in common pathways or processes, networks were constructed on the basis of known and predicted physical interactions (23) and clustered into modules based on connection density (**Fig. 1D**; ref. 24). This produced a network with six modules containing a mix of known in ICC, COSMIC, and novel genes. Gene ontology analysis was performed on the modules to ascertain the biological processes in which each module may participate (Supplementary Table S6). Finally, we sought to determine whether, based on our mutations as defined through IntOgen, we could define the genetic interactions between mutations in this patient dataset. Within our group of canonical driver mutations, we found that (with the exception of *KRAS* and *TP53*) there is a consistent and high level of mutual exclusivity between canonical driver genes, suggesting that the recurrent mutations found across studies are not interdependent for ICC initiation and growth (**Fig. 1E**). Interestingly, when we expanded this analysis to include COSMIC mutations and novel mutations (those that recur ≥15 times within our dataset *BRAF*, *PIK3CA*, *EPHA2*, and *SMAD4* are individually identified and those with ≤15 occurrences within this set are grouped as "cosmic" and novel are defined as "other"), we found distinct patterns of co-occurrence and mutual exclusivity that support the idea that there are *bona fide* genetic interactions between canonical driver genes (e.g., *KRAS*) and the large group of novel, infrequently mutated genes identified by IntOgen (**Fig. 1F**).

In vivo CRISPR-SpCas9 screening identifies novel tumor suppressors in RAS-driven ICC

Clonal analysis of patient with ICC has failed to identify a consensus mutational route through which tumors progress (25) or epistatic mutations that functionally interact to drive tumor initiation and growth. Relatively low sample number and high genetic heterogeneity in ICC exacerbate the difficulties with this type of associative analysis.

To overcome these limitations and to define which candidate drivers are functionally capable of initiating ICC, we developed an *in vivo* screening approach that allowed us to functionally prioritize ICC driver mutations (**Fig. 2A**). Previous work using multiplex-mutagenesis in the liver has demonstrated that editing specific genomic loci in hepatocytes can give rise to ICC, albeit using a relatively limited pool of gRNAs, targeting 10 genes (26). In this system, naked DNA is delivered to the liver using a high pressure, hydrodynamic injection into the lateral tail-vein of mice, an approach that has been used in the past to deliver cDNAs coding oncogenes (27, 28). When paired with CRISPR-SB plasmids containing sgRNAs and SpCas9 flanked between two Sleeping Beauty (SB) inverted terminal repeats, and a plasmid expressing Sleeping Beauty (SB) transposase, it is possible to edit *in vivo* endogenous genes in hepatocytes in a mosaic manner (29). Using the CRISPR-SB system as a starting point, we generated a large-scale multiplexed CRISPR-SpCas9 plasmid library (known hereafter as ICC^{Lib}) containing triplicate gRNAs targeting 91 mouse homologues of our putative, patient-derived ICC driver genes identified through our *in silico* approach. Five genes (*KRTAP12–1*, *RBMXL1*, *RBMX*, *SPANXD*, and *ZNF181*) from our human dataset had no identifiable murine orthologue, and so were excluded from further analysis (Supplementary Figs. S3A and S3B).

Using this CRISPR-SpCas9 system, we randomly introduced CRISPR-SpCas9 targeted mutations into these candidate "patient-led" ICC genes in otherwise wild-type mice to determine which loss-of-function mutations are necessary for tumor initiation and determine whether any loss-of-function mutations interact to drive ICC formation. The ICC^{Lib} alone failed to induce any tumors in mice after 10 weeks, suggesting that within this timeframe these loss-of-function mutations alone are insufficient to initiate cancer. Gain-of-function mutations in *KRAS* and *NRAS* have been previously described in ICC (8) and through our IntOGen analysis, we similarly identified recurrent mutations in both of these genes (*KRAS* 18.05% and *NRAS* 2.88%, Supplementary Table S4). In experimental models, expression of mutant RAS in the adult liver is weakly oncogenic and normally insufficient to initiate ICC formation; rather, mutant cells undergo oncogene-induced senescence and are removed from the liver by immune clearance (30). Therefore, we co-expressed cDNAs of GFP tagged $KRAS^{G12D}$ or $NRAS^{G12V}$ with our ICC^{Lib} to determine whether any of the loss-of-function mutations introduced via the ICC^{Lib} synergize with mutant RAS proteins to promote ICC initiation. Within 10 weeks, mice that received either $KRAS^{G12D}$ or $NRAS^{G12V}$ and loss-of-function mutations induced by the CRISPR-SpCas9 containing ICC^{Lib} developed macroscopic and multifocal cancer; this was accelerated in $KRAS^{G12D}$ mice, which developed symptomatic liver cancer in 8 weeks (**Fig. 2B**). In those mice that developed cancer, multiple tumors formed per mouse (**Fig. 2C**), which were histologically aggressive adenocarcinoma with a poorly differentiated cholangiocellular morphology. Importantly, these tumors express GFP (Supplementary Fig. S4A), denoting that they continue to express the $KRAS^{G12D}$ or $NRAS^{G12V}$ constructs, and the cholangiocyte marker Keratin-19, which is constrained to the biliary epithelium in normal livers (Supplementary Fig. S4B). Together, these data demonstrate that mutant Ras ($NRAS^{G12V}$ or $KRAS^{G12D}$) can interact with at least one loss-of-function mutation generated by the gRNAs contain in the ICC^{Lib} to initiate ICC *in vivo*.

To determine which CRISPR-SpCas9 mutational events cooperated with $KRAS^{G12D}$ or $NRAS^{G12V}$ and lead to the emergence of liver cancer, whole exomes were sequenced from 14 $KRAS^{G12D}$ and 10 $NRAS^{G12V}$ driven tumors. All indels within 50bp of a sgRNA target site were manually inspected to determine whether they were Cas9-induced or of spontaneous origin. Almost all indels had start or end positions approximately 3bp upstream of the SpCas9 protospacer adjacent motif (PAM) sequence, strongly indicating that they are a consequence of CRISPR-SpCas9 editing (Supplementary Fig. S5A). Our data shows that tumors acquired multiple CRISPR-SpCas9-induced lesions; $KRAS^{G12D}$ tumors contained an average of 7.5 ± 1.19 mutations and in $NRAS^{G12V}$ tumors there were on average 7.7 ± 2.59 CRISPR-SpCas9 induced mutations (Supplementary Fig. S5B). Across both $Kras^{G12D}$ and $Nras^{G12V}$ screens, 66 of the 91 predicted drivers targeted with the ICC^{Lib} were mutated and 27 of these were shared between $NRAS^{G12V}$ and $KRAS^{G12D}$ tumors (**Fig. 2D**; Supplementary Fig. S6). The most common CRISPR-SpCas9 mutation we identified was unsurprisingly in *Trp53*, reiterating the ability of cells with *Trp53* loss-of-function mutations to overcome RAS-induced senescence (31). We also found recurrent CRISPR-induced indels from both $NRAS^{G12V}$ and $KRAS^{G12D}$ screens in genes whose loss has been linked to ICC but that have not previously been shown to genetically interact with mutant RAS in this cancer including *Brca2*, *Nf2*, and *Plk2* (**Fig. 2D**).

