PhD degree in Systems Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples “Federico II”

Settore disciplinare: BIO/13

EPIGENETIC MODIFICATIONS REGULATING LIVER HOMEOSTASIS AND TUMORIGENESIS

D’Ambrosio Alessandro

CIBIO, Trento

Matricola n. R12413

Tutor: Prof. Fulvio Chiacchiera, CIBIO, Trento

Co-Tutor: Prof. Diego Pasini, IEO, Milano

PhD Coordinator: Prof. Saverio Minucci

Anno accademico 2022/2023

TABLE OF CONTENT

LIST OF ABBREVIATIONS

Biliary epithelial cells	BECs
BRG1- or BRM-associated factors	BAF
Chromatin immunoprecipitation	ChIP
chromatin remodelling complexes	CRCs
Colorectal cancer	CRC
Cre-recombinase Estrogen Receptor T2	CreERT2
double Knock-Out	DKO
Hepatocellularcarcinoma	HCC
Histone H2A lysine K119 mono ubiquitin	H2AK119ub
Histone H3 lysine K27 acetylated	H3K27ac
Histone H3 lysine K27 monomethylated	H3K27me1
Histone H3 lysine K27 trimethylated	H3K27me3
Histone H3 lysine K4 monomethylated	H3K4me1
Histone H3 lysine K4 trimethylated	H3K4me3
Immunofluorescence	IF
Immunohistochemistry	IHC
Intraperitoneal injection	IP injection
Intrahepatic cholangiocarcinoma	iCCA
Knock-out	KO
Liver progenitor-like cells	LPLCs
Optimal Cutting Temperature	OCT
Polycomb Repressive Complex 1	PRC1
Polycomb Repressive Complex 2	PRC2

Polymerase Chain Reaction	PCR
Post translational modification	PTM
primary liver cancer	PLC
Read Per Kilobases of Million Mapped Reads	RPKM
Real time quantitative PCR	qRT-PCR
RNA sequencing	RNA-seq
SWItch/Sucrose Non-Fermentable	SWI/SNF
Tamoxifen	TAM
TCPOBOP	TC
Transcription Factor	TF
Western blot	WB
Wild-type	WT

1. ABSTRACT

Hepatocellular carcinoma (HCC) is one of the most common tumors worldwide and the fourth cause of death. The major risk factors predisposing to HCC are well known although the lack of reliable markers and efficient pharmacological treatments for HCC patients remains a major problem. In the recent years, several driver mutations have been identified, including signaling molecules and epigenetic modifiers. Specifically, mutations that affect the activity of SWI/SNF chromatin remodeling complex have been commonly reported. Mutational studies clearly indicated a role for epigenetic modifications in HCC development. However, it is still unclear whether SWI/SNF mutations have a role in tumor formation. Our data suggest a new role of Arid1A as one of the key players in HCC tumor development.

In our study, we show that ARID1A predispose to HCC formation with the establishment of chronic inflammation. HCC derived from Arid1A deprivation are characterized by the selection of mutations able to induce immune escape. Conversely, the abrogation of all the cBAF complexes, targeting both Arid1A and Arid1B subunits, led to synthetic lethality in a couple of weeks. Moreover, loss of all the Arid1 proteins caused a strong effect on transcriptional profile with loss of all liver specific metabolic pathway. Furthermore, we observed how the absence of both the ARID1 proteins characterized a complete loss of H3K4me1 and H3K27Ac histone modifications on all the FoxA2 targets both on promoter and enhancers. Ultimately, cBAF complex is necessary to maintain the liver metabolism and its homeostasis. Mechanistically, ARID1A loss induced chronic inflammation and, probably indirectly, genomic instability with the upregulation of the interferon pathway due to an increase of micronuclei. This statement imposed a selective pressure which deeply predispose to HCC formation by promoting viral mimicry. All these findings led to actively select mutations to bypass this block induced by the innate immune response activation, among these mutations were found MYC, YAP1, PTEN and S33A β-Catenin mutation.

2. INTRODUCTION

2.1. HCC OVERVIEW

Two distinct subtypes of liver cancer have been described in the literature, and the differences between them depend on specific parenchymal liver cells from which the tumor originates. One of the main liver tumors is Hepatocellular Carcinoma (HCC, 80% of the cases), it is characterized by malignant transformation of the principal, and more abundant, cell types that characterized liver parenchyma, hepatocytes. The second, by incidence, is intrahepatic Cholangiocarcinoma (iCCA, 20% of the cases) resulting from cholangiocytes transformation. Despite differences in etiological factors and origins, they both can be originate from chronic liver damage, in which the coexistence of damage and regeneration are able to induce a pro-oncogenic environment, promoting selection of specific mutations[1] that could lead to tumor development. These tumors could derive from the three different resident cells types (hepatocytes, cholangiocytes, and progenitor cells). Depending on the single target of malignant transformation, liver oncogenic development is characterized by a heterogeneous spectrum of histological patterns that define the type of tumor: HCC, iCCA, or the combination of the two HCC–iCCA[2]. Although, common oncogenic pathways are activated, an intrinsic heterogeneity due to different cells of origin can be retrieved in human primary liver cancers, highlighting the necessity of individualized therapeutic approaches[3].

Hepatocellular carcinoma is the most common primary malignancy of the liver, the fourth most common cause of cancer-related death worldwide. Moreover the 80% of HCC cases occur in low-resource and middle-resource countries, in particular in Eastern Asia and sub-Saharan Africa, where medical care resources are often constrained[4]. HCC is more common among males, with a male-female ratio of 2.4 worldwide, and the most common age of appearance is usually between 30 and 50 years[5].

This type of cancer is considered an unmet medical challenge due to several reasons: the most prevalent oncogenic mutations are undruggable, where a molecular subclass or a pathway can be identified in a patient, it is difficult to translate into clinical trials. Tumor heterogeneity and the role of the immune system in HCC development make the patients stratification difficult in the field of precision medicine[6].

HCC has a multitude of etiological risk factors, and in particular, it was demonstrated that some of them have a strong association with its development[5]. HBV, HCV, and hepatitis D virus (HDV) have a strong association with development of HCC. HBV chronic infection caused an increase of the relative risk for developing HCC of 15-fold with a mortality rate of approximately 30%–50% among all cases of chronic HBV infection. Moreover, another important risk that determines more than 85% of HCC cases is the presence of advanced hepatic fibrosis and/or cirrhosis[1]. Lastly, the other major risk factors are alcohol abuse, non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH)[5].

HCC is an invasive tumor that metastasize hematogenously and lymphogenously reaching other organs, even after local recurrence. The metastases are mainly localized in lungs, bone, brain, and lymph nodes but this kind of tumor are able to form intra-organ metastasis[7,8].

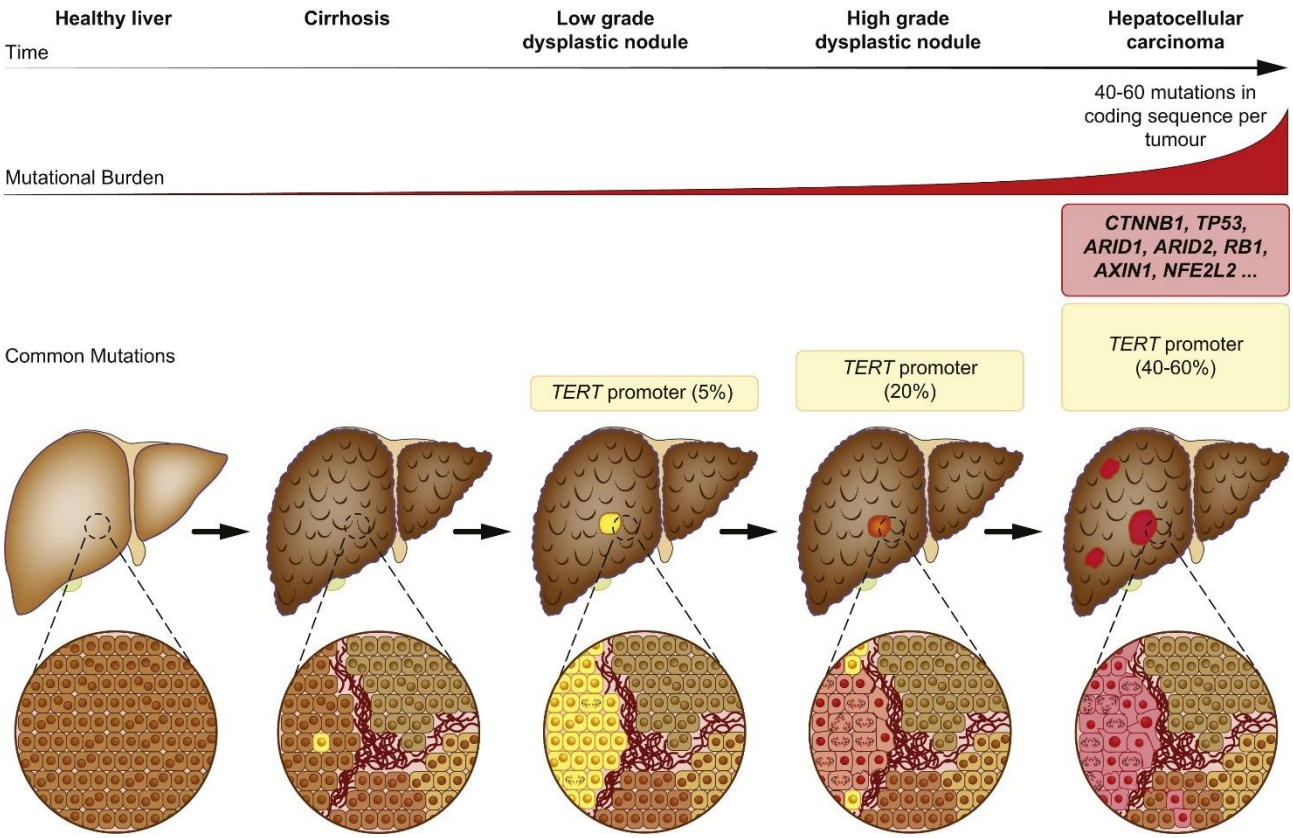


Figure 2.1 HCC development is a multistep process

Schematic representation of liver cancer transformation from healthy liver to HCC. The first step is a establishment of cirrhosis with accumulation of fibrosis; subsequent mutations are responsible of the tumor initiation and development[9].

HCC transformation is a complex process characterized by an early step which is defined by an increase of fibrosis[9]. Indeed cirrhosis is a pre-tumoral stage characterized by massive inflammation, and this state is found in about the 90% of patients with HCC; instead in the remaining 10% of the patients, noncirrhotic mechanisms of carcinogenesis are responsible for HCC development[10]. It is clear that HCC developed in a setting of underlying cirrhosis results from a complex interplay among genetic and non-genetic factors, exposure to environmental carcinogens and viruses[5]. This is a multistep process, starting from cirrhosis, followed by steatohepatitis and then a precise sequence of lesions: low grade dysplastic nodules (LGDN), high grade dysplastic nodules (HGDN), early HCC, progressed HCC and

advanced HCC (with the possibility of metastases dissemination)[11]. Within this context hepatocytes switch to high proliferative cells vascularized by the portal system, and they acquire an invasive and metastatic potential[12]. There are evidences showing that telomere and telomerase play a critical role in liver cirrhosis and HCC progression. Indeed, somatic mutations in TERT promoter is one of the most common triggering events found in the HCC progression from the LGDN to the early HCC increasing the risk of lesion progression also in aggressiveness[13].

2.2. GENETIC LANDSCAPE

Due to the identification of around 28000 different somatic mutations, Hepatocellular carcinoma shows an highly heterogeneous nature. This renders early detection and prevention extremely challenging[14]. Most HCCs generally present 21 silent and 64 non-silent mutations, which affect from 8 to 11 main molecular pathways and biological processes[15]. Initially 30 putative driver genes and 11 core pathways were identified for HCC progression, with TERT activation as a central factor[16]. Subsequent studies highlighted that HBV integration at TERT locus can promote hepatocarcinogenesis[17]. Moreover, TERT genetic alterations were associated with activating of CTNNB1 mutations[18].

Recently, four oncogenic networks have been analyzed and are considered as paramount in HCC development: metabolic processes, WNT—β-catenin pathway, chromatin remodeling complexes and PI3K-AKT-mTOR activation[19]. The most frequently mutated gene is TERT, in particular in its promoter (60%) associated with an increased telomerase expression, followed by TP53 and CTNNB1 mutated in the 25-30% of HCC patients[6].

Regarding chromatin modifiers, their role in cancer development acquired relevance in the last decades, due to their activity involved in altering nucleosome structure and accessibility.

The biological consequences of their alteration are directly linked with differentiation, proliferation, and DNA repair. Inactivating mutations and homozygous deletions in ARID1A, ARID1B and ARID2, subunits of the SWI/SNF complexes, suggests their role as tumor suppressor genes in HCC development[20]. Recurrent somatic alterations in the histone methylation writer family are frequent in HCC, in particular in MLL, MLL2, MLL3 and MLL4 genes[19].

Mutations in key genes of the oxidative stress pathway have been directly linked to HCC tumorigenesis. Specifically, somatic mutations of NRF2, IRF2, and KEAP1 were described. It has been observed that NFE2L2 and KEAP1 mutations occur in 14% of HCC specimens[15]. By amplification of the FGF19/CCND1 locus, PI3K/AKT/MTOR and RAS/RAF/mitogen-activated protein kinase pathway are activated in 5-10% of HCC. Moreover, it has been observed that activating mutations in PIK3CA and inactivating mutations of TSC1 or TSC2 are associated with the triggering of the AKT pathway in a context of HCC. Additionally, homozygous deletion of PTEN, PI3K inhibitor, has been detected in a subset of HCC patients[19].

Overall, the most affected pathway in HCC involves gene associated with WNT-β-catenin signaling, counting for more than 60% of the cases[21]. Genetic alterations are mainly found in CTNNB1, but also in APC and AXIN1. Deregulation of this specific signaling pathway results in uncontrolled cellular growth by activation of MYC and cyclin D1[1]. Indeed, it was recently shown how CTNNB1 and MYC cooperate to the transcriptional activation of YAP/TAZ to induced HCC formation[22].

Hepatocellular Carcinoma can also arise in a non-cirrhotic context, from the malignant transformation of hepatocellular adenoma, which is a monoclonal benign hepatocytes proliferation. Hepatocellular adenoma can be classified in 4 molecular subclasses according to mutations in specific pathways[23]. In this context, the earlier steps of tumorigenesis are characterized by activating mutation in exon 3 of CTNNB1, which is the locus of its

phosphorylation to induce its degradation, followed by TERT promoter mutations and global hypermethylation in the last steps of transformation[24].

3. LIVER

3.1. LIVER ANATOMY

Liver is the largest solid organ and gland in the human body and it is part of the digestive system. Among the many roles of the liver there are xenobiotic and endobiotic detoxification and the production of chemicals like bile to complete the food digestion process, fats and proteins catabolism, absorption of fats, cholesterol, and vitamins. Bile is essential for vitamin K production, crucial in blood clots formation.

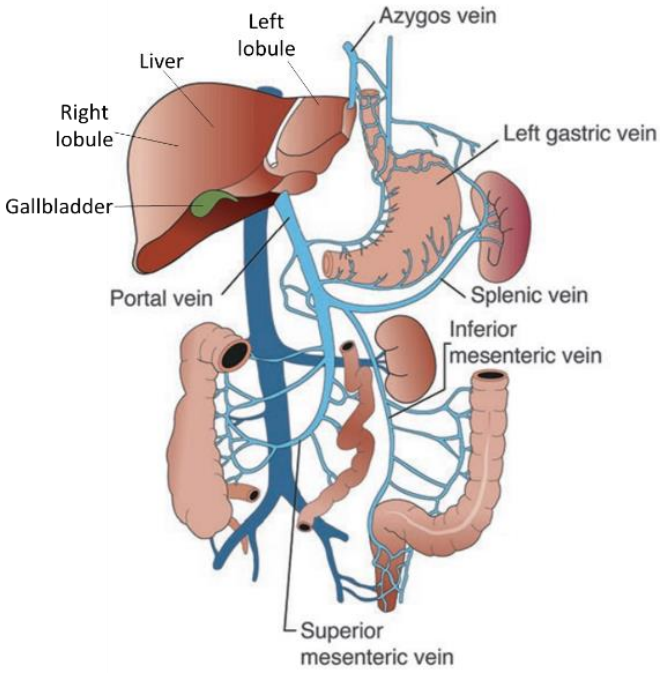


Figure 3.1 liver and portal vein circulation [24]

It is located in the upper right-hand part of the abdomen, in particular under the diaphragm, and on top of the stomach, right kidney, and intestines (Fig. 3.1). It exhibits a reddish-brown color with a rubbery texture. It is roughly triangular and consists of two lobes: a larger right lobe and a smaller left lobe, separated by the falciform ligament, a band of tissue that keeps it anchored to the diaphragm. The lobes are constituted by 8 segments that have 1 thousand of repetitive functional unit called lobules[25,26]. The liver is composed by distinct types of

resident cell: hepatocytes, biliary epithelial cells (BECs; also known as cholangiocytes), progenitor cells (LPCs), and stellate cells[27].

The lobules are connected to small ducts that lead to larger ducts, forming the common hepatic duct, which sends the bile produced by the hepatocytes to the gallbladder and then to the duodenum through the common bile duct[28]. The lobule is reached by two major sources of blood: the portal vein which brings inside nutrient-rich blood from the digestive system, and the hepatic artery carries oxygenated blood from the heart. The blood vessels branch out into small capillaries, each of which ending in a lobule. Blood flows from the liver through three hepatic veins. A layer of fibrous tissue called Glisson's capsule covers the outside of the liver. This capsule is further covered by the peritoneum, a membrane that forms the lining of the abdominal cavity and protects the liver from physical damage[29].

3.2. LIVER ARCHITECTURE AND ROLES

The basic model of the repeat unit of the liver include: the central vein (CV)- which is centered in the liver lobule, the portal triad (PT) and the hepatic acinus[28,27]. In general, a lobule consists of a single centrally located CV and three to six PT blood sources located at the vertexes of a roughly hexagonal region centered at the CV. Each PT contains a branch of the artery, a branch of the portal vein and bile duct[30]. Either of these units can construct a liver lobule in which blood flows from the hepatic artery and portal veins (which delivers venous blood from the small intestine) of the PT, across a complex network of sinusoids draining into the central vein[27]. In the CV-centered liver lobule, blood flow converges from

peripheral to central regions. In contrast, in the PT-centered model, blood flow diverges from the PTs[28,26] (Fig 3.2).

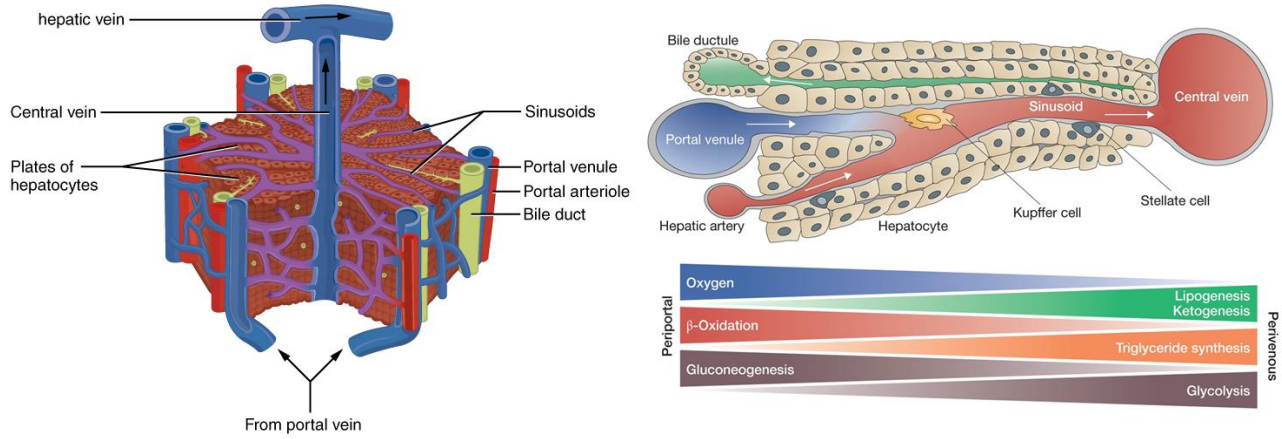


Figure 3.2 schematic representation of liver lobule (left) and liver sinusoid (right) with related fluid circulation, adapted from[25,31].

In the lobules is possible to retrieve specific subregions termed zones, not visible in standard histological sections. In order to visualize the specific zones in liver-tissue sections, either staining for zonally restricted proteins or the observation of localized necrosis following exposure to a liver toxicant are required[27,26,31,32].

The zones visualized in tissue sections appear as concentric rings centered on either the CV or a PT. In general, the lobule is defined on the basis of three main zones: Zone I is periportal and includes the PTs and post terminal hepatic venules; Zone III is the region around the CV; Zone II includes the region intermediate between the periportal and pericentral regions[27,30]. Zone I is primarily responsible for glucose metabolism and ammonia detoxification processes; Zone III carries out most lipogenesis and glutamine synthesis and expresses high levels of xenobiotic metabolizing cytochrome P450s such as CYP1a2 and CYP2e1. Zone II is a transitional region whose capabilities are generally not characterized but presumably has capabilities intermediate between those of Zone I and Zone III[33]. The decreasing blood oxygen concentration from the periportal to pericentral regions may allow

individual hepatocytes to determine their position along the portal to central axis, hence their zonal position, and to change behavior accordingly[27,34].

Hepatocytes are the functional cells of the liver and perform an astonishing number of metabolic, endocrine and secretory functions. Moreover, they contribute roughly to 80% of the liver mass. Hepatocyte nuclei are distinctly round, with one or two prominent nucleoli. A majority of cells have a single nucleus, however in rat the 80%-90% of adult hepatocytes are polyploid, compared to 30-40% in humans [35,36]. The cells possess a polygonal shape, and their sides can be in contact either with sinusoids (sinusoidal face) or neighboring hepatocytes (lateral faces). A portion of the lateral faces of hepatocytes is modified to form bile canaliculi. Hepatocytes are exceptionally active in synthesis of protein and lipids for export. Consequently, examination of hepatocytes reveals remarkable quantities of both rough and smooth endoplasmic reticulum. In contrast to most glandular epithelial cells that contain a single Golgi organelle, hepatocytes typically contain plenty of Golgi membranes[37]. The liver is classified as a gland and is the paramount site of metabolism and clearance of xenobiotics from the blood[38,39]. Despite the difficulties in retrieving a precise number, it is thought that the liver carries out 500 distinct roles for the proper functioning of the organism[39]. The liver is irrigated by the portal vein which carries molecules obtained after digestion in the intestine. Carbohydrates are stored in the hepatocytes, and consequently released in the bloodstream when a quick burst of energy is required[40]. Bile is synthetized inside the hepatocytes and released in the intestine, and is responsible for plenty of liver functions: fats and proteins catabolism, absorption of fats, cholesterol, and vitamins. Bile is essential for vitamin K production, crucial in blood clots formation[41]. Liver is the organ responsible for the synthesis of numerous molecules: lipoproteins for lipids transport in the bloodstream; fibrinogen; plasma albumin; angiotensinogen; urea as a byproduct of proteins degradation; heme group for hemoglobin synthesis. Liver is involved in the storage of a plenty of molecules: vitamins, iron in form of ferritin, copper, and is pivotal for the removal

of compounds from the body, such as hormones, alcohol and drugs[37,38,39]. Remarkably, liver is considered part of the mononuclear phagocyte system, indeed it is characterized by a strong presence of Kupffer cells that are involved in immune activity. These cells are able to destroy any disease-causing agents that might enter the liver through the gut, by removing foreign particles from the portal system through phagocytosis and pinocytosis, and also by removing apoptotic cells from circulation[35].

