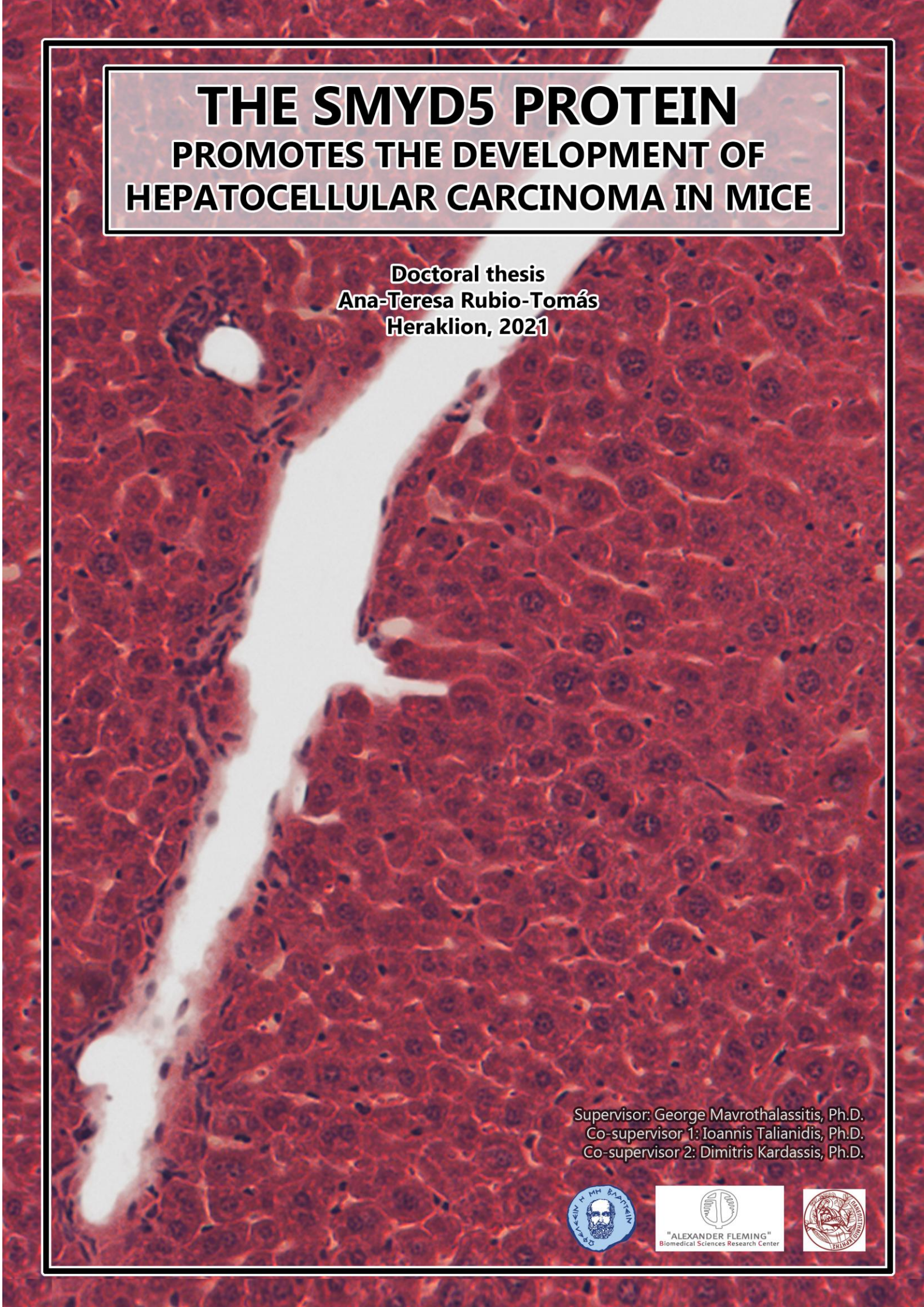
THE SMYD5 PROTEIN PROMOTES THE DEVELOPMENT OF HEPATOCELLULAR CARCINOMA IN MICE

Doctoral thesis
Ana-Teresa Rubio-Tomás
Heraklion, 2021



Supervisor: George Mavrothalassitis, Ph.D.
Co-supervisor 1: Ioannis Talianidis, Ph.D.
Co-supervisor 2: Dimitris Kardassis, Ph.D.

"ALEXANDER FLEMING"
Biomedical Sciences Research Center

Ο ΡΟΛΟΣ ΤΗΣ ΠΡΩΤΕΪΝΗΣ SMYD5 ΣΤΗΝ ΚΥΤΤΑΡΙΚΗ ΔΙΑΦΟΡΟΠΟΙΗΣΗ ΚΑΙ ΚΑΘΟΡΙΣΜΟ ΣΤΟ ΗΠΑΡ ΚΑΙ ΤΟ ΕΝΤΕΡΟ

THE SMYD5 PROTEIN PROMOTES THE DEVELOPMENT OF HEPATOCELLULAR CARCINOMA IN MICE

Acknowledgments

I would like to acknowledge Professor Ioannis Talianidis for giving me the opportunity to do a Ph.D. under his supervision. I am very grateful for his mentoring and the quality of the research performed in his laboratory. I am also very grateful to Professor Dimitris Kardassis and Professor George Mavrothalassitis for their great job as members of my thesis committee. I also want to thank other senior scientists who guided me during my Ph.D.; Mihalis Verykokakis, Pantelis Hatzis, Antonis Giakountis and Panagiotis Moulos, as well as all the members of Talianidis' research group for helping me with my project; Anastasia Antonoglou, Stella Balliou, Sofia Gargani, Anna Haroniti, Kostas Nikolaou, Michalis Sarris, and, especially, Evaggelia Papageorgopoulou for her technical and emotional support at the end of my Ph.D., Dimitra Saviolaki for taking care of me as if I was her daughter and, last but not least, Panagiota Karagianni for being a wonderful scientist and a better friend. I will always be indebted to all the members of Professor Talianidis' research group for at least 10 hours of fun per day. We managed to laugh at absolutely everything.

I also want to thank everybody else I met in the Biomedical Sciences Research Center "Alexander Fleming" for their advice regarding the experiments, for their advice regarding my career, for their advice regarding life, for all the time spent inside and outside that building and for making me feel at home.

And, of course, thanks go to all my friends in Athens (including, obviously, Vicky Navarro) and other parts of Greece and to all of those who went with me for coffee, for a drink, to the beach or on an excursion, saving me from loneliness and helping me discover such a wonderful country.

This amazing personal and professional experience would not have been possible without Professor Kriton Kalantidis, who first accepted me in his laboratory without knowing me at all, and, afterwards, suggested that I apply for a position in the research group of Professor Ioannis Talianidis.

I will always be grateful for having had the opportunity of being part of the Marie Skłodowska-Curie Nuclear Receptor-Network, where I met very nice people and impressive scientists. We were always looking forward to our meetings.

As an accidental and unexpected part of my Ph.D., I want to thank Dr. Pau Sancho-Bru and all the members of the "Liver cell plasticity and tissue repair" laboratory for their

help with the article and their support during a difficult period of my life. I would also like to mention Anabel Álvarez and thank her for her understanding and for being my second, unofficial work team.

I want to thank all my friends in Barcelona, all around Spain and all around the world, who visited me, called me, supported me and never forgot about me during all these years abroad.

I would like to dedicate some words to Carlos; thank you for designing such a beautiful front cover with liver histology images.

It is impossible not to thank a person who might know this project better than me, because he has listened to all the stories about my Ph.D. in Athens and in Barcelona, and understood why it was so important for me. Thank you so much, Tuán.

Finally, I acknowledge my family for encouraging me to do everything I have achieved in my life, and I would like to dedicate this Ph.D. to them.

I would like to say sorry to everybody who was a victim of my high levels of stress throughout the tough process of becoming a doctor. It was not on purpose! Remember that, salt-peppered-haired guy.

I apologize if I have forgotten anybody. This Ph.D. was such an important experience in my professional and personal life that it is impossible to mention everybody that was part of it, but, for sure, all these years will always be part of me, as they made me the person who I am now.

And I will always remember that “εύκολη είναι μόνο η δουλειά του άλλου”.

This work was supported by the European Union, ERC Advanced Investigator Grant (ERC-2011-AdG294464) and the “Nuclear Receptor-Network” Consortium Grant PITN-GA-2013-606806 from the European Union Seventh Framework Programme Marie Skłodowska-Curie Initial Training Network and as part of the PEOPLE-2013 program.

This doctoral thesis has been defended at the School of Medicine, University of Crete.

Index

Tables

Figures

Περίληψη

ΙΣΤΟΡΙΚΟ: Η πρωτεΐνη SMYD5 είναι μια μεθυλοτρανσφεράση των ιστονών που περιέχει τις δομικές περιοχές Su(Var)3-9, Enhancer-of-zeste και Trithorax (SET) και Myeloid, Nervy, και DEAF-1 (MYND) και η οποία αρχικά περιγράφηκε ως συστατικό του συμπλόκου πυρηνικού υποδοχέα (NCoR) που καταλύει την τριμεθυλίωση της λυσίνης 20 της ιστόνης 4 (H4K20me3).

ΛΟΓΙΚΗ: Με βάση τη δομή και τα στοιχεία που έχουν δημοσιευτεί στο παρελθόν που δείχνουν το ρόλο της πρωτεΐνης SMYD5 ως μεθυλοτρανσφεράση, επιδιώξαμε να εντοπίσουμε εν δυνάμει υποστρώματα μεθυλίωσης του μορίου SMYD5 *in vitro*. Επιπλέον, δεδομένου ότι άλλα μέλη της οικογένειας SMYD, όπως η SMYD2 και η SMYD3 πρωτεΐνες, έχει δειχθεί ότι παίζουν σημαντικό ρόλο στην ανάπτυξη του καρκίνου, συμπεριλαμβανομένου του καρκίνου του ήπατος και του παχέος εντέρου, διερευνήσαμε την πιθανή επίπτωση του SMYD5 στην καρκινογένεση χρησιμοποιώντας μοντέλα ποντικών και ανθρώπινα δείγματα.

ΠΡΟΣΕΓΓΙΣΗ: Πραγματοποιήσαμε μια δοκιμασία *in vitro* μεθυλίωσης χρησιμοποιώντας την ανασυνδυασμένη ανθρώπινη πρωτεΐνη SMYD5 με υποστρώματα ιστονών και μη ιστονικά υποστρώματα. Επιπλέον, δημιουργήσαμε ανασυνδυασμένους ποντικούς με έλλειψη του *Smyd5* γονιδίου, στους οποίους το γονίδιο *Smyd5* απενεργοποιήθηκε είτε σε όλους τους ιστούς (*Smyd5*^{flox}/MX-Cre) είτε συγκεκριμένα στο ήπαρ (*Smyd5*^{flox}/Alfp-Cre and *Smyd5*^{flox}/Alb-Cre). Επιπλέον, δημιουργήσαμε ποντίκια με έλλειψη *Smyd5* ειδικά για το έντερο (*Smyd5*^{flox}/Villin-Cre). Η ανάλυση της έκφρασης του γονιδίου *SMYD5* σε ανθρώπινο ηπατοκυτταρικό καρκίνωμα (HCC) και καρκίνο του παχέος εντέρου πραγματοποιήθηκε χρησιμοποιώντας τα σύνολα δεδομένων The Human Cancer Genome Atlas (TCGA).

ΑΠΟΤΕΛΕΣΜΑΤΑ: Στην παρούσα μελέτη παρουσιάζουμε ότι η υπερεκφρασμένη με βακιλοϊό και καθαρισμένη SMYD5 πρωτεΐνη μεθυλιώνει, κατά προτίμηση, τις ιστόνες 2A και 2B (H2A και H2B) *in vitro*.
Και τα τέσσερα στελέχη SMYD5 knockout (KO) ποντικών που δημιουργήθηκαν αναπτύχθηκαν κανονικά και δεν εμφάνισαν ορατό φαινότυπο. Ωστόσο, όταν υποβλήθηκε σε πρωτόκολλο χημικώς επαγόμενης καρκινογένεσης diethylnitrosamine (DEN), τα ποντίκια με πλήρη έλλειψη του γονιδίου *Smyd5* από όλους τους ιστούς καθώς και μεμονωμένα στο ήπαρ, παρουσίασαν καθυστερημένη ανάπτυξη όγκου του ήπατος σε σύγκριση με ποντίκια άγριου τύπου (WT) που εκτέθηκαν στην ίδια

θεραπεία. Τα ποντίκια υποβλήθηκαν σε μακροσκοπική και ιστολογική ανάλυση, συμπεριλαμβανομένων των κηλίδων για την αξιολόγηση της ηπατικής βλάβης του DNA και του πολλαπλασιασμού, καθώς και σε ανάλυση του προτύπου έκφρασης του μεταγραφώματος με αλληλούχιση RNA (RNA-seq) και με ποσοτική αλυσιδωτή αντίδραση πολυμεράσης αντίστροφης μεταγραφάσης (qPCR). Ορισμένα γονίδια που σχετίζονται με τον καρκίνο και τον πολλαπλασιασμό έδειξαν μειωμένη έκφραση στο ήπαρ ποντικών με έλλειψη του *Smyd5* γονιδίου σε σύγκριση με ποντίκια WT.

Ανάλυση ποντικών *Smyd5*^{flox}/Villin-Cre που υποβλήθηκαν σε χημική θεραπεία προκαλώντας καρκινογένεση του παχέος εντέρου δεν έδειξε εμφανείς φαινοτυπικές διαφορές σε σύγκριση με τα ποντίκια WT που έλαβαν την ίδια θεραπεία.

Σύμφωνα με τα δεδομένα από την ανάλυση των ανασυνδυασμένων ποντικών, βρήκαμε μια σημαντική συσχέτιση μεταξύ της έκφρασης του *SMYD5* και της επίπτωσης στον ηπατοκυτταρικό καρκίνο. Η έκφραση του *SMYD5* γονιδίου σε δείγματα από όγκο ανθρώπινου ήπατος σχετίζεται αρνητικά με τη συνολική επιβίωση του ασθενούς και την πιθανότητα επιβίωσης “χωρίς όγκο” μετά από χημειοθεραπεία για ένα δεδομένο χρονικό πλαίσιο, αλλά σχετίζεται θετικά με την πιθανότητα προόδου σε όγκους υψηλού βαθμού (G3/G4). Επιπλέον, βάσει των πειραμάτων ποντικών και των δειγμάτων καρκίνου του παχέος εντέρου από την TCGA, καταλήξαμε στο συμπέρασμα ότι το γονίδιο *SMYD5* δεν φαίνεται να παίζει ρόλο στον καρκίνο του παχέος εντέρου, χρησιμοποιώντας τουλάχιστον τις συγκεκριμένες παραμέτρους της ανάλυσής μας.