Nf2 loss interacts with $KRAS^{G12D}$ and *Trp53* loss to promote sarcomatoid phenotypes in ICC

Our data demonstrate that loss of numerous genes mutated at low frequency in human ICC potentially interact with activating

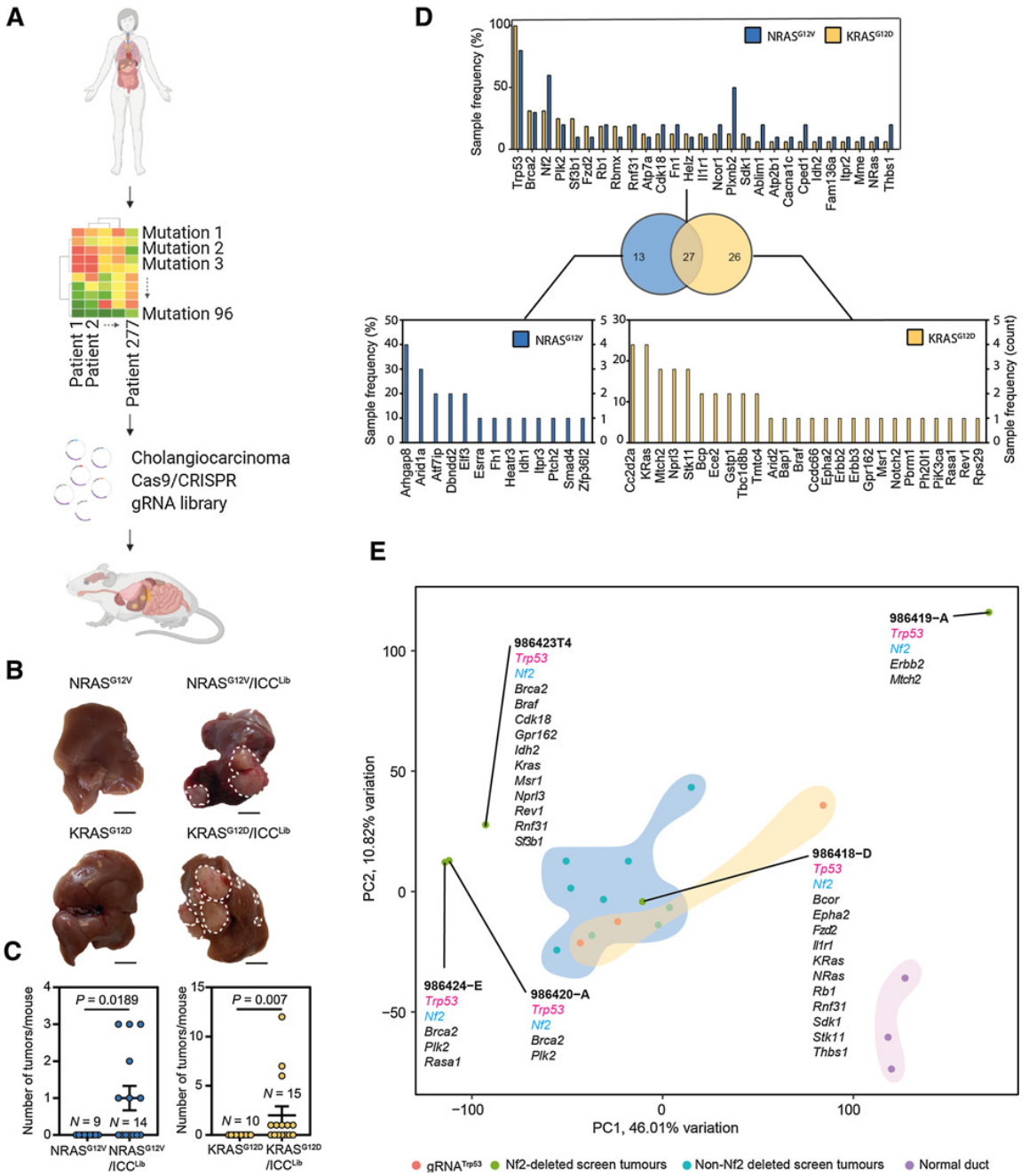


Figure 2.

In vivo CRISPR-Cas9 screening identifies transforming mutations that interact with mutant Ras. **A,** Schematic of this study in which high-content sequencing data are collated from patients with ICC and the mutational profile of these tumors rationalized to identify novel, high confidence drivers of ICC. These putative drivers are used as input for an *in vivo* SpCas9/CRISPR screen to identify novel functional processes that drive ICC growth. **B,** Macroscopic images of the livers following injection with either $NRAS^{G12V}$ or $KRAS^{G12D}$ alone (left) or in combination with ICC^{Lib} (right; dotted line, tumor). Scale bar, 1 cm. **C,** Quantification of macroscopic tumors per mouse at 10 weeks in mice bearing $Nras^{G12V}$ -expressing tumors and 8 weeks in those with $Kras^{G12D}$-driven cancer. Each circle represents a different animal. **D,** The number of samples containing Indels in a particular gene following whole exome sequencing. Top graph lists those mutations found in both $NRAS^{G12V}$ and $KRAS^{G12D}$ tumors, and bottom graphs denote those mutations that are found only in $KRAS^{G12D}$ - or $NRAS^{G12V}$ -expressing tumors. Sample frequency (%) denotes the proportion of tumors containing any given mutation, whereas (count) is absolute number. (*N* represents anatomically discreet tumors recovered from at least four individual animals, $KRAS^{G12D}$ $N = 14$ and $NRAS^{G12V}$ $N = 10$.) **E,** PCA showing how samples group based on their transcriptomic signature and gRNA-induced mutations associated with each tumor type.