3.3. DETOXIFICATION PATHWAY

Hepatocytes are the cell type deputed to receive potentially toxic substances, in order to perform the detoxifying activity of endobiotic and xenobiotic compounds[39]. The constitutive androstane receptor (CAR), through modulating the transcription of its numerous target genes, exerts a crucial function in the regulation of drug metabolism, energy homeostasis, and cancer development[42].

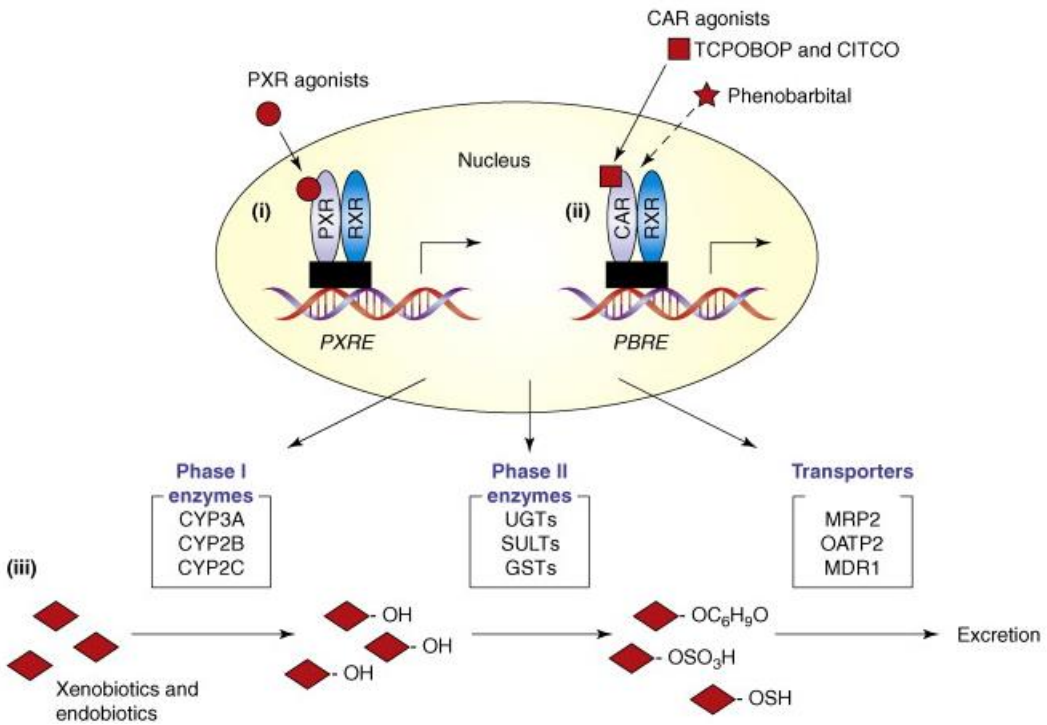


Figure 3.3 Schematic representation of CAR pathway target genes, showing how its agonist works and what are its target[42]

Differently from prototypical nuclear receptors, CAR can be activated by either direct ligand binding or ligand-independent binding, both mechanisms starting with nuclear translocation of CAR from the cytoplasm into the nucleus[42]. Once translocated, CAR can interact with other co-activator proteins and maintain a constitutive activated status. The highly conserved DBD domains in CAR contain unique structures able to recognize and bind

specific promoter regions in target genes, namely xenobiotic response elements such as the PBREM and XREM in the CYP2B6 promoter. CAR can regulate gene expression by forming heterodimers with the retinoid X receptor and subsequently bind to xenobiotic responsive elements present in the promoter regions of drug-metabolizing enzyme and transporter genes.

Liver detoxification pathway is based on two types of enzymatic modifications which have been divided in two phases, Phase I and Phase II, which are characterized by different enzymatic involvement. The Phase I is exerted by cytochrome P450 superfamily (CYP450). It is the first step of the body to transform xenobiotics, steroid hormones, and drugs. In particular, the CYP450 add a reactive group such as a hydroxyl, carboxyl, or an amino group through oxidation, reduction, and/or hydrolysis reactions[38,39]. This first step of the detoxification reaction could create oxidative damage within cell systems with formation of reactive species. After a xenobiotic has gone through the phase I becoming hydrophilic through reactions overseen by CYP450 enzymes, its reactive site can be conjugated with an endogenous hydrophilic substance. With the phase II of this pathway there is the transfer of hydrophilic compounds (via their corresponding enzymes), including glucuronic acid (glucuronyl transferases), sulfate (sulfotransferases), glutathione (glutathione transferases), amino acids (amino acid transferases), an acetyl group (N-acetyl transferases), and a methyl group (N- and O-methyltransferases) [38,43].

3.4. LIVER REGENERATION

In contrast to all other organs, liver's regenerative capacity is remarkable. The complex mechanisms involved in this uniquely hepatic regenerative process has been called "hepatostat"[44,45]. Another characteristic of the liver is the lack of a stem cell pool able to re-populate the liver tissue[46]. While this could be considered a limitation in regenerative capabilities, it is completely coherent with liver cell physiology. DNA labelling studies have shown that the percentage of hepatocytes in DNA synthesis in normal liver is very low (0.2%), corresponding to one or two cell division per year needed mostly to maintain the tissue[47]. This turnover is not compatible with stem cell like regeneration since it requires high tissue turnover (like the intestine) and is kind of slow. After injury, liver must achieve fast regeneration otherwise the body will suffer of metabolic and endocrine dysfunction[47,48]. The origin of regeneration depends on the type of insult: the entity of the damage could activate different type of cells; indeed, after a small resection, about one third, the hepatocytes go to hypertrophy. If the partial hepatectomy was of two third, the hepatocytes start to proliferate (Fig.3.4.1). In mammals, mouse model, when the resection is 80-90% of its mass, BECs are activated and can dedifferentiate in LPLCs. After the amplification, these cells are able to differentiate in cholagiocytes and hepatocytes (Fig.3.4.2)[48,49].

The regeneration is due to compensatory growth of the hepatocytes. This is the most studied form of liver regeneration occurring after loss of hepatocytes in a single acute injury, such as partial hepatectomy or administration of damaging chemicals (CCl4, acetaminophen)[44,45].

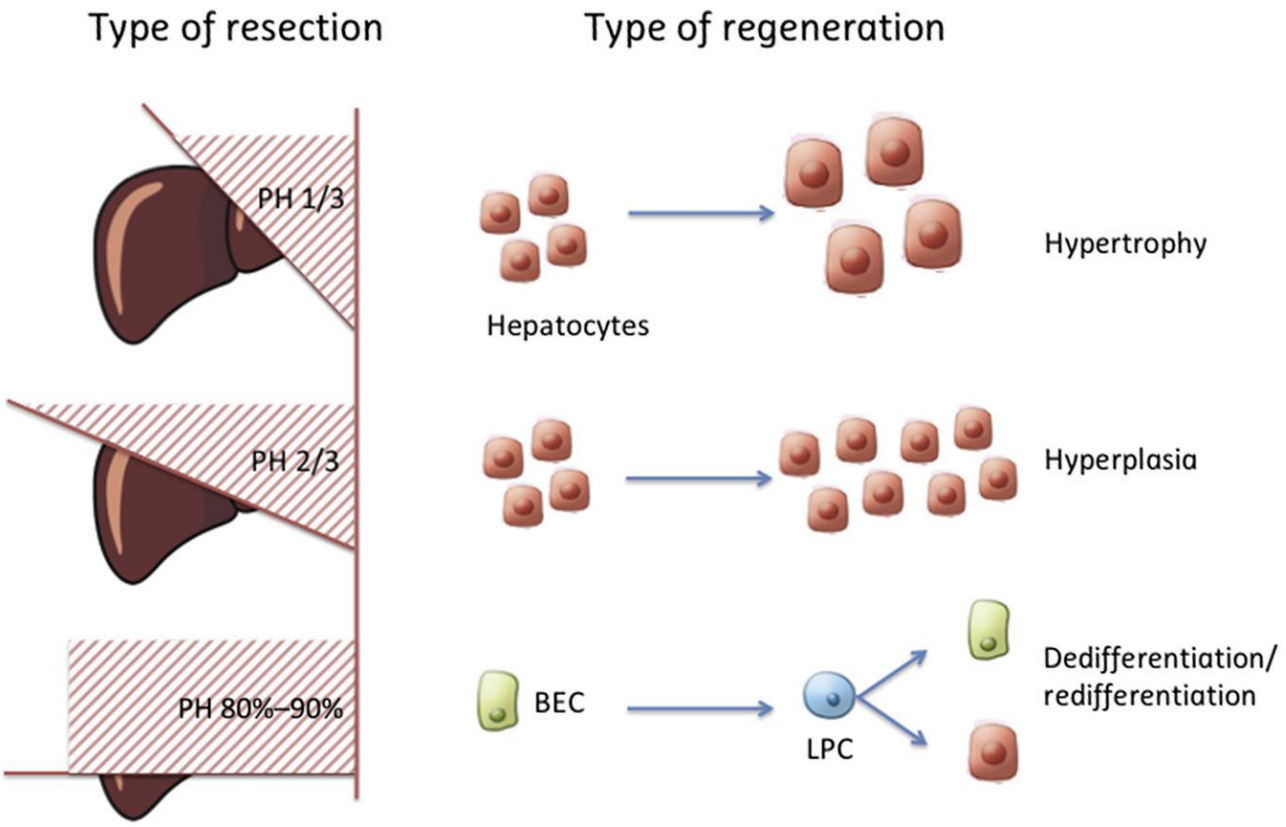


Figure 3.4.1 Liver regeneration

Schematic representation of the different regeneration mechanism of the liver due to a different type of surgery[47].

This normally happen also under physiological condition and is mediated by a complex interplay of signals between hepatocytes and surrounding cells, starting from fundamental mitogenic signals required for proliferation like hepatocyte and Epithelial growth factors (HGF, EGF), complex signals for organ self-perception and homeostasis like YAP, WNT and hedgehog (HGG), and auxiliary mitogen signals like interleukin-6 (IL6), noradrenaline (NA), tumor necrosis factor (TNF), vascular endothelial growth factor (VEGF), that are not essential but necessary to accelerate the regenerative process[47,50]. This process occurs because hepatocytes and cholangiocytes are very plastic and they can transdifferentiate one in each other, acquiring properties as “facultative stem cells” (Fig.3.4.2)[49].

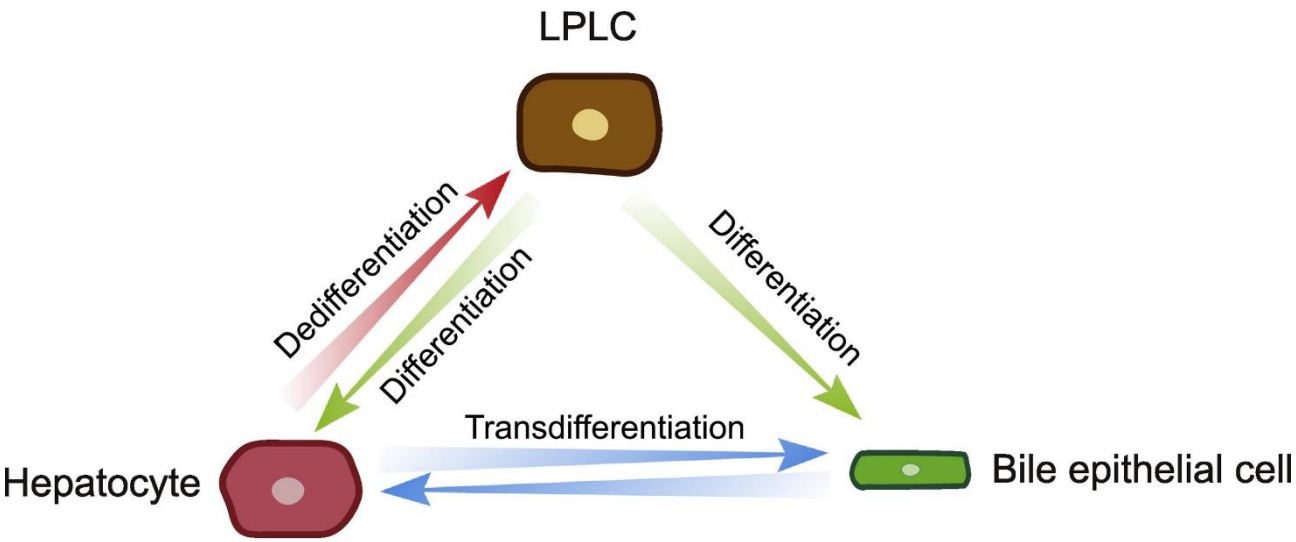


Figure 3.4.2 Liver progenitors like cells

Hepatocytes and cholangiocytes are very plastic. Due to a give stimulus, they can differentiate, dedifferentiate and transdifferentiate to trigger the regeneration if liver parenchyma in order to response to a specific damage[48].

Whether this transformation is transient representing metaplasia, or a stable process of transdifferentiation, it remained unclear. Recent reports have shown successful mutual transdifferentiation, dependent on the extent of injury and the cell type predominantly affected by that injury. Different pathway previously involved in hepatocytes regeneration has also been interrogated using molecular engineering in vivo models[51] but unfortunately there still some controversy.

Finally, the last method that the liver adopts for regeneration is based on a pool of cells called liver progenitor-like cells (LPLCs). They can originate from cholangiocytes or from the dedifferentiation of hepatocytes[52]. These cells are characterized by the expression of both hepatocyte and cholangiocytes marker, HNF4α+/SOX9+, and increase under sustained and extended liver injury[53,54]. The importance of these cells is the capability to regenerate both type of cell, and they could be useful for therapeutical approach.

4. SWI/SNF COMPLEXES

4.1. CHROMATIN STRUCTURE

In every eukaryotic cell the genomic DNA is compartmentalized within the nucleus and it take interaction with histone proteins in order to assume a complex structure, the chromatin. This structure allows packaging of the genetic material and it is pivotal for regulation of gene expression. The chromatin is organized in nucleosomes which are the chromatin functional unit. These are formed by 147 base pairs wrapped around a histone octamer, which consists of two molecules of histones H2A, H2B, H3 and H4 (Fig. 4.1)[55,56]. Chromatin is commonly described as nucleosome core particles connected by linker DNA, creating the peculiar "beads on string" structure, in which the histone protein H1 provides stability by binding to the core at the DNA entry and exit sites[57]. According to the transcriptional status, chromatin can assume two different state: heterochromatin, densely packed and not accessible for transcription factors; or euchromatin, relaxed and actively transcribed[58].

Chromatin can be regulated by several processes such as remodeling of its nuclear architecture, DNA modifications, and histones modifications. All these mechanisms modulate gene expression, thereby regulating cell differentiation, cell division and development[59]. In the past decades, disruption of the chromatin regulatory processes emerged as pivotal for the development of several diseases, including cancer[60]. Enzymes that modify chromatin structure are emerging as paramount regulators of the tumorigenic process, and the inhibition of these enzymes may have interesting therapeutic implications in cancer[61].

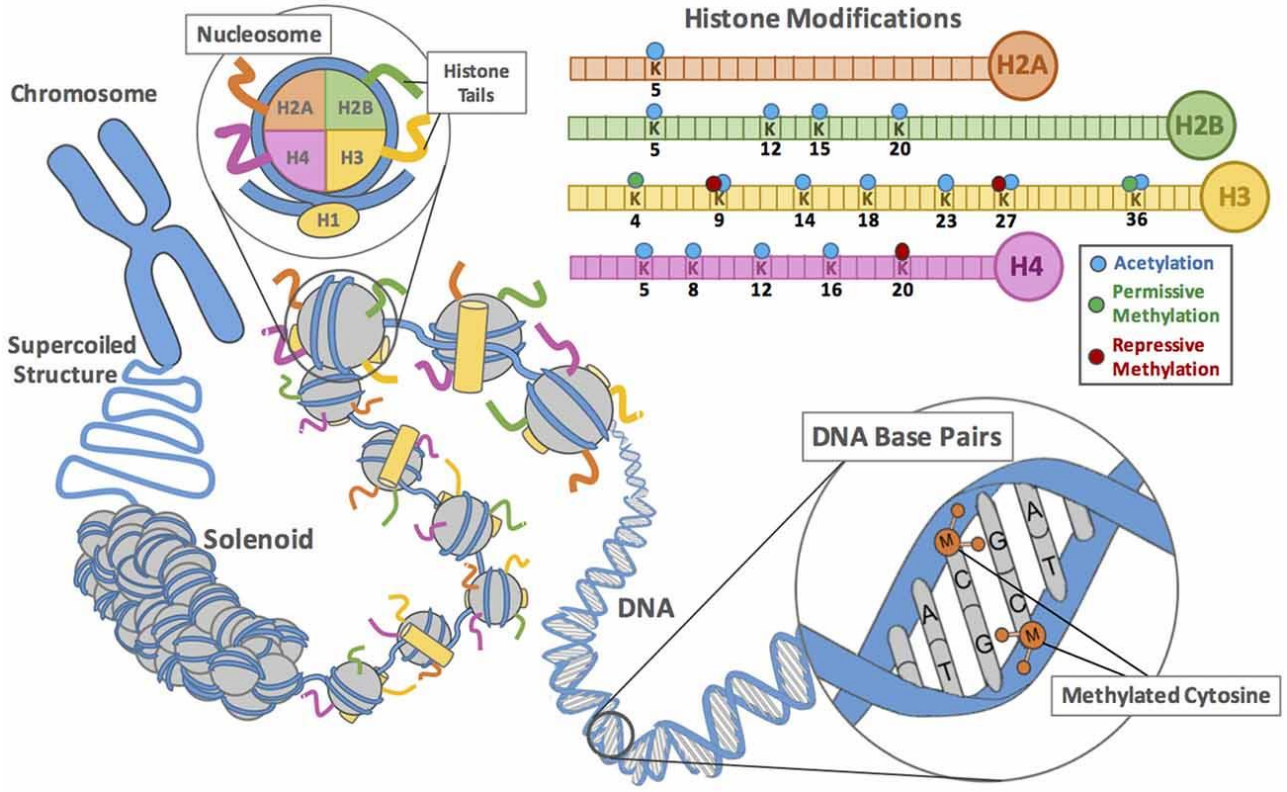


Figure 4.1 DNA structures and modifications

Representation of the DNA organization in nucleosomes in which they are connected by linker DNA. The histones exposed out of the DNA structure tails, these residues could be modified. In the upper part of the representation there is a detailed overview of the common modification on the H2, H3 and H4[56].

4.2. SWI/SNF CHROMATIN MODIFIER

Chromatin structure is basically regulated by three classes of protein complexes: DNA-modifying proteins, histone modifying proteins, and ATP-dependent chromatin remodelling complexes (CRCs). Taken together, these complexes cooperate to control chromatin accessibility and then gene expression[62].

The ATP-dependent remodelers can be further divided into several subclasses based on biochemical activity and subunit composition: SWI/SNF, ISWI, INO80, SWRI, NURD/Mi2/CHD complexes. They are multi-subunit complexes using the energy derived from ATP hydrolysis to eject, slide or alter the composition of nucleosome octamers, allowing the access of transcription factors and facilitating or repressing gene expression[59]. Nucleosome sliding mechanism of SWI/SNF includes the following steps: binding of the complex to a fixed position on DNA, disruption of histone-DNA contacts, then the ATPase

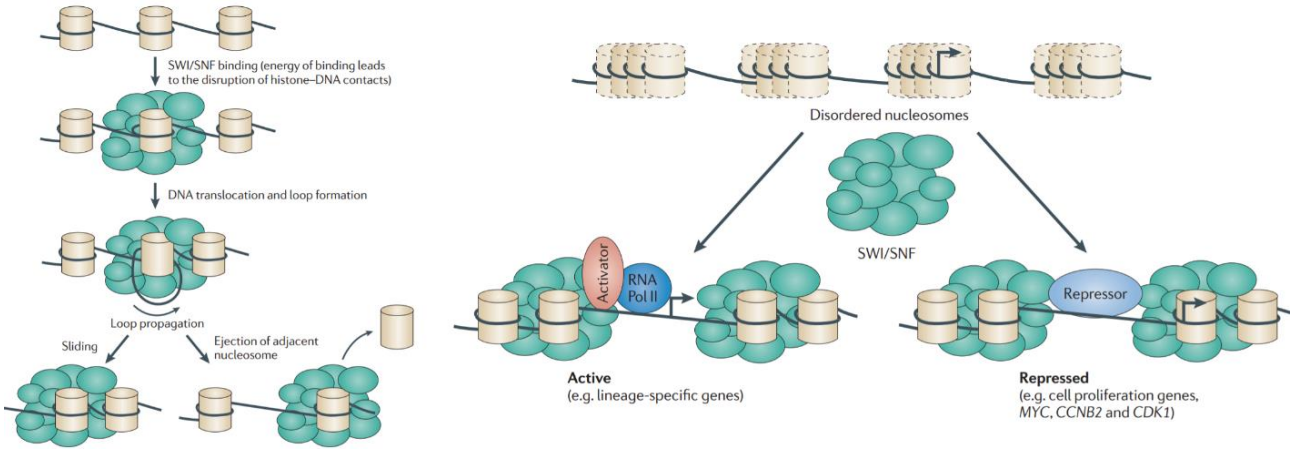


Figure 4.2 SWI/SNF remodeling mechanism

The steps of remodeling include SWI/SNF binding, disruption of histon-DNA contacts, creation of loop of DNA that propagates around the nucleosome in a wave-like manner and the repositioning of DNA[62].

subunit induced the translocation of the DNA, forming a DNA loop. Finally, the DNA loop is able to propagate around the nucleosome, creating sites more accessible for DNA-binding proteins (Figure 4.2)[62,63]. Even though nucleosome remodeling is considered the paramount function of SWI/SNF complex, other activities may be retrieved due to the interaction with others chromatin proteins[64].

Although, S. cerevisiae SWI/SNF complexes were identified because of their roles in the activation of transcription, there are evidence indicating that mammalian SWI/SNF complexes collectively contribute to both genes' repression and activation. SWI/SNF complexes are capable of recruiting histone deacetylases (HDACs), which remove activating acetyl marks from histone tails, contributing to the repression mechanism[65].

SWI/SNF genes were identified through two genetic screening in Saccharomyces cerevisiae: the first allowed to identify genes for sucrose metabolism (sucrose non-fermenting (SNF) mutants)[66], while the second revealed genes for mating type switching (switch (SWI) mutants)[67]. further studies led to the discovery of a 9 to 12 proteins complex encoded by at least 29 SWI and SNF genes, with a molecular weight of 1.14 MDa[68]. SWI/SNF is considered the most extensively mutated class of CRC protein families[60], and more than 20% of human cancers bear a mutation in one of its subunits[69]. Throughout evolution, SWI/SNF complexes have gained or lost subunits, together with the increased complexity of the genome[70]. Biochemical studies allowed the identification of homologous subunits with a similar biochemical activity, such as the RSC complex in S. cerevisiae, BAP and PBAP Drosophila melanogaster[71]. In mammals, the BRG1-associated factor (BAF or SWI/SNF-A) and the polybromo BRG1-associated factor (PBAF or SWI/SNF-B) have been characterized[72]. The mammalian complexes BAF and PBAF are constituted by: one of the two mutually exclusive catalytic ATPase subunits BRM (also named SMARCA2) or BRG1 (also named SMARCA4); a group of highly conserved core subunits, SNF5 (also named SMARCB1), BAF155, and BAF170; and variant subunits, which possess a role in complex lineage-specific function[72].