ΣΥΜΠΕΡΑΣΜΑΤΑ: Στην παρούσα διδακτορική διατριβή διαπιστώσαμε ότι η πρωτεΐνη SMYD5 παρουσιάζει δραστικότητα μεθυλοτρανσφεράσης των ιστονών *in vitro*, με τα *in vivo* δεδομένα από μοντέλα SMYD5 KO ποντικών και δείγματα από καρκίνο στον άνθρωπο να υποστηρίζουν το ρόλο της και τη σημασία της στην καρκινογένεση του ήπατος.

Abstract

BACKGROUND: SMYD5 is a Su(Var)3-9, Enhancer-of-zeste and Trithorax (SET) and Myeloid, Nervy, and DEAF-1 (MYND) domain-containing histone methyltransferase, which was originally described as a component of the nuclear receptor corepressor (NCoR) complex catalyzing histone 4 lysine 20 trimethylation (H4K20me3).

RATIONALE: Based on its domain composition and on previously published data indicating its role as a methyltransferase, we sought to identify putative methylation substrates of SMYD5 *in vitro.* In addition, given that other members of the SMYD family, such as SMYD2 and SMYD3, have shown to play important roles in cancer promotion, including liver and colon cancer, we investigated the potential implication of SMYD5 in carcinogenesis using murine models and human samples.

APPROACH: We performed an *in vitro* methylation assay using the recombinant human SMYD5 protein with histone and non-histone substrates. In addition, we generated *Smyd5*-deficient mice, in which the *Smyd5* gene was inactivated either in all tissues (*Smyd5*^{flox}/MX-Cre) or specifically in the liver (*Smyd5*^{flox}/Alfp-Cre and *Smyd5*^{flox}/Alb-Cre). Moreover, we produced intestine-specific *Smyd5*-deficient mice (*Smyd5*^{flox}/Villin-Cre). Analysis of *SMYD5* expression in human hepatocellular carcinoma (HCC) and colon cancer was conducted using The Human Cancer Genome Atlas (TCGA) datasets.

RESULTS: We show that purified, baculovirus-expressed SMYD5 preferentially methylates histones 2A and 2B (H2A and H2B) *in vitro*.
All four SMYD5 mouse knockout (KO) strains generated developed normally and did not display any visible phenotype. However, when subjected to a protocol of chemically induced carcinogenesis by diethylnitrosamine (DEN) injection, whole-tissue and liver *Smyd5*-deficient mice exhibited delayed liver tumor growth compared to wild-type (WT) mice exposed to the same treatment. The mice were subjected to macroscopic and histological analysis, including stainings to evaluate hepatic DNA damage and proliferation, as well as expression profiling by RNA-sequencing (RNA-seq) and by quantitative polymerase chain reaction (qPCR). A number of cancer- and proliferation-related genes showed reduced expression in the liver of *Smyd5*-deficient mice compared to WT mice.

Analysis of *Smyd5*^{flox}/Villin-Cre mice subjected to chemical treatment inducing colon carcinogenesis did not indicate any obvious phenotypic differences compared to WT mice receiving the same treatment.

Consistent with the mice data, we found a significant correlation between *SMYD5* expression and the incidence of hepatocellular cancer. *SMYD5* expression in human liver tumor samples negatively correlated with overall patient survival and the probability of "tumor-free" survival after chemotherapy for a given time frame, but positively correlated with the probability of progressing to high-grade (G3/G4) tumors. Furthermore, based on mouse experiments and human colon cancer samples from TCGA, we concluded that SMYD5 does not seem to play a role in colon cancer, at least using the specific parameters of our analysis.

CONCLUSIONS: We found that SMYD5 presents histone methyltransferase activity *in vitro,* with *in vivo* data from SMYD5 KO mouse models and human cancer samples supporting its role in liver carcinogenesis.

Introduction

Liver and colon physiology

Both the liver and the colon, the carcinogenesis of which is studied in this thesis, are part of the **gastrointestinal system**. The gastrointestinal system consists of the gastrointestinal tract, a tube including the oral cavity, pharynx, esophagus, stomach, small intestine, large intestine (which comprises the cecum, colon, and rectum), and anus, as well as accessory digestive organs, including the salivary glands, liver, gallbladder and pancreas, that are not part of the tract but secrete substances into it via connecting ducts. The main function of the gastrointestinal system is the processing and conversion of ingested food into molecules that can be transferred along with salts and water to the body's internal environment and delivered to the cells through the circulatory system. More specifically, the gastrointestinal tract performs four important processes; 1) **digestion** of macromolecules, 2) **secretion** of digestive enzymes, 3) **absorption**, meaning the transport of molecules produced by digestion from the lumen of the gastrointestinal tract across a layer of epithelial cells to enter the blood or lymph, and, 4) **motility**, which refers to the contractions of smooth muscles in the gastrointestinal tract wall that mix the luminal contents with the secretions and move the contents through the tract from the mouth to the anus (1) (**Figure 1**).

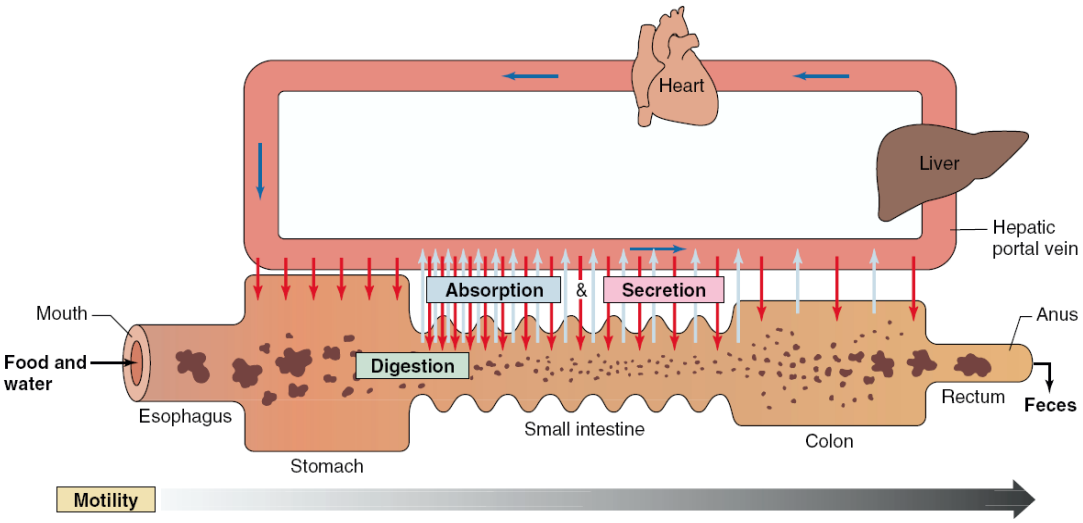


Figure 1. The functions of the gastrointestinal tract. Extracted from (1)

The liver

The liver is the largest gland of the body and it is located in the right upper quadrant of the human abdomen. The mass of a healthy human liver is approximately 1.5 kg (around 2% body weight), while murine liver typically weights 2–3 g (3-5% body weight) and occupies a larger proportion of abdominal cavity (2–4).

The liver has a variety of functions, that are similar in humans and mice. The liver **exocrine** functions are related to digestion and involve the synthesis and secretion of bile salts, which are necessary for digestion and absorption of lipids, as well as the secretion of a bicarbonate-rich solution into the bile, which assists in the neutralization of acids in the initial segment of the small intestine, called the duodenum. The liver also has many **endocrine** functions, including secretion of insulin-like growth factor I (IGF-I) in response to growth hormone, participation in the activation of vitamin D, formation of triiodothyronine (T3) from thyroxine (T4), secretion of angiotensinogen (the substrate used by renin to form angiotensin I), metabolization of hormones, and secretion of cytokines, which are involved in immune response. Moreover, the liver plays a role in blood **clotting** because it produces numerous plasma clotting factors, such as prothrombin and fibrinogen, and also because bile salts are essential for the gastrointestinal absorption of vitamin K, which is necessary for the synthesis of the clotting factors. Additionally, many **plasma proteins**, such as plasma albumin, acute phase proteins, binding proteins for various hormones and trace elements, and lipoproteins, are synthesized and secreted by the liver. Likewise, the liver plays a key role in **organic metabolism**. The liver converts plasma glucose into glycogen and triacylglycerols during the absorptive period as well as plasma amino acids to fatty acids, which can be incorporated into triacylglycerols during the absorptive period. Therefore, synthesis of triacylglycerols is also performed by the liver, which then secretes them as lipoproteins during the absorptive period. In addition, the liver produces glucose from glycogen (glycogenolysis) and other sources (gluconeogenesis) during the postabsorptive period and releases the glucose into the blood. This organ also converts fatty acids into ketones during fasting, and produces urea, the major end product of amino acid catabolism, and releases it into the blood. The liver participates in **cholesterol metabolism** by synthesizing cholesterol and releasing it into the blood, as well as secreting plasma cholesterol into the bile and converting plasma cholesterol into bile salts. Finally, the **excretory and degradative** functions of the liver include secretion of bilirubin and other bile pigments into the bile, excretion of many endogenous and exogenous organic molecules as well as trace metals via the bile, biotransformation of many endogenous and exogenous organic molecules and degradation of old erythrocytes (1).

The liver is formed by four **lobes**, called the right, left, caudate and quadrate in humans, and the right, median, left and caudate in rodents (2). Each liver lobe is divided into thousands of small units known as **lobules**, which constitute the functional units. Each lobule is roughly hexagonal in shape, with a **central vein** located in the center, from which **hepatocyte** cords radiate towards portal triads. The **portal triads** are situated at the vertices of each hexagon and consist of a portal vein, hepatic artery and biliary duct. Hepatocyte cords are single-cell sheets of hepatocytes separated by **sinusoids** (discontinuous vessels built from specialized fenestrated endothelial cells of the liver) that carry blood from the portal triads to the central vein (5). The structure of the lobules is shown in **Figure 2B**.

Liver blood supply ocurrs as depicted in **Figure 2A**. The blood carried by the **portal vein**, which comes from the intestine, spleen and pancreas, and the blood carried by the **hepatic artery** coming from the abdominal aorta, flow into the liver, while the bile carried by the **bile duct** flows from the liver to the intestine. These vessels are divided into **capillaries** ending up in the lobules. The portal vein supplies 75% of the total liver supply and carries all the nutrients and digestion products from the intestine that will be metabolized or detoxified in the liver. The hepatic artery supplies the remaining 25% of the blood supply and carries oxygen-rich arterial blood. Oxygen is also supplied by the portal vein, but to a lesser extent. Blood flows though the sinusoids until the central vein of each lobule and all the central vein ramifications are joined into hepatic veins that transport blood out of the liver. Moreover, the liver produces bile, which is collected in the **bile canaliculi** that merge and form the bile duct. The bile is either secreted into the duodenum via the bile duct, or temporarily stored in the gallbladder (1) (**Figure 2**).

With respect to its cellular composition, the liver consists of 60-80% of **hepatocytes** and 20-40% of non-hepatic populations. Non-hepatic populations include biliary epithelium cells or **cholangiocytes** (~3%), **liver sinusoidal endothelial cells** (~50%), resident macrophages called **Kupffer cells** (~20%), **hepatic stellate cells** (~8%) and **resident lymphocytes** (~25%) such as T cells, B cells, natural killer (NK) cells and natural killer T (NKT) cells (6–8). The space of Disse, located between hepatocytes and liver sinusoidal endothelial cells, has traditionally been pointed out as the source of liver stem cells. However, plasticity of preexisting hepatocytes and cholangiocytes can also contribute to liver regeneration without the involvement of the stem cell compartment (9) (**Figure 3**). Hepatocytes and biliary epithelium cells share a common progenitor, called hepatoblast, and are derived from the embryonic endoderm, while liver sinusoidal endothelial cells, Kupffer cells and hepatic stellate cells are of mesodermal origin (10).

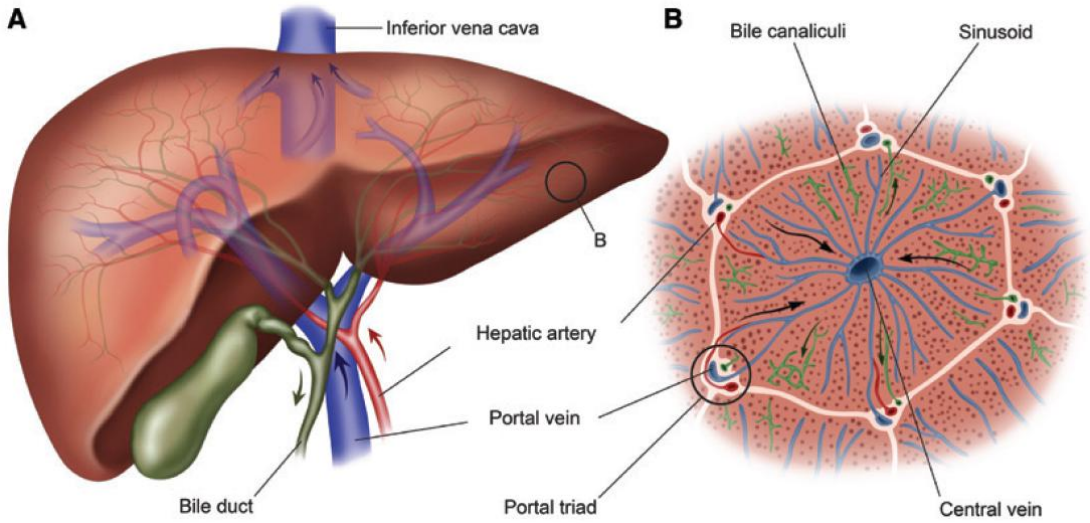


Figure 2. Liver anatomy (**A**) Entire organ and blood supply. Blue indicates venous blood, red indicates arterial blood, green indicates bile, and arrows indicate flow direction. (**B**) Liver lobule showing rows of hepatocytes radiating out from the central vein towards the portal triad. Extracted from (11)

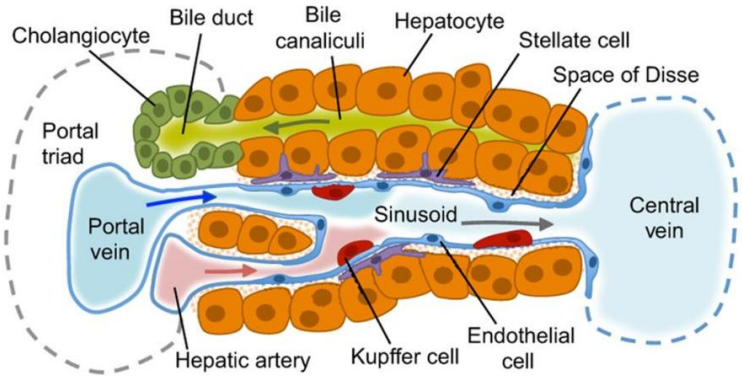


Figure 3. Liver structure and cell types. Extracted from (5)

The colon

The colon is the portion of the large intestine extending from the cecum to the rectum and is derived from the embryonic endoderm. The average human colon size is 1-1.5 m in length and 5 cm in diameter and it consists of the ascending colon, transverse colon, descending colon and sigmoid colon. The main function of the large intestine is the **storage and concentration of fecal material** before defecation. The large intestine lacks digestive enzymes and its secretions are limited to mucus and fluid containing bicarbonate and potassium ions. Absorption is also scarce in the large intestine compared to the small intestine. Products absorbed from the lumen to blood mainly include sodium by active transport and accompanying water by osmosis, and some of the products generated by the intestinal bacteria, such as short-chain fatty acids resulting from undigested polysaccharide (fiber) metabolism and low quantities of vitamins, especially vitamin K (1,12).