$NRAS^{G12V}$ and $KRAS^{G12D}$ mutations to promote ICC initiation *in vivo*. These data and those of others has demonstrated that *KRAS* mutations occur more frequently in ICC than those in *NRAS* (9), therefore we prioritized validating loss-of-function mutations that genetically interact with $KRAS^{G12D}$. From our screen, we identified which mutations in patients have the potential to initiate the formation of Ras-induced ICC, therefore we set about to define whether any particular mutation (or set of mutations) alters the phenotype of these cancers. RNA sequencing analysis of our ICC^{Lib} screen tumors showed that, on the whole, tumors transcriptionally clustered closely to each other by principal component analysis (PCA) and were transcriptionally similar to tumors generated by expressing both $KRAS^{G12D}$ and deleting *Trp53* (**Fig. 2E**). Furthermore, our screen tumors were transcriptionally distinct from normal bile ducts. However four tumors in our screen transcriptionally segregated from all other tumors; all contained SpCas9/CRISPR-induced mutations in both *Trp53* and *Nf2*. In fact, of the 14 cancers from our screen that we exome sequenced, four of the five containing *Nf2* mutations segregated away from the main cluster (**Fig. 2E**), suggesting that the addition of a mutation in *Nf2* can functionally cooperate with $Kras^{G12D}$ and *Trp53* mutations and affects the phenotype of ICC. NF2 is also known as Merlin and has a well-defined role in the Hippo/LATS signaling pathway, where it negatively regulates pathway activation through the phosphorylation of *Mst1/2*; however, NF2/Merlin is also known to interact with a number of other signaling pathways including PI3K and Wnt signaling (32). We elected to investigate the genetic interaction of *Nf2*, *Trp53*, and $Kras^{G12D}$ further, by generating gRNAs to specifically target and disrupt the *Nf2* and *Trp53* loci (or a nontargeting control, $gRNA^{scrm}$), which were then coinjected hydrodynamically with our $KRAS^{G12D}$ expressing construct to define whether loss of these tumor suppressors can specifically cooperate with $KRAS^{G12D}$ to promote tumor initiation. *Trp53* loss and *Nf2* loss were both capable of overcoming the senescence inducing effects of $Kras^{G12D}$ expression in the liver and mice developed lethal tumors within 8 weeks following injection (**Fig. 3A**). In the presence of $KRAS^{G12D}$, the singular deletion of either *Trp53* or *Nf2* resulted in large discrete tumors. However, dual loss of *Trp53* and *Nf2* resulted in cancers that were highly aggressive and invasive, which had a median survival of 14 days compared with 39 and 41 days in singular *Nf2*-deleted and *Trp53*-deleted tumors, respectively (**Fig. 3A**). Tumors that lacked both *Trp53* and *Nf2* were highly diffuse and covered less liver area compared with both single gene deletions; however, the number of tumors that formed was significantly higher in *Trp53*;*Nf2* codeleted tumors, suggesting that mutations in these two tumor suppressors may synergize in cancer and are not functionally redundant (**Fig. 3B** and **C**). Rather, the codeletion of *Trp53* and *Nf2* enables the increased retention of mutant $Kras^{G12D}$-expressing cells within the liver. This retention leads to more cancers forming and that the accelerated mortality seen in animals bearing *Trp53*;*Nf2* codeleted tumors is likely due to the number of tumors that are able to form, rather than due to their size. Histopathologically, the *Trp53*;*Nf2* codeleted cancers are highly invasive with CK19 and GFP ($KRas^{G12D}$-expressing) immunopositive cells migrating throughout the liver (**Fig. 3D**; Supplementary Fig. S7) and represent a model to study the biology of invasive, sarcomatoid ICCs, which migrate along the ducts and invade the liver (33). Rather than those ICCs, which are mass forming (and which have been previously modeled in mice), sarcomatoid ICC, although rare, has a very poor prognosis with a survival of weeks to months following diagnosis (34). In mouse, these *Trp53*;*Nf2* codeleted sarcomatoid cancers grow rapidly, however, we did not observe perineural or microvascular invasion, nor did we see enlarged lymph nodes in these mice, suggesting the main cause of lethality is loss of functional liver capacity rather than the exit of cells from the tumor mass.

To determine the transcriptomic differences driving this aggressive phenotype in *Trp53*;*Nf2* codeleted cancers when compared with individual *Trp53* or *Nf2* deleted tumors, we undertook bulk RNA-seq of tumors with $Kras^{G12D}$;$Trp53^{KO}$, $Kras^{G12D}$;$Nf2^{KO}$ or $Kras^{G12D}$;$Trp53^{KO}$;$Nf2^{KO}$ genetic profiles. Sarcomatoid, *Trp53*;*Nf2* codeleted tumors are transcriptionally distinct from cancers containing either *Trp53* or *Nf2* deletions alone based on PCA (Supplementary Fig. S8A; Supplementary Table S7). By comparing the up- and downregulated genes in *Trp53* versus *Trp53*;*Nf2* codeleted tumors against the changes found in *Nf2* versus *Trp53*;*Nf2* codeleted tumors, we identified enrichment in signatures for Wnt signaling and Cadherin signaling using PANTHER (Supplementary Fig. S8B) in our model of sarcomatoid ICC. However, we did not find a transcriptional signature for Hippo signaling (which we would anticipate based on NF2′s canonical role in this pathway), nor could we identify YAP-positive cells within the *Nf2*-deleted tumors, rather YAP-positive cells are only found adjacent to the tumor mass. Together, these data indicate that *Nf2* loss in these tumors fails to activate Hippo signaling (**Fig. 3E**; Supplementary Figs. S8B and S8C). Previous work from our group has shown that Wnt signaling promotes ICC growth in the absence of classical Wnt pathway "activating mutations" (35). In *Trp53*;*Nf2* codeleted cancers, we observed upregulation of *Wnt9b* and suppression of inhibitors of Wnt signaling *Nkd2* and *Sfrp2*, suggesting that alterations in ligand levels and negative regulators of Wnt signaling are important mediators of ICC progression and that these cancers form independent of core Wnt-activating mutations in *Apc*, *Axin2*, and *Ctnnb1* (36). Interestingly, sarcomatoid ICC in patients displays changes in cadherin expression, although it is not clear whether these changes are causative for the aggressive sarcomatoid phenotype (37). In our model, we found transcriptional suppression of a range of cadherins and proto-cadherins that have been also implicated in Wnt regulation (38).

As there are few targeted treatments for poorly differentiated ICC, we screened our models driven by $KRAS^{G12D}$-expression and either *Tp53* loss, *Nf2* loss, or a combination of the two for activated and pharmacologically targetable signaling pathways using highly multiplexed RPPAs. Deletion of *Nf2* (either alone or in combination with *Trp53*) results in increased inhibitory phosphorylation of GSK3β and activation of β-catenin signaling when compared with $KRAS^{G12D}$ driven tumors lacking *Trp53* alone. These changes demonstrate that in *Nf2-mutant* cancers the canonical Wnt signaling pathway is highly activated and that the alterations in levels of Wnt ligands and inhibitors found at the transcriptional level translate into increased pathway activation (**Fig. 3F**). Furthermore, when *Trp53* and *Nf2* are concurrently deleted, the proportion of AKT that is phosphorylated at Serine-473 significantly increases (**Fig. 3G**), indicating that within these $Kras^{G12D}$;$Trp53^{KO}$;$Nf2^{KO}$ cancers signaling via AKT is also elevated. Histologically, β-catenin is found within all KRAS-driven cancers (**Fig. 3G**) but the dephosphorylated (active) form of the protein is increased in protein lysates from $Kras^{G12D}$;$Trp53^{KO}$;$Nf2^{KO}$ tumors (**Fig. 3H**). These data suggest that concurrent pAKT and Wnt activity promote the development of sarcomatoid ICC and that this combination of mutations is also sufficient to suppress apoptosis in these cancer cells (Supplementary Figs. S9A and S9B), thereby promoting the development of cancer in the liver, which rapidly becomes lethal. To test whether the aggressiveness of *Trp53*;*Nf2* mutated ICC is dependent on Wnt and AKT signaling directly, we treated mice bearing $KRAS^{G12D}$;*Nf2*;*Trp53*-KO tumors with an inhibitor of Porcupine (LGK974), which reduces Wnt ligand secretion by preventing