Regarding these variant subunits, the AT-rich interactive domain-containing protein 1A (ARID1A; also known as BAF250A and SMARCF1) and ARID1B are mutually exclusive and present only in BAF complexes, while BAF180, BAF200 and BRD7 subunits are peculiar of PBAF complexes, and finally GLTSCR1/1L and BRD9 are specific subunits of the ncBAF complex (Figure 4.2.1)[73,74].

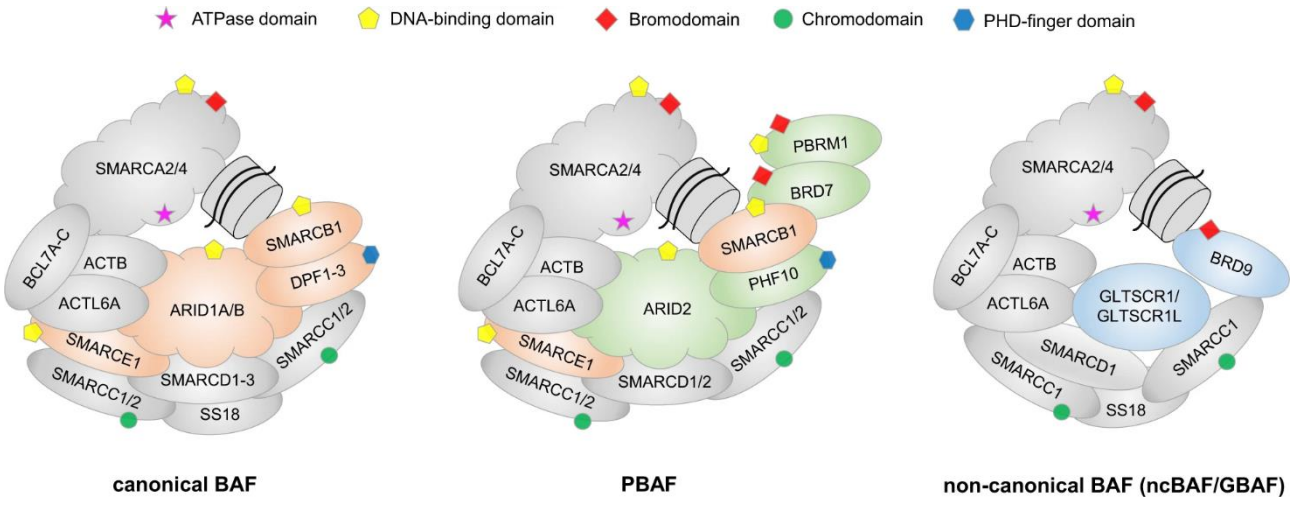


Figure 4.2.1 Mammalian SWI/SNF complexes

Three main mSWI/SNF complexes and the way in which they bind nucleosome, in grey are represented they subunits shared between all the complexes in green are highlighted the PBAF specific subunits. In orange, with the exception of ARID1 and DPF, are the subunits in common between BAF and PBAF, and at least in blue the subunits that's are specific for the ncBAF[74].

4.3. SWI/SNF IN CANCER

SWI/SNF complexes are involved in the control of lineage specification, in particular processes as neurogenesis, myogenesis, adipogenesis, osteogenesis, and hematopoiesis. During differentiation, these complexes cooperate with tissue-specific transcription factors in order to activate lineage-specific genes and repress genes for cell proliferation[75]. Combinatorial assembly of SWI/SNF members seems to regulate pluripotency and self-renewal of ES cells. Moreover, the inactivation of BRG1 or ARID1A leads to self-renewal defects in ES cells. Taken together, these evidences highlight the involvement of SWI/SNF in the oncogenic process, due to unbalance between self-renewal and differentiation[76].

Although SWI/SNF complexes are involved in the transcriptional regulation of a wide range of genes, insights of a pivotal tumor suppressive activity begun to emerge in the last years. Several studies linked SWI/SNF with different oncosuppressor or oncogenes. For instance, SWI/SNF complex cooperate with RB to repress E2F target genes and induce cell cycle

arrest[77]. Nuclear hormone receptors (NHRs) facilitate hormone-induced gene expression and have a role in control of cell proliferation and differentiation. It has been observed that they are directly involved in oncogenic transformation and SWI/SNF seem to interact with NHRs during transformation. The inactivation of SNF5 leads to the activation of Hedgehog signalling pathway, a mechanism which occurs in several cancer types[78]. Moreover, Knockdown of SNF5 increases RHOA activity, whose overexpression stimulate migration by inducing stress fiber formation in cancer[79].

Insights from SWI/SNF-mutant cancers highlight the activation of pathways involved in cellular motility and could contribute to invasion and metastases.

The presence of genome instability in many cancers makes it difficult to determine whether epigenetic changes are cancer formation drivers or secondary events. Several studies report that SWI/SNF inactivation leads to increased sensitivity to DNA damage, and suggest how the complexes have roles in DNA damage response[80]. It has been observed that TopoIIα's chromatin binding is dependent on the ATPase activity of Brg1, compromised in oncogenic Brg1 mutants. This support the role of BAF subunits as tumor suppressors[81].

BAF complexes are one of the pivotal tumor suppressors in humans, with an incidence of mutation of more than 20% of sequenced human cancers. Exome sequencing studies point out how specific cancers carry mutations in different BAF subunits (Fig. 4.3)[82]. In some cases, ubiquitously expressed subunits do not show a detectable genetic contribution to cancer, however are physically located near subunits with a significant role[69]. The genomic observation shown that three main mechanism highlighting tumor suppressive by BAF complexe. The first is the dosage sensitive. The second, cannot be detected through in-vitro chromatin remodeling assays. The third, while these mechanism are general, they are visible in a highly tumor specific and tissue specific manner[83].

Biallelic inactivation of BAF47 causes the 100% of rare childhood cancer MRT, and these tumors show a classic loss of heterozygosity[84]. SMARCA4 (or BAF190, BRG1) is frequently mutated in cancer, and most of the missense mutations found in the highly conserved ATPase domain. Similar mutations have been identified in the homologous BRM protein (encoded by SMARCA2 gene)[85]. Regarding BAF57 and BAF45d, the first is mutated in 100% of spinal meningiomas[86,87,88], while the second is amplified in breast cancer[69].

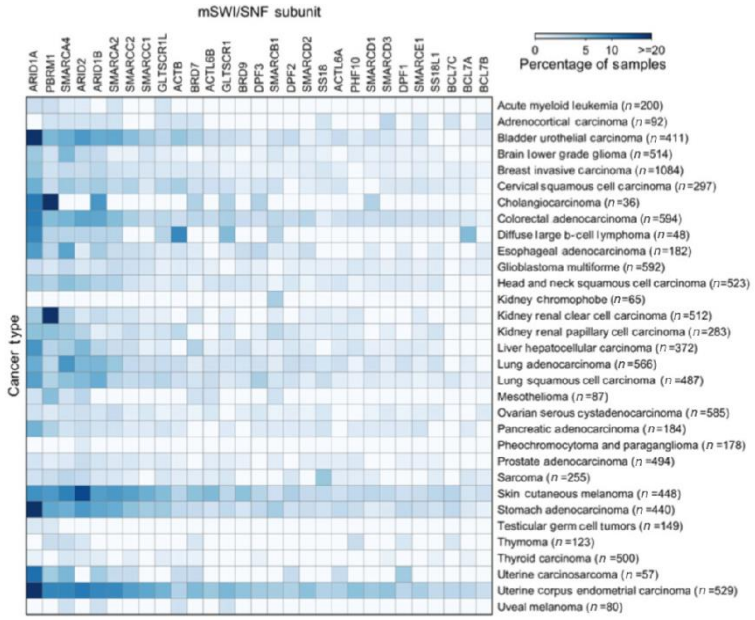


Figure 4.3 SWI/SNF in cancer

Heatmap of the frequency of mSWI/SNF complex subunits mutations across cancer types, adapted from[82].

5. ARID1A and ARID1B

5.1. GENE OVERVIEW

The AT-rich interaction domain 1 A and B are two proteins which are a constitutive, mutually exclusive core part of cBAF complex. They are part of a seven-member family protein who share the ARID consensus sequence, that spans about 100 amino acid residues and within their family multiple copies of the sequence motif, LXXLL. Especially at the C-terminal domain who is the site for protein-protein interactions (Fig. 5.1)[89], which has been shown to facilitate the interaction of various proteins with nuclear hormone receptors[90,89].

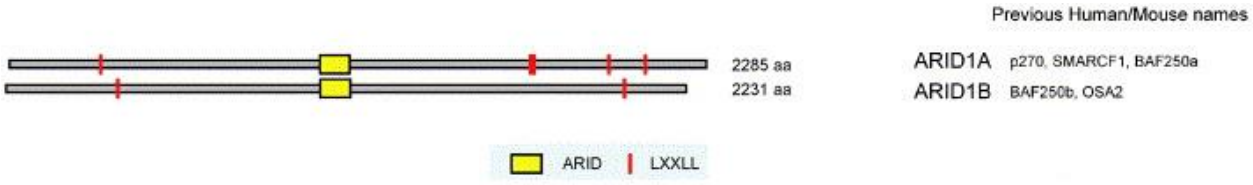


Figure 5.1 ARID1 proteins

Structure of Arid1A and Arid1B comparing their domains with all they domains, adapted from[89].

Between them there is a 60% of sequence identity, perhaps they share the same function as being the non-specific DNA binding module of the cBAF complex, who drive complex assembly and identity. These two subunits are fundamental for cBAF assembly, without them, cBAF complex would not be able to assemble[70]. However, between these two mutually exclusive proteins there is also a different distribution inside the mouse embryonic tissue and inside the cell cycle, ARID1A repress the cell cycle and accumulates in G0/G1 phase and is downregulated in S and G2/M phases[90,91]. While Arid1B expression remains constant during the cell cycle and promote the cell cycle progression[90,91]. This cyclic expression was justified by the finding that in absence of Arid1A, Arid1B facilitates expression of proliferative genes involved in PI3K/AKT/mTOR and ERBB receptor tyrosine kinase signalling in colorectal carcinoma[92]. Arid1B induce

proliferation via MYC and Cyclin E activation compared to Arid1A[91]. Recent work highlighted how Arib1B are able to induce bypass of senescence and promoting tumor formation[93]. Arid1A loss results in dramatic changes in chromatin accessibility, while ARID1B knockdown has no effect[92].

Arid1B mutations are less prevalent in cancer compared with Arid1A, but when Arid1B was found mutated it is characterized by copy number amplification, while Arid1A is more susceptible to missense or non-sense mutation[94,95]. In literature, Arid1A role is more studied as a predictive and prognostic biomarker, on the contrary Arid1B is more associated with neurodevelopmental diseases[96,97]. Those observation underlings the dominant effect of Arid1A during tumorigenesis. There was evidences that never occur biallelic co-mutations of Arid1A and Arid1B in both genes, suggesting how some Arid1 function is essential for cell survival both in cancer and normal cells[98]. Indeed, there are many studies suggesting that in Arid1A mutated cancer there is a possible compensation due to Arid1B and vice versa, underling how this complex is essential for cell viability and they concomitant loss led to synthetic lethality[98,99,100]. Unfortunately, the molecular basis of the synthetic lethality is not clear and seems to be controversial. Recently, it was report that the dual loss of Arid1A/Arid1B was not only a tumor prone events but also affect the assembly of the other SWI/SNF complexes, the remain cBAF sub-complexes leftover could interfere with assembly/functions of other mSWI/SNF complexes[95].

5.2. ARID1A in cancer

So far, Arid1A is the most affected BAF subunit in cancer (Fig. 5.2)[74], indeed it has been found mutated in a wide range of primary human tumors, including the 46% of endometrial carcinoma and the 57% of ovarian clear cell carcinoma (one of the most lethal subtypes of ovarian cancer)[62].

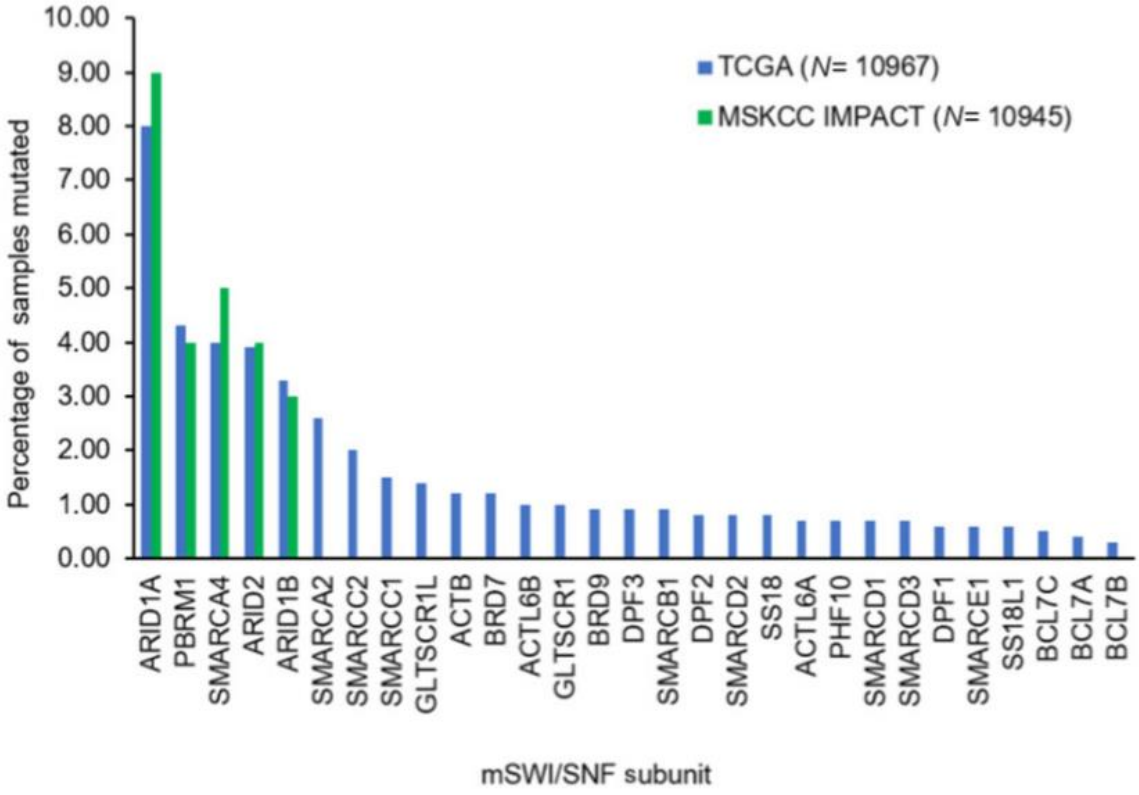


Figure 5.2 Impact of mSWI/SNF subuntis mutations on carcer

Percentage of mutations of single subunits among all cancer types, adapted from[82]

Moreover, this gene is mutated in colon-rectal cancers[101]; gastric cancers[102,103]; pancreatic cancers[104]; transitional cell carcinoma of the bladder; cholangiocarcinomas[105,106]; childhood neuroblastoma[107]; and hepatocellular carcinoma[108,109]. The genetic alterations frequently detected in Arid1A gene are basically missense, nonsense and frameshift mutations, these mutations lead to protein

downregulation[110]. Several studies demonstrate that neither global nor liver-specific Arid1A deficient mice spontaneously develop tumors by one year, indicating that Arid1A loss is not able per se to induce cancer development[111]. Indeed, in ovarian cancer, PIK3CA co-activation is required together with Arid1A deficiency to promote tumorigenesis[112,113].

In vitro experiments support the hypothesis that Arid1A exerts a tumor suppressive role. Indeed, Arid1A knockdown promotes cell cycle progression, cell proliferation, migration, invasion, and metastasis, while Arid1A overexpression inhibits cell proliferation and tumor growth[114]. *In vivo* studies confirm how lack of Arid1A could promote tumor formation[106,113].

In Hepatocellular Carcinoma, more than 20% of the tumors have alterations in SWI/SNF components, of those Aird1A remains the pivotal mutated gene (from 10 to 15% of HCC carry Arid1A mutations)[17]. Arid1A has been suggested to have a role as a tumor suppressor gene in many tumors. However, *in vivo* studies point out that Arid1A works as a tumor suppressor in HCC have not been provided yet, and its effective role in the development of HCC remains unclear[115].

Fang and colleagues demonstrate that liver specific Arid1A-deficient mice undergo chronic liver damage and inflammation, steatohepatitis, hepatocyte dysplasia and the animals developed HCC. They show how Arid1A loss of function mutations could be a triggering event required for HCC initiation[116]. Other evidences regarding the onco-suppressive role of Arid1A in HCC come from the study of *Fe and colleagues*, in which it is shown how Arid1A knockdown promotes HCC cell proliferation, while overexpression of Arid1A inhibits proliferation and impairs clonal formation of HCC cells. Moreover, they elucidate a mechanism according to which decreased Arid1A expression is associated with development of HCC metastases, especially in the lungs[117].

In a recent work published by *Zhu and colleagues*, the authors came to a completely different hypothesis regarding Arid1A role in tumorigenesis. They propose an oncogenic tumor suppressor role of Arid1A depending on the context. In particular in mice, liver-specific Arid1A loss prevent HCC formation while a gain of Arid1A function increases CYP450 mediated oxidative stress, promoting HCC initiation (underlying oncogenic role). On the contrary, loss of Arid1A in established tumors increases transcription of genes associated with migration, invasion, and metastases, promoting tumor aggressiveness[61].

6. AIMS OF THE PROJECT

HCC is one of the most common causes of death for cancer worldwide. Nearly 1 million of newly cases are diagnosed every year. However due to its silent nature the vast majority of HCCs are diagnosed at advanced stages when only palliative cures can be applied. HCCs develop in a context of chronic cellular damage, typically triggered by proinflammatory cues, which induces a chronic regenerative response increasing the chance to acquire oncogenic mutations.

During the last ten years several genetic and expression profiling studies described the mutational spectrum of HCC. Oncogenic mutations have been found in members of Wnt, Akt, TGFβ signaling pathways, in known tumor suppressors (TP53, AXIN1, RB1)[118] and several chromatin-modifying complexes. Inactivating mutations and homozygous deletions have been found in IDH1, BAP1 and in members of SWI/SNF complexes ARID1A, ARID1B and ARID2[118,119].

Through their mutually exclusive, catalytic subunits (BRG1 and BRM), SWI/SNF complexes are able to locally modify chromatin structure. In particular they catalyze sliding, eviction and insertion of the histone octamers, at promoters and enhancers, modulating DNA accessibility and promoting transcriptional activation or repression. By associating with cell-type specific transcription factors (TFs) they contribute to the transcriptional activation of the specific differentiation programs and to the transcriptional repression of pro-proliferative genes. Mutations affecting the activity of different SWI/SNF subunits have been reported in various cancers[82]. Arid1A subunit is mutated in 14% of HCC[118,119]. Very recently, some important works have been published suggesting a role of Arid1A in promoting liver regeneration[77] and HCC invasion and spreading[61].

Surprisingly, Arid1A mutations prevent Myc- and RAS-dependent HCC formation while

promoting metastasis formation when lost in established tumors[61]. These data suggest that Arid1A mutations could be a late event during HCC formation. The analysis of normal and cirrhotic liver samples, suggest that Arid1A mutation could be instead an early mutational event, in a cirrhotic background, which could represent an adaptive response required to restrain fibrosis and transformation[120]. A tumorigenic role of Arid1A mutations have been also demonstrated in a context of active-YAP1. Arid1A mutations unleash nuclear YAP1, potentiating its activity and promoting tumor formation[121].

From all these works a pivotal role in maintaining tissue homeostasis clearly emerges for Arid1A-containing SWI/SNF complexes. However, the consequences of Arid1A mutations in promoting tumor formation and the epigenetic mechanisms through which Arid1A preserve tissue homeostasis remains largely unclear.

In this context, the aim of this project is to shed light on the role of Arid1A in liver homeostasis and tumorigenesis focusing on the changes induced by its loss, but also to evaluate the role of all cBAF complexes in the maintenance of liver homeostasis by targeting both Arid1A and Arid1B subunits.

7. MATERIAL AND METHOD

7.1. CELL LINES

Arid1A proficient and deficient (Q456*/Q456*) HCT116 cell lines (HD 104-049, parental cell line CCL-247) were purchased from Horizon Discovery Ltd. The cell lines were grown in RPMI supplemented with Penicillin/Streptomycin, L-glutamine at 95% of O2 and 5% of CO2. The cells were split when 70/80% of confluency was reached.

7.2. MOUSE MODELS

Hepatocytes specific conditional knockout mice were generated by crossing ARID1A $^{fl/fl}$ (these mice have loxP sites flanking exon 8)[76] with Albumin-CreERT2 mice (from Pierre Chambon)[122]. Genotype of mice was checked by PCR from DNA extracted from hear punch. The recombination due to Cre activation induced by intraperitoneal injection of tamoxifen (Sigma-Aldrich) at 75 mg/kg.

Albumin-CreERT2/ ARID1A $^{fl/fl}$ were crossed with Albumin-CreERT2 B-catenin EXON3 in order to obtain AlbuminCREERT2 Arid1A$^{fl/fl}$ B-catenin EXON3 mouse line; this model was used taken advantages from the leakiness, stochastic CRE activation, of the Albumin-CREERT2.

Prom1-CreERT2 mice were crossed with ARID1A fl/fl to generate Prom1-CreERT2/ ARID1A $^{fl/fl}$ in order to delete ARID1A in different compartment, but we focus on stomach, intestine and colon of these mice.

Albumin-CreERT2/ ARID1A $^{fl/fl}$ were crossed with Albumin-CreERT2/Arid1B$^{fl/fl}$ to target all the cBAF complex in hepatocytes.

Mice were treated with a non-genotoxic agonist 1,4-Bis [2-(3,5-Dichloropyridyloxy)] benzene (TCPOBOP, TC) that was solubilized in DMSO and diluted in corn oil at the concentration of 0,75ug/ul. After 7 days from the last tamoxifen injection, TCPOBOP were administrated with intraperitoneal injection, at the concentration of 3mg/kg, and after 60 days a second injection were performed. Mice were sacrificed at 2, 30, 150 days for all the analysis.

To mimic liver regeneration, animals were treated, after 7 days from tamoxifen injection, with a specific diet enriched in *3,5-diethoxycarbonyl-1,4-dihydrocollidine* (DDC). The control mice were feed with control diet (chow diet). After two weeks of treatment the animals were sacrificed.

7.3. HISTOLOGY AND IHC

For histological analysis, livers were harvested at the described time points and fixed in 10% buffered formaldehyde overnight at 4°C, the day after formaldehyde was changed with 70% ethanol. The harvested tissues were dehydrated and paraffin-embedded. Embedded tissues were cut 5µm in thickness with microtome (Leica). Hematoxylin and eosin Y, and picrosirius red staining were performed. For IHC, the sections were dewaxed, rehydrated and antigen retrieval was performed using sodium citrate buffer pH 6 or EDTA buffer pH 8. Before the incubation with the primary antibody, Mouse-on-mouse (Abcam) or 5% normal goat serum (Sigma-Aldrich) has been used as blocking solution. The primary antibodies were leave on the sections at 4°C overnight; secondary antibody (Dako EnVision+ System- HRP Labelled Polymer anti-rabbit, mouse on mouse kit Abcam) was used for 1h at RT, counterstained with Harris hematoxylin, dehydrated and mounted with Eukitt (Bio-Optica). Images were acquired using Nikon Eclipse i90 widefield microscope.