Histologically, the colon has essentially the same structure as the whole gastrointestinal tract, with few peculiarities. The colon is formed by the following four layers or tunics, named from the inner part of the tube (lumen) to the outer part: 1) the **mucosa**, a mucous membrane consisting of a lining epithelium, a connective tissue layer called lamina propria and a smooth muscle layer called muscularis mucosa, 2) the **submucosa**, a layer of fibroelastic connective tissue irrigated by blood and lymphatic vessels, and nerves, 3) the **muscularis externa**, formed by two thick layers of smooth muscle, called inner circular muscle and longitudinal muscle, respectively, and 4) the **adventitia or serosa**, a dense irregular connective tissue, named according to its structure, which varies in each part of the colon. It is separated from the muscularis externa by a thin layer of connective tissue called **subserosa**. In contrast to the small intestine, the large intestine does not have finger-like protrusions (villae) in the mucosa epithelium, but only invaginations (crypts) (13) (**Figure 4**).

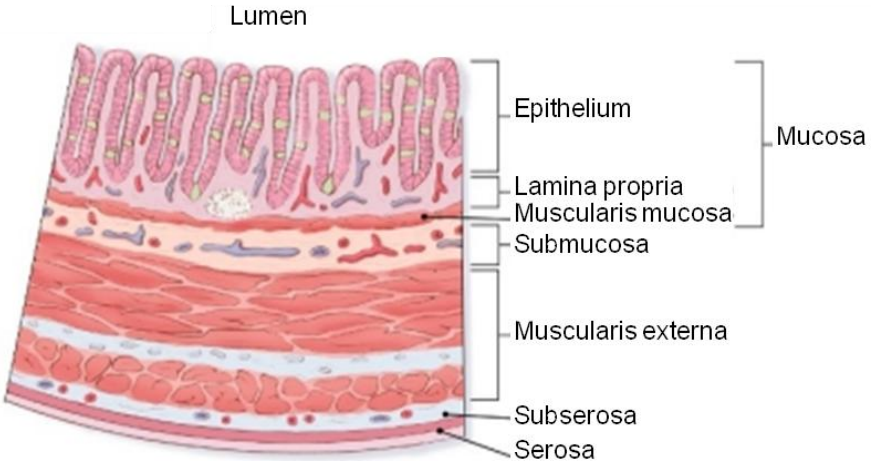


Figure 4. Representation of a longitudinal section of normal colon, with each layer labelled. Modified from (14)

The mucosa epithelium is a simple columnar epithelium populated by six cell types; 1) **enterocytes**, the most abundant cell type, involved in nutrient absorption, 2) **goblet cells,** producing mucins that form mucus, 3) **enteroendocrine cells**, which synthesize a wide range of hormones, 4) **tuft cells**, secreting prostanoids and opiods, 5) Microfold cells (**M cells**) located over lymphoid nodules and involved in antigen sampling, and 6) **Paneth cells** situated at the crypt base, releasing antimicrobial substances such as lysozyme and defensin but also supporting intestinal stem cell population renewal. **Intestinal stem cells** reside at the crypt base, adjacent to Paneth cells, and are positive for leucine-rich-repeat-containing G-protein-coupled receptor 5 (LGR5). Since intestine differentiated cell types are short lived, LGR5-positive intestinal stem cells continuously generate proliferating **transient amplifying cells** that commit to an intestinal cell lineage while moving upwards thought the crypt. The **+4 label retaining**

cells are believed to be able to restore the Lgr5+ stem cell compartment upon injury (13,15,16) (**Figure 5**).

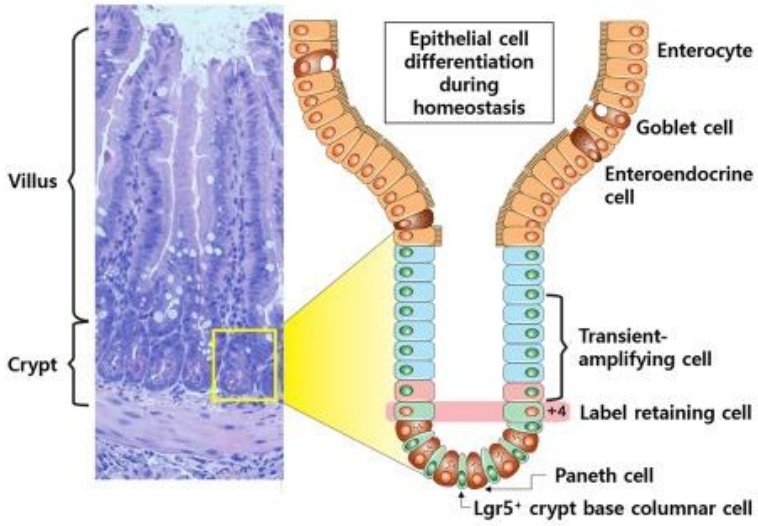


Figure 5. Representation of an intestinal crypt. Despite the absence of villae, the colon crypt organization is essentially the same as in the small intestine. Extracted from (16)

Carcinogenesis

The term **cancer**, also referred to as malignant neoplasm or malignancy, is defined by the National Institute of Cancer (NIH) as a group of diseases in which abnormal cells divide without control and can invade nearby tissues. Cancer can also spread to other parts of the body through the blood and lymph systems, thereby givising rise to metastases. In contrast, benign (nonmalignant) tumors may grow larger but do not spread to other parts of the body. Of note, cancer is a heterogeneous group of diseases, with different clinical manifestations and prognosis, but also different genetic and epigenetic aberrations (17,18).

Nomenclature of cancer types

Types of cancer are generally named for the organs, tissues or cell types from which the original (primary) tumors form, in line with the cancer stem cell hypothesis, which describes cancer stem cells as a subset of tumoral cells with stem cell features (See "The relevance of cancer stem cells and tumor microenvironment" section). Metastatic tumors are considered to be the same type of cancer as primary tumors. **Carcinoma** is the most common type of cancer and is formed by epithelial cells. Adenocarcinoma is the specific name for a carcinoma arising from epithelial cells that produce fluids or mucus. Other relevant categories are the following: 1) sarcoma, which develops in bone or in the soft tissues of the body, including cartilage, muscle, fat, blood or lymph vessels, fibrous tissue (tendons and ligaments), or other connective or supportive tissue, 2) hematopoietic cancer, including leukemia, lymphoma and multiple myeloma, deriving from immune system cells, 3) melanoma, beginning in melanocytes usually located in the skin, 4) brain and spinal cord tumors, which are named according to the cell of origin, 5) germ cell tumors, that appear in the testes or ovaries, 6) neuroendocrine tumors, which arise from cells capable of releasing hormones into the blood, and 7) blastoma, coming from immature precursor cells or embryonic tissues and is more common during childhood (17). The present study is focused on liver cancer, specifically **hepatocellular carcinoma (HCC)**, and, to a lesser extent, colon cancer.

Contrary to the traditional reductionist view of tumors as homogeneous masses of rapidly dividing cells, cancer is now considered to have a remarkable intratumoral heterogenicity. Indeed, each tumor is formed by different cell subpopulations that interact with each other and with normal stromal cells or other cell types surrounding the tumor. In fact, these interactions have a key role in the multistep process of carcinogenesis, involving **cell autonomous** as well as **non-cell autonomous**

mechanisms. Recruitment of normal stromal cells and also cellular plasticity of tumoral and normal cell types is involved in the development of tumors (18).

Two important concepts in cancer biology are oncogenes and tumor suppressor genes. **Oncogenes** are genes encoding for proteins that promote cancer development, while **tumor suppressor genes** encode negative regulators of tumorigenesis. However, not all the genetic or epigenetic alterations that cancer cells carry are the cause of tumor evolution. In this sense, the distinction between cancer drivers and passenger mutations is critical for deciphering cancer molecular pathology. A **cancer driver** is a molecular alteration that influences cancer development at any stage. Conversely, **passenger mutations** are genetic events not associated with tumor initiation or progression that occur randomly in the genome (19,20).

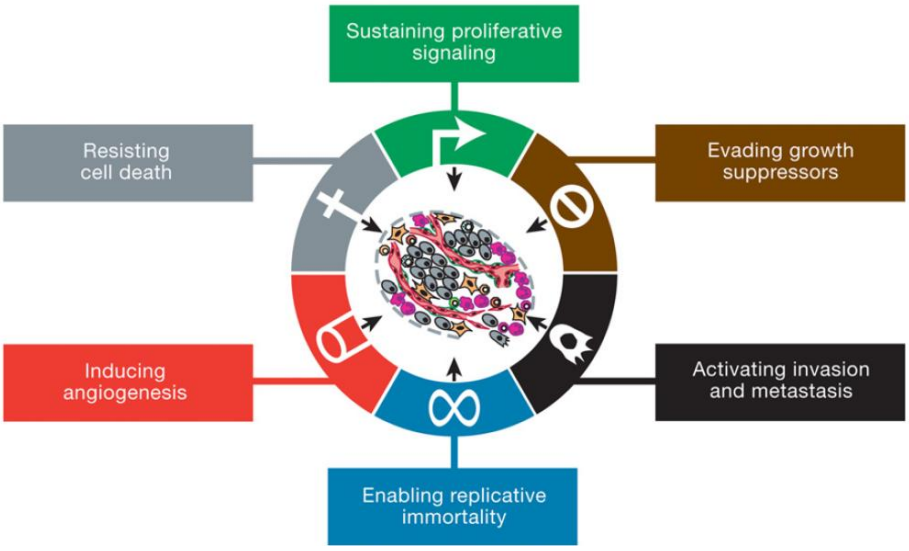


Figure 6. The six hallmark capabilities of cancer. Extracted from (18)

In 2000, Hanahan and Weinberg attempted to rationalize the complex biology of cancer by establishing **six hallmarks of cancer** (21). Hallmarks are defined as unique and complementary functional capabilities acquired at different points during the multistep development of human tumors that enable tumor growth and metastatic dissemination and colonization. Later in 2011 the same authors revisited and complemented the concept of hallmarks of cancer, in light of new advances in the field (18) (**Figure 6**) (each hallmark is explained below). Remarkably, when a hallmark is abolished by therapeutic targeting, cancer cells can rely on other hallmarks to survive. Importantly, some of the hallmarks of cancer may share common signaling pathways. This redundancy should be considered when designing new therapies (18).

Cancer cells sustain proliferative signaling

The most basic characteristic of cancer cells is their ability to induce and sustain chronic proliferation. Cancer cells manage to deregulate proliferative signals usually transmitted by **growth factors** that bind to cell surface receptors containing intracellular tyrosin kinase domains. Cancer cells achieve this by: 1) autocrine proliferative stimulation, in which cancer cells producte ligands for their own cell surface growth factor receptors, 2) stimulation of tumor-associated normal stroma cells to release growth factors, 3) overexpression of receptor proteins that make them hyperresponsive to even low amounts of growth factors, 4) ligand-independent firing of their receptors resulting from structural alterations in the receptor proteins or constitutive activation of components of the signaling pathways downstream of the receptors, such as the B-Raf protein, involved in the Ras/Raf/mitogen-activated protein kinase (MAPK) pathway, and phosphoinositide 3 (PI3) kinase, or 5) disruption of negative feedback loops. The latter can be due to the hyporesponsiveness of components of downstream pathways that attenuate proliferative signaling as a result of mutations compromising Ras GTPase activity, promoter methylation or loss of function mutations in phosphatase and tensin homolog (PTEN) phospatase or inhibition of mTOR kinase, among other alterations. Of note, excessive proliferative signaling, such as the Ras oncoprotein, can induce **cellular senescence** (a specific type of almost-permanent cell cycle arrest) and/or **apoptosis**. These mechanisms may represent an intrinsic cellular defense to eliminate cells hyperactivating growth pathways. In order to overcome this cellular defense, growth factor signaling in cancerous cells may be within a range to sustain proliferation without triggering senescence or apoptosis, or cancer cells may develop mechanisms to impair senescence or apoptosis program activation (18,22).

Cancer cells evade growth suppressors

In order to sustain proliferation, cancer cells must not only activate proliferation programs but also bypass growth suppression programs that mainly rely on the action of **tumor suppressor genes**. Tumor suppressors function in various ways to limit cell growth and proliferation. The most well-characterized tumor suppressor genes are those encoding retinoblastoma-associated (RB) and TP53. These two proteins sense and transduce extracellular and intracellular **growth inhibition signals**, respectively, although they most likely belong to a network of proteins with a certain degree of functional redundancy. TP53 can halt cell-cycle progression or trigger apoptosis in response to DNA damage, hypoxia and nucleotide deprivation. Lack of tumor suppressor activation enables cancer cells to proliferate even in suboptimal conditions

and therefore accumulate DNA errors and chromosomal aberrations. Genes encoding for proteins regulating growth suppression due to cell-to-cell contact (contact inhibition) in normal cells, such as Merlin (the product of the *NF2* gene) and LKB1 epithelial polarity protein, can also act as tumor suppressors and be lost in cancer cells. For example, Merlin can promote contact inhibition via coupling cell-surface adhesion molecules such as E-cadherin to transmembrane receptor tyrosine kinases such as epidermal growth factor receptor (EGFR). In addition, and despite its well-known anti-proliferative effects, transforming growth factor beta (TGF-β) can activate **epithelial-to-mesenchymal transition (EMT)** in cancer cells (see below). EMT can be activated by TGF-β in cooperation with oncogenic RAS signaling, among other players (18,23).