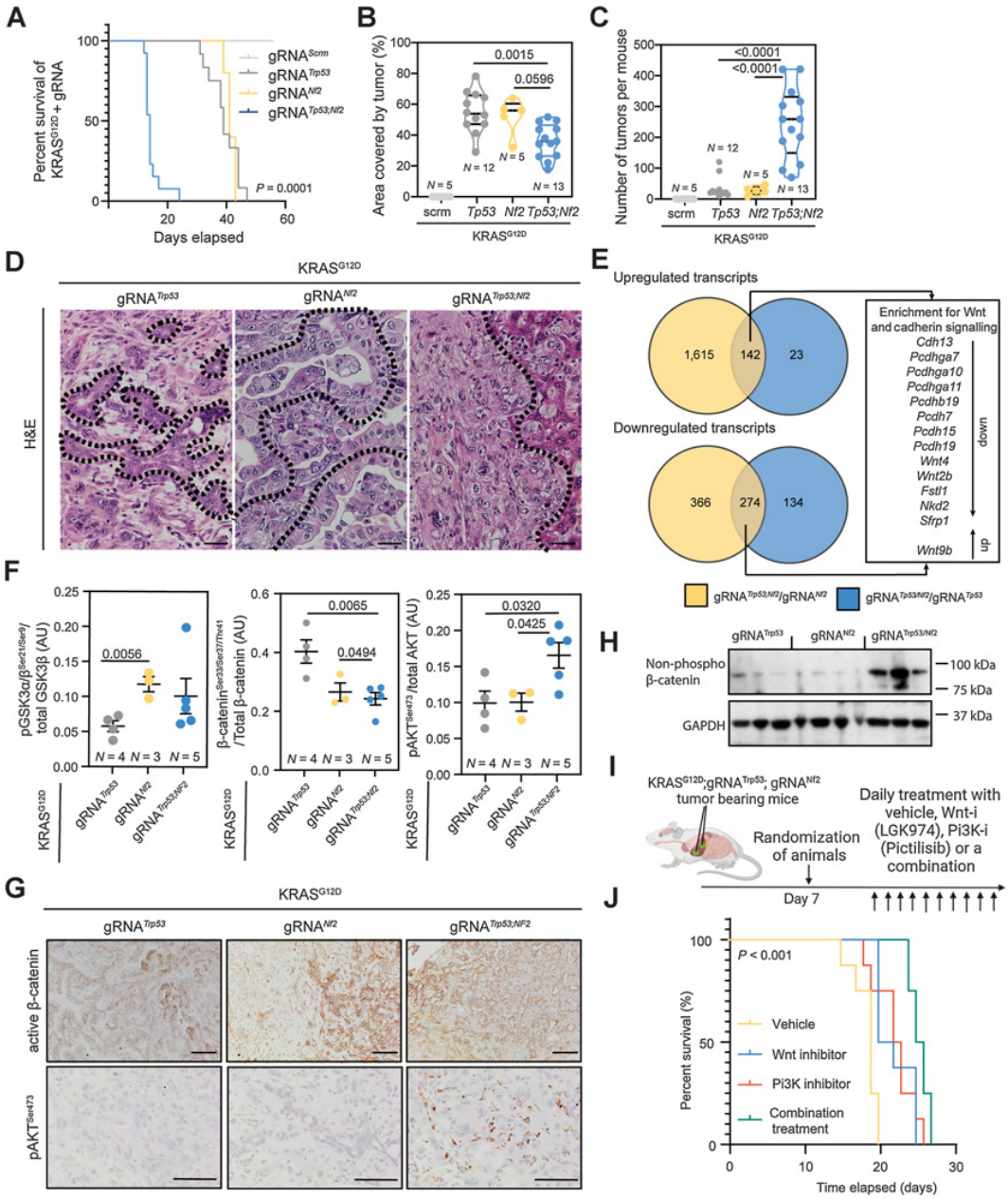
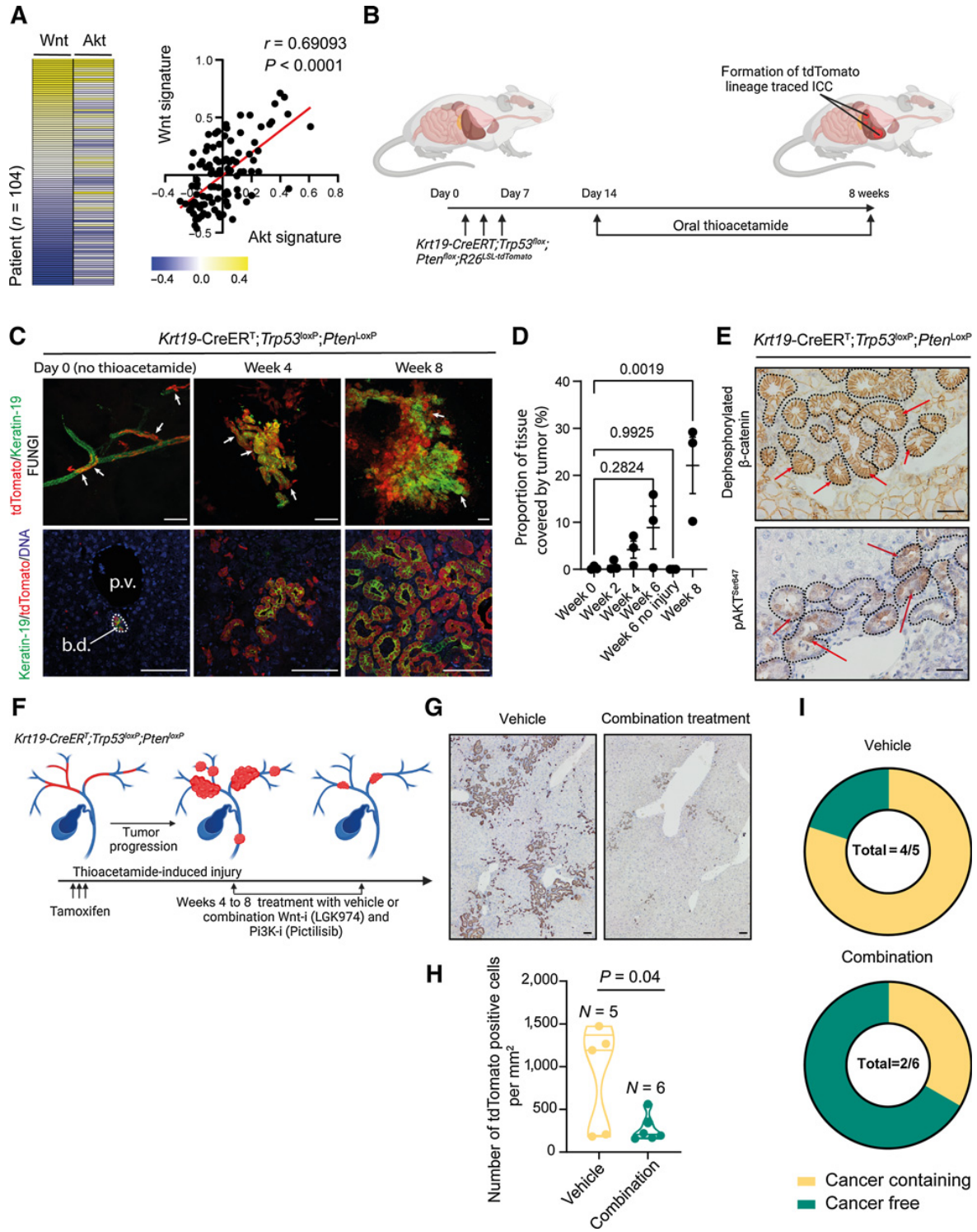


Figure 3.