7.4. RNA EXTRACTION AND REAL TIME-QPCR

Liver specimens were mechanically disrupted in 1mL of Trizol with metallic beads using tissue lyser (Qiagen). After the homogenization, total RNA was purified using RNeasy purification kit (Qiagen). 500/1000 ng of RNA of purified RNA was use for each Retro-transcription reaction. RT were performed using Promega ImProm-II kit Promega standard transcription protocol. Real-time qPCR was carried out using CFX384 PCR machine (BioRad) or Quantstudio 5 Applied (Thermofisher), reaction mix was prepared using GoTaq qPCR master mix (Promega) following manufacturer instructions

7.5. LIBRARY PREPARATION AND RNA SEQUENCING

Starting from total RNA extracted with Trizol, 1µg of total RNA will be used for RNA-seq library preparation using TruSeq Stranded Total RNA Library Prep Kit (Illumina) following the manufacturer instructions. Amplified DNA was sequenced using Illumina NovaSeq2000.

7.6. HEPATOCYTES PURIFICATION AND CHIP SEQUENCING

Primary murine hepatocytes were purified by perfusing the liver with collagenase (7mg) and trypsin inhibitor(2mg). After the perfusion, the rupture of the capsule that allowed hepatocytes release were performed with a second digestion buffer containing collagenase (5mg) and trypsin inhibitor(2mg). After the two digestion, the solution was passed through 70-µm cell strainer. The cells that pass through the filter were centrifuge

at 50g for 5min and washed 3 times with ice-cold EBSS. Then, the isolated hepatocytes were fixed in 1% FA for 10 min, the formaldehyde were quenched with Glycine for 5 min. Two washed were performed with PBS to eliminate all the formaldehyde residues and lysed in SDS Buffer (50 mM Tris (pH 8.1), 0.5% SDS, 100 mM NaCl, 5 mM EDTA, and protease inhibitors). Chromatin was sonicated in IP Buffer (100 mM Tris (pH 8.6), 0.3% SDS, 1.7% Triton X-100, and 5 mM EDTA) to an average length of 500–1000 bp. For histone modification, 250µg of sonicated chromatin was incubated with 5µg of H3K4me1, H3K27Ac, H3K4me3, while for TFs 1mg of sonicated chromatin was incubated with 10µg Foxa2, HNF4α antibodies at 4°C overnight, and the immunocomplexes recovered using protein A/G–conjugated Sepharose beads. Purified chromatin was decross-linked overnight in 0.1 M NaHCO3 and 1% SDS and purified at 65°C, checked for qPCR and then sequenced.

7.7. WHOLE EXOME SEQUENCING

Tumors were dissected from freshly dissected organs, pulverized in liquid nitrogen, genomic DNA was isolated and used to generate library using SureSelect XT - low input (Agilent) and SureSelect XT Mouse All Exon library for post-capture processing (Agilent). Libraries have been sequenced using Illumina NovaSeq6000.

7.8. WES DATA ANALYSIS

Whole exome sequencing data were profiled for somatic point mutations and copy number status. Firstly, the raw reads were trimmed with Trimmomatic v0.32[123]. Secondly, bwa v0.7.12 was used for reads alignment[124]. Thirdly, SAM files were merged with Picard v2.17.4 (Broad Institute) and converted to sorted BAM files with Samtools [125]. Fourtly,

duplicated reads were removed with the picard function MarkDuplicates, followed by realignment of indels and base quality score recalibration with GATK tools v3.1[126]. MuTect v1.1.7 [127] was employed to detect single nucleotide variants (SNVs), which were annotated with SnpEff 5.0[128]. DNA copy number analysis was performed with the R package CopywriteR[129], version 2.10.0. Furthermore, the segmentation of the copy number profiles was obtained with CGHcall 2.54.0[130].

7.9. RNA-SEQ DATA ANALYSIS

To investigate the transcriptional profile of ARID1A KO, we performed RNA-seq analysis. First, the pair-end raw reads were trimmed using fastp v0.20.0[131] with parameters "-t 1 -A -Q -L". Afterwards, the trimmed reads were aligned with STAR v2.7.3[132] against the mouse genome (mm10 - GRCm38.p6). Duplicated reads were removed with samblaster[133]. At this point, we quantified the gene expression using featureCounts v1.6.4[134] with parameters "-s 0 -t exon -p -g gene_name". The annotation file (gtf) for the counting was downloaded from Gencode (Release M21 - GRCm38.p6). Once we obtained the gene expression table, differential expression analysis was performed with the R package DESeq2 v1.32[135]. Moreover, shrinkage of fold changes for lowly expressed genes was achieved with the apeglm package[136], and p-values were FDR corrected and weighted based on the base mean through the IHW package[137].

The thresholds for significance were set to 0.05 for adjusted p-values and 1.5 for the fold change. Finally, gene set enrichment analysis was carried out with the R package clusterProfiler[138].

7.10. CHIP-SEQ DATA ANALYSIS

For the ChIP-Rx-seq data analysis, we first trimmed the pair-end reads with fastp v0.20.0[132], setting as parameters "-t 1 -A -Q -L". Next, the trimmed reads were aligned with bowtie[139] against the mouse genome (mm10 - GRCm38.p6) and against the drosophila melanogaster genome (dm6 - Release 6 plus ISO1 MT). The parameters used for bowtie are: "--chunkmbs 1024 -m 1 --best -S --no-unal -q -I 10 -X 1000". From the two generated bam files we removed the ambiguous reads, i.e. those aligning on both the genomes. At this point, peak calling was performed with MACS2 v2.2.7.1. in narrow mode[140], setting the parameters to "--keep-dup all -m 3 30 --format BAMPE --pvalue 1e-05". The called peaks were further filtered with a p-value threshold of 1e-10 and annotated using the package ChIPseeker v1.28.3[141], with a range to define a promoter peak of ±-2.5 kb.

To generate the bigwig files, bamCoverage from deepTools v3.5.1[142] was employed with parameters "--binSize 50 --extendReads". The scaling factor was calculated as described in[143]. Finally, regions' intensities values for the boxplots were retrieved from the bigwigs through the deepTools function multiBigwigSummary in BED-file mode. To split the enhancers in quantiles, we computed the log2 ratio of H3K27Ac levels between Arid1A KO and Arid1A WT samples.

7.11. TCGA ANALYSIS

For the survival analysis we employed cBioPortal (https://www.cbioportal.org/). Specifically, we selected the Liver Hepatocellular Carcinoma dataset (TCGA, Firehose Legacy) and divided the samples into those with mutations in both Arid1A and Ctnnb1 and those without. Moreover, to build the upset plot we computed the co-occurrence frequencies of Arid1A mutations with mutations in other genes involved in immune escape. This was also done in the LIHC TCGA firehose legacy.

7.12. ANTIBODIES

Antibody	Company	Host	Application	Concentration
H3K27Ac	Abcam	Rabbit	ChIP	5ug/mL
H3K4Me1	Abcam	Rabbit	ChIP	5ug/mL
H3K4Me3	Abcam	Rabbit	ChIP	5ug/mL
H3K27Me3	CST	Rabbit	ChIP/IHC	5ug/mL
IBA1	Fujifilm	Rabbit	IHC	1:1000
CDKN1A	Abcam	Rabbit	IHC	1:1000
HNF4α	Abcam	Rabbit	IHC	1:1000
HNF4α	sc	Goat	ChIP	10ug/mL
FOXA2	sc	Rabbit	ChIP	10ug/mL
CK19	Abcam	Rabbit	IHC	1:1000
SOX9	Millipore	Rabbit	IF	1:2500
HNF4α	Abcam	Mouse	IF	1:2500
γH2AX	CST	Rabbit	IHC	1:1000
ARID1A	Abcam	Rabbit	IHC	1:1000
KI67	Abcam	Rabbit	IHC	1:1000
CTNNB1	Sigma	Rabbit	IHC	1:500

7.13. PRIMERS

MOUSE		
CCL5	Forward	ACCATATGGCTCGGACACCACT
	Reverse	ACCCACTTCTTCTCTGGGTTGG
GBP2	Forward	AAGACTCTGTGTGGTGGCAT
	Reverse	ATGCAGGGGAGAGACCCATT

GBP3	Forward	TGTGTGGAGGGAAAGGAGGA
	Reverse	CTCCATGGTCCACTCGGAAG
OAS1A	Forward	GAAGAGTCTCATCCGCCTGG
	Reverse	CAGTGAGCAACTCTAGGGCG
OASL1	Forward	ACTGGACCAAGCACTACACG
	Reverse	CCACATTGTTGGTGGGGTCT
RSAD2	Forward	TGGAGCGTCACAAAGAGGTG
	Reverse	CCGGTACAGTTCAGAAAGCG
CXCL10	Forward	CCAAGTGCTGCCGTCATTTT
	Reverse	AGCTTCCCTATGGCCCTCAT
IRF7	Forward	GGGACCTCTTGCTTCAGGTT
	Reverse	AGGGTTCCTCGTAAACACGG
ISG15	Forward	TAATTCCAGGGGACCTAGAGC
	Reverse	GAGTTAGTCACGGACACCAG
IFI44	Forward	TTCTGAGCTGGTGGCAAAGAT
	Reverse	CCACGTGTGTAAGTAAAGCCAA
IFIT1	Forward	AACCCAGAGAACAGCTACCAC
	Reverse	CTGAAGCAGATTCTCCATGACCT
IFIT3	Forward	AGATTTCTGAACTGCTCAGCCC
	Reverse	CAGAGATTCCCGGTTGACCTC
HUMAN		
CXCL10	Forward	AGCAGAGGAACCTCCAGTCT
	Reverse	ATGCAGGTACAGCGTACAGT
ISG15	Forward	CAGCGAACTCATCTTTGCCAG
	Reverse	GGACACCTGGAATTCGTTGC
IFI44	Forward	TGGGAGCTGGACCCTGTAAA
	Reverse	CCTCCCTTAGATTCCCTATTTGCT
IFIT3	Forward	AAGAACAAATCAGCCTGGTCAC
	Reverse	GACCTCACTCATGACTGCCC

8. RESULTS

8.1. ARID1A LOSS INDUCE H3K27Ac DECREASE ON A SUBSET OF ENHANCERS

In the last decades a considerable amount of data has been generated by cancer genomic studies showing that chromatin modifiers are the most frequently mutated genes. Among them, SWI/SNF activity is frequently affected by loss of function due to mutations involving several subunits. ARID1A, a cBAF subunit, is the most frequently mutated. While ARID1A has been considered a tumor suppressor gene, recent works suggests its oncogenic role in promoting hepatocellular carcinoma (HCC). We generated tamoxifen-inducible liver-specific ARID1A knockout mice, by crossing AlbCre-ERT2 mice with ARID1Afl/fl mice, to further analyze the effect of ARID1A loss in liver homeostasis. ARID1A expression was efficiently abrogate by tamoxifen injection. The deletion was restricted to hepatocytes without any effect on other cell types. Given the extensive literature indicating its recruitment at distal regulatory elements to promote chromatin accessibility, we confirmed its localization at distal elements, and then decided to analyze the role of ARID1A in preserving enhancers activity. Accordingly, in adult hepatocytes can be found ARID1A peaks preferentially at distal elements and only few peaks at promoters (Fig.8.1 A) and we decided to focus our attention on those distal elements. To better appreciate the effect on these regions, we divided them in four groups (quantiles). The quantiles are characterized by H3K27Ac intensity ratio of the WT hepatocytes. Although to a different extent, they are all active regulatory elements as shown by the high H3K27Ac levels. Surprisingly, ARID1A loss do not results in a global inactivation of these regulatory elements (Fig.8.1 B) but after its loss only a subset of enhancer loose H3K27Ac. On the first and the second quantiles can be observed a considerable H3K27Ac reduction, a modest decrease for the third, and a gain of acetylation can be observed in the last quantile (Fig.8.1 B, C). These data show that ARID1A loss has

relatively modest effects at distal regions, indeed only those regions with the lowest H3K27Ac levels (i.e. less active), belonging to the first quantile, are significantly affected by its abrogation. Similar conclusions can be also drawn for those regions were H3K27Ac accumulates, since they are among the mostly active (Fig.8.1 B, C). Seen the very few peaks on promoters no changes are observed on these regions (Fig. 8.1 D).

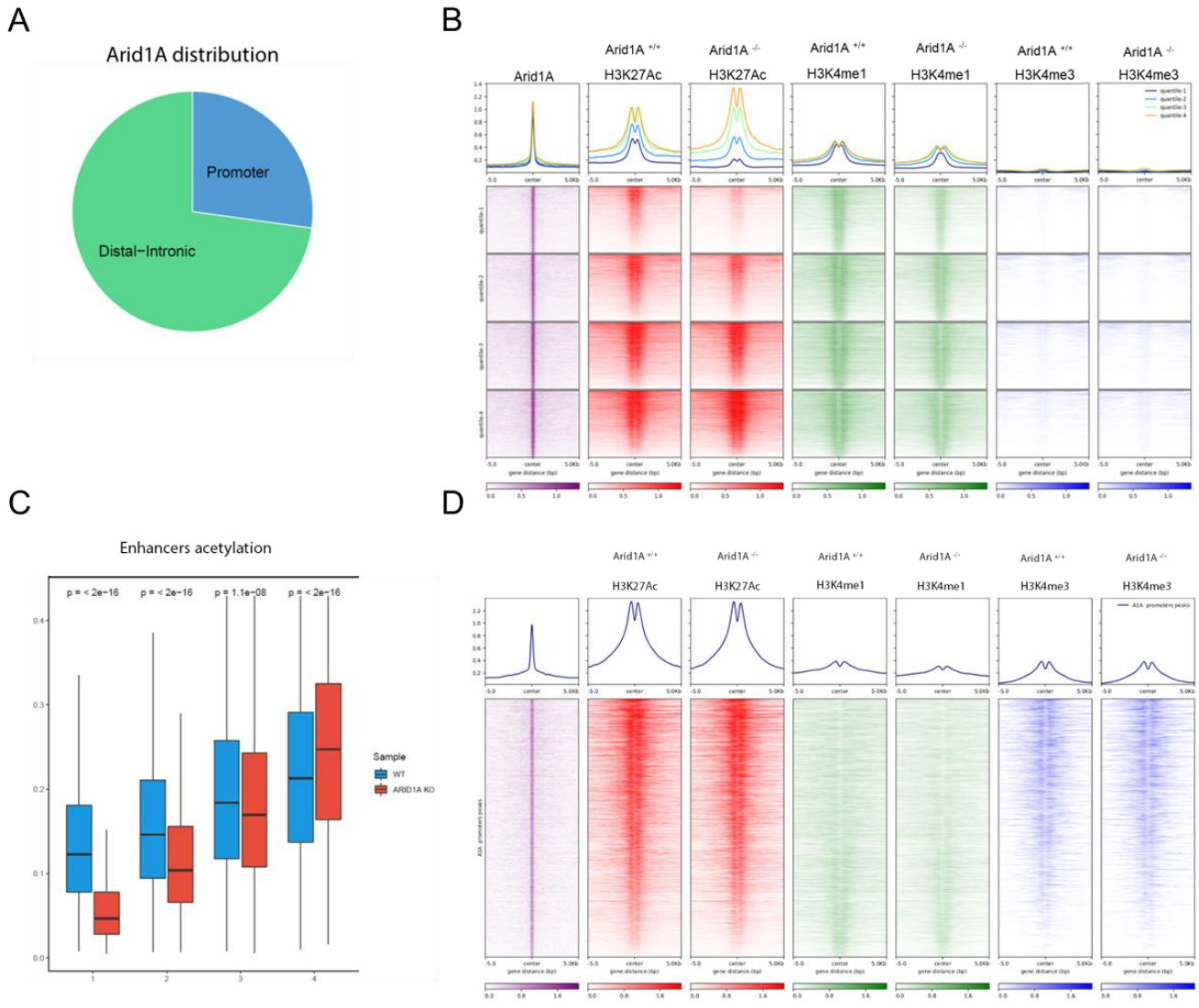


Figure 8.1 Arid1A loss has a strong effect on enhancers without effect on the promoters.

A) Representation of Arid1A occupancy on hepatocytes chromatin in homeostasis. **B)** Heatmaps display Chip-seq H3K4me1 and H3K27Ac divided on quantiles based on acetylation ratio between KO and WT sample, compared with the Arid1A peaks on enhancer on the left of the heatmap. **C)** Enhancers H3K27Ac quantification starting from the heatmaps. **E)** Heatmaps display ChIP-seq of H3K4me3 and H3K27Ac. down- and upregulated genes.

8.2. ARID1A AND ARID1B DOUBLE KNOCK OUT INCREASE INFLAMMATION, CELL DEATH AND CAUSE A TRANSCRIPTION FACTORS RELOCALIZATION

Since cBAF has two subcomplexes with unique ARID1 protein that characterized it, we decided to investigate whether ARID1B could compensate ARID1A loss. To understand if the two subcomplexes have compensatory effect in hepatocytes, we generated AlbCre-ERT2Arid1Afl/flArid1Bfl/fl mice. The analysis of ARID1A occupied enhancers reveal that

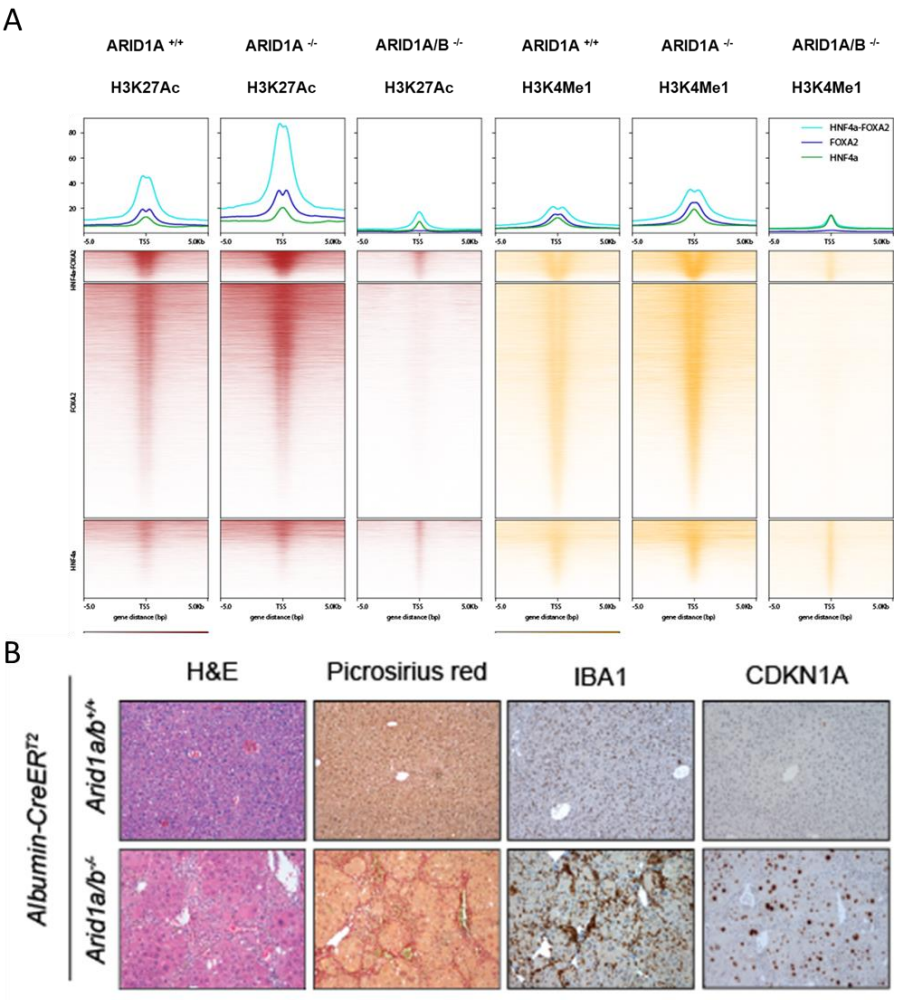


Figure 8.2.1 CBAF loss cause H3K27Ac and H3K4Me1 loss and trigger fibrosis

A) Heatmaps display Chip-seq in which the H3K4me1 and H3K27Ac were completely loss after cBAF loss on HNF4a and FOXA2 target enhancers. **B)** Representative images of IHC showing HE, and stainings for ARID1A, ARID1B, and α-SMA.

loss of both ARID1 proteins leads to complete loss of H3K27Ac and H3K4Me1 (Fig. 8.2.1 A), confirming the redundant activity of ARID1B-containing cBAF complex.

While concomitant ARID1A/B loss is efficiently achieved using these mice, between 15 and 20 days the animals must be sacrificed for ethical reasons. Histological analysis of liver samples revealed the presence of aggressive hepatitis characterized by extensive fibrosis (Fig. 8.2.1 B), immune infiltrate, and cell death. Notably, knockout cells are rapidly counter

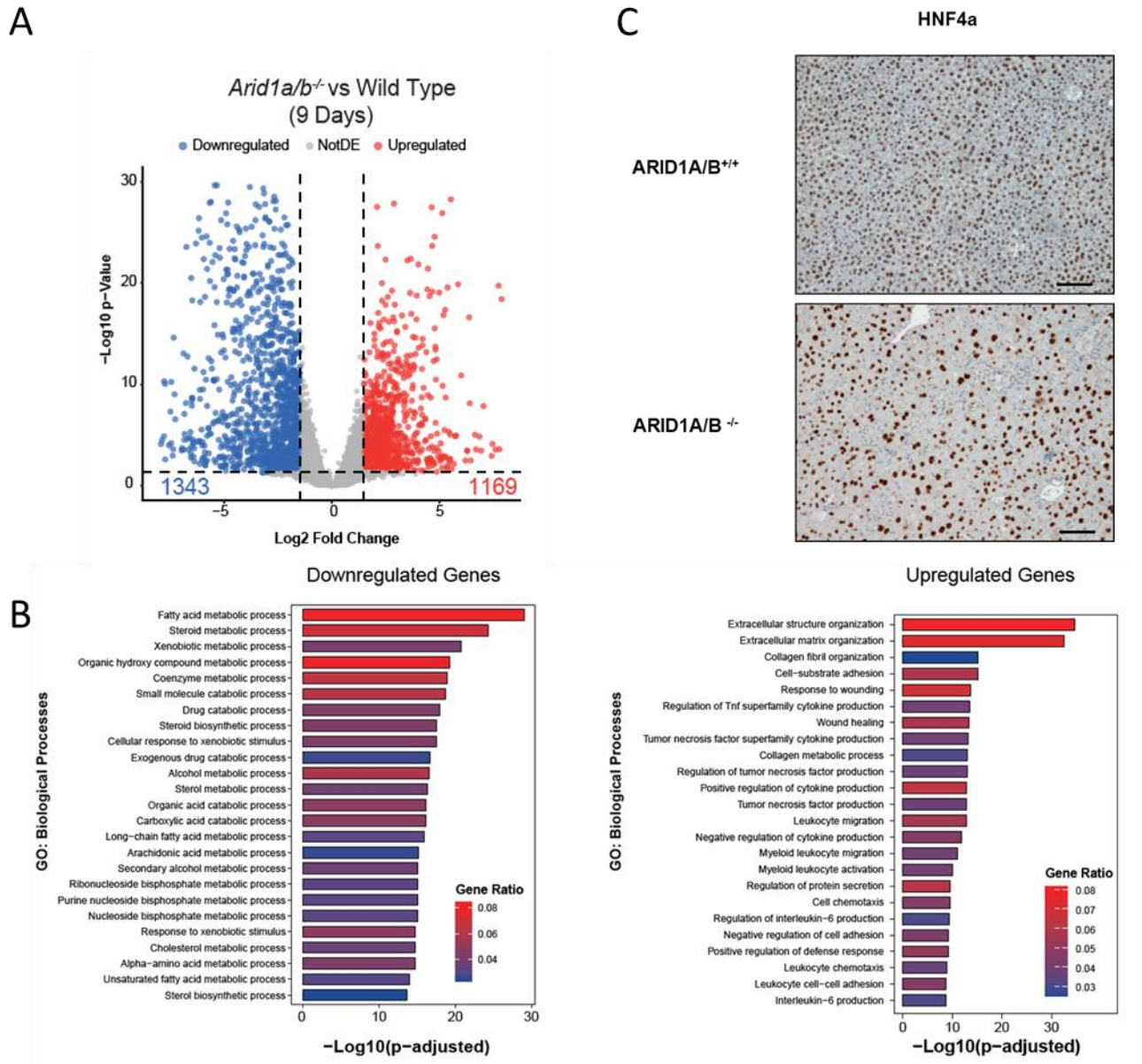


Figure 8.2.2 CBAF destabilization has an impact on transcriptional profile and liver tissue homeostasis

A) Volcano plot comparing Arid1A/Arid1B double KO and WT transcription. **B)** GO analyses of up- and downregulated genes pointed out the loss liver metabolic processes and the upregulation of wound healing, collagen deposition. **C)** representative images of IHC showing that HNF4a is not affected by cBAF loss.

selected and only sporadic ARID1A or ARID1B single knockout patches of cells can be identified after 20 days from Cre activation (Fig. 8.2.3).