Cancer cells resist cell death

Apoptosis

As mentioned above, apoptosis is considered a natural barrier to tumorigenesis that activates in response to physiologic stresses, such as an increase in oncogen signaling or DNA damage resulting from excessive proliferation, or anti-cancer therapy. Cancer cells are able to circumvent this barrier and become resistant to cell death. The apoptosis program involves the action of many regulator and effector proteins. Classically, upstream regulators have been associated with either the extrinsic (mediated by receptors) or intrinsic apoptotic program (mediated by the mitochondria), which leads to the activation of caspase 8 and 9, respectively and then converges at the execution phase, performed mainly by the effector caspases 3, 6 and 7. Mitochondrial signaling in apoptosis ocurrs via the Bcl-2 family of regulatory proteins. Apoptosis is triggered when proapoptotic proteins (Bax and Bak) are released from apoptotic inhibitors (Bcl-2, Bcl-xL, Bcl-w, Mcl-1, A1). Tumor suppressor gene TP53 and hyperactivation of oncogenes such as Myc can regulate the apoptosis program to induce cell death. Cancer cells are capable of escaping apoptosis by losing TP53, overexpressing anti-apoptotic factors, downregulating proapoptotic proteins or interfering with the extrinsic ligand-induced apoptotic program (18).

Autophagy

Autophagy is an intracellular degradative process that also plays a role in cancer development. Similarly to apoptosis, autophagy can be activated in response to cell stress, especially nutrient deprivation. Autophagy is mediated by autophagosomes, which are double-membrane intracellular vesicles that envelope intracellular organelles and terminate in the lysosomal compartment. Autophagy has a dual role in cancer cell death, since it can promote cancer cell survival or trigger cell death. This is due to the

existence of common regulatory circuits controlling autophagy, apoptosis and cellular homeostasis, including PI3 kinase, protein kinase B (AKT), mTOR kinases and the apoptotic protein Beclin-1. Therefore, on one hand, autophagy can be a barrier to carcinogenesis, acting independently or in coordination with apoptosis, while on the other hand, activation of autophagy can lead to reversible dormancy in response to many stressors, including anti-cancer drugs, enabling the survival of this latent cells after treatment (24,25).

Necrosis

Although necrosis is expected to have a beneficial role as a barrier to cancer cell hyperproliferation, a certain degree of necrosis seems beneficial for tumor development. In contrast to apoptosis and autophagy, cells undergoing necrosis release their cellular content as well as pro-inflammatory molecules and proliferative factors, such as interleukin (IL)-1α, into the sourrounding microenvironment. Pro-inflammatory signals recruit immune inflammatory cells, which, in this case, have been shown to actively promote tumor growth (18).

Emerging types of cell death

Other recently described types of cell death may play a role in carcinogenesis (26). **Necroptosis** combines features of apoptosis and necrosis and is present in many solid tumors. This type of regulated cell death may promote carcinogenesis and metastasis, probably by inducing inflammatory responses (27,28). **Pyroptosis** is an alternative mode of inflammatory programmed cell death. Pyroptosis is mediated by inflammasomes that activate the secretion of cytokines such as IL-18 and IL-1β. It is generally considered to promote cancer development, since it is accompanied by inflammatory and immune responses, but it might have an anti-tumoral effect in specific contexts (29). **Ferroptosis** is an iron-dependent programmed cell death. It is characterized by the presence of reactive oxygen species (ROS) and the accumulation of lipid peroxides. In some types of cancer, p53 can trigger or suppress ferroptosis in a context dependent-manner (30). However, the general view is that cancer cells display aberrant ferroptosis. Many anti-tumoral therapies enhance ferroptosis, reinforcing the notion that activation of ferroptosis is detrimental for cancer cells (31). Of note, the inducers and downstream regulators of these emerging types of cell death may overlap with each other and with those regulating apoptosis, necrosis and autophagy (26).

Cancer cells enable replicative immortality

Cancer cells have overcome natural barriers to proliferation, which are cellular senescence and crisis, thereby exhibiting unlimited replicative potential. When a cell

population is subjected to many cycles of cell division, most of the cells enter nonreplicative senescence and the subpopulation that continues to proliferate enters a crisis phase, generally resulting in cell death. Nevertheless, some of these cells in crisis accummulate changes in their genome and epigenome that lead to **immortalization**. Many of these alterations are due to telomere shortening. **Telomeres** are repetitive sequences of non-coding DNA situated at the end part of the chromosomes that shorten each time the cell divides, provoking end-to-end fusions at the ends of chromosomes. This karyotypic instability jeopardizes cell viability. **Telomerase** is a DNA polymerase that adds telomere repeat segments to the ends of the chromosomes to avoid their shortening, Therefore, during immortalization, cells activate expression of telomerase or recombination mechanisms for the maintainance of sufficient telomere length. Indeed, activation of telomerase may happen after complete telomere failure, due to chromosome aberrations, playing a role in promoting tumorigenesis (18).

In turn, telomere shortening can lead to a subtype of senescence called **replicative senescence**. Excessive oncogenic signaling due to hyperactivation of an oncogene or inactivation of a tumor suppressor gene can also render cells into **oncogene-induced senescence** (32).

In parallel, the telomerase reverse transcriptase (TERT) subunit seems to have many non-canonical roles, related to strengthening of Wnt pathway signaling, cell proliferation or DNA damage repair, among others. The physiological relevance of these telomere-unrelated functions has been questioned (33).

Cancer cells induce angiogenesis

Angiogenesis is the sprouting of new vessels from pre-existing vessels. Early during carcinogenesis, angiogenesis is activated to supply cancer cells with vessels to transport nutrients and oxygen, as well as to eliminate the products of their metabolism. Regulation of angiogenesis involves the binding of inducers, including vascular endothelial growth factor (VEGF)-A (VEGF-A), or inhibitors, such as thrombospondin-1 (TSP-1), to membrane receptors displayed by vascular endothelial cells. In tumors, angiogenic factors can be produced directly by immune inflammatory cells or their expression can be upregulated by oncogenes such as *Ras* and *Myc*, while inhibitors of angiogenesis, such as TSP-1, agiostatin and endostatin, can be produced as natural barriers to tumor angiogenesis. Pericytes, the supportive cells that are associated with normal vasculature, and bone-marrow derived vascular progenitor cells, also play a role in assisting tumor angiogenesis (18,34)

Cancer cells activate invasion and metastasis

Both local invasion and distant metastasis require modifications in cancer cell shape and in their attachment to other cells and to the extracellular matrix. The most relevant cell-to-cell adhesion molecule involved in adherens junctions between epithelial cells is E-cadherin, which is typically lost in carcinomas. In contrast, adhesion proteins involved in cell migrations in the context of embryonic development and inflammation, such as N-cadherin, are re-activated in cancer cells . The so called invasion-metastasis cascade includes many steps and can be broadly dissected into two processes; 1) **dissemination** of cancer cells from the primary tumor by intravasation into blood and lymphatic vessels and subsequent extravasation to remote tissues, and 2) **colonization**, involving the growth of emigrated cancer cells in the new tissue microenvironment. During the colonization step cancer cells need to adapt to the new tissue microenvironment. Each microenvironment may be more or less supportive for the invading cells. In some cases, metastatic cells must overcome difficulties such as systemic suppressor factors released by some primary tumors. Consequently, gene expression programs related to metastasis vary broadly among cancer types and metastatic sites (18,35).

Epithelial-to-mesenchymal transition

EMT is a program normally occuring during embryogenesis and wound healing. Cancer cells take advantage of EMT in order to adquire invasion, resistance to apoptosis and dissemination properties. Indeed, the EMT program results in loss of adherens junctions, suppression of apical-basal cell polarity, cytoskeletal rearrangements and subsequent changes in morphology, expression of matrix-degrading proteases, enhanced motility, and resistance to apoptosis. EMT is regulated by many transcription factors, including Twist, Snail, Slug/Snail2, ZEB1 and ZEB2, that can inhibit E-cadherin gene expression. Expectedly, cancer cells at the margins of carcinomas undergo EMT more frequently than cells in the core of the tumor. Importantly, signals from the tumor-associated stromal cells are necessary for the activation of the EMT program by tumor cells. In this sense, mesenchymal stem cells can supply CCL5/RANTES, and macrophages can secrete matrix-degrading enzymes such as metalloproteases and cysteine cathepsin proteases in response to tumoral production of IL-4, or epidermal growth factor (EGF) in response to tumoral production of colony stimulating factor 1 (CSF-1) (18,36).

Importantly, cancer cells can undergo the opposite process, **mesenchymal-to-epithelial transition (MET)**, once they are established in distant tissues. Indeed, it has been hypothesized that cancer cells enter EMT partially, acquiring some of the

capabilities, and that this process is dynamic and reversible, reflecting a high degree of cellular plasticity (36).

Moreover, other types of invasion have been described, apart from mesenchymal invasion related to the EMT program. In opposition to EMT single-cell migration, collective invasion is a form of spreading in which masses of cancer cells penetrate the peritumoral stroma while maintaining cell-to-cell contact. Conversely, like EMT, the ameboid form of invasion involves single-cell migration and loss of cell-to-cell contacts but, in contrast to EMT, takes advantage of existing interstices in the extracellular matrix, without significant turnover of cell-matrix adhesions (18,37).

In their seminal review article in 2011, Hanahan and Weinberg proposed two mechanisms for the adquisition of the hallmarks of cancer by transformed cells. The so-called **enabling characteristics** are genome instability and mutation, and tumor-promoting inflammation. The authors also introduced two new hallmarks of rising importance in light of new discoveries in the field. These **emerging hallmarks** are the capability of evading immune destruction and reprogramming energy metabolism. The tumor-promoting role of inflammation and the simultaneous necessity of cancer cells to evade immune destruction reflects the dual role of the immune system in carcinogenesis (**Figure 7**).

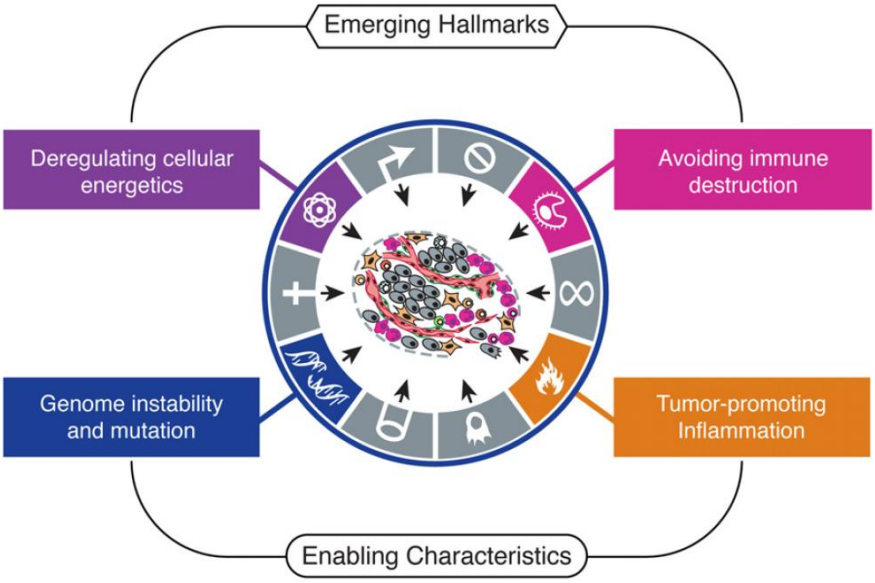


Figure 7. Emerging hallmarks and enabling characteristics of cancer. Extracted from (18)

Genome instability and mutation in cancer cells

Genome instability is an increased rate of genome alteration during cell division that gives rise to genomic heterogeneity, including mutations, single nucleotide variations, small insertions or deletions, copy number alterations, and chromosome structural rearrangements. Genome instability can be achieved by disruption of one or more components of the **genomic maintenance machinery (caretaker genes)** and/or by enhanced sensitivity to mutagenic agents. Surveillance systems, such as tumor suppressor TP53, which usually force the entry of genetically damaged cells into senescence or apoptosis, can also be affected. As mentioned previously, the loss of telomeres can generate karyoptic instability. Importantly, some specific chromosomic aberrations and gene mutations seem to be recurrent in cancer cells, suggesting that they have a causal role in carcinogenesis. Indeed, the so-called **mutator hypothesis** attributes genome instability in precancerous lesions to mutations in caretaker genes. Although this mutator hypothesis is widely accepted for hereditary cancers, its validity for sporadic (non-hereditary) cancers is under discussion. For sporadic cancers, the **oncogene-induced DNA replication stress model** has been proposed. According to this model, oncogenes induce DNA replication stress and DNA double-strand breaks, which activate the ataxia telangiectasia mutated (ATM), p53, p16INK4A or p14ARF, among other proteins. Since the genes encoding for these proteins are frequently mutated in cancer, these pathways are deregulated, resulting in genome instability. In any case, these defects are beneficial for cancer cells and participate in the multistep process of carcinogenesis. Notwithstanding, **epigenetic** alterations also drive cancer development by affecting gene expression (38,39).

Tumor-promoting inflammation

Although the immune system can initially be a barrier for tumorigenesis, paradoxically, chronic inflammation plays a key role in tumor progression. Of note, even at early stages inflammatory immune cells can release ROS that can be mutagenic for surrounding cancer cells, thereby accelerating their transformation. Chronic inflammation can enhance carcinogenesis by enabling the adquisition of cancer hallmarks by supplying the tumor microenvironment with growth factors, survival signals, proangiogenic factors, extracellular matrix-modifying and inductive signals that lead to activation of EMT. This view of **the tumor as a wound that doesn't heal** has been the rationale for targeting pro-inflammatory mediators (40,41). Unexpectedly, senescence, usually perceived as a cancer barrier, can be associated with tumor-promoting inflammation mainly due to the production of the senescence-associated secretory phenotype (SASP) (32,42–44).

Cancer cells evade immune destruction

As mentioned before, immune cells promote carcinogenesis via inflammation. At the same time, due to the dual role of the immune system, cancer cells have to avoid detection by the anti-tumor immunity, mainly coordinated by CD8+ cytotoxic T lymphocytes and NK cells. Cancer cells evade immune killing by undergoing **immunoediting**. The process of immunoediting involves the elimination of highly immunogenic cancer cell clones by the immune system and, therefore, the selection of weakly immunogenic variants. Immunoediting can also arise as an adaptation to therapies directed at potentiating anti-tumor immunity, for example by disrupting the negative regulators of immune activation (45). Another mechanism employed by cancer cells to prevent their elimination is the recruitment of actively immunosuppressive inflammatory cells, including regulatory T cells and myeloid-derived suppressor cells (18,46).