Nf2 loss results in Ras^{G12D}-induced oncogenesis and cooperates with *Trp53* loss to accelerate ICC formation. **A,** Kaplan–Meier curve demonstrating the relative survival proportions of mice with KRAS^{G12D} and gRNAs targeting *Trp53* ($N = 12$), *Nf2* ($N = 5$), *Nf2*;*Trp53* ($N = 13$), or nontargeting control (scrm, $N = 5$). **B** and **C,** Proportion of liver occupied by tumor (**B**) and number of tumors per mouse (**C**). **D,** Hematoxylin and eosin (H&E) staining of KRAS^{G12D} tumors with *Trp53*, *Nf2*, or *Trp53*;*Nf2* loss. Scale bar, 100 μm. Dotted line, tumor-stroma boundary. **E,** Comparison of RNA-seq analysis when the transcriptomes from *Nf2*;*Trp53* versus *Trp53* alone tumors (blue) are compared with transcripts from *Nf2*;*Trp53* versus *Nf2* alone (yellow) tumors. Each group contains $N = 4$ regionally distinct tumors. **F,** Analysis of RPPA data demonstrating the changes in the proportion of phosphorylated GSK3α/β, β-catenin, and pAKT relative to total protein levels in KRAS^{G12D};*Trp53*^{KO} (gray points), KRAS^{G12D};*Nf2*^{KO} (yellow points), KRAS^{G12D};*Trp53*^{KO};*Nf2*^{KO} (blue points). **G,** IHC of active, dephosphorylated β-catenin (top) and phosphorylated AKT^{Ser647} (bottom) in KRAS^{G12D};*Trp53*^{KO}, KRAS^{G12D};*Nf2*^{KO}, KRAS^{G12D};*Trp53*^{KO}/*Nf2*^{KO} tumors. Scale bar, 50 μm. **H.** Immunoblot for dephosphorylated (active) β-catenin (β-catenin$^{Ser33/37/Thr41}$) in tumors isolated from mice baring Kras^{G12D} -driven ICC with *Trp53*, *Nf2*, or *Trp53*;*Nf2* co-loss. GAPDH was used as a loading control. **I,** Schematic representing our dosing approach to determine whether Wnt inhibition, PI3K inhibition, or a combination of the two is effective in improving the survival of mice with KRAS^{G12D};*Trp53*^{KO};*Nf2*^{KO} ICC. **J,** Kaplan–Meier curve demonstrating the survival changes when KRAS^{G12D};*Trp53*^{KO};*Nf2*^{KO} animals are treated with vehicle (yellow line), LGK974 (Wnt-inhibitor; blue line), pictilisib (PI3K inhibitor; orange line), or a combination (green line; $N = 5$ per group).



palmitoylation of Wnt ligands and a PI3K inhibitor, pictilisib, which prevents the conversion of PIP_2 into PIP_3 and thereby reduces AKT phosphorylation. Animals were given sarcomatoid ICC (by a hydrodynamic injection of $KRAS^{G12D}$;$gRNA^{Nf2}$;$gRNA^{Trp53}$) and at 7 days following tumor initiation were randomized to receive vehicle, LGK974, pictilisib or a combination of the two compounds (combination treatment), summarized in **Fig. 3I**. Both LGK974 and pictilisib significantly improve survival of tumor bearing mice compared with animals treated with vehicle alone [median survival of 19 days in vehicle treated vs. 21 days in LGK974-treated ($P = 0.0015$) and 22.5 days in pictilisib-treated animals ($P = 0.0058$), respectively]. Wnt inhibition and PI3K inhibition performed similarly, with no statistically significant difference in survival outcomes when used as single agents (**Fig. 3J**; Supplementary Figs. S10A–S10C; Supplementary Table S8). In combination, LGK974 and pictilisib improve median survival from 19 to 25.5 days (**Fig. 3J**), significantly ($P < 0.001$) reducing mortality compared with single treatments (Supplementary Figs. S10D–S10F) and demonstrating that coinhibition of Wnt and PI3K signaling is an effective treatment in sarcomatoid tumors that lack classical Wnt and PI3K activating mutations (i.e., mutations in *APC*, *CTNNB1*, and *PI3KCA*).

Wnt and PI3K/AKT represent a conserved mechanism by which distinct pathologic subtypes of ICC grow

From previous transcriptomic studies of ICC, it is possible to identify a subgroup of patients with high Notch pathway activity and who would be sensitive to treatment with γ-secretase-inhibitors (39). We therefore sought to identify whether there is a group of patients with ICC who could be sensitive to coinhibition of Wnt and PI3K signaling. In 104 transcriptomes from ICC patients (21), there is a high level of correlation ($r = 0.690$; $P < 0.0001$) between those with a high expression of genes associated with Wnt signaling and high expression of genes associated with AKT signaling (**Fig. 4A**; Supplementary Table S9). Therefore, we considered whether the activation of Wnt and PI3K signaling is a more universal process in ICC formation and sought to address whether these pathways are recurrently activated in ICC lacking *RAS* mutations, particularly given the majority of ICC patient cancers (79.07%) do not carry a mutation in either *Kras* or *Nras* (Supplementary Table S4).