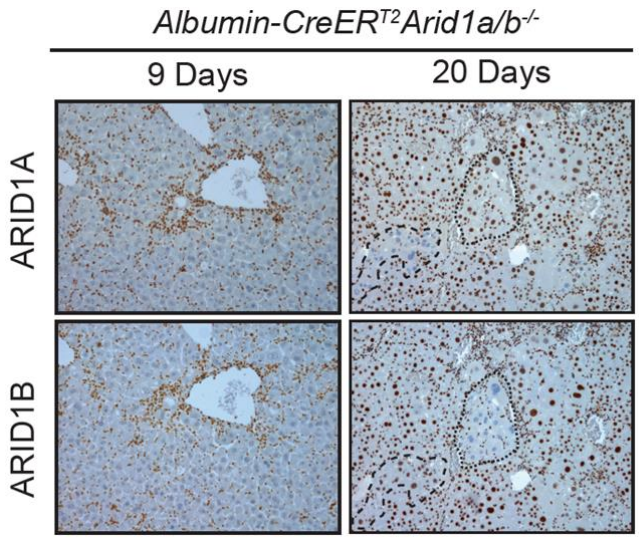


Figure 8.2.3 Deletion efficiency of both Arid1 proteins

Representative images of IHC showing Arid1A and Arid1B staining at 9 and 20 days after Tamoxifen injection.

Transcriptional analysis performed 9 days after tamoxifen administration shown that loss of both ARID1A and ARID1B dramatically alter hepatocytes transcriptome (Fig. 8.2.2 A) with downregulation of genes involved in liver-specific metabolic functions (Fig. 8.2.2 B), while there is a strong upregulation of genes involved in wound healing and matrix deposition (fibrosis) (Fig. 8.2.2 B). No differences have been observed in both HNF4α (Fig. 8.2.2 C) and FOXA2 levels suggesting that cBAF defective cells preserve cell identity. Moreover, the ChIP-seq data show that ARID1A loss leads to a rearrangement with the relocalization of HNF4α and FOXA2 comparing their peaks with the WT (Fig. 8.2.3). A switch of the two transcription factors is found both on enhancers and promoters, FOXA2 gains peaks on the target of HNF4a both on promoters and enhancers, while HNF4a was found on the enhancers target of FOXA2. These changes are maintained also after the loss of both ARID1 proteins with a completely loss of FOXA2 peaks. Taken together these data

demonstrate that ARID1B could compensate the activity on ARID1A distal elements preserving FOXA2 binding on enhancers and hepatocytes-specific transcriptional programs.

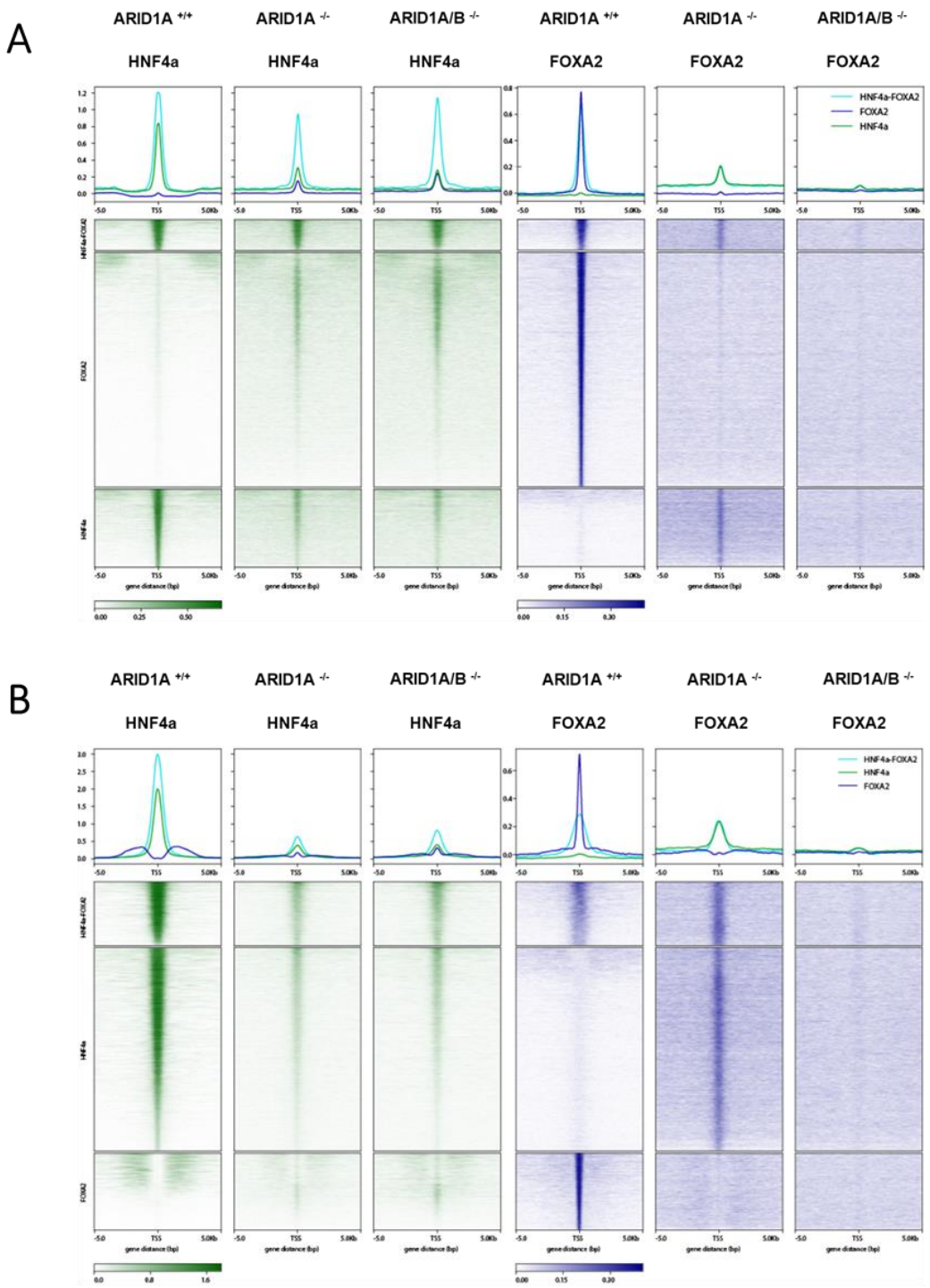


Figure 8.2.3 FOXA2 and HNF4α relocalization in both at the enhancers and promoters

A) Heatmaps showing distal peaks for HNF4α, FOXA2 and the shared regions, between WT, Arid1a, and Arid1A/B Double KO. **B)** Heatmaps showing promoters peaks for HNF4α, FOXA2 and the shared regions of both TFs comparing WT, Arid1a, and Arid1A/B Double KO.

8.3. ARID1A LOSS INDUCE AN ECTOPIC INTERFERON PATHWAY ACTIVATION

To shed light on the early events following ARID1A loss, we analyzed the transcriptional profile induced after nine and thirty days from tamoxifen administration by performing RNA-seq analysis. Despite Arid1A-containing cBAF complex is predominantly considered to favor transcriptional activation a significant number of up- and down-regulated genes can be identified at both 9 and 30 days after tamoxifen injection (Fig. 8.3.1 A). Recently an oncogenic activity of Arid1A has been described mechanistically linked to an upregulation of cytochrome genes and increase oxidative stress, but we did not observe significant differences in the expression of cytochrome genes. The very few overlaps between Arid1A-

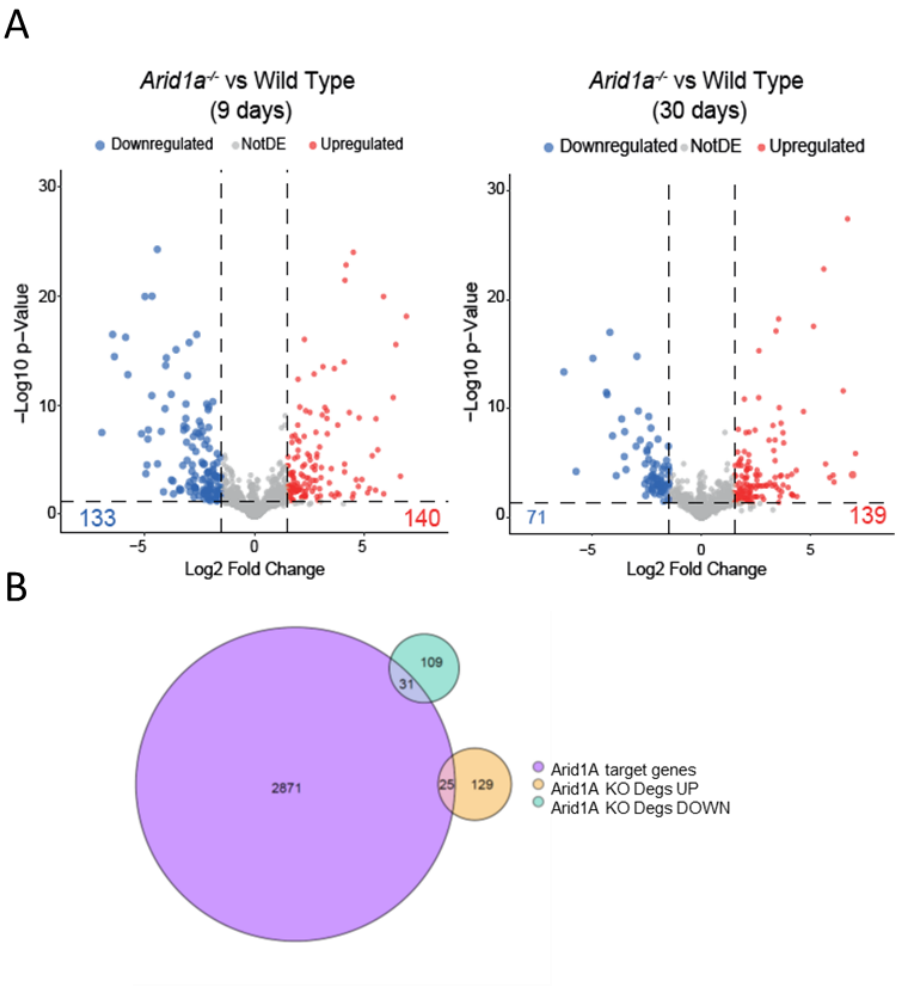


Figure 8.3.1 Transcription profile of Arid1A KO livers

A) Volcano plot showing differential expressed genes of Arid1A KO vs WT mice obtained by RNA-seq at 9 and 30 days. **B)** Overlap of the Arid1a peaks with the down- and upregulated genes.

occupied promoters and the differentially expressed genes (Fig. 8.3.1 B) further suggest that the transcriptional alterations observed may be largely indirect. To understand weather the changes observed on enhancer's acetylation could justify the differences observed in the transcription of the differential expressed genes. We decided to analyze the transcription of genes located at +/-20Kb from the differentially acetylated enhancers without any correlation between modifications and gene expression (Fig. 8.3.2).

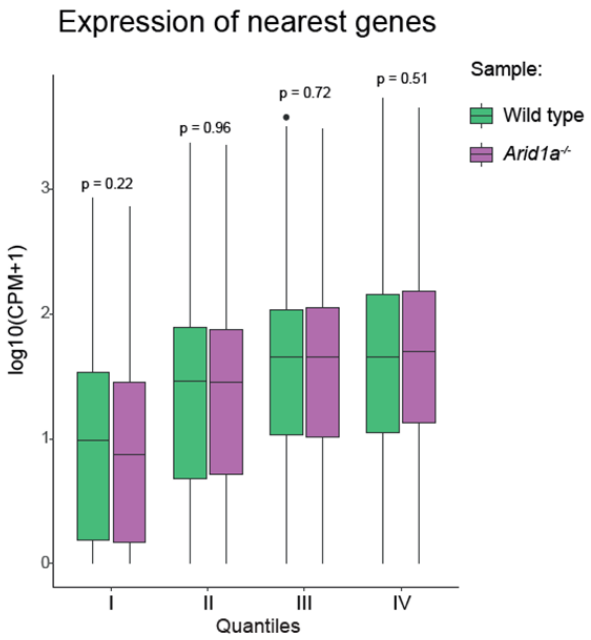


Figure 8.3.2 Transcriptional activity of the genes located near the modified enhancers

Boxplot showing the expression level of the genes located at +/- 20Kb from the enhancers that loose acetylation.

Gene Ontology (GO) terms analysis performed using up- and down-regulated genes reveal that no defined pathway was significantly enriched between downregulated genes. Surprisingly, we found upregulation of genes involved in viral response and interferon response genes at both time points (Fig. 8.3.3 A).

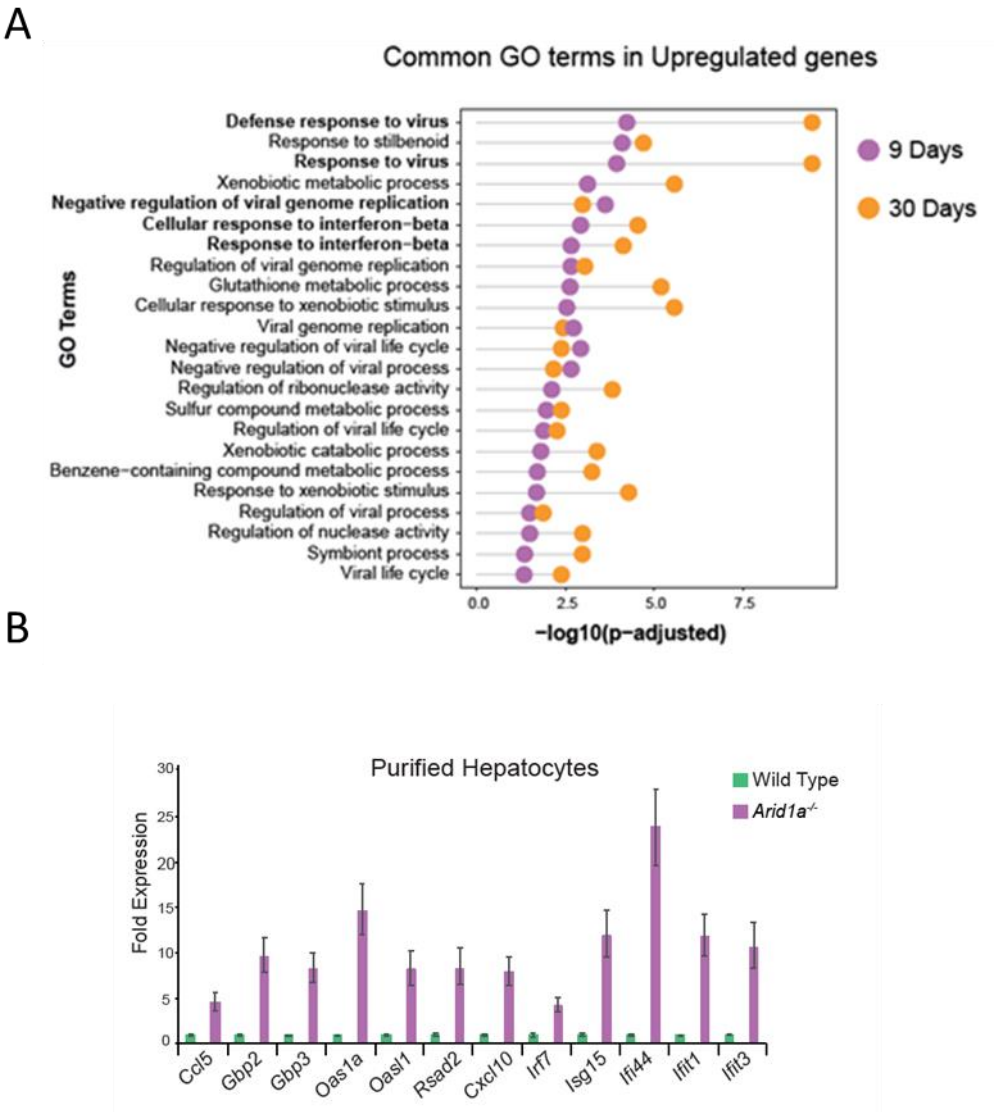


Figure 8.3.3 Upregulation of interferon pathway after Arid1A loss

A) Overlap of the GO terms of the upregulated genes at 2 and 30 days. **B)** Quantitative PCR analysis of interferon target genes in purified hepatocytes.

Among these genes, we found several interferon-stimulated genes (ISGs), pro-inflammatory chemokines such as Cxcl10 and Ccl5, Usp18 and master transcription factors of this pathway which is Irf7 (Fig. 8.3.3 B). To evaluate if transcriptional activation of interferon target genes was a tissue-specific effect we reanalyzed published datasets obtained from ARID1A proficient and deficient: small intestinal crypts[144], HCT116[145] and MCF7[146] cell lines. GO terms enrichment analysis of upregulated genes revealed that pathways involved in interferon response and response to virus were shared between all the reanalyzed datasets (Fig. 8.3.4 A).

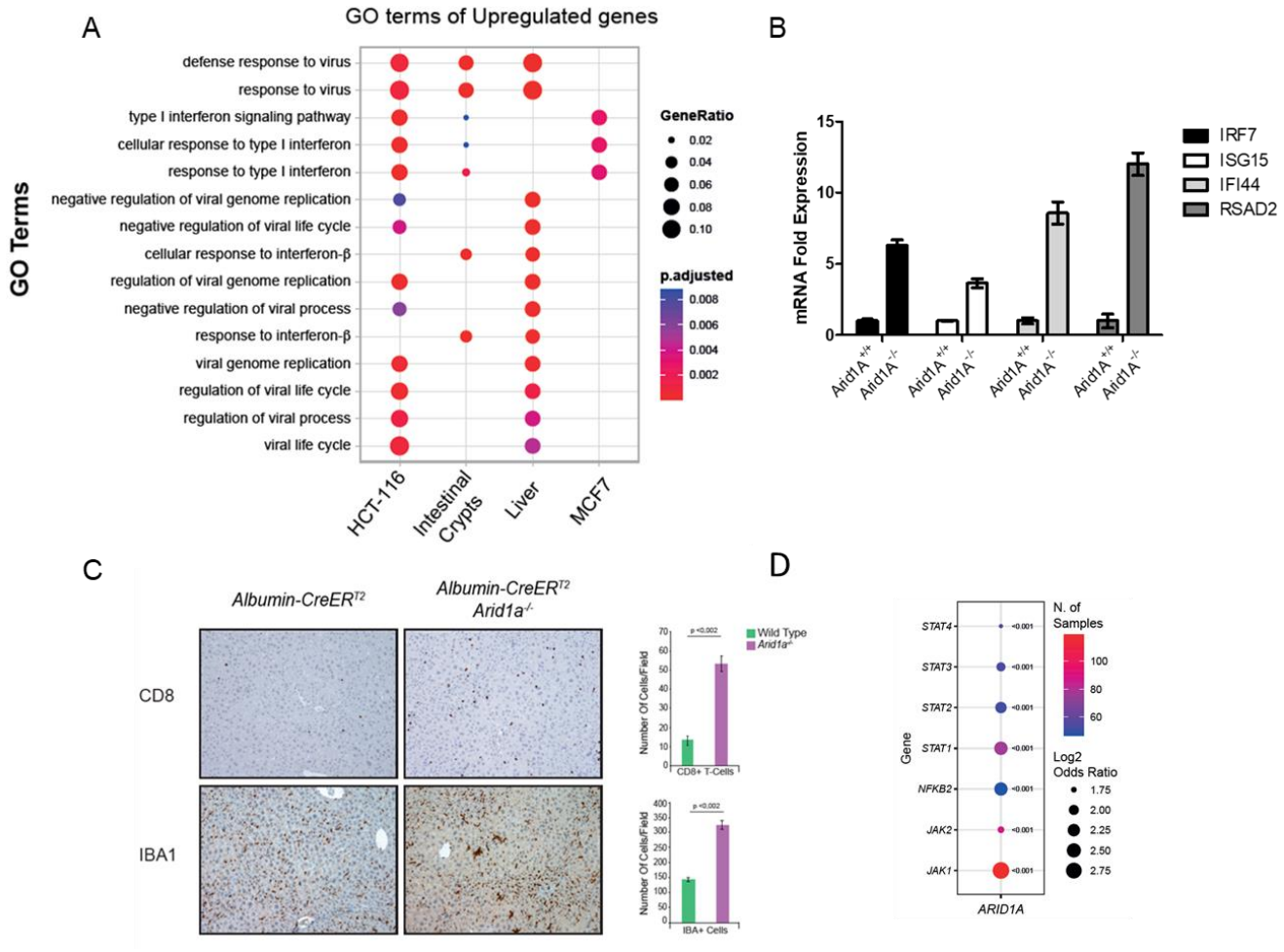


Figure 8.3.4

A) comparison of the GO terms of reanalyzed published data compared with our data (REF). **B)** Quantitative PCR analysis of interferon target genes in HCT116 cell line proficient and deficient for Arid1A. **C)** IHC for IBA1 and CD8 positive cells in Arid1A ko liver and WT with their quantification. **D)** co-occurrence analysis, using TCGA, between the Arid1a KO tumors and inactivating mutations of key players of the innate immune response.

We decided to confirm the activation of this pathway in HCT116 proficient and deficient for ARID1A to understand the possible mechanism. This isogenic model show the same genes activation (Fig. 8.3.4 B).

To investigate the biological effect of the upregulation of pro-inflammatory cytokines we performed immunohistochemistry analysis using specific immune cell markers. This analysis shown increased T-cells (CD8+ cells) and macrophages (IBA1 + cells) recruitment in ARID1A defective livers after thirty days from tamoxifen administration (Fig. 8.3.4 C) suggesting chronic inflammation promoted by ARID1A loss. Chronic activation of the interferon response may provide the ability to restrict microbial infections[147], which is a selective advantage already characterized to preserve adult stem cells[148]. This mechanism could be an advantage for normal cells, on one hand it could be potent tumor suppressor mechanism but it could select mutations able to affect interferon in order to positively select tumors. To verify this hypothesis, we analyzed ARID1A mutant tumors, of different origin from TCGA datasets. Our analysis reveals that inactivating mutations affecting at least one of the keys signaling molecules involved in immune response cooccur with ARID1A mutations, with JAK1 being the most frequently altered (Fig. 8.3.4 D) further corroborating the link between ARID1A mutations and innate immune response.