Cancer cells reprogram their energy metabolism

Cancer cells suffer an active metabolic reprogramming required to support sustained cancer cell proliferation. In the presence of oxygen (aerobic conditions), normal cells catabolize glucose into pyruvate via glycolysis in the cytosol and subsequently process it to carbon dioxide in the mitochrondria. In hypoxia, cells rely mainly on glycolysis for energy production and limit mitochondrial oxidative phosphorylation to minimize oxygen consumption. Cancer cell metabolic reprogramming typically consists in a preference for glycolysis even in the presence of sufficient oxygen; the so-called **aerobic glycolysis or Warburg Effect**. This switch is associated with activation of *Ras* and *Myc* oncogenes and mutations in the *TP53* tumor suppressor. The Ras oncoprotein, as well as hypoxia, can upregulate hypoxia inducible factor (HIF)-1α and HIF-2α, leading to glycolysis upregulation (47). The Warburg effect seems counterintuitive, since rapidly-dividing cancer cells are energy-demanding and glucose catabolism into lactate is less efficient than mitochondrial oxidative phosphorylation. One explanation for this phenomenon may be the presence of two tumoral subpopulations, one that has access to sufficient oxygen and another suffering from hypoxia. This may be the result of the temporal and regional fluctuations in oxygen concentrations in the tumor, probably due to the chaotic organization of its neovasculature. In this context, cancer cells with less access to oxygen rely on aerobic glycolysis to obtain energy and produce lactate that can be used by nearby well-oxygenated cancer cells. Besides this symbiotic theory, the Warburg Effect has been proposed to be an adaptation to sustain the biosynthetic requirements of uncontrolled proliferation, since glycolytic intermediates can be used for the biosynthesis of nucleosides and amino acids (48,49).

The relevance of cancer stem cells and tumor microenvironment

Understanding tumor biology concerns studying the diverse cell populations within a tumor and their reciprocal interactions, as well as the **tumor microenvironment**. The tumor microenvironment consists of a myriad of cell types sourrounding and crosstalking with cancer cells, and the extracellular matrix. On one hand, the most critical cell types for tumor support and progression are immune inflammatory cells, cancer-associated fibroblasts, endothelial cells, pericytes and stem and progenitor cells of the tumor stroma (**Figure 8**). On the other hand, the most remarkable specialized cancer cell type are **cancer stem cells,** which are a subset of cells with stem cell features. Cancer stem cells are considered key drivers of tumor initiation and progression and are also responsible for tumor persistence, relapse after therapy, and metastasis (50–52).

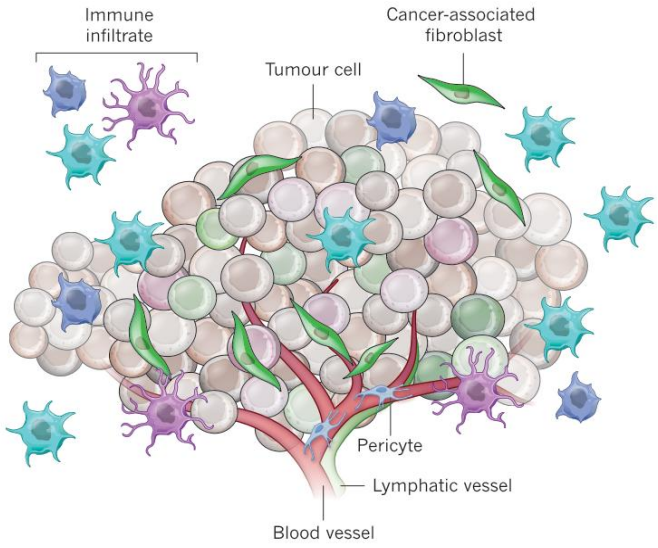


Figure 8. Cancer cells and other cell types forming the tumor microenvironment. Extracted from (53)

Liver carcinogenesis

Epidemiology and risk factors

Liver cancer is the fourth most common cause of cancer-related death worldwide. Its incidence has risen over the last 20 years in Western countries and is predicted to continue increasing in the future. Indeed, the World Health Organization estimates that more than 1 million patients will die from liver cancer in 2030 (54,55). The incidence and mortality by liver cancer for both sexes around the world in 2018 are depicted in **Figure 9A** and **9B**, respectively. Of note, both incidence and mortality rates are 2-3 fold higher among men in most world regions.

HCC is the most frequent primary liver cancer and typically arises in patients with underlying chronic liver disease. Therefore, its worldwide incidence is the consequence of the heterogeneous prevalence of its risk factors, which, in fact, are the etiologies for sustained liver injury. In accordance with this idea, most HCC cases (80%) occur in sub-Saharan Africa and eastern Asia, where there is a high prevalence of hepatitis B and population exposure to aflatoxin B1 (AFB1)-contaminated food. In the United States of America (USA), Europe and Japan, hepatitis C and excessive alcohol intake are the main risk factors. Interestingly, the expected progressive increase in HCC incidence until 2030 is due to an increase in some countries such as the USA, while in other countries, including Japan, the incidence has started to decline. These different temporal trends can be attributed at least, in part, to the time of spread of hepatitis C virus (HCV) among regions and, more recently, to the rise of non-alcoholic fatty liver disease (NAFLD) (56). Since the spread of HCV occurred earlier in Europe and Japan than in the USA, the incidence of HCC in these regions has reached a plateau and is declining in some areas. In the USA the incidence of HCC is increasing and will foreseeably stabilize by 2020 (57). NAFLD is an important cause of HCC. Its incidence is increasing in developed regions and may partially overlap with alcohol-related liver disease. Metabolic syndrome, diabetes and obesity have been associated with the risk of developing HCC in patients with NAFLD (58). Tobacco use and co-infection of human immunodeficiency virus (HIV) with either hepatitis B virus (HBV) or HCV are considered risk factors, while coffee consumption is associated with decreased risk (56). Recently, a pathogenic role for adeno-associated virus type 2 (AAV2) in HCC development has been proposed (59).

In conclusion, there are many HCC risk factors heterogeneously distributed around the world, but the most prominent are chronic HBV and HCV viral infection, chronic alcohol consumption, AFB1-contaminated food and all cirrhosis-inducing conditions (56). Possible mechanisms of hepatocarcinogenesis for these risk factors are shown in **Figure 10**.

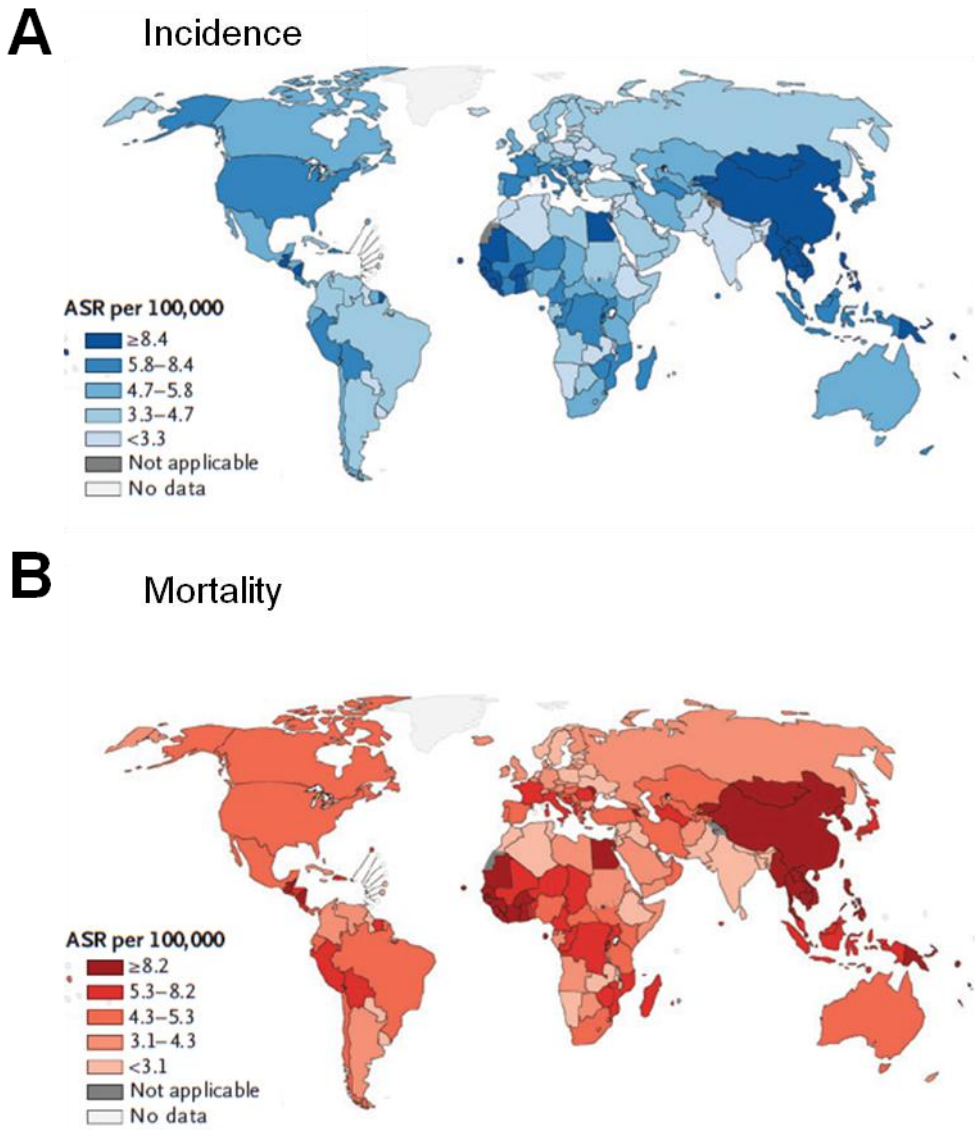


Figure 9. Worldwide epidemiology of liver cancer in 2018 (55). (**A**) Incidence of liver cancer. (**B**) Liver cancer-associated deaths. Data are expressed as the age-standardized rate (ASR) per 100,000 population. Extracted from (60)

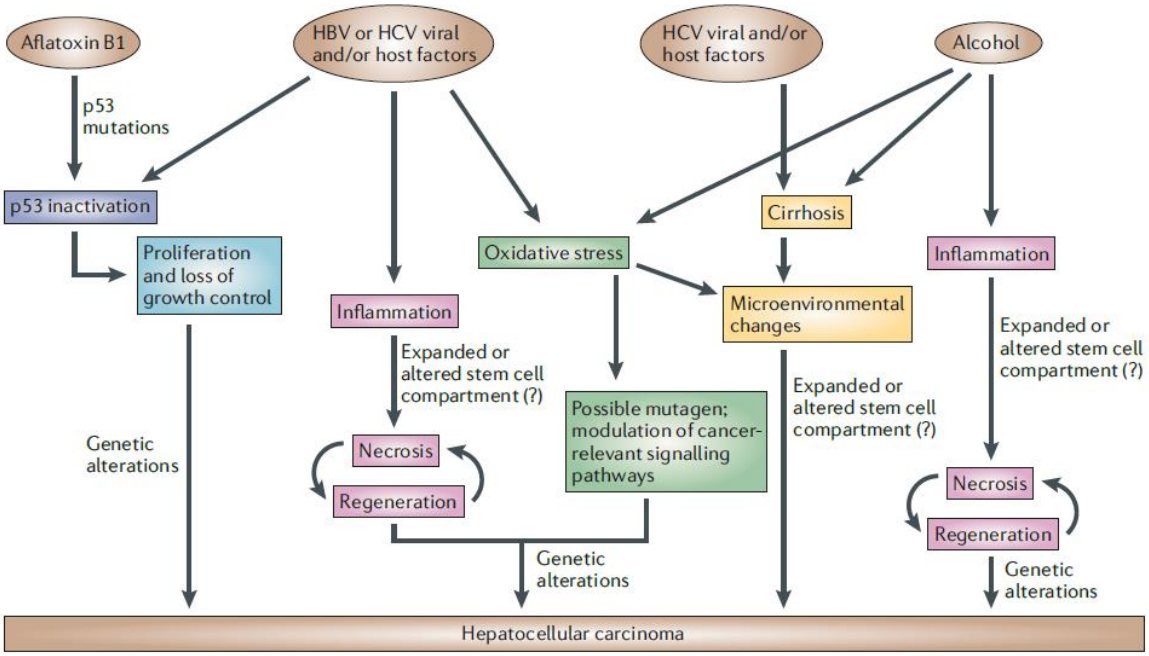


Figure 10. Mechanisms of hepatocarcinogenesis. Abbreviations: hepatitis B virus (HBV), hepatitis C virus (HCV). Extracted from (61)

Histopathology

The development of HCC is a complex multistep process and is closely related to the presence of chronic liver disease and progressive liver function impairment (**Figure 11**). Patients with chronic liver disease have sustained hepatic inflammation, fibrotic deposition and aberrant hepatocyte regeneration, which can lead to cirrhosis. Dysplastic lesions generally develop in cirrhotic livers and are *bona fide* preneoplastic lesions. The accumulation of genetic and epigenetic aberrations can make dysplastic cells proliferative and invasive, leading to HCC establishment (60). Remarkably, HCC frequently occurs in the form of multiple nodules that may reflect multi-centric carcinogenesis or intrahepatic metastasis, the latter category being more aggressive and more poorly differentiated (62,63).

The classical histomorphologic features of HCC are intense vascularization, wide trabeculae, distinguishable acinar pattern, small cell changes, cytologic atypia, mitotic activity, absence of Kupffer cells, vascular invasion and the loss of the reticulin network. Bile production and Mallory bodies can also be observed (64). Mallory bodies are irregular rope-like eosinophilic cytoplasmic structures found in hepatocytes that are composed of a complex of misfolded cytoskeletal elements, including aggregates of cytokeratin intermediate filaments (mainly cytokeratins 8 [CK8] and 18 [CK18]), ubiquinated cytokeratins, heat shock proteins and other peptides (65).

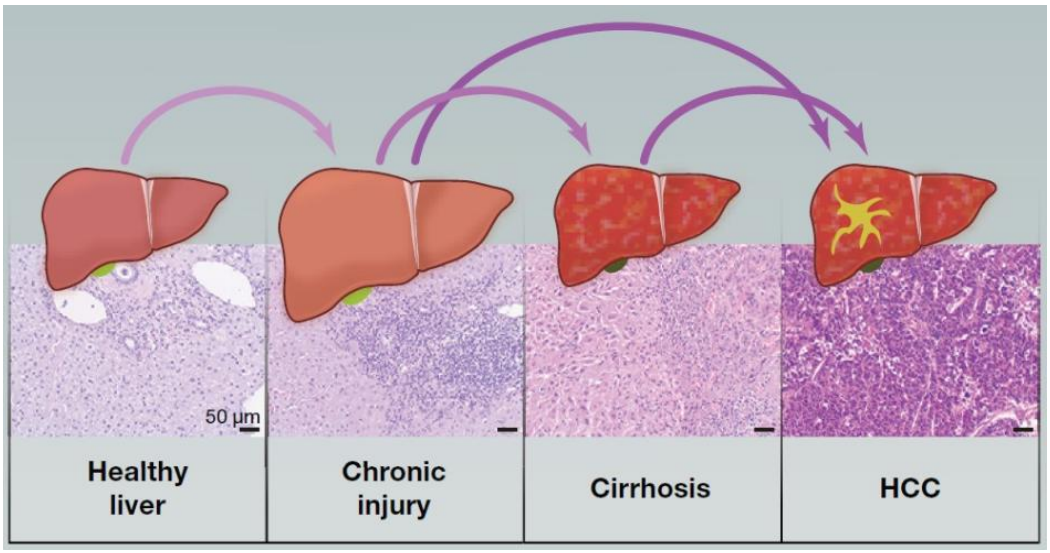


Figure 11. Clinical progression of liver cancer. Abbreviations: HCC, hepatocellular carcinoma. Extracted from (66)

Molecular pathogenesis

Indeed, HCC has an enormous molecular heterogeneity. However, there are currently no molecular classifications of HCC able to predict disease progression or recurrence (67).The Barcelona Clinic Liver Cancer (BCLC)(68) algorithm is the most commonly used staging system for HCC. It defines five stages that are in accordance with treatment recommendations, but, unfortunately, it does not take in account the molecular biology of the tumors (56).