To test this, we deleted *Trp53*, which is the most common mutation in our patient cohort, mutated in 19.13% of cases (**Fig. 1C**) and *Pten* specifically in biliary epithelial cells using a *Keratin19-*Cre^{ERT} knock-in mouse line. Although *Pten* mutations were not found in our computational analysis of ICC, loss of *Pten* recapitulates the effects of PI3KCA mutations (which is mutated in 5.4% of patients in our cohort, **Fig. 1C**; ref. 40), resulting in the accumulation of PIP_3 and AKT activation. Furthermore, previous studies have demonstrated that *Pten* loss potentiates ICC formation on the background of steatosis (fatty liver) and interacts with *Smad4*-loss to initiate liver cancer (41, 42). Concurrently, with *Trp53* and *Pten* deletion, recombined cells were labeled with tdTomato (from here on in this line is known as the KPPTom line, **Fig. 4B**). These animals are then challenged with chronic administration of Thioacetamide, a hepatotoxin that has been used to induce ICC in other models but is not itself mutagenic (43). Within 4 weeks of Thioacetamide treatment, mice develop moderate to well-differentiated cholangiocellular neoplasms around the portal tracts within the liver, which are lineage traced with tdTomato, confirming that they have arisen from the biliary epithelium (**Fig. 4C**). Using FUnGI imaging (44), we found that small, tdTomato-positive ICC microfoci (∼200 μm across) can be identified forming from the ducts by 4 weeks and in normal 2D histology appear as small clusters of lineage traced ducts, with an atypical luminal structure. These cancerous microfoci grow progressively over 8 weeks and occupy approximately 22% of the liver (**Fig. 4D**). In this model, the deletion of *Trp53* and *Pten* alone are insufficient to promote tumor initiation, instead the addition of a damaging agent (Thioacetamide) and the induction of ductular regeneration are essential for tumor initiation (**Fig. 4D**). Tumors from KPPTom mice are positive for dephosphorylated (active) β-catenin, which localizes to the cytoplasm and nucleus of cancer cells, as well as phosphorylated AKT^{Ser647} (**Fig. 4E**). As KPPTom mice show both Wnt and PI3K/AKT activity, we treated them with LGK974 (to inhibit Wnt signaling) and pictilisib (to inhibit Pi3K signaling) at 4 weeks following tumor induction (i.e., the start of Thioacetamide treatment), and a time point at which we know that there is ICC forming in the liver (**Fig. 4F**), to determine whether a model of well-differentiated ICC is susceptible to coinhibition of these pathways. Following treatment, the number of tdTomato-positive cancer cells was significantly reduced (by 68.3%) when compared with control vehicle treated animals (**Fig. 4G** and **H**). Indeed, when treated with LGK974 and pictilisib, only 33% of KPPTom mice developed ICC, whereas in the vehicle treated cohort, 80% of KPPTom mice contained cancerous lesions (**Fig. 4I**).

Discussion

ICC is highly complex at the genetic (25) and cellular levels. Although a number of candidate mutations have been identified that can be pharmacologically targeted (19), these have not led to a broadly applicable treatment. Moreover, the presence of these mutations does not necessarily predict therapeutic responsiveness in patients with ICC and only a subset of these patients respond to targeted therapy (20). Consequently, there is a clinical necessity to identify the mechanisms that ICC uses to grow and define whether these processes can be used in patient stratification and targeted treatment. To define a targeted therapeutic approach that can be used in treating a specific cohort of patients with ICC, we need to understand whether the genetic complexity found in ICC ultimately translates to phenotypic diversity or whether ICC relies on a limited number of signaling cascades to grow in spite of this genetic complexity.

Figure 4.

Therapeutic coinhibition of Wnt and PI3K signaling reduces tumor growth in ICC. **A,** RNA-seq data of human ICC demonstrating a positive correlation between the activity of canonical Wnt signaling and Akt signaling. **B,** Schematic representation of the KPPTom cholangiocarcinoma model where Cre^{ERT} expression in Keratin-19–positive cholangiocytes results in the inactivation of *Trp53* and *Pten*, whereas labeling transformed cells with tdTomato. **C,** Representative IHC staining of KPPTom model following tamoxifen administration (day 0) and following 4 and 8 weeks of thioacetamide administration. tdTomato (red) denotes recombined cholangiocytes (denoted by Keratin-19; green). Blue, DNA. Top, whole mount FUNGI images; bottom, 2D histologic sections. Scale bar, 200 μm. White arrows, tdTomato-positive cells. **D,** Quantification of liver tissue occupied by tumor in the KPPTom ICC model. **E,** IHC showing that KPPTom ICC has activated canonical Wnt signaling [by staining for dephosphorylated (active) β-catenin] and PI3K activity (through $pAKT^{Ser647}$ positivity). Red arrows, positive cells. Scale bar, 100 μm. **F,** A schematic representation of how the KPPTom model was used to test the effectiveness of Wnt and Pi3K inhibitor combinations on ICC progression. **G,** IHC staining for tdTomato-positive cancer cells in vehicle-treated animals compared with those treated with a combination of LGK974 and pictilisib. Scale bar, 100 μm. **H,** Number of tdTomato-positive cells in KPPTom animals given vehicle or LGK974 and pictilisib in combination. **I,** Proportion of KPPTom animals containing macroscopic tumors in KPPTom animals treated with vehicle versus combination treatment. b.d., bile duct; p.v., portal vein.

The identification of causative mutations in ICC is fraught with complications. Historically these studies relied heavily on identification of genes with recurrent consensus mutations in patients. This approach has been severely limited by the relatively small number of ICC samples available. Increasing sample size by pooling data from across published cohorts and combining this with a driver prediction pipeline that puts less weight on mutation recurrence within a population, but rather concentrates on the patterns and predicted effects of mutations (10) enabled us to identify an expanded set of candidate mutations in ICC that have the potential to act as oncogenes.

Although able to predict novel candidate oncogenic mutations, the approach described here is not able to infer which mutations act in epistasis to promote tumor formation. To overcome this limitation we developed an *in vivo*, highly multiplexed CRISPR-SpCas9 screening approach. Our strategy builds on previous work demonstrating that smaller gene-editing libraries can be used to define which tumor suppressors are important for ICC initiation and build on a substantial body of work identifying which oncogenes are essential for the formation of cancer cells in the liver (26), particularly KRas, which has previously been used to define the responsiveness to MEK inhibition in ICC (45). Rather than using whole genome screening, we biased our libraries to ensure that they targeted genes mutated in patients with ICC. Using this strategy has enabled us to define a range of putative tumor suppressors that cooperate with gain-of-function *Ras*-mutations to overcome RAS-induced senescence and initiate tumor formation. We validated that one of these, *Nf2* or Merlin, can act as an important factor in the initiation of RAS-driven ICC and we show increased penetrance and aggression of RAS-driven ICC when *Nf2* loss is combined with *Trp53* loss. The phenotypes of these *Nf2*-mutant tumors are highly sarcomatoid, the cancer cells have a more spindle-like morphology and are dispersed throughout the liver. Perhaps this is unsurprising given the role of NF2 in maintaining contact inhibition and restricting cell migration in the presence of Ras mutations, as NF2 has been implicated in the fine tuning of Ras signaling in Schwann cells (46). However, loss of *Nf2* has not been shown to be sufficient to cooperate with Ras and drive tumor formation independent of other accessory mutations and whereas loss of *Nf2* has been previously shown to promote the formation of mixed HCC and ICC when it is deleted from the liver during development (47), it was unclear whether *de novo* somatic *Nf2* mutations interact with other ICC-relevant mutations *in vivo*. In addition to Hippo signaling, NF2 has also been implicated in the regulation of a number of other signaling pathways (48). We found that Wnt and PI3K signaling are recurrently deregulated following *Nf2* loss in our mouse model, leading us to identify coinhibition of Wnt and PI3K as a potent therapeutic combination in reducing ICC growth. In cancer, Wnt signaling is classically activated by mutations in a core set of genes (49). In hepatocellular carcinoma for example, the Wnt signaling pathway is activated via gain-of-function mutations in β-catenin and is known to directly promote tumor progression in mice (50). No such classical activating mutations have been found in ICC genome sequencing (with the exception of rare RNF43 mutations (8)) despite a number of descriptions of ICC being a Wnt-high tumor (35). Our data indicates that other mutations can similarly potentiate Wnt signaling likely by enabling cancer cells to enter a state that is permissive to ligand reception. Critically, we show that Wnt signaling is part of a signaling network that also involves the activation of AKT, which in ICC can be activated through a number of mechanisms and is known to affect the stability of β-catenin by inhibitory phosphorylation of GSK3β. Furthermore, Wnt (porcupine) inhibitors and PI3K inhibitors are currently in clinical trials for other solid malignancies and our data supports recent findings in pancreatic cancer (51), that treating patients with Wnt and PI3K inhibitors provides an attractive therapeutic route to reduce tumor burden.