8.4. ARID1A LOSS AFFECT LIVER REGENERATION

To investigate the regenerative capacity of Arid1A deficient livers, we decided to feed the mice with the DDC diet. This diet is largely used in literature as chronic stimulus that mimic a pathology which is sclerosis cholangitis. DDC diet provokes the inhibition of ferrochelatase, this is a mitochondrial enzyme that catalyzes the insertion of Fe2+ into protoporphyrin IX to form heme. Protoporphyrin accumulates in the cytoplasm of parenchymal cells; its excess exits the liver through biliary excretion forming crystals plugs in bile canaliculi and bile ducts [38]. Bile duct damage is associated with cholangiocytes proliferation in portal area; DR expands forming small pseudo-ducts next to pre-existing bile ducts but remaining enclosed within the portal mesenchyme. Moreover, there is the formation of periductular fibrosis and inflammatory cells infiltration.

The use of DDC diet is a strong and reproducible model having minor impact on animal welfare. Using this approach, we observed the impaired capacity of ARID1A deficient liver to promote expansion of bile ducts (Fig. 8.4.1 A).

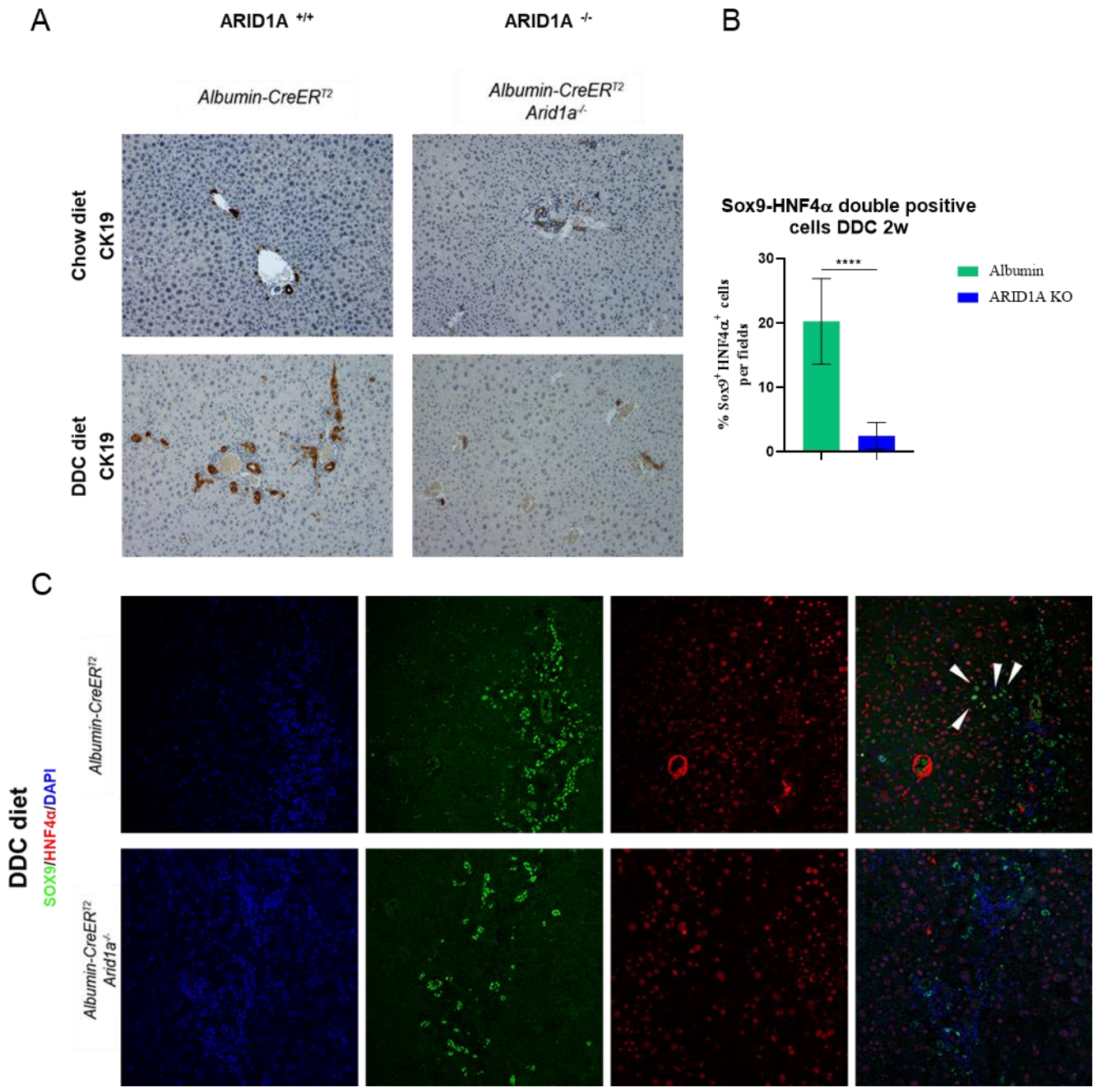


Figure 8.4.1 Arid1A affect the regeneration capacity of the liver

A) IHC showing the strong decrease of the cholangiocyte marker CK19 after DDC administration in the Arid1A KO livers. **B)** Quantification of SOX9/HNF4α double positive cells after 2 weeks of DDC diet. **C)** representative images of immunofluorescence of SOX9 (green) and HNF4α (red). The arrow head indicate the LPLCs.

In literature is reported that after DDC administration there is an expansion of the liver progenitor cells, but the loss of Arid1A dramatically impact on the regenerative capacity by

preventing the amplification of LPLCs which are characterized by the co-expression of SOX9 and HNF4α (Fig. 8.4.1 B).

Interestingly, we observed an increase deposition of H3K27Me3 level on the SOX9 promoter in the Arid1A KO hepatocytes compared to the control hepatocytes, also with loss of HNF4a and histone modification H3K27Ac (Fig. 8.4.2).

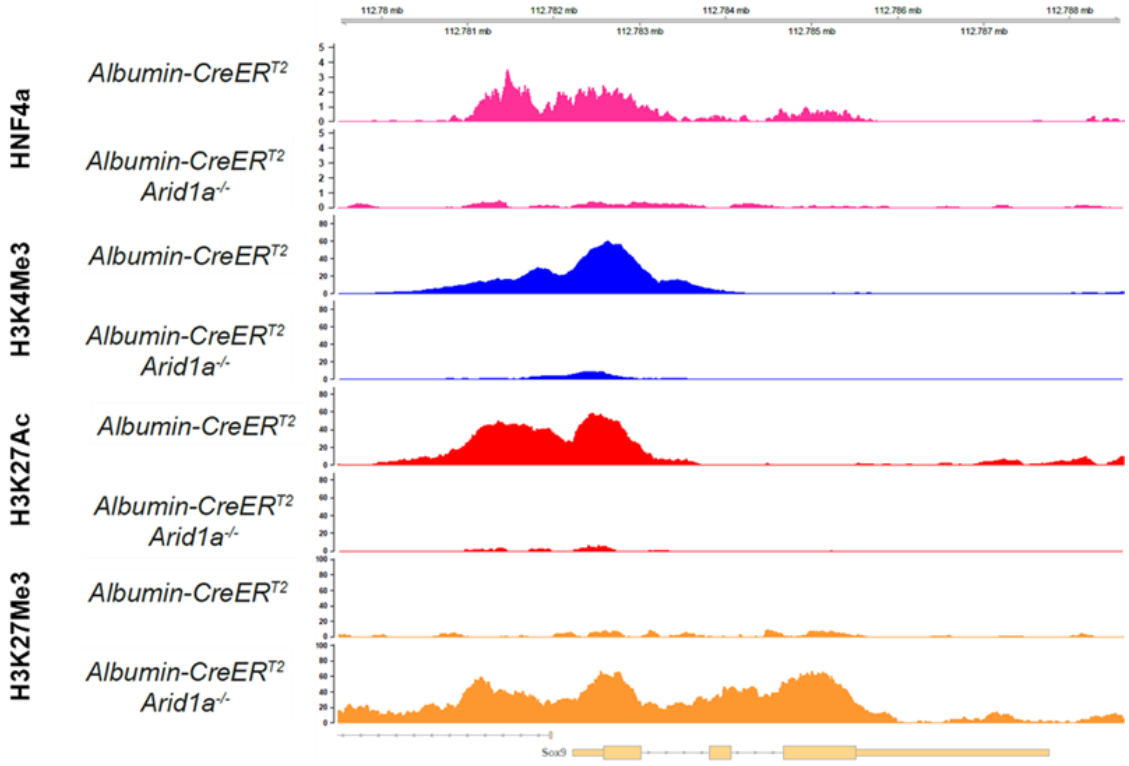


Figure 8.4.2 SOX9 promoter lose H3K27Ac, HNF4a and reach H3K27Me3

Snapshot of sox9 promoter region in which there is accumulation of H3K27Me3 with loss of the other modification / TF.

8.5. ARID1A LOSS INDUCE DNA DAMAGE ACCUMULATION PROMOTING GENOMIC INSTABILITY

Interferon response activation in the absence of viral infections is typically observed upon DNA damage[149]. We decided to performed immunohistochemistry analysis, observing an increase signal for γH2AX. Looking at the γH2AX staining we observed an increase of positive cells, number of nuclear malformations, and a strong presence of micronuclei in the ARID1A loss samples. All micronuclei were heavily stained by H3K27Me3 and accumulate DNA damage as shown by γH2AX (Fig. 8.5.1 A).

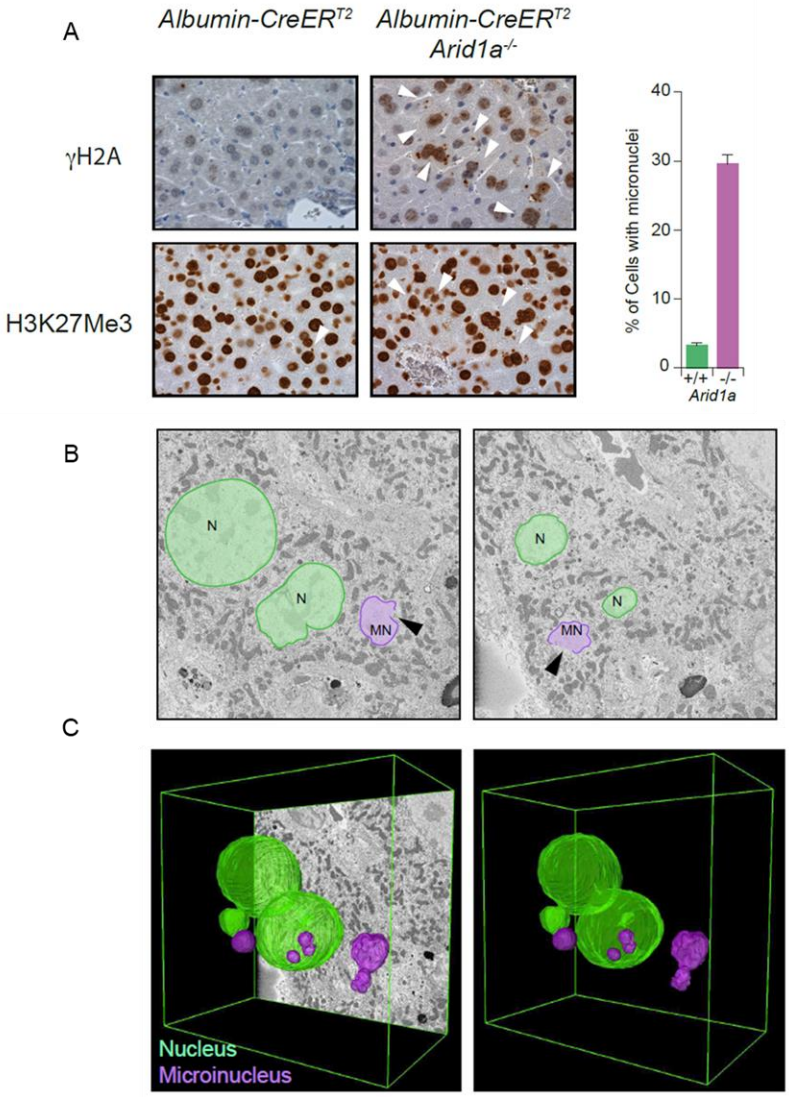


Figure 8.5.1

A) IHC showing γ-H2Ax and H3K27Me3 in Arid1A KO liver and WT, with the count of the percentage of cells with micronuclei. **B-C)** Electron microscopy images showing micronuclei. The arrowhead highlights a rupture of micronucleus membrane. 3D reconstruction of the micronuclei.

Micronuclei are small DNA-containing nuclear structures that are isolated from the nucleus, and characterized by nuclear membrane in order to isolate it from the cytoplasm. These structures are consequence of DNA damage occurring before or during mitosis due to chromosomal aberrations. The micronuclei are characterized by a particularly fragile membrane and they are prone to undergo ruptures with the exposure of DNA in the cytoplasm. We carefully analyzed ARID1A deficient liver samples for the presence of micronuclei. A strong increase of micronuclei-containing cells is observed after 9 days from tamoxifen administration (Fig. 8.5.1 A). To verify the integrity of micronuclei membrane we performed serial block face electron microscopy. Ultrastructural 3D analysis confirmed the presence of micronuclei-containing cells and reveal the presence of membrane lesions in around 15% of them (Fig. 8.5.1 B, C).

The micronuclei accumulation is not tissue specific, but we observed that HCT116 deficient for Arid1A show an increase of micronuclei if compared with the WT (Fig. 8.5.2). All these data suggest a common role of Arid1A in the maintenance of genome integrity.

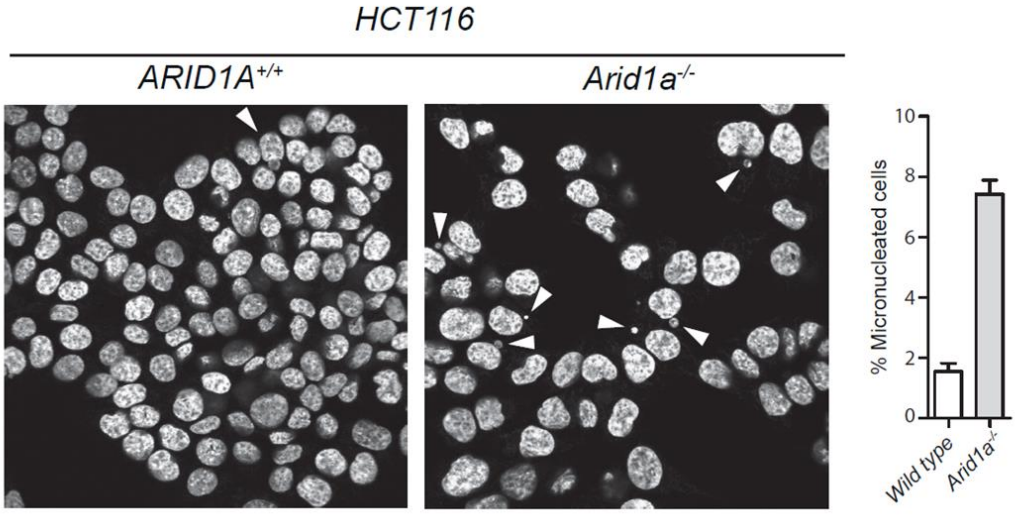


Figure 8.5.2 HCT116 deficient for Arid1a are characterized by micronuclei accumulation

Immunofluorescence using DAPI to visualized the micronuclei. To visualized clearly the micronuclei 63X objective was used.

To verify whether DNA damage accumulation could be a common consequence of ARID1A loss that can be observed in different tissues, we generated *Prom1CreER*^{T2}*Arid1A*$^{fl/fl}$ mouse line. In these mice Cre expression can be observed in the stem/progenitor compartment of several different tissues, including small and large intestinal crypts, gastric glands, prostate epithelium, lung and kidney among others. Using this model, we confirmed transcriptional activation of interferon target genes in purified *Arid1A*$^{-/-}$ epithelial cells from small intestine and colon crypts after 7 days from tamoxifen administration (Fig. 8.5.2 A). we also found an increase of p21 and γ-H2Ax (Fig. 8.5.2 B) in stomach glands.

These data indicate that ARID1A loss has as a common consequence in different compartment the increase of DNA damage with the upregulation of the upregulation of the interferon target genes that lead to a chronic inflammation.

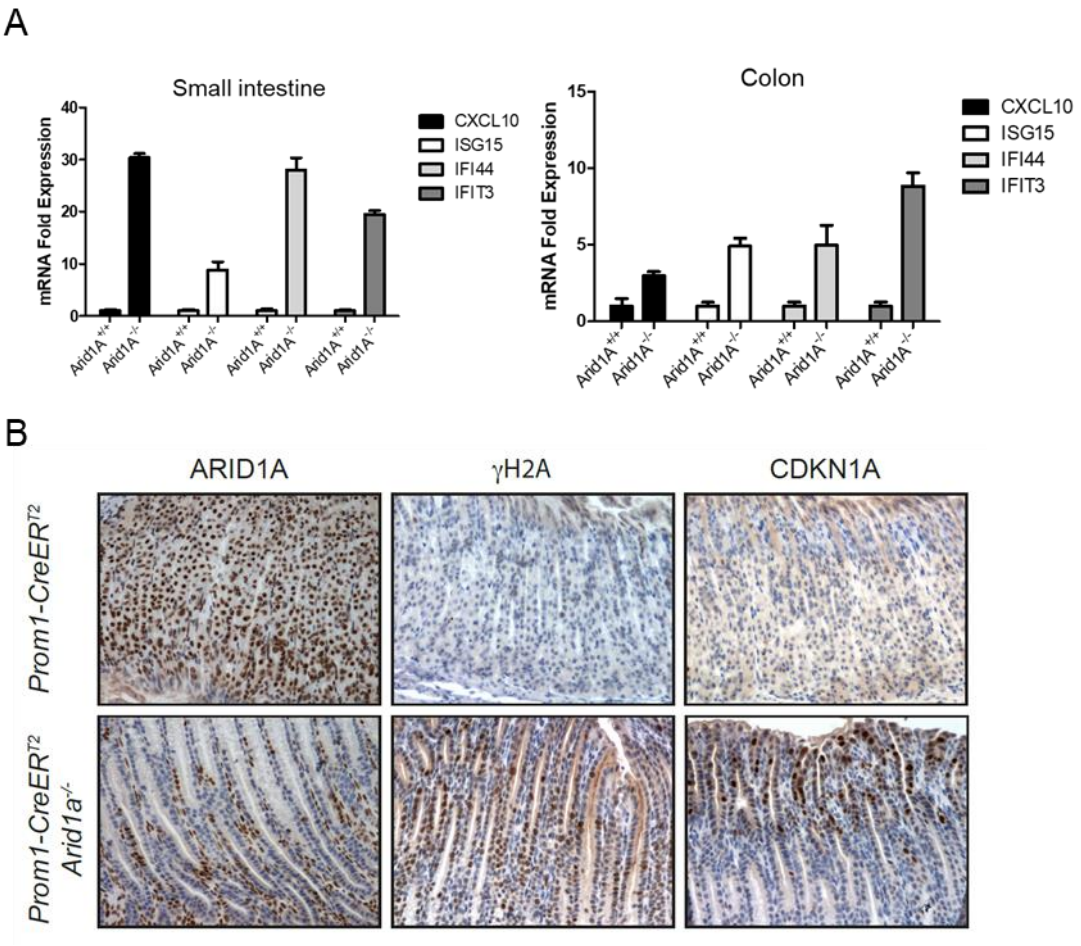


Figure 8.5.2

A) Quantitative PCR of interferon target genes in small intestine and colon **B)** Representative IHC performed on gastric epithelium of Prom1CREERT2 and Prom1CREERT2-Arid1A Fl/Fl for γ-H2Ax and CDKN1A.

8.6. ARID1A LOSS IS SUFFICIENT TO PREDISPOSE TO HCC

One of the main functions of the liver is to detoxify potentially toxic substances, such as drugs, steroids compounds and bile acids. To investigate the role of ARID1A loss in HCC, after the induction of ARID1A deletion by intraperitoneal tamoxifen (TAM) injection, we decide to challenge the animal with an acute stimulus to activate the xenobiotic response. After seven days from deletion, we treated mice with TCPOBOP (or DMSO as negative control) to activate the Constitutive androstane receptor (CAR) (Fig. 8.6.1 B). Performing a time-course experiments with TCPO we observed a transient increase of liver-to-body weight ratio in WT mice, while ARID1A loss liver mice showed liver overgrowth (Fig. 8.6.1 A). To understand if ARID1A loss is sufficient to promote liver tumorigenesis, we sacrificed treated mice with TCPOBOP after 5 months. Surprisingly, only TCPOBOP treated ARID1A KO mice developed tumors, we found multiple macroscopic nodules in eight out of ten animals (Fig. 8.6.1 C). On the contrary, ARID1A KO mice did not show any nodules, but they are characterized by a clear state of fibrosis and hepatitis (Fig. 8.6.1 D).

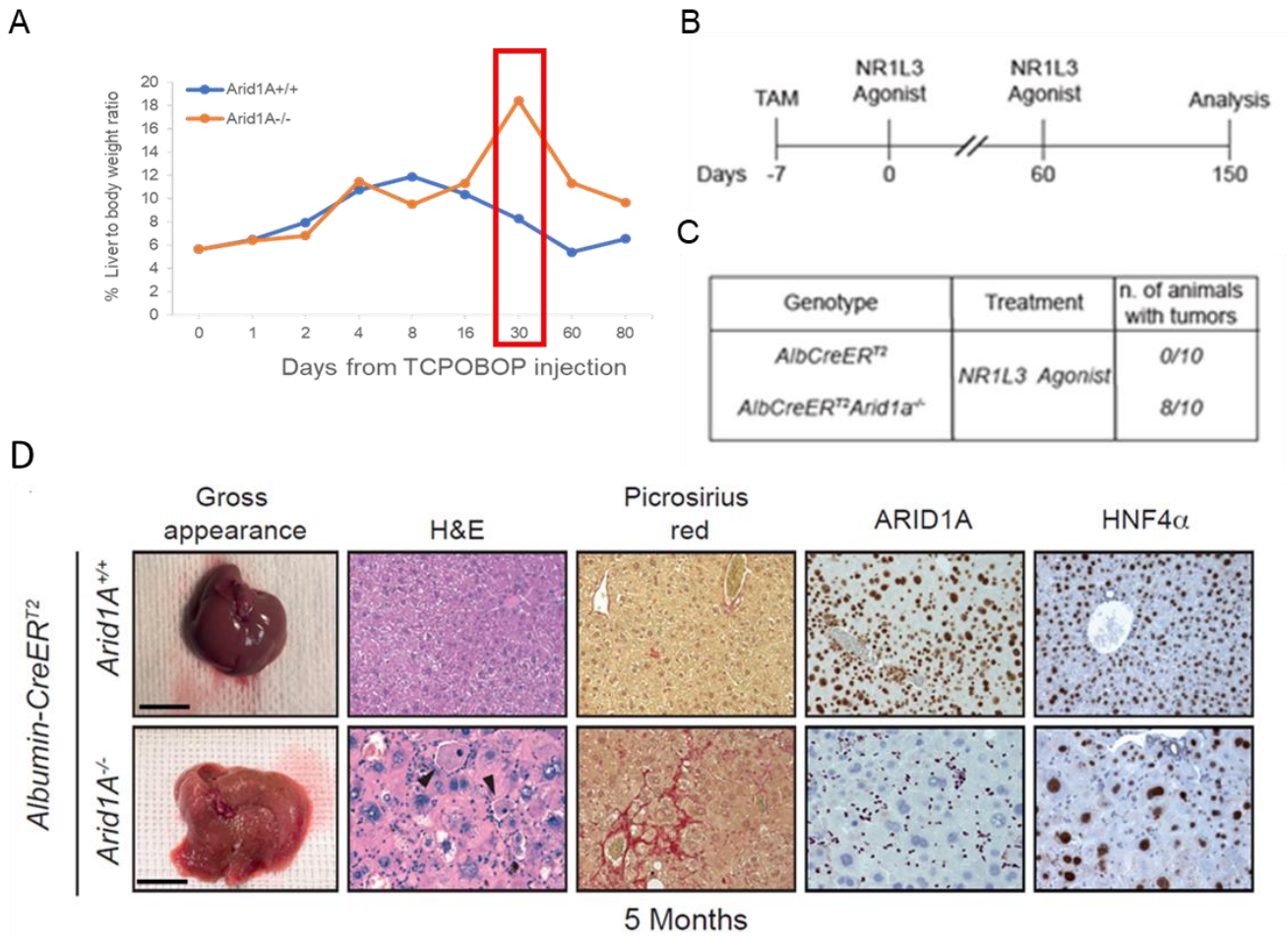


Figure 8.6.1

A) Time-couse after TC injection show percentage of liver/body weight ratio of WT (blue) and KO (red). Red box highlights the overgrowth of the KO liver compared to the WT. **B)** mice were injected with tamoxifen and treated with two shot of TCPOBOP/DMSO after seven days from Arid1A abrogation and 60 days from first administration. **C)** Table showing the incidence on animal treated with the CAR agonist. **D)** IHC comparing Arid1A KO and to WT liver showing cell death and nuclear malformation (H&E), increase of extracellular matrix (picrosirius red)

To understand which kind of mutation can synergized with ARID1A loss and the potential mutational spectrum of the TCPOBOP treatment-derived tumors, we performed whole exome sequencing. From this kind of analysis, we found Myc and YAP1 amplifications, PTEN deletion and β-CATENIN (or CTNNB1) point mutation (Fig. 8.6.2 A). We performed immunohistochemistry (IHC) experiments to better characterize these tumors. To understand which kinds of tumors they are we started to stain the tumoral tissues with HNF4α, which is distinctive of HCC, observing its expression in our samples. Then, we

scored a strong increase of proliferation by using Ki-67 marker. And finally, we observed the nuclear translocation of β-CATENIN in some nodules that confirm the whole exome sequencing data (Fig. 8.6.2 C). Furthermore, some of these tumors showed an increased expression of the stemness-related marker Cytokeratin 19 (Krt19 or CK19), which is the typical feature of more aggressive HCC with poor overall survival (38) (Fig.8.6.2 B). All together, these data suggest that ARID1A loss favor HCC formation upon an acute stimulus activating detoxification, but also it can synergize with mutations such as β-Catenin.