Nevertheless, the accumulation of alterations in cancer driver genes and associated pathways have been detected at variable mutational rates, probably reflecting clinical heterogeneity, and are major triggers of hepatocellular carcinogenesis and tumor progression. TERT reactivation caused by promoter mutations, viral (mainly HBV and AAV2) integrations or focal amplifications is the most common somatic alteration observed in HCC (70% frequency). It is present in a proportion of premalignant lesions developed in cirrhosis and early HCC, suggesting its involvement in tumor initiation (69). Along with TERT alterations, somatic mutations affecting other genes involved in cell cycle control, such as *TP53* (the gene encoding tumor suppressor protein p53, 30%) and *CTNNB1* (30%, coding for β-catenin), as well as other members of the Wnt pathway are also recurrently described in HCC samples. Moreover, inactivating mutations in chromatin remodeling proteins have emerged as a major deregulated pathway in HCC, as well as mutations of the *KMT2* gene family members encoding for histone methyltransferases. Similarly to most solid tumors, besides this set of relatively frequent mutations, next generation sequencing has uncovered a large number of low-

frequency somatic mutations that affect multiple genes implicated in cell cycle control, PI3K/mTOR signaling, MAP kinase signaling, apoptosis, hepatic differentiation, oxidative stress, TGFβ signaling and epigenetic regulation. Epigenetic modifications contributing to HCC development include dysregulated methylation of multiple gene regions, histone modifications and chromatin remodeling (70–77).

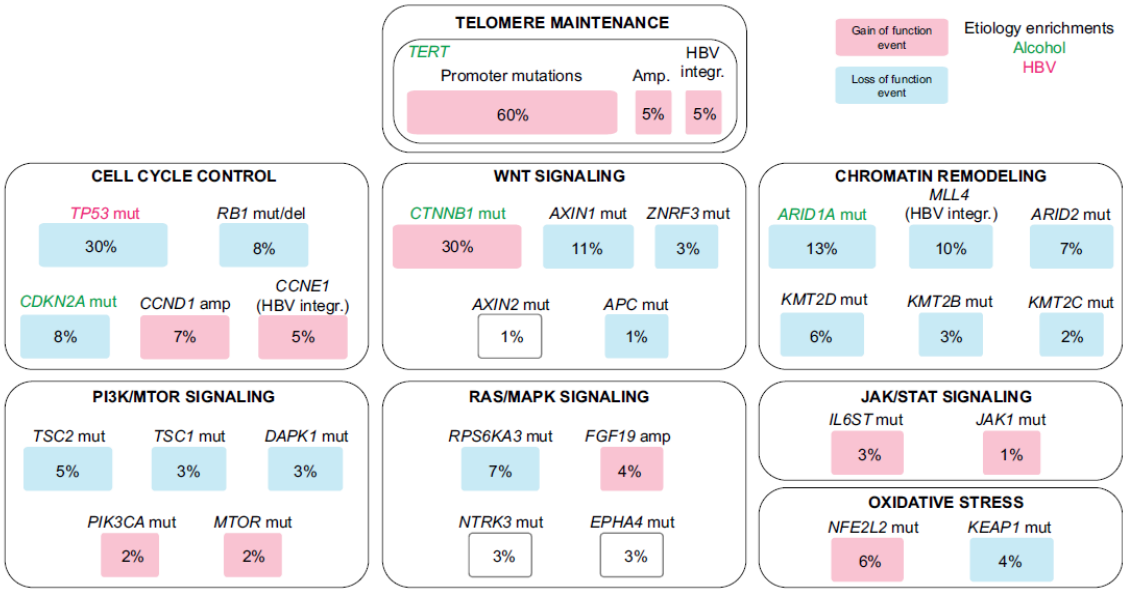


Figure 12. Mutational landscape of HCC. Graphical summary of the main mutated genes in hepatocellular carcinoma (HCC) and their reported frequency. Some mutations may co-exist in the same patient. Data suggest that background etiology impacts the mutation rate. Most mutations affect 3 regions: *TERT* promoter, *CTNNB1* gene and *TP53* gene. Abbreviations: amplification (amp.), integration (integr.), mutation (mut.), deletion (del.). Exactred from (63)

Besides mutations, DNA copy number alterations are frequent genetic events in HCC. Broad genomic deletions and gains affecting 1p, 4p-q, 6q, 8p, 13p-q, 16p-q, 17p, 21p-q, 22q and at 1q, 5p, 6p, 8q, 17q, 20q, Xq have been described (71,73–76). Recently, there have been important efforts in the field of HCC to develop the so-called **nucleotide or mutational signatures** in order to link intrinsic biological processes (for example, defective DNA repair or aging) or exposure to external genotoxic agents (such as AFB1, tobacco or ultraviolet light) with specific global mutational patterns at the nucleotide level (74,76,78). **Figure 12** summarizes the main mutated genes in HCC, their reported frequency and etiology enrichments.

Sex differences in liver cancer

In humans, males account for a higher percentage of HCC cases (61). This phenomenon is partly explained by the increased ethanol consumption in males compared to females, but also by hormonal differences between males and females, that are also present in mice and are therefore relevant to the current study. While the androgen receptor binds to testosterone and other male hormones and is more abundant in male individuals, the estrogen receptor regulates the female reproductive cycle and mainly binds to estrogen that occurs at higher levels in females. The androgen and the estrogen receptor seem to promote HBV-associated liver cancer and to protect against it, respectively. Moreover, the estrogen receptor inhibits liver carcinogenesis by repressing pro-inflammatory cytokine IL-6 and thereby attenuating chronic liver inflammation, that is related to HCC development, as well as by interacting with FOXA1 and FOXA2, that act as transcription factors in liver cells. Additionally, prolactin, a hormone involved in milk secretion and more abundant in females, plays a protective role in liver cancer development due to its anti-inflammatory effects (79). Low adiponectin levels also contribute to the increased risk of liver cancer in males (80). In conclusion, the aforementioned differences and other unknown parameters that differ between males and females might play a role in HCC development.

Treatment

Multiple treatment options are available for HCC and the suitability of each treatment depends on tumor stage, liver function and patient performance status. Patients in an early stage (BCLC stage 0 or A) can benefit from surgical therapies such as resection or transplantation, and tumor ablation, which includes different methods that produce tumor necrosis. The most common type of ablation is image-guided percutaneous radiofrequency ablation, but there are others such as microwave ablation, cryoablation and ethanol injection. Transarterial therapies, including transarterial chemoembolization and internal radiation therapy, are frequently used in patients with intermediate-stage tumors (BCLC stage B). Transarterial chemoembolization requires intraarterial infusion of a cytotoxic agent followed by embolization of the vessels that feed the tumor but not healthy liver tissue, which is irrigated mainly from the portal vein. Internal radiation therapy entails intraarterial infusion of microspheres with the radioisotope yttrium-90. Finally, systemic therapies include drugs with anti-proliferative and anti-angiogenic effects and are recommended for patients with advanced disease (BCLC stage C) or intermediate-stage tumors (BCLC stage B) after transarterial therapies (60). The gold standard drug is sorafenib. Sorafenib is an oral inhibitor of multiple kinases, and it was the first systemic drug approved for the treatment of HCC

as a result of two positive randomized placebo-controlled trials (81,82). Two other multikinase inhibitors, lenvatinib and regorafenib, have recently been approved as first-line (83) and second-line treatment (84), respectively. Furthermore, the inhibitor of receptor tyrosine kinases, cabozantinib (85), and the antibody against the VEGF receptor 2, ramucirumab (86), are also available for second-line treatment.

Cancer stem cells in hepatocellular carcinoma

Cancer stem cells are key drivers of liver carcinogenesis (see "The relevance of cancer stem cells and tumor microenvironment" section). In the case of HCC, cancer stem cells are resistant to radiotherapy and chemotherapy and are related to sorafenib relapse (87). Cancer stem cells are subjected to epigenetic regulation in liver cancer, including DNA epigenetic marks, histone marks, expression of epigenetic regulators, and non-coding RNAs (88).

The cell of origin of hepatic cancer stem cells has not yet been clarified. Cancer stem cells are broadly assumed to derive from bipotent hepatic progenitor cells (89). In addition, adult hepatocytes and bone marrow cells have also been proposed as the potential origin of hepatic cancer stem cells (90). Previous work by our research group (91) supports the concept that hepatic cancer stem cells may originate from normally existing progenitor cell populations rather than via lineage dedifferentiation of hepatocytes. PR/SET domain-containing protein 7 (PR-SET7) liver-specific knockout (KO) mice, in which PR-SET7 inactivation occurs at postnatal stages, spontaneously develop HCC. However, PR-SET7-deficient hepatocytes either die due to G2 arrest or undergo necrosis once they enter the cell cycle during compensatory proliferation. Thus, in this model the liver is replenished by newly-generated hepatocytes derived from the activation, proliferation and differentiation of $Sox9^+$ and $A6^+$ ductal progenitor cells. These hepatocytes provide a source of highly-proliferating tumoral cells that would dedifferentiate to cancer stem cells (91). In conclusion, hepatic cancer stem cells may originate from all three cell types (hepatic progenitor cells, adult hepatocytes and bone marrow cells), with an important degree of cell plasticity, depending on the biological context.

Murine models of hepatocellular carcinoma

A wide range of murine models, including chemical treatments, genetically engineered mouse models and xenografts, are available to study HCC (**Table 1**).

Although tumor xenografts models are the priority in preclinical studies, they were not suitable for the purpose of the present study because they require immunodeficient mice, e.g. nude, severe combined immune deficient (SCID) or nonobese diabetic-severe combined immunodeficiency disease (NOD/SCID) mice (92). Neither were genetically engineered mouse models suitable for our scope since we wanted to examine the role of a specific protein in hepatocellular carcinogenesis by inducing carcinogenesis in KO mice lacking this protein, and the presence of multiple transgenes may act as a confounding factor as well as delay the breeding process of the mouse colonies.

	Properties	Latency	Notes
Chemical			
DEN/Phenobarbital DEN (single injection)	Genotoxic	5–10 m	When DEN is injected to adults promotion is needed, used in combination with genetic/dietary/environmental models
Aflatoxin	Genotoxic	>22 m	Often combined with genetic models
CDE diet, TAA, CCl4, peroxisome proliferators etc.	Associated with steatohepatitis, fibrosis, etc.	>12 m	Variability due to different experimental protocols, used for context-specific modeling. Often combined with DEN or genetic models
GEMMs			
HBV-derived (sAg, HBx)	ER stress,focal necrosis, proliferation	12–24 m	Several lines/backgrounds with different penetrance/latency
HCV-derived (Core)	Steatosis	12–24 m	Several lines/backgrounds with different penetrance/latency
Mdr2 knock-out	Cholangitis	>12 m	Strong inflammatory component, strain dependent
Mosaic GEMMs			
Somatic gene/molecules delivery: Virus-based, RCAs/TVA, Hydrodynamic	Fast, cost-effective	Few weeks	Can be combined with GEMMs, suitable for imaging
Implantation models: p53KO; mycTg hepatoblasts	Fast, cost-effective	Few weeks	Suitable for imaging and large-scale screens

Table 1. Mouse models to study HCC development. Abbreviations: DEN, diethylnitrosamine; CDE, choline-deficient, ethionine-supplemented; TAA, thioacetamide; CCl_4, carbon tetrachloride; GEMMs, genetically engineered mouse models; HBV, hepatitis B virus; SAg, hepatitis B surface antigen; HBx, HBV X protein; HCV, hepatitis C virus; Mdr2, ATP-binding cassette sub-family B (MDR/TAP) member 4; RCAS/TVA system, replication-competent avian sarcoma-leukosis virus long terminal repeat with splice acceptor/tumor virus A; KO, knockout; Tg, transgenic; ER, doplasmic reticulum. Extracted from (93)

The diethylnitrosamine mouse model

The diethylnitrosamine (DEN) mouse model is a well established model of liver cancer. Various nitrosamines including DEN are metabolically activated in the liver, principally in hepatocytes, by cytochrome P450 enzymes including P450 2E1 (CYP2E1). Reactive metabolites may bind as electrophiles to DNA, giving rise to DNA adducts, which are a form of DNA damage and can lead to initiation of carcinogenesis. Typical adducts formed in DNA are N^7-methylguanine and 0^6-methylguanine, The most prominent effects of DEN on the liver are intense neutrophilic infiltration, extensive centrilobular haemorrhagic necrosis, bile duct proliferation, fibrosis, and bridging necrosis that ends in hepatocarcinogenesis (94,95).

Colon carcinogenesis

Epidemiology and risk factors

Colorectal cancer (CRC), also known as colon cancer, is the most common carcinoma of the digestive tract. Worldwide, CRCs are the third most common cause of cancer-related death, being colon adenocarcinoma the most frequent form of CRC. CRC is directly linked to Western diet and its incidence tends to rise in countries undergoing major development transition. Therefore, the incidence of CRC can be considered a marker of socioeconomic development. Although temporal patterns of incidence and mortality largely vary worldwide and even among Western countries, the global burden of CRC is expected to increase. The World Health Organization estimates that around 1.1 million patients with die from CRC in 2030 (55,96). The incidence and mortality of CRC around the world in 2018 are depicted in **Figure 13A** and **13B**, respectively.

Similarly to cancers affecting other organs, CRC development is influenced by non-modificable factors and modificable (non-hereditary) factors, that have a positive or negative impact on the development of the disease. On one hand, the most relevant non-modificable factors that increase the risk of CRC are male gender, age, type 2 diabetes, inflammatory bowel disease, hereditary CRC syndromes and a positive family history, while modificable risk factors are tobacco, excessive alcohol intake, processed and red meat, low intake of vegetables and fruits, and body fat and obesity. On the other hand, the most prominent protective factors are modificable factors, including physical activity and regular consumption of whole grains, dietary fiber, tree nuts, fish, vitamins (D, C, and others), dairy products and calcium supplements. Some therapies, such as aspirin or nonsteroidal anti-inflammatory drugs (for exemple ibuprofen), statin and menopausal hormone therapy, may also have a protective effect against CRC (97).