Conclusions

Our data demonstrate the importance of understanding the function of rare mutations in ICC and show that these low-frequency mutations not only influence the outcome of more common driver mutations, but also can lead us to define applicable therapeutic strategies that can be used to develop personalized approaches that could be applied clinically to stratify patients and to treat ICC of divergent genotypes based on the signaling pathways that are deregulated in these cancers.

Authors' Disclosures

E. Jarman reports grants from Leap Therapeutics, BASL, and CCF outside the submitted work. P. Tennant reports grants from Medical Research Council during the conduct of the study. T.J. Kendall reports personal fees from Incyte Corporation outside the submitted work. M.S. Taylor reports grants from Medical Research Council during the conduct of the study and grants from Wellcome Trust outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

N.T. Younger: Conceptualization, formal analysis, investigation, writing–original draft. **M.L. Wilson:** Conceptualization, formal analysis, validation, investigation, writing–original draft. **A. Martinez Lyons:** Formal analysis, investigation. **E.J. Jarman:** Formal analysis, methodology. **A.M. Meynert:** Data curation, formal analysis. **G.R. Grimes:** Data curation, formal analysis. **K. Gournopanos:** Investigation, methodology. **S.H. Waddell:** Formal analysis, investigation, writing–original draft. **P.A. Tennant:** Resources, methodology. **D.H. Wilson:** Investigation. **R.V. Guest:** Conceptualization. **S.J. Wigmore:** Conceptualization. **J.C. Acosta:** Resources. **T.J. Kendall:** Conceptualization, investigation. **M.S. Taylor:** Formal analysis, investigation. **D. Sproul:** Conceptualization, supervision. **P. Mill:** Conceptualization, investigation, methodology, writing–original draft. **L. Boulter:** Conceptualization, resources, formal analysis, supervision, funding acquisition, investigation, writing–original draft, project administration, writing–review and editing.

Acknowledgments

The authors thank Prof. Roland Rad for his provision of the SB-CRISPR plasmid and Dr. Diego Calvesi for providing the $Kras^{G12D}$-GFP plasmids. L. Boulter was funded by The Wellcome Trust (207793/Z/17/Z), AMMF (2016/108, 2017/115), and Cancer Research UK (C52499/A27948). L. Boulter is also supported by an MRC university grant to the MRC Human Genetics Unit.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 2, 2021; revised December 14, 2021; accepted January 18, 2022; published first January 24, 2022.

References

1. Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol 2020;17:557–88.
2. Valle JW, Kelley RK, Nervi B, Oh D-Y, Zhu AX. Biliary tract cancer. Lancet 2021; 397:428–44.
3. Adeva J, Sangro B, Salati M, Edeline J, La Casta A, Bittoni A, et al. Medical treatment for cholangiocarcinoma. Liver Int 2019;39:123–42.
4. Petmitr S, Pinlaor S, Thousungnoen A, Karalak A, Migasena P. K-ras oncogene and p53 gene mutations in cholangiocarcinoma from Thai patients. Southeast Asian J Trop Med Public Health 1998;29:71–75.