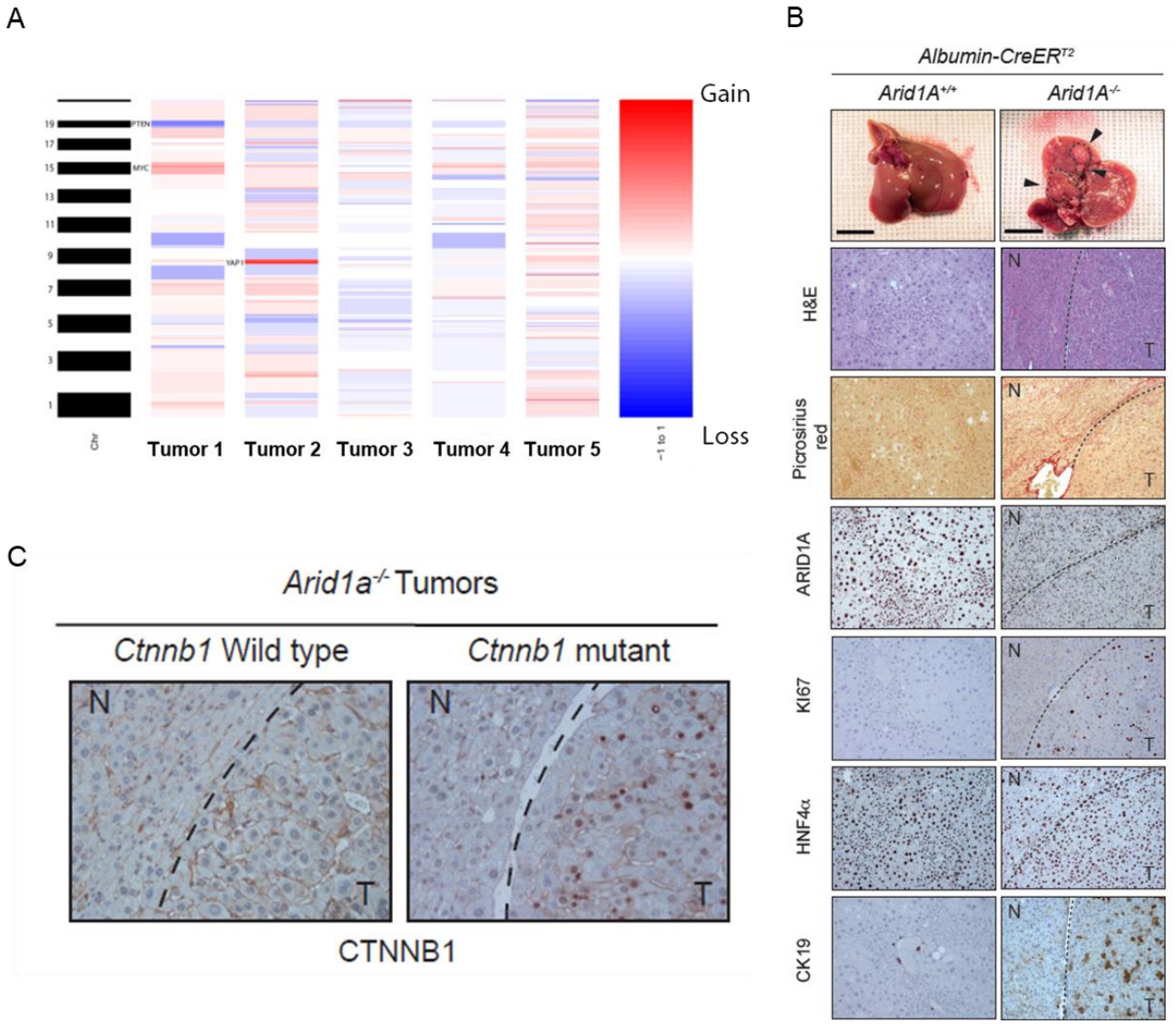


Figure 8.6.2

A) Whole exome sequencing of Arid1A ko tumors showing the mutations identified. **B)** Gross appearance of the nodules with the IHC performed on HCC derived from Arid1a KO mice for Arid1A, Krt19, Ki67, HNF4α, CTNNB1. **C)** representative characterization of Ctnnb1 mutations in Arid1A KO tumors.

To identify the transcriptional program possibly involved in the phenotype observed we performed RNAseq analysis after 30 days from CAR activation. We identified 876 differentially expressed genes (Fig. 8.6.3 A). GO terms enrichment analysis reveals that, together with genes involved in the interferon response, several genes involved in G2/M checkpoint and chromosome segregation were upregulated further supporting a role of ARID1A in preserving genomic integrity during G2/M transition (Fig. 8.6.3 B). These data indicate a central role of ARID1A in preserving genome integrity and show how ARID1A loss is sufficient to sensitize hepatocytes to endogenous and exogenous insults, favoring the formation of liver tumors.

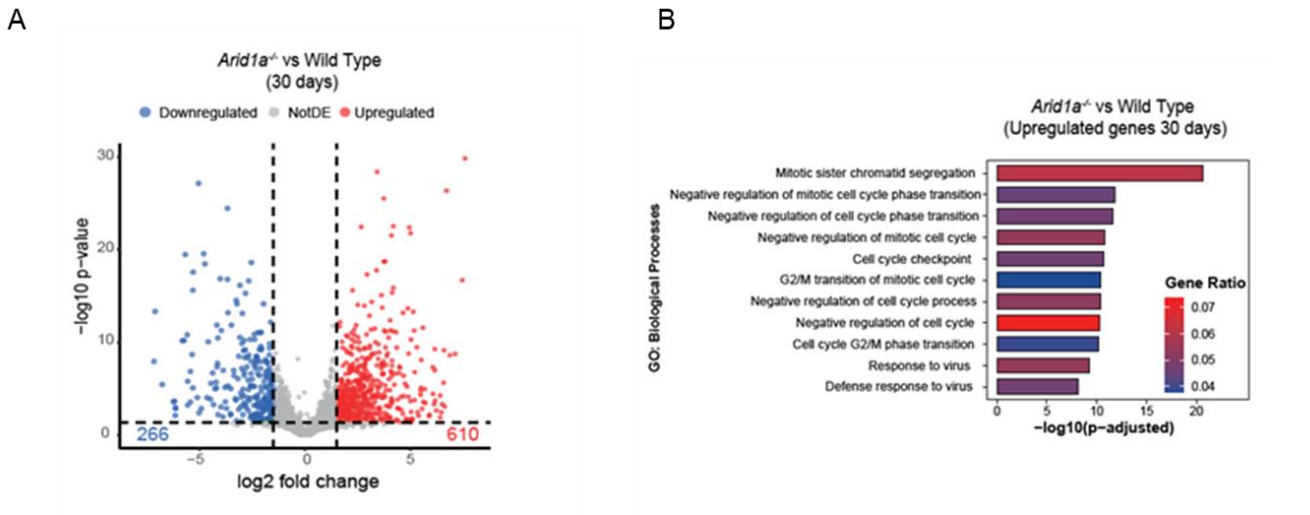


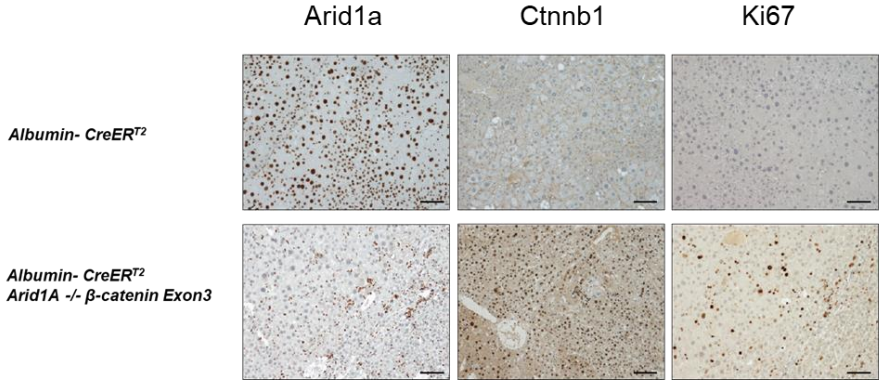
Figure 8.6.3

A) Volcano plot showing differential expressed genes of the Arid1A KO liver and the WT liver both treated with TCPOBOP obtained by RNA-seq. **B)** GO terms of the upregulated genes of the RNA-seq.

8.7. ARID1A COOPERATES WITH B-CATENIN TO INDUCE AGGRESSIVE TUMORS

Our previous findings suggested that ARID1A and β-CATENIN cooperate in hepatocarcinogenesis. To investigate this cooperation, we generated Arid1A KO mice harbored an oncogenic mutated form of β-CATENIN (Exon 3-/-). When these mice were injected with tamoxifen led to the deletion of the exon 3. Exon 3 is a key exon encoding serine-threonine phosphorylation sites for GSK-3β that activates degradation of β-catenin forming. We decided to use the leakiness of the model that caused a stochastic activation of the Cre-recombinase. β-CATENIN has an important role in maintaining liver zonation. The activation in whole liver induce liver zonation loss with consequent loss of the urea metabolism, and in a couple of weeks led the mouse to death by intoxication of ammonium ions (39).

A



B

Mouseline	n. of mice with tumors
Albcreert2	0/7
Albcreert2 Arid1a fl/fl	0/7
Albcreert2 β-catenin Exon3	0/7
Albcreert2 Arid1A-β-catenin Exon3	20/20

C

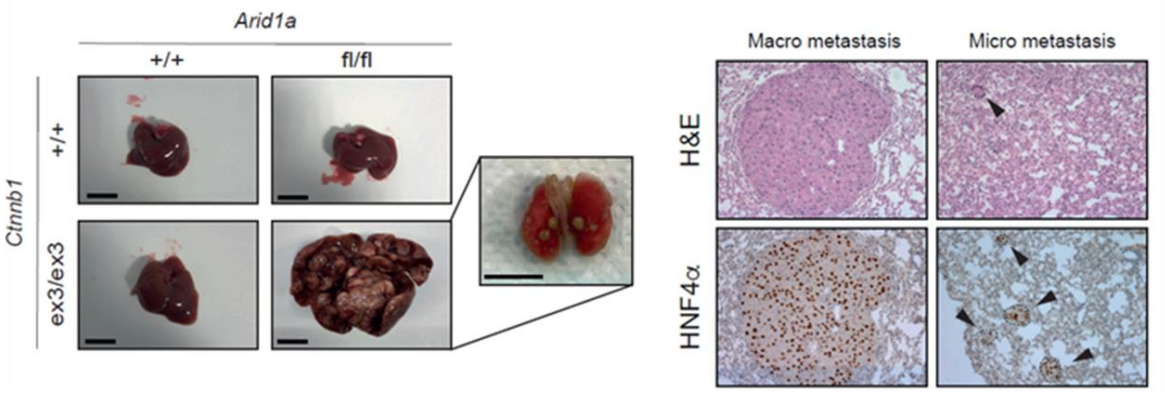


Figure 8.7.1

A) IHC showing Arid1A depletion, B-Catenin activation and Ki67 increase in the derived tumors. **B)** table indicate the stochastic tumor developed from the different mousseline. **C)** Representative images of liver after 10 months with lung metastasis positive to HNF4α

After 10 months mice were sacrificed and all the animals with the double mutations show multiple liver tumors (Fig. 8.7.1 C), while in single Arid1A fl/fl and WT mice, sacrificed at the same time point, no nodules were found and they appear completely normal and indistinguishable from wild type animals. (Fig. 8.7.1 B).

IHC showed that tumor derived from the stochastic activation of the Cre-recombinase are characterized by a complete Arid1A loss and a nuclear translocation of β-CATENIN with an increase of Ki-67 as proliferation marker (Fig. 8.7.1 A). Moreover, 20% of these animals developed lung metastasis which are characterized by nodules lacking Arid1A and expressing HNF4α (Fig. 8.7.1 C).

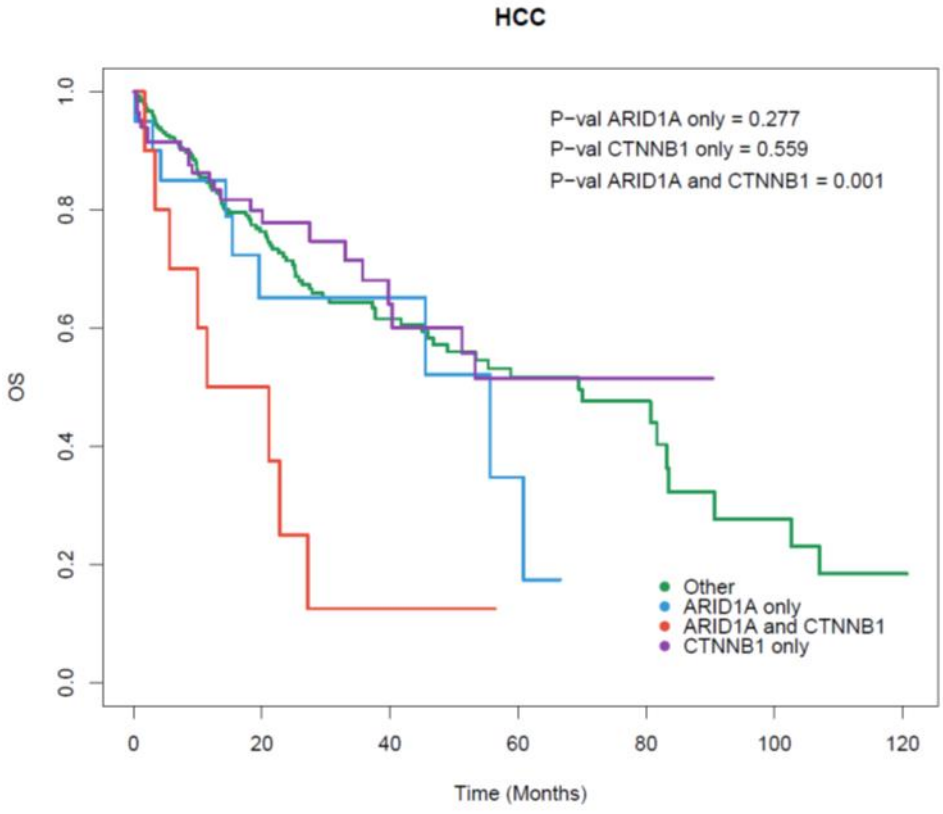


Figure 8.7.2 Overall survival of Arid1A and β- catenin mutations

Survival plot from the TCGA comparing double mutant patients to the single mutations with all the other HCC.

Considering these features, we decided to compare these mutations on human HCC by using TCGA. Hence, we interrogated the TCGA for survival of patients carrying these

mutations comparing them to all the other HCC patients. The worst prognosis was observed for the double mutant HCC (ARID1A and β-CATENIN) followed by the single ARID1A mutant and then all the other HCC (Fig. 8.7.2).

9. Discussion

The involvement of Arid1A in tumorigenesis was largely studied in the past years, however its role was investigated mainly *in vitro* using cancer cell lines and only recently using *in vivo* models with controversial results on the base of the context and the tissue involved. In the present study, we shed light on the role of ARID1A in liver homeostasis and tumorigenesis. Its contribution has been recently discussed by different works. Sun X. and colleagues argued that ARID1A abrogation improves liver regeneration after partial hepatectomy or after chemicals administration[77]. However, our work showed that ARID1A loss blocks proliferation and induce hepatocytes to a senescence state blocking them in a cell cycle checkpoint. The same authors[61] described ARID1A as an oncogene underlying its requirement to tumor initiation and its loss for progression and metastasis in established tumor. Conversely, our data showed that ARID1A loss induces a pro-tumoral environment and when an acute activation of the detoxification pathway occurs it favors tumorigenesis. Furthermore, in the last work from the same laboratory it was described how ARID1A loss synergized with the mutation of ARID1B inducing rapid carcinogenesis caused by the disruptive redistribution of all BAF complexes[150]. We did not find any ARID1B mutation in our whole exome sequencing data but instead MYC and YAP1 amplifications, PTEN deletion, and point mutation of β-CATENIN.

In liver, loss of ARID1A produces a deregulation of a subset of enhancer, half of them, with loss acetylation while only one quantile presented a gain of this modification. Of note, when both the Arid1 proteins where deleted, there was a strong H3K4me1 loss with the loss also of H3K27Ac on all the Arid1A enhancers. Looking at the transcriptional profile of the Arid1A/B double knock out, there be found the downregulation of genes all involved in the liver metabolic process with the consequence loss of its functionality. Among the upregulated genes there be found genes involved in wound healing, inflammation, and

fibrotic processes. This highlighted the redundancy of Arid1B which is involved in the maintenance expression of liver specific genes in the absence of Arid1A.

We also characterized the synergistic effect of Arid1A and β-CATENIN mutations. These tumors are very similar to the tumors observed in human, where tumors harboring both mutations display worsen prognosis if compared with the tumor that harbor single mutation. Moreover, constitutively stable β-CATENIN and ARID1A loss synergize to form aggressive, metastatic HCCs.

An analysis performed in human liver samples by *Zhu M. and colleagues* suggested how ARID1A mutations could be an early mutational event. Indeed, its mutations was found in pre-tumoral stage with HBV-HCV infection and no one of them developed tumors[120]. We showed that the activation of the interferon pathway caused by ARID1A abrogation leads to an increase of immune cells recruitment. This findings are confirmed by re-analyzing published datas[144,145,146]. The upregulation of this pathway was common among different tissues and models upon ARID1A loss underling a common effect. Moreover, we have shown that upregulation of these genes is cell autonomous by using a cell line proficient or deficient for ARID1A in which the interferon pathway was upregulated in cells lacking ARID1A.

Our data suggest that Arid1A play a dual role in tumorigenesis, by increasing genomic instability with the formation of micronuclei, and favoring chronic inflammation by the activation of the interferon pathway. The lesion on the membrane of the micronuclei exposed the DNA to the cytoplasm. This event mimics a viral infection and probably is detected from the cGAS-STING pathway, pathway involved in the sense of cytoplasmic DNA, which contribute upregulation of the interferon response genes. The chronic activation of the innate immune response could favor the immune editing with the selection of specific mutations able to bypass the immune surveillance. Indeed, both β-Catenin and MYC were identified

as genes able to directly promote immune escape[152] or as genes that directly suppress immune response[153].

Recently, *Weiping Li and colleagues* attributed the inability of the liver to regenerate to Yap which prevents the amplification of the LPLCs[151]. In our work, after Arid1A depletion we link the inability to amplify the LPLCs to an increase level of H3K27Me3 on the SOX9 promoter with a concomitant loss of HNF4a binding and H3K27Ac.

We propose a possible model in which ARID1A loss does not promote directly HCC but it induces a pro-tumoral environment. Indeed, an acute stimulation of the detoxification pathway with a non-genotoxic agonist of Constitutive androstane receptor, is sufficient to promote HCC formation.

In this work we characterized different aspects that could be improve to develop treatment for HCC lacking Arid1A. First the possible involvement of TOPIIA, and second the possible role of the cGAS-STING pathway.

Nevertheless, further investigation is required to understand the interplay between the different BAF complexes in orchestrating nucleosomes repositioning and their role to maintain the transcription programs.

10. BIBLIOGRAPHY

1. Marquardt, J. U., Andersen, J. B. & Thorgeirsson, S. S. Functional and genetic deconstruction of the cellular origin in liver cancer. *Nat. Publ. Gr.* (2015) doi:10.1038/nrc4017.

2. Yamashita, T. & Wang, X. W. The Journal of Clinical Investigation Cancer stem cells in the development of liver cancer. **123**, (2013).

3. Llovet, J. M. & Bruix, J. Molecular targeted therapies in hepatocellular carcinoma. *Hepatology* **48**, 1312–1327 (2008).

4. Yang, J. D. *et al.* A global view of hepatocellular carcinoma: trends, risk, prevention and management HHS Public Access. *Nat Rev Gastroenterol Hepatol* **16**, 589–604 (2019).

5. Ghouri, Y. A., Mian, I. & Rowe, J. H. Review of hepatocellular carcinoma: Epidemiology, etiology, and carcinogenesis. (2017) doi:10.4103/jcar.JCar_9_16.

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

7. Uka, K. *et al.* Clinical features and prognosis of patients with extrahepatic metastases from hepatocellular carcinoma RAPID COMMUNICATION. *World J Gastroenterol* **13**, 414–420 (2007).

8. Li, X. *et al.* Animal model of intrahepatic metastasis of hepatocellular carcinoma: establishment and characteristic. *Sci. Reports 2020 101* **10**, 1–6 (2020).

9. Müller, M., Bird, T. G. & Nault, J.-C. The landscape of gene mutations in cirrhosis and hepatocellular carcinoma. (2020) doi:10.1016/j.jhep.2020.01.019.

10. Zhang, D. Y. & Friedman, S. L. Fibrosis-dependent mechanisms of hepatocarcinogenesis. *Hepatology* **56**, 769–775 (2012).

11. Kojiro, M. *et al.* Pathologic diagnosis of early hepatocellular carcinoma: A report of the International Consensus Group for Hepatocellular Neoplasia. *Hepatology* **49**, 658–664 (2009).

12. Farazi, P. A. & DePinho, R. A. Hepatocellular carcinoma pathogenesis: From genes to environment. *Nat. Rev. Cancer* **6**, 674–687 (2006).