Pathogenesis

CRC normally arises from a neoplastic precursor lesion (a polyp) that has derived from an aberrant crypt. Normal intestinal stem cells, which reside in the crypts and express the Wnt/β-catenin potentiator LGR5, have been assumed to be the cell of origin of the majority of CRCs, consistent with the well characterized deregulation of Wnt/β-catenin signaling during colon carcinogenesis. Therefore, LGR5-positive cancer stem cells presumably evolve from normal LGR5-expressing stem cells as a result of progressive accumulation of genetic and epigenetic alterations that repress tumor suppressor genes and hyperactivate oncogenes (97,98). The most frequently mutated gene is the tumor suppressor adenomatous polyposis coli (*APC)* (85% frequency), and its

inactivation results in excessive Wnt signaling due to impaired β-catenin oncoprotein degradation. Other examples of silencing of tumor suppressors that is commonly associated with CRC are: 1) genetic or epigenetic inactivation of mutL homolog 1 (*MLH1)*, MutS homolog 2 (*MSH2*) and MutS homolog 6 (*MSH6*) (15-25%), 2) mutations in *TP53* (35-55%), 4) mutations in TGFβ receptor 2 (*TGFBR2*), a receptor triggering growth arrest and apoptosis (25-30% of all CRC cases, but >90% of tumors with microsatellite instability), 5) mutations in SMAD family member 4 (*SMAD4*), a

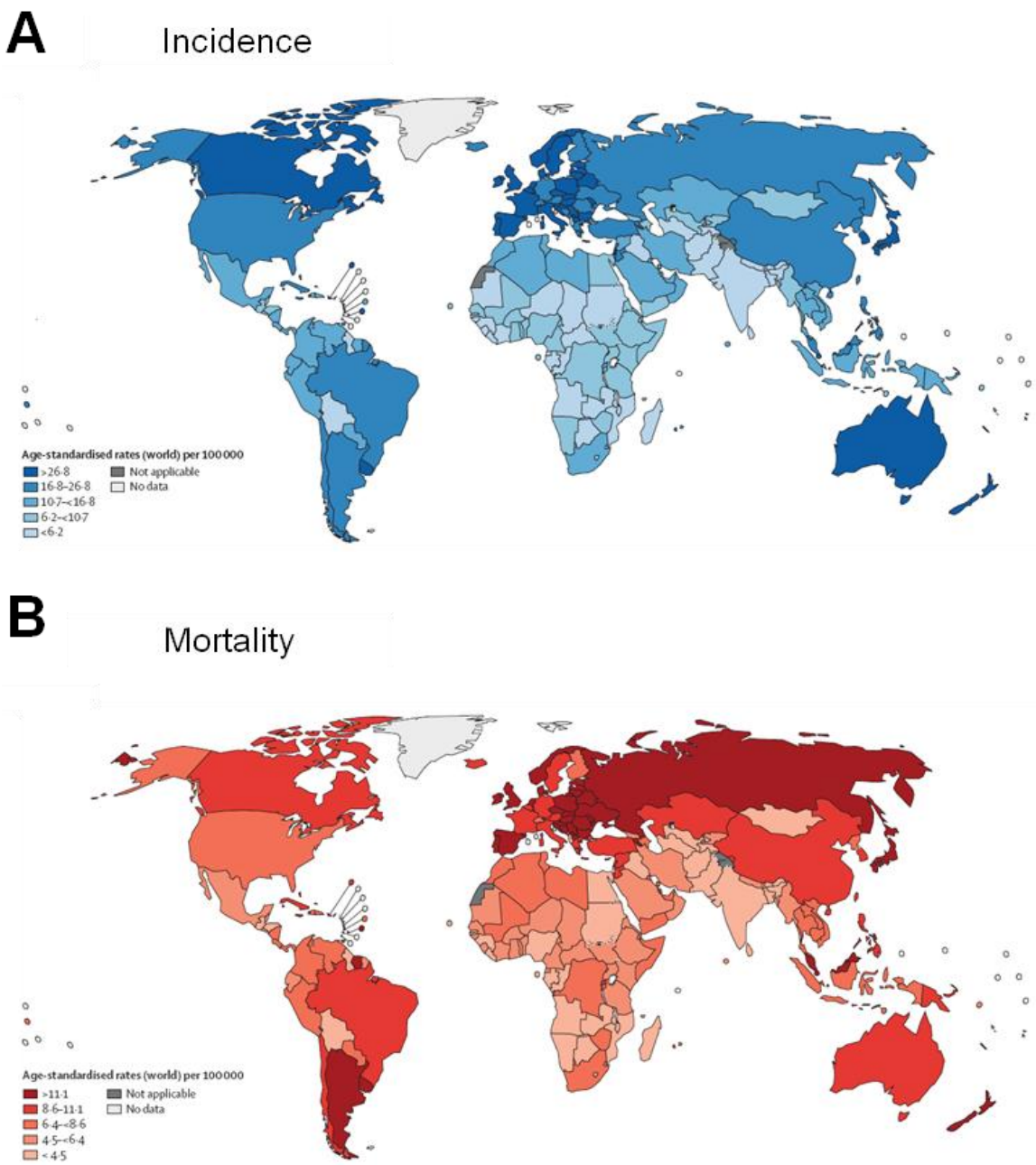


Figure 13. Worldwide epidemiology of colorectal cancer (CRC) in 2018 (55). (**A**) Incidence of CRC. (**B**) CRC-associated deaths. Data are expressed as the age-standardized rate per 100,000 population. Extracted from (97)

component of the TGFβ signaling pathway (10-35%), and 6) mutations in *PTEN* (10-15%). Regarding oncogenes, the most frequently mutations affect KRAS proto-oncogene GTPase (*KRAS*) (35-45%) and B-Raf proto-oncogene serine/threonine kinase (*BRAF*) (8-12%). Growth factor pathways also play key roles in CRC development; the enzyme cyclooxygenase-2 (COX-2) and EGFR are activated in CRC and may mediate tumorigenesis, while 15-prostaglandin dehydrogenase (15-PGDH) and TGFβ are inactivated in CRC, probably due to their tumor suppressor properties (99).

The polyp to colorectal cancer sequences

CRC pathogenesis is currently viewed as a multistep sequential mutational pathway, in which the acquisition of each histological feature is the consequence of a molecular alteration. Overall, there are two major normal colon to CRC sequences (**Figure 14**); 1) the classical or traditional **adenoma-carcinoma pathway** or chromosomal instability sequence (accounting for 70-90% of CRC cases), and 2) the **serrated pathway** (10–20% of CRC). The first sequence involves the development of tubular adenomas, whereas the second one involves serrated polyps. Both types of lessions can progress to adenocarcinomas. Some genes are mutated or epigenetically altered in both pathways, while others are involved in only one of the sequences. The adenoma-carcinoma pathway is associated with chromosomal instability (CIN). CIN phenotypes typically arise upon APC regulator of Wnt signaling pathway (*APC*) mutation, followed by *KRAS* activation and *SMAD4* or *TP53* inactivation. Microsatellite instability (MSI) occurs during the adenoma-carcinoma pathway in the case of Lynch syndrome which is a hereditary non-polyposis type of CRC in which mutations affecting the DNA mismatch repair machinery lead to MSI. Lynch syndrome-associated cancers exhibit MSI, but are generally negative for the CpG island methylator phenotype (CIMP). Unlike the adenoma-carcinoma pathway, the serrated pathway is associated with CIMP and *BRAF* mutations. CIMP involves silencing of tumor suppressor genes by promoter CpG island methylation, which is a major cancer driver event. The epigenetic instability associated with CIMP leads to microsatellite stable and unstable cancers. Most sporadic microsatellite unstable colon tumors are positive for CIMP and develop through the serrated pathway (97,100–102).

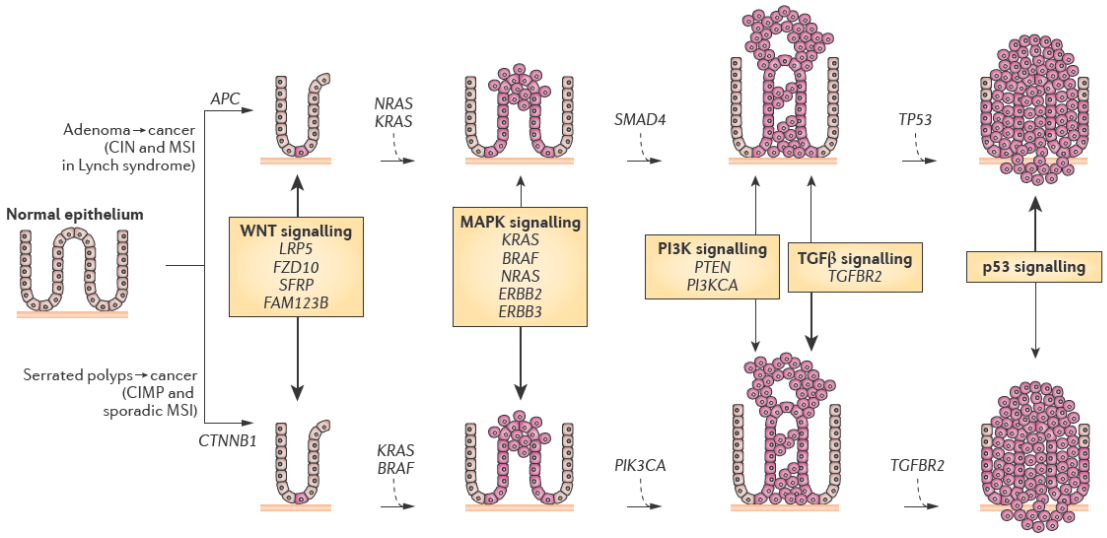


Figure 14. The polyp to colorectal cancer (CRC) sequences. The classical or traditional adenoma-carcinoma pathway or chromosomal instability (CIN) sequence (top) and the serrated pathway (bottom). Abbreviations: MSI, microsatellite instability; *APC*, adenomatous polyposis coli; *NRAS*, NRAS proto-oncogene GTPase; *KRAS*, KRAS proto-oncogene GTPase; *SMAD4*, SMAD family member 4; *TP53*, tumor protein p53; CIMP, CpG island methylator phenotype; *CTNNB1*, catenin-β1; *BRAF*, B-Raf proto-oncogene serine/threonine kinase; *PI3KCA*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α; *TGFBR2*, TGFβ receptor 2; *LRP5*, low-density lipoprotein receptor-related protein 5; *FZD10*, frizzled class receptor 10; *SFRP*, secreted frizzled-related protein; *FAM123B*, family with sequence similarity 123B (also known as *AMER1*); MAPK, mitogen-activated protein kinase; *ERBB2*, erb-b2 receptor tyrosine kinase 2; *ERBB3*, erb-b2 receptor tyrosine kinase 3; PI3K, phosphatidylinositol 3-kinase; *PTEN*, phosphatase and tensin homolog; TGFβ, transforming growth factor-β. Extracted from (100)

CRC are classified into four consensus molecular subtypes (CMS 1-4), according to differences in gene expression. CMS1, CMS2, CMS3 and CMS4 involve alterations in unique immune, canonical, metabolic and mesenchymal pathways, respectively. Diagnosis and clinical management decisions are based on clinical symptoms, endoscopy results, carcinoembryonic antigen blood concentrations and histology. Treatment options include chemotherapy and targeted therapies such as anti-VEGF or anti-EGFR agents, BRAF or BRAF plus MEK inhibitors, the multi-kinase inhibitor regorafenib, and immunotherapy (97).

A murine model of colorectal cancer: the 1,2-dimethylhydrazine and dextran sodium sulfate murine model

Similarly to liver cancer models, CRC development in mice can be achieved by means of chemically induced animal models, including treatments with: 1) azoxymethane (AOM), 2) 1,2-dimethylhydrazine (DMH), 3) heterocyclic amines (e.g. 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine [PhIP], 2-amino-33-methylimidazo[4,5-*f*]quinoline [IQ]), 4) aromatic amines (e.g. 3,2´-dimethyl-4-aminobiphenyl [DMAB]) and

5) alkylnitrosamide compounds (e.g. methylnitrosourea [MNU], N-methyl-N´-nitro-N-nitrosoguanidine [MNNG]), or using genetic models such as *APC* KO mice (103,104)..
DMH and its metabolite AOM are genotoxic colonic procarcinogens that require metabolic activation to form DNA reactive products. DMH metabolism involves a variety of xenobiotic-metabolizing enzymes that perform several *N*-oxidation and hydroxylation reactions, finally leading to the formation of methylazoxymethanol (MAM), the reactive metabolite. The main enzyme responsible for the hydroxylation of AOM and subsequent formation of MAM is alcohol-inducible cytochrome P-450 isoform, CYP2E1. MAM is a relatively stable compound, with a half life of 12 hours, which is sufficient time to be transported to the colon via the bloodstream. It may be further metabolized in the blood and the colon, finally leading to DNA reactive products (103).

Dextran sodium sulfate (DSS) is a water-soluble chemical colitogen with anti-coagulant properties widely used to generate colitis in mice. DSS induces intestinal inflammation probably by producing damage to the intestinal epithelium and allowing the diffusion of pro-inflammatory intestinal contents, such as bacteria and its products, into the underlying tissue (105).

The combination of DMH and DSS is a commonly used treatment for the induction of colitis-associated colon cancer in mice (106).

Epigenetics

The term epigenetics is commonly defined as the study of modifications of chromatin, that can be heritable or not, and occur without alterations in DNA sequences. These changes in chromatin can modify its function, which may affect gene expression, replication or DNA repair, among other biological processes. Hence, **epigenetic modifications** are another layer of gene expression control in addition to the genotype and can affect the final output, that is the phenotype. **Epigenetic mechanisms** include changes in DNA methylation (which in mammals occur predominantly at position 5 of cytosines in the context of cytosine guanine dinucleotides [CpG]), posttranslational modifications (PTMs) of histones and replacement of canonical isoforms with histone variants such as macroH2A, H2AX or H2A.Z, as well as non-coding RNAs. All these epigenetic mechanisms have been involved in the pathogenesis of many diseases, including cancer, where they can contribute to and also be the consequence of the acquisition of cancer hallmarks (107–110).