5. Hill MA, Alexander WB, Guo B, Kato Y, Patra K, O'Dell MR, et al. Kras and Tp53 mutations cause cholangiocyte- and hepatocyte-derived cholangiocarcinoma. Cancer Res 2018;78:4445–51.
6. Goeppert B, Folseraas T, Roessler S, Kloor M, Volckmar A-L, Endris V, et al. Genomic characterization of cholangiocarcinoma in primary sclerosing cholangitis reveals therapeutic opportunities. Hepatology 2020;72:1253–66.
7. Zou S, Li J, Zhou H, Frech C, Jiang X, Chu JSC, et al. Mutational landscape of intrahepatic cholangiocarcinoma. Nat Commun 2014;5:5696.
8. Chan-On W, Nairismägi M-L, Ong CK, Lim WK, Dima S, Pairojkul C, et al. Exome sequencing identifies distinct mutational patterns in liver fluke-related and non-infection-related bile duct cancers. Nat Genet 2013;45:1474–8.
9. Jusakul A, Cutcutache I, Yong CH, Lim JQ, Huang MN, Padmanabhan N, et al. Whole-genome and epigenomic landscapes of etiologically distinct subtypes of cholangiocarcinoma. Cancer Discov 2017;7:1116–35.
10. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A, et al. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods 2013;10:1081–2.
11. Sia D, Losic B, Moeini A, Cabellos L, Hao K, Revill K, et al. Massive parallel sequencing uncovers actionable FGFR2-PPHLN1 fusion and ARAF mutations in intrahepatic cholangiocarcinoma. Nat Commun 2015;6:6087.
12. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014;343:84–87.
13. Read A, Gao S, Batchelor E, Luo J. Flexible CRISPR library construction using parallel oligonucleotide retrieval. Nucleic Acids Res 2017;45:e101.
14. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. Nat Biotechnol 2020;38:276–8.
15. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinformatics 2013;14:178–92.
16. Rios AC, Capaldo BD, Vaillant F, Pal B, van Ineveld R, Dawson CA, et al. Intraclonal plasticity in mammary tumors revealed through large-scale single-cell resolution 3D imaging. Cancer Cell 2019;35:618–32.
17. Sriskandarajah P, De Haven Brandon A, MacLeod K, Carragher NO, Kirkin V, Kaiser M, et al. Combined targeting of MEK and the glucocorticoid receptor for the treatment of RAS-mutant multiple myeloma. BMC Cancer 2020;20:269.
18. Lowery MA, Burris HA, Janku F, Shroff RT, Cleary JM, Azad NS, et al. Safety and activity of ivosidenib in patients with IDH1-mutant advanced cholangiocarcinoma: a phase 1 study. Lancet Gastroenterol Hepatol 2019;4:711–20.
19. Javle MM, Borbath I, Clarke SJ, Hitre E, Louvet C, Mercade TM, et al. Infigratinib versus gemcitabine plus cisplatin multicenter, open-label, randomized, phase 3 study in patients with advanced cholangiocarcinoma with FGFR2 gene fusions/translocations: the PROOF trial. JCO 2019;37:TPS4155.
20. Bekaii-Saab TS, Bridgewater J, Normanno N. Practical considerations in screening for genetic alterations in cholangiocarcinoma. Ann Oncol 2021;32:1111–26.
21. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, et al. Genomic spectra of biliary tract cancer. Nat Genet 2015;47:1003–10.
22. Farshidfar F, Zheng S, Gingras M-C, Newton Y, Shih J, Robertson AG, et al. Integrative genomic analysis of cholangiocarcinoma identifies distinct IDH-mutant molecular profiles. Cell Rep 2017;19:2878–80.
23. Wu G, Feng X, Stein L. A human functional protein interaction network and its application to cancer data analysis. Genome Biol 2010;11:R53.
24. Newman MEJ. Modularity and community structure in networks. Proc Natl Acad Sci USA 2006;103:8577–82.
25. Dong L-Q, Shi Y, Ma L-J, Yang L-X, Wang X-Y, Zhang S, et al. Spatial and temporal clonal evolution of intrahepatic cholangiocarcinoma. J Hepatol 2018;69:89–98.
26. Weber J, Öllinger R, Friedrich M, Ehmer U, Barenboim M, Steiger K, et al. CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. Proc Natl Acad Sci USA 2015;112:13982–7.
27. Wang J, Dong M, Xu Z, Song X, Zhang S, Qiao Y, et al. Notch2 controls hepatocyte-derived cholangiocarcinoma formation in mice. Oncogene 2018;37:3229–42.
28. Fan B, Malato Y, Calvisi DF, Naqvi S, Razumilava N, Ribback S, et al. Cholangiocarcinomas can originate from hepatocytes in mice. J Clin Invest 2012;122:2911–5.
29. Tennant PA, Foster RG, Dodd DO, Sou IF, McPhie F, Younger N, et al. Fluorescent *in vivo* editing reporter (FIVER): a novel multispectral reporter of *in vivo* genome editing. BioRxiv 2020.
30. Kang T-W, Yevsa T, Woller N, Hoenicke L, Wuestefeld T, Dauch D, et al. Senescence surveillance of pre-malignant hepatocytes limits liver cancer development. Nature 2011;479:547–51.
31. Morton JP, Timpson P, Karim SA, Ridgway RA, Athineos D, Doyle B, et al. Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. Proc Natl Acad Sci USA 2010;107:246–51.
32. Kim M, Kim S, Lee SH, Kim W, Sohn MJ, Kim HS, et al. Merlin inhibits Wnt/β-catenin signaling by blocking LRP6 phosphorylation. Cell Death Differ 2016;23:1638–47.
33. Malhotra S, Wood J, Mansy T, Singh R, Zaitoun A, Madhusudan S. Intrahepatic sarcomatoid cholangiocarcinoma. J Oncol 2010;2010:701476.
34. Kim DK, Kim BR, Jeong JS, Baek YH. Analysis of intrahepatic sarcomatoid cholangiocarcinoma: experience from 11 cases within 17 years. World J Gastroenterol 2019;25:608–21.
35. Boulter L, Guest RV, Kendall TJ, Wilson DH, Wojtacha D, Robson AJ, et al. WNT signaling drives cholangiocarcinoma growth and can be pharmacologically inhibited. J Clin Invest 2015;125:1269–85.
36. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad IB, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet 2012;44:694–8.
37. Sato K, Murai H, Ueda Y, Katsuda S. Intrahepatic sarcomatoid cholangiocarcinoma of round cell variant: a case report and immunohistochemical studies. Virchows Arch 2006;449:585–90.
38. Mah KM, Houston DW, Weiner JA. The γ-Protocadherin-C3 isoform inhibits canonical Wnt signalling by binding to and stabilizing Axin1 at the membrane. Sci Rep 2016;6:31665.
39. O'Rourke CJ, Matter MS, Nepal C, Caetano-Oliveira R, Ton PT, Factor VM, et al. Identification of a pan-gamma-secretase inhibitor response signature for notch-driven cholangiocarcinoma. Hepatology 2020;71:196–213.
40. Carnero A, Blanco-Aparicio C, Renner O, Link W, Leal JFM. The PTEN/PI3K/AKT signalling pathway in cancer, therapeutic implications. Curr Cancer Drug Targets 2008;8:187–98.
41. Xu X, Kobayashi S, Qiao W, Li C, Xiao C, Radaeva S, et al. Induction of intrahepatic cholangiocellular carcinoma by liver-specific disruption of Smad4 and Pten in mice. J Clin Invest 2006;116:1843–52.
42. Chen J, Debebe A, Zeng N, Kopp J, He L, Sander M, et al. Transformation of SOX9+ cells by Pten deletion synergizes with steatotic liver injury to drive development of hepatocellular and cholangiocarcinoma. Sci Rep 2021;11:11823.
43. Guest RV, Boulter L, Dwyer BJ, Kendall TJ, Man T-Y, Minnis-Lyons SE, et al. Notch3 drives development and progression of cholangiocarcinoma. Proc Natl Acad Sci USA 2016;113:12250–5.
44. Dawson CA, Pal B, Vaillant F, Gandolfo LC, Liu Z, Bleriot C, et al. Tissue-resident ductal macrophages survey the mammary epithelium and facilitate tissue remodelling. Nat Cell Biol 2020;22:546–58.
45. Wang P, Song X, Utpatel K, Shang R, Yang YM, Xu M, et al. MEK inhibition suppresses K-Ras wild-type cholangiocarcinoma *in vitro* and *in vivo* via inhibiting cell proliferation and modulating tumor microenvironment. Cell Death Dis 2019;10:120.
46. Cui Y, Groth S, Troutman S, Carlstedt A, Sperka T, Riecken LB, et al. The NF2 tumor suppressor merlin interacts with Ras and RasGAP, which may modulate Ras signaling. Oncogene 2019;38:6370–81.
47. Benhamouche-Trouillet S, O'Loughlin E, Liu C-H, Polacheck W, Fitamant J, McKee M, et al. Proliferation-independent role of NF2 (merlin) in limiting biliary morphogenesis. Development 2018;145:dev162123.
48. Attisano L, Wrana JL. Signal integration in TGF-β, WNT, and Hippo pathways. F1000Prime Rep 2013;5:17.
49. Monga SP. β-Catenin signaling and roles in liver homeostasis, injury, and tumorigenesis. Gastroenterology 2015;148:1294–310.
50. Qiao Y, Wang J, Karagoz E, Liang B, Song X, Shang R, et al. Axis inhibition protein 1 (Axin1) deletion-induced hepatocarcinogenesis requires intact β-catenin but not notch cascade in mice. Hepatology 2019;70:2003–17.
51. Zhong Z, Sepramaniam S, Chew XH, Wood K, Lee MA, Madan B, et al. PORCN inhibition synergizes with PI3K/mTOR inhibition in Wnt-addicted cancers. Oncogene 2019;38:6662–77.