13. Nault, J. C. *et al.* Telomerase reverse transcriptase promoter mutation is an early somatic genetic alteration in the transformation of premalignant nodules in hepatocellular carcinoma on cirrhosis. *Hepatology* **60**, 1983–1992 (2014).

14. Umemura, A. *et al.* p62, Upregulated during Preneoplasia, Induces Hepatocellular Carcinogenesis by Maintaining Survival of Stressed HCC-Initiating Cells. *Cancer Cell* **29**, 935–948 (2016).

15. Schulze, K. *et al.* Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets Europe PMC Funders Group. *Nat Genet* **47**, 505–511 (2015).

16. Chen, Y., Wang, L., Xu, H., Liu, X. & Zhao, Y. Exome capture sequencing reveals new insights into hepatitis B virus-induced hepatocellular carcinoma at the early stage of tumorigenesis. *Oncol. Rep.* **30**, 1906–1912 (2013).

17. Fujimoto, A. *et al.* Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. (2012) doi:10.1038/ng.2291.

18. Nault, J. C. *et al.* ARTICLE High frequency of telomerase reverse-transcriptase promoter somatic mutations in hepatocellular carcinoma and preneoplastic lesions. (2013) doi:10.1038/ncomms3218.

19. Totoki, Y. *et al.* Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. *Nat. Publ. Gr.* **46**, (2014).

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

21. Nahon, P. *et al.* A variant in myeloperoxidase promoter hastens the emergence of hepatocellular carcinoma in patients with HCV-related cirrhosis. *J. Hepatol.* **56**, 426–432 (2012).

22. Bisso, A. *et al.* Cooperation Between MYC and β-Catenin in Liver Tumorigenesis Requires Yap/Taz. *Hepatology* **72**, 1430–1443 (2020).

23. Nault, J. C., Bioulac-Sage, P. & Zucman-Rossi, J. Hepatocellular Benign Tumors—From Molecular Classification to Personalized Clinical Care. *Gastroenterology* **144**, 888–902 (2013).

24. Zucman-Rossi, J. *et al.* Genotype-phenotype correlation in hepatocellular adenoma: New classification and relationship with HCC. *Hepatology* **43**, 515–524 (2006).

25. Morini, S., Carpino, G., Carotti, S. & Gaudio, E. Anatomy and Embryology of the Liver. *Liver Dis.* 3–16 (2020) doi:10.1007/978-3-030-24432-3_1.

26. 23.6 Accessory Organs in Digestion: The Liver, Pancreas, and Gallbladder - Anatomy and Physiology 2e | OpenStax. https://openstax.org/books/anatomy-and-physiology-2e/pages/23-6-accessory-organs-in-digestion-the-liver-pancreas-and-gallbladder.

27. Trefts, E., Gannon, M., Wasserman, D. H. & Author, C. B. The liver HHS Public Access Author manuscript. (2018) doi:10.1016/j.cub.2017.09.019.

28. Fu, X. *et al.* Modeling of xenobiotic transport and metabolism in virtual hepatic lobule models. (2018) doi:10.1371/journal.pone.0198060.

29. Abdel-Misih, S. R. Z., Bloomston, M. & Bismuth, H. Liver Anatomy. (2010) doi:10.1016/j.suc.2010.04.017.

30. Rani, H. P., Sheu, T. W. H., Chang, T. M. & Liang, P. C. Numerical investigation of non-Newtonian microcirculatory blood flow in hepatic lobule. *J. Biomech.* **39**, 551–563 (2006).

31. Alva, V. & Lupas, A. N. The TULIP superfamily of eukaryotic lipid-binding proteins as a mediator of lipid sensing and transport. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1861**, 913–923 (2016).

32. Birchmeier, W. Orchestrating Wnt signalling for metabolic liver zonation. *Nat. Cell Biol. 2016 185* **18**, 463–465 (2016).

33. Chen, M. *et al.* The liver toxicity knowledge base: A systems approach to a complex end point. *Clin. Pharmacol. Ther.* **93**, 409–412 (2013).

34. Jungermann, K. & Kietzmann, T. Oxygen: Modulator of metabolic zonation and disease of the liver. *Hepatology* **31**, 255–260 (2000).

35. Gissen, P. & Arias, I. M. Structural and functional hepatocyte polarity and liver disease. *J. Hepatol.* **63**, 1023–1037 (2015).

36. Kreutz, C. *et al.* Hepatocyte ploidy is a diversity factor for liver homeostasis. *Front. Physiol.* **8**, 862 (2017).

37. Schulze, R. J., Schott, M. B., Casey, C. A., Tuma, P. L. & Mcniven, M. A. Beyond the Cell The cell biology of the hepatocyte: A membrane trafficking machine. (2019) doi:10.1083/jcb.201903090.

38. Grant, D. M. Detoxification pathways in the liver. *J. Inherit. Metab. Dis.* **14**, 421–430 (1991).

39. Sevior, D. K., Pelkonen, O. & Ahokas, J. T. Hepatocytes: The powerhouse of biotransformation. *Int. J. Biochem. Cell Biol.* **44**, 257–261 (2012).

40. Barker, N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat. Rev.* | *Mol. CELL Biol.* **15**, (2013).

41. Sultana, H., Komai, M. & Shirakawa, H. The Role of Vitamin K in Cholestatic Liver Disease. (2021) doi:10.3390/nu13082515.

42. Wada, T., Gao, J. & Xie, W. PXR and CAR in energy metabolism. *Trends Endocrinol. Metab.* **20**, 273–279 (2009).

43. Hodges, R. E. & Minich, D. M. Modulation of Metabolic Detoxification Pathways Using Foods and Food-Derived Components: A Scientific Review with Clinical

Application. *J. Nutr. Metab.* **2015**, (2015).

44. Michalopoulos, G. K. Hepatostat: Liver regeneration and normal liver tissue maintenance. *Hepatology* **65**, 1384–1392 (2017).

45. Michalopoulos, G. K. & Bhushan, B. Liver regeneration: biological and pathological mechanisms and implications. *Nat. Rev. Gastroenterol. Hepatol.* **18**, 40–55 (2021).

46. Yanger, K. *et al.* Adult hepatocytes are generated by self-duplication rather than stem cell differentiation. *Cell Stem Cell* **15**, 340–349 (2014).

47. Monga, S. P. Updates on hepatic homeostasis and the many tiers of hepatobiliary repair. doi:10.1038/s41575-018-0090-x.

48. Gilgenkrantz, H. & Collin de l'Hortet, A. Understanding Liver Regeneration: From Mechanisms to Regenerative Medicine. *Am. J. Pathol.* **188**, 1316–1327 (2018).

49. Li, W., Li, L. & Hui, L. Cell Plasticity in Liver Regeneration. *Trends Cell Biol.* **30**, 329–338 (2020).

50. Mao, S. A., Glorioso, J. M. & Nyberg, S. L. Liver regeneration. *Transl. Res.* **163**, 352–362 (2014).

51. Michalopoulos, G. K., Barua, L. & Bowen, W. C. Transdifferentiation of rat hepatocytes into biliary cells after bile duct ligation and toxic biliary injury. *Hepatology* **41**, 535–544 (2005).

52. Han, X. *et al.* Lineage Tracing Reveals the Bipotency of SOX9 + Hepatocytes during Liver Regeneration. *Stem Cell Reports* **12**, 624–638 (2019).

53. Lin, Y., Zhang, F., Zhang, L., Chen, L. & Zheng, S. Characteristics of SOX9-positive progenitor-like cells during cholestatic liver regeneration in biliary atresia. *Stem Cell Res. Ther.* **13**, 1–11 (2022).

54. Miyajima, A., Tanaka, M. & Itoh, T. Stem/Progenitor Cells in Liver Development, Homeostasis, Regeneration, and Reprogramming. *Cell Stem Cell* **14**, 561–574 (2014).

55. Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol. 2015 163* **16**, 178–189 (2015).

56. Weaver, I. C. G. *et al.* Stress and the Emerging Roles of Chromatin Remodeling in Signal Integration and Stable Transmission of Reversible Phenotypes. (2017) doi:10.3389/fnbeh.2017.00041.

57. Fyodorov, D. V, Zhou, B.-R., Skoultchi, A. I. & Bai, Y. Emerging roles of linker histones in regulating chromatin structure and function HHS Public Access. *Nat Rev Mol Cell Biol* **19**, 192–206 (2018).

58. Hergeth, S. P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* **16**, 1439–1453 (2015).

59. Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).

60. Valencia, A. M. & Kadoch, C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer HHS Public Access. *Nat Cell Biol* **21**, 152–161 (2019).

61. Sun, X. *et al.* Arid1a Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer Cell* **32**, 574-589.e6 (2017).

62. Wilson, B. G. & Roberts, C. W. M. SWI/SNF nucleosome remodellers and cancer. *Nat. Rev.* | *CANCER* **11**, 481 (2011).

63. Saha, A., Wittmeyer, J. & Cairns, B. R. Chromatin remodelling: The industrial revolution of DNA around histones. *Nat. Rev. Mol. Cell Biol.* **7**, 437–447 (2006).

64. Dechassa, M. L. *et al.* SWI/SNF Has Intrinsic Nucleosome Disassembly Activity that Is Dependent on Adjacent Nucleosomes. *Mol. Cell* **38**, 590–602 (2010).

65. Zhang, Z.-K. *et al.* Cell Cycle Arrest and Repression of Cyclin D1 Transcription by INI1/hSNF5. *Mol. Cell. Biol.* **22**, 5975 (2002).

66. Neigeborn, L. & Carlson, M. Genes Affecting the Regulation of SUC2 Gene Expression by Glucose Repression in SACCHAROMYCES CEREVISIAE. *Genetics* **108**, 845 (1984).

67. Breeden, L. & Nasmyth, K. Cell cycle control of the yeast HO gene: Cis- and Trans-acting regulators. *Cell* **48**, 389–397 (1987).

68. Smith, C. L., Horowitz-Scherer, R., Flanagan, J. F., Woodcock, C. L. & Peterson, C. L. Structural analysis of the yeast SWI/SNF chromatin remodeling complex. (2003) doi:10.1038/nsb888.

69. Kadoch, C. & Crabtree, G. R. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, (2015).

70. Mashtalir, N. *et al.* Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. *Cell* **175**, 1272-1288.e20 (2018).

71. Mohrmann, L. & Verrijzer, C. P. Composition and functional specificity of SWI2/SNF2 class chromatin remodeling complexes. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1681**, 59–73 (2005).

72. Wang, W. *et al.* Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**, 5370–5382 (1996).

73. Kaeser, M. D., Aslanian, A., Dong, M. Q., Yates, J. R. & Emerson, B. M. BRD7, a Novel PBAF-specific SWI/SNF Subunit, Is Required for Target Gene Activation and Repression in Embryonic Stem Cells. *J. Biol. Chem.* **283**, 32254 (2008).

74. Wanior, M., Krämer, A., Knapp, S. & Joerger, A. C. Exploiting vulnerabilities of SWI/SNF chromatin remodelling complexes for cancer therapy. *Oncogene* **40**, 3637–3654 (2021).

75. Flowers, S., Nagl, N. G., Beck, G. R. & Moran, E. Antagonistic roles for BRM and BRG1 SWI/SNF complexes in differentiation. *J. Biol. Chem.* **284**, 10067–10075 (2009).

76. Gao, X. *et al.* ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 6656 (2008).

77. Sun, X. *et al.* Suppression of the SWI/SNF component Arid1a promotes mammalian

regeneration HHS Public Access. *Cell Stem Cell* **18**, 456–466 (2016).

78. Jagani, Z. *et al.* Loss of the tumor suppressor Snf5 leads to aberrant activation of the Hedgehog-Gli pathway. *Nat. Med. 2010 1612* **16**, 1429–1433 (2010).

79. Caramel, J., Quignon, F. & Delattre, O. RhoA-Dependent Regulation of Cell Migration by the Tumor Suppressor hSNF5/INI1. *Cancer Res.* **68**, 6154–6161 (2008).

80. Park, J. H. *et al.* Mammalian SWI/SNF complexes facilitate DNA double-strand break repair by promoting γ-H2AX induction. *EMBO J.* **25**, 3986–3997 (2006).

81. Dykhuizen, E. C. *et al.* BAF complexes facilitate decatenation of DNA by topoisomerase IIα. *Nat. 2013 4977451* **497**, 624–627 (2013).

82. Centore, R. C. *et al.* Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends Genet.* **36**, 936–950 (2020).

83. Ramos, P. *et al.* Small cell carcinoma of the ovary, hypercalcemic type, displays frequent inactivating germline and somatic mutations in SMARCA4. *Nat. Genet. 2014 465* **46**, 427–429 (2014).

84. Roberts, C. W. M., Galusha, S. A., McMenamin, M. E., Fletcher, C. D. M. & Orkin, S. H. Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13796–13800 (2000).

85. Khavari, P. A., Peterson, C. L., Tamkun, J. W., Mendel, D. B. & Crabtree, G. R. BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nat. 1993 3666451* **366**, 170–174 (1993).

86. Smith, M. J. *et al.* Loss-of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nat. Genet. 2013 453* **45**, 295–298 (2013).

87. St Pierre, R. *et al.* SMARCE1 deficiency generates a targetable mSWI/SNF dependency in clear cell meningioma. doi:10.1038/s41588-022-01077-0.

88. Smith, M. J. *et al.* Germline SMARCE1 mutations predispose to both spinal and cranial clear cell meningiomas. *J. Pathol.* **234**, 436–440 (2014).

89. Wilsker, D. *et al.* Nomenclature of the ARID family of DNA-binding proteins. *Genomics* **86**, 242–251 (2005).

90. Wu, R.-C., Wang, T.-L. & Shih, M. The emerging roles of ARID1A in tumor suppression. *Cancer Biol. Ther.* **655**, 655–664 (2014).

91. Nagl, N. G., Wang, X., Patsialou, A., Scoy, M. Van & Moran, E. Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J.* **26**, 752–763 (2007).

92. Kelso, T. W. R. *et al.* Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. doi:10.7554/eLife.30506.001.

93. Tordella, L. *et al.* SWI/SNF regulates a transcriptional program that induces senescence to prevent liver cancer. *Genes Dev.* **30**, 2187–2198 (2016).

94. Caumanns, J. J., Wisman, G. B. A., Berns, K., van der Zee, A. G. J. & de Jong, S. ARID1A mutant ovarian clear cell carcinoma: A clear target for synthetic lethal

strategies. *Biochim. Biophys. Acta - Rev. Cancer* **1870**, 176–184 (2018).

95. Wang, Z. *et al.* Dual ARID1A/ARID1B loss leads to rapid carcinogenesis and disruptive redistribution of BAF complexes. *Nat. Cancer* doi:10.1038/s43018-020-00109-0.

96. Santen, G. W. E. *et al.* The ARID1B phenotype: What we have learned so far. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **166**, 276–289 (2014).

97. Moffat, J. J., Smith, A. L., Jung, E. M., Ka, M. & Kim, W. Y. NEUROBIOLOGY OF ARID1B HAPLOINSUFFICIENCY RELATED TO NEURODEVELOPMENTAL AND PSYCHIATRIC DISORDERS. *Mol. Psychiatry* **27**, 476 (2022).

98. Helming, K. C. *et al.* ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat. Med.* (2014) doi:10.1038/nm.3480.

99. Sasaki, M. & Ogiwara, H. Synthetic lethal therapy based on targeting the vulnerability of SWI/SNF chromatin remodeling complex-deficient cancers. *Cancer Sci.* **111**, 774–782 (2020).

100. Helming, K. C., Wang, X. & Roberts, C. W. M. Vulnerabilities of mutant SWI/SNF complexes in cancer. *Cancer Cell* **26**, 309–317 (2014).

101. Tokunaga, R. *et al.* The impact of ARID1A mutation on molecular characteristics in colorectal cancer. *Eur. J. Cancer* **140**, 119–129 (2020).

102. Ashizawa, M. *et al.* Prognostic role of ARID1A negative expression in gastric cancer. *Sci. Rep.* **9**, (2019).

103. Zafra, M. P. & Dow, L. E. Revealing ARID1A Function in Gastric Cancer from the Bottom Up. *Cancer Discov.* **11**, 1327–1329 (2021).

104. Wang, S. *et al.* SWI/SNF component ARID1A restrains pancreatic neoplasia formation. *Gut* **68**, 1259–1270 (2019).

105. Zhao, S. *et al.* ARID1A Variations in Cholangiocarcinoma: Clinical Significances and Molecular Mechanisms. *Front. Oncol.* **11**, (2021).

106. Guo, B. *et al.* Arid1a mutation suppresses TGF-β signaling and induces cholangiocarcinoma. *Cell Rep.* **40**, (2022).

107. Shi, H. *et al.* ARID1A loss in neuroblastoma promotes the adrenergic-to-mesenchymal transition by regulating enhancer-mediated gene expression. *Sci. Adv.* **6**, (2020).

108. Xiao, Y. *et al.* Loss of ARID1A Promotes Hepatocellular Carcinoma Progression via Up-regulation of MYC Transcription. *J. Clin. Transl. Hepatol.* **9**, 528 (2021).

109. Fang, J.-Z. *et al.* Hepatocyte-Specific Arid1a Deficiency Initiates Mouse Steatohepatitis and Hepatocellular Carcinoma. (2015) doi:10.1371/journal.pone.0143042.

110. Wiegand, K. C. *et al.* ARID1A Mutations in Endometriosis-Associated Ovarian Carcinomas. *N. Engl. J. Med.* **363**, 1532 (2010).

111. Chandler, R. L. *et al.* ARID1a-DNA Interactions Are Required for Promoter Occupancy by SWI/SNF. *Mol. Cell. Biol.* **33**, 265 (2013).

112. Guan, B. *et al.* Roles of Deletion of Arid1a, a Tumor Suppressor, in Mouse Ovarian Tumorigenesis. *JNCI J. Natl. Cancer Inst.* **106**, (2014).

113. Chandler, R. L. *et al.* Coexistent ARID1A-PIK3CA mutations promote ovarian clear-cell tumorigenesis through pro-tumorigenic inflammatory cytokine signaling. *Nat. Commun.* **6**, 6118 (2015).

114. Guan, B., Gao, M., Wu, C. H., Wang, T. L. & Shih, I. M. Functional Analysis of In-frame Indel ARID1A Mutations Reveals New Regulatory Mechanisms of Its Tumor Suppressor Functions. *Neoplasia* **14**, 986 (2012).

115. Amaddeo, G., Guichard, C., Imbeaud, S. & Zucman-Rossi, J. Next-generation sequencing identified new oncogenes and tumor suppressor genes in human hepatic tumors. *Oncoimmunology* **1**, 1612 (2012).

116. Fang, J. Z. *et al.* Hepatocyte-Specific Arid1a Deficiency Initiates Mouse Steatohepatitis and Hepatocellular Carcinoma. *PLoS One* **10**, e0143042 (2015).

117. He, F. *et al.* Decreased expression of ARID1A associates with poor prognosis and promotes metastases of hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* **34**, (2015).

118. Chaisaingmongkol, J. *et al.* Common Molecular Subtypes Among Asian Hepatocellular Carcinoma and Cholangiocarcinoma. *Cancer Cell* **32**, 57-70.e3 (2017).

119. Jiao, Y. *et al.* Exome sequencing identifies frequent inactivating mutations in BAP1, ARID1A and PBRM1 in intrahepatic cholangiocarcinomas. *Nat. Genet. 2013 4512* **45**, 1470–1473 (2013).

120. Zhu, M. *et al.* Somatic mutations increase hepatic clonal fitness and regeneration in chronic liver disease. *Cell* **177**, 608 (2019).

121. Chang, L. *et al.* The SWI/SNF complex is a mechanoregulated inhibitor of YAP and TAZ. (2018) doi:10.1038/s41586-018-0658-1.

122. Schuler, M., Dierich, A., Chambon, P. & Metzger, D. Efficient temporally controlled targeted somatic mutagenesis in hepatocytes of the mouse. *Genesis* **39**, 167–172 (2004).

123. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

124. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. (2013).

125. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinforma. Appl. NOTE* **25**, 2078–2079 (2009).

126. Mckenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. doi:10.1101/gr.107524.110.

127. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* (2013) doi:10.1038/nbt.2514.

128. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster

strain w1118; iso-2; iso-3. *Fly (Austin).* **6**, 80–92 (2012).

129. Kuilman, T. *et al.* CopywriteR: DNA copy number detection from off-target sequence data. *Genome Biol.* **16**, 1–15 (2015).

130. van de Wie, M. A. *et al.* CGHcall: calling aberrations for array CGH tumor profiles. *Bioinformatics* **23**, 892–894 (2007).

131. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. doi:10.1093/bioinformatics/bty560.

132. Dobin, A. *et al.* Sequence analysis STAR: ultrafast universal RNA-seq aligner. **29**, 15–21 (2013).

133. Faust, G. G. & Hall, I. M. SAMBLASTER: fast duplicate marking and structural variant read extraction. **30**, 2503–2505 (2014).

134. Liao, Y., Smyth, G. K. & Shi, W. Sequence analysis featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. **30**, 923–930 (2014).

135. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

136. Zhu, A., Ibrahim, J. G., Love, M. I. & Stegle, O. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. doi:10.1093/bioinformatics/bty895.

137. Ignatiadis, N., Klaus, B., Zaugg, J. B. & Huber, W. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nat. Methods* **13**, 577–580 (2016).

138. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov.* **2**, 100141 (2021).

139. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, 1–10 (2009).

140. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). (2008) doi:10.1186/gb-2008-9-9-r137.

141. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. doi:10.1093/bioinformatics/btv145.

142. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Web Serv. issue Publ. online* **44**, (2016).

143. Orlando, D. A. *et al.* Quantitative ChIP-Seq normalization reveals global modulation of the epigenome. *Cell Rep.* **9**, 1163–1170 (2014).

144. Hiramatsu, Y., Designed Research, A. F., Hiramatsu, ; Y & Hanyu, Y. Arid1a is essential for intestinal stem cells through Sox9 regulation. **116**, 1704–1713 (2019).

145. Mathur, R. *et al.* ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat. Publ. Gr.* **49**, (2016).

146. Xu, G. *et al.* ARID1A determines luminal identity and therapeutic response in

estrogen-receptor-positive breast cancer. *Nat. Genet. 2020 522* **52**, 198–207 (2020).

147. Ding, S. *et al.* STAG2 deficiency induces interferon responses via cGAS-STING pathway and restricts virus infection. (1234) doi:10.1038/s41467-018-03782-z.

148. Wu, X. *et al.* Intrinsic Immunity Shapes Viral Resistance of Stem Cells HHS Public Access. *Cell* **172**, 423–438 (2018).

149. Brzostek-Racine, S., Gordon, C., Van Scoy, S. & Reich, N. C. The DNA damage response induces IFN. *J. Immunol.* **187**, 5336–5345 (2011).

150. Wang, Z. *et al.* Dual ARID1A/ARID1B loss leads to rapid carcinogenesis and disruptive redistribution of BAF complexes HHS Public Access. doi:10.1038/s43018-020-00109-0.

151. Li, W. *et al.* A Homeostatic Arid1a-Dependent Permissive Chromatin State Licenses Hepatocyte Responsiveness to Liver-Injury-Associated YAP Signaling. *Cell Stem Cell* **25**, 54-68.e5 (2019).

152. de Galarreta, M. R. *et al.* β-catenin activation promotes immune escape and resistance to anti–PD-1 therapy in hepatocellular carcinoma. *Cancer Discov.* **9**, 1124–1141 (2019).

153. Wu, S. Y. *et al.* MYC suppresses STING-dependent innate immunity by transcriptionally upregulating DNMT1 in triple-negative breast cancer. *J. Immunother. cancer* **9**, (2021).