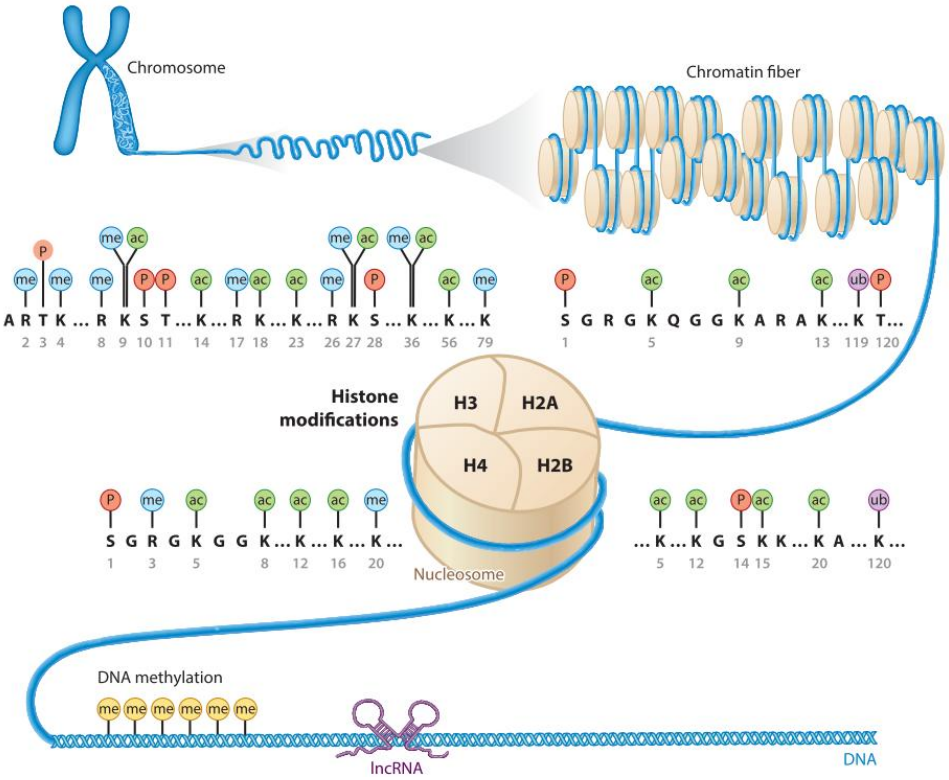


Figure 15. Overview of chromosome structure and epigenetic modifications. DNA binds to core histone octamers in order to be packed into nucleosomes and subsequently into chromatin fibers and chromosomes. Abbreviations: ac, acetylation; me, methylation; P, phosphorylation; ub, ubiquitination; lncRNAs, long noncoding RNAs. Extracted from (111)

Chromatin structure and posttranslational modifications of histones

The basic structural and functional unit of chromatin is the **nucleosome**, which is composed of a histone octamer complex and a 147-bp DNA segment wrapped around it. There are four classes of core histones, histone 2A (H2A), histone 2B (H2B), histone 3 (H3), histone 4 (H4), and one linker histone, histone 1 (H1). The histone octamer is assembled from a histone H3:H4 tetramer and two H2A:H2B dimers. Octamers are separated by linker DNA. The linker histone H1 binds to the nucleosome core particle at the entry and exit sites of the linker DNA to facilitate nucleosomal stabilization and formation of higher order structure (112). The N-termini of core histones are referred to as histone tails because they project out of the nucleosome core. Importantly, the histone tails, as well as the globular domains (located within the nucleosome) of all four core histones, are subjected to a variety of PTMs, acting as signal integration platforms. Additionally to histones, non-histone protein substrates can also be subjected to PTMs. Well-established PTMs include: 1) **acetylation**, the addition of an acetyl group to lysine residues, 2) **methylation**, the addition of a methyl group to lysine and arginine residues, 3) **phosphorylation**, the addition of a phosphoryl group to serine and threonine residues, 4) **ubiquitylation**, the ligation of a ubiquitin protein tag to lysine residues, 5) **sumoylation**, the attachment of a small ubiquitin-like modifier (SUMO) protein to lysine residues, 6) **adenosine diphosphate (ADP) ribosylation**, the marking of glutamic acid residues with an ADP-ribose molecule, 7) **deimination** involving the conversion of an arginine to a citrulline, 8) **proline isomerization,** implying the exchange between proline *cis* and *trans* conformation, 9) **crotonylation**, the addition of a crotonyl group to lysine residues, 10) **succinylation**, the addition of a succinyl group to lysine residues, 11) **malonylation**, the addition of a malonyl group to lysine residues, 12) **propionylation**, the addition of a propionyl group to lysine residues, and 13) **butyrylation**, the addition of a butyryl group to lysine residues. (107,111,113–120). DNA packing and epigenetic modifications, including most well-known core histone PTMs and the specific amino acid residue modified, are shown in **Figure 15**. The classes of histone PTMs identified, the type of residue modified and the functions regulated by each PTM are summarized in **Table 2**.

Chromatin Modifications	Residues Modified	Functions Regulated
Acetylation	**K**-ac	Transcription, Repair, Replication, Condensation
Methylation (lysines)	**K**-me1 **K**-me2 **K**-me3	Transcription, Repair
Methylation (arginines)	**R**-me1 **R**-me2a **R**-me2s	Transcription
Phosphorylation	**S**-ph **T**-ph	Transcription, Repair, Condensation
Ubiquitylation	**K**-ub	Transcription, Repair
Sumoylation	**K**-su	Transcription
ADP ribosylation	**E**-ar	Transcription
Deimination	**R > Cit**	Transcription
Proline Isomerization	**P-cis > P-trans**	Transcription

Table 2. Overview of different classes of histone posttranslational modifications (PTMs), the residues that are modified and the biological functions associated with each modification. Extracted from (114)

Histone PTMs control the accessibility of DNA to components of the transcriptional apparatus, such as transcription factors, as well as other regulators, including components of the DNA replication and repair machinery. This regulatory process involves the interplay between three categories of epigenetic regulators; 1) **writers**, the enzymes that apply PTMs to specific amino acid residues in histones, such as histone lysine methyltransferases, histone acetyltransferases and ubiquitin ligases, 2) **readers**, the proteins that recognize and interpret specific PTMs at defined sites in the protein backbone of histones, and 3) **erasers**, the enzymes that remove specific PTMs at defined positions in histones, including histone demethylases and histone deacetylases. Of note, DNA methylation, as well as PTMs in non-histone proteins, are also regulated by the joint action of writers, readers and erasers (111,121) Remarkably, readers contain binding domains with specialized protein-protein interaction motifs that allow them to recognize and discriminate between diverse PTMs (115) (**Figure 16**).

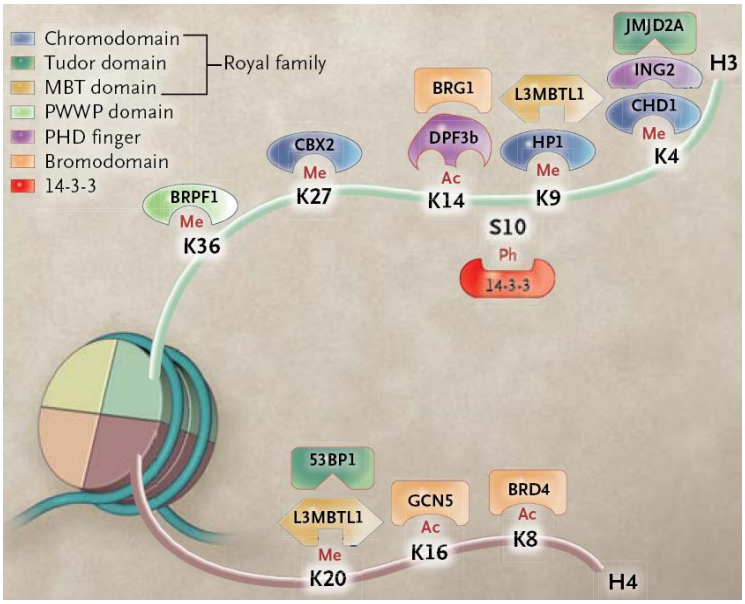


Figure 16. Classes or families of specialized binding domains present in readers. Domains within a class or family can have subtle variations that modify their substrate preference. Extracted from (115)

Altogether, this epigenetic regulation orchestrated by the PTMs at the histone level is built upon the notion of an existing **histone code**, meaning that multiple histone PTMs act in a combinatorial or sequential fashion in one or multiple histone tails to specify unique downstream functions and biological outcomes. Therefore, histone PTMs can positively or negatively affect other PTMs (**Figure 17**). Indeed, histone modification crosstalk can occur via multiple mechanisms: 1) **competitive antagonism** between PTMs when targeting the same site, for example a particular lysine acetylated, methylated or ubiquitylated by different enzymes, 2) **positive transregulation** when one PTM relies on another PTM, for example, methylation of H3 lysine 4 (H3K4) by scCOMPASS and of H3 lysine 79 (H3K79) by scDot1 are completely dependent on the ubiquitylation of H2B lysine 123 (H2BK123) by scRad6/Bre1 and this mechanism is conserved from yeast to humans, 3) **negative regulation**, when the binding of a protein factor to a specific PTM can be abrogated by a contiguous or even distant PTM, such as HP1 binding to H3 lysine 9 di- or trimethylation (H3K9me2/3), which is inhibited during mitosis as a result of the phosphorylation of H3 serine 10 (H3S10) (the so-called phospho switch), 4) **alteration of enzymatic activity when its substrate has been modified**, for example, in yeast, the conformational changes in H3 proline 38 (H3P38) caused by the activity of scFpr4 proline isomerise affect the ability of scSet2 methyltransferase to recognize and methylate H3 lysine 36 (H3K36), and 5) **cooperation between PTMs for the recruitment of specific factors**, for example, PHF8 binding to H3 lysine 4 trimethylation (H3K4me3) via its PHD finger is stabilized when H3 lysine 9 (H3K9) and H3 lysine 14 (H3K14) are acetylated on the same tail of H3, probably due to the recruitment of additional factors in complex with PHF8 (122).

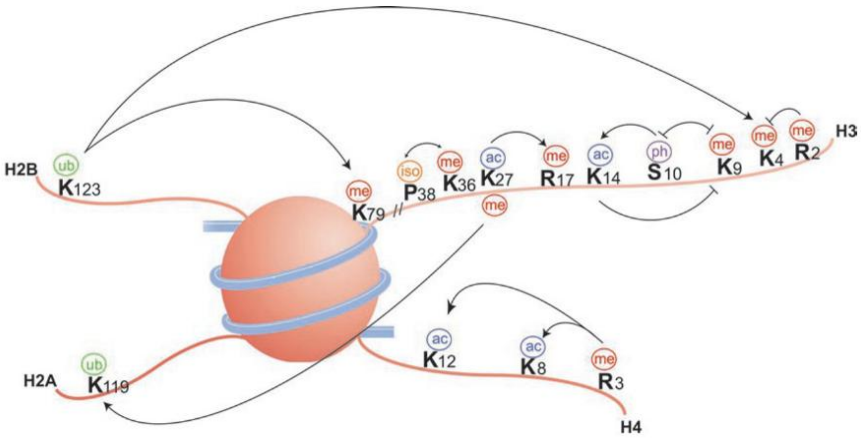


Figure 17. Histone modification crosstalk. A positive effect is indicated by an arrow head and a negative effect is indicated by a flat head. Extracted from (122)

In fact, the reciprocal interaction between histone PTMs is a specific case of a much broader phenomenon; the crosstalk of a network of epigenetic mechanisms, also including the methylation status of DNA and non-coding RNAs. DNA methylation can have a positive or negative effect on the action of writers, readers and erasers. For

example, UHRF1 ubiquitin ligase binding to nucleosomes carrying H3 lysine 9 trimethylation (H3K9me3) is stronger when the nucleosomal DNA is CpG-methylated. In contrast, DNA methylation disrupts the binding of KDM2A lysine demethylase to nucleosomes bearing H3K9me3 (122). Non-coding RNAs can also interact with histone PTMs. For example, the non-coding RNA BORDERLINE controls the spatial localization of the chromatin-associated protein Swi6 and histone H3K9 methylation in yeast (123).

Figure 18. Epigenetic marks in transcriptionally active versus silenced genes. The active gene (top) and silenced gene (bottom) representations combine an overview of their epigenetic marks and chromatin immunoprecipitation sequencing (ChIP-seq) data. Endothelial nitric oxide synthase (eNOS) and myosin heavy chain (MHC) expression in human umbilical vein endothelial cells (HUVECs) serve as examples of active and silenced genes, respectively. Extracted from (111)

Importantly, although the effect of each epigenetic mark may be dependent on the biological context as well as on the interaction with other epigenetic marks, some epigenetic marks in specific genomic locations have been associated with gene activation or silencing (**Figure 18**). These specific genomic elements are principally: 1) the **gene body** itself, including exons and introns, 2) **enhancers**, regulatory DNA sequences in which specific transcription factors bind to facilitate or increase the

transcription of a gene, and 1) **promoters**, regulatory DNA sequences that define where transcription of a gene starts (111,124).

Active genes (**Figure 18**, upper panel) are typically enriched in H3 lysine 4 monomethylation (H3K4me1) and H3 lysine 27 acetylation (H3K27ac) in the enhancer regions; H3K4me3, RNA polymerase II (Pol II), and histone variant H2A.Z in the promoter regions; and H3 lysine 36 trimethylation (H3K36me3) in the gene bodies. At the same time, transcriptionally active genes present DNA hypomethylation in their promoters and hypermethylation in their gene bodies (111). Enhancer-derived RNAs (eRNAs) can also be found in active genes and may have a role in transcriptional regulation (125). Silenced genes (**Figure 18**, lower panel) are marked by H3 lysine 27 trimethylation (H3K27me3) in the enhancer and promoter regions, H3K9me3 in the gene bodies, DNA hypermethylation in the promoters, and hypomethylation in the gene bodies (111).

Histone methylation

Histone methylation, as well as non-histone protein methylation, can occur in all basic residues: lysines, arginines and histidines. While lysine methylation occurs in their ε-amine group, arginines are methylated in their guanidinyl group. Remarkably, methylated lysines exist in all three methylation states (monomethylation [me1], dimethylation [me2] and trimethylation [me3]), whereas arginines can be monomethylated, symmetrically dimethylated (me2s, with methyl groups on each of the two nitrogens) or asymmetrically dimethylated (me2a, with two methyl groups on a single nitrogen). Although there is evidence of the occurrence of histidine monomethylation, it seems to be rare and is not well characterized. The main histone methylation sites are lysines H3K4, H3K9, H3K27, H3K36, H3K79 and H4K20, as well as arginines H3R2, H3R8, H3R17, H3R26 and H4R3. Additionally, many other methylation sites have been reported, but their biological relevance remains to be clarified (126). Notably, the recently-discovered H2B lysine 5 monomethylation (H2BK5me1) is associated with active promoters downstream of the transcriptional start site and highly-expressed exons in the gene bodies, suggesting that it is a novel activation mark (127,128). Histone methylation is involved in the regulation of many basic biological processes, including transcription, RNA processing, gene silencing, DNA repair and subsequent genomic stability, DNA replication and DNA recombination, as exemplified in **Figure 19** for the major lysine methylation sites. Arginine methylation sites listed in **Table 3** illustrate the abundance of histone methylation sites identified to date.

Similarly to other PTMs, regulation of histone methylation is orchestrated by methyltransferases (writers), which transfer a methyl group to the substrate using *S*-adenosylmethionine (SAM) as a donor, demethylases (erasers) and readers (126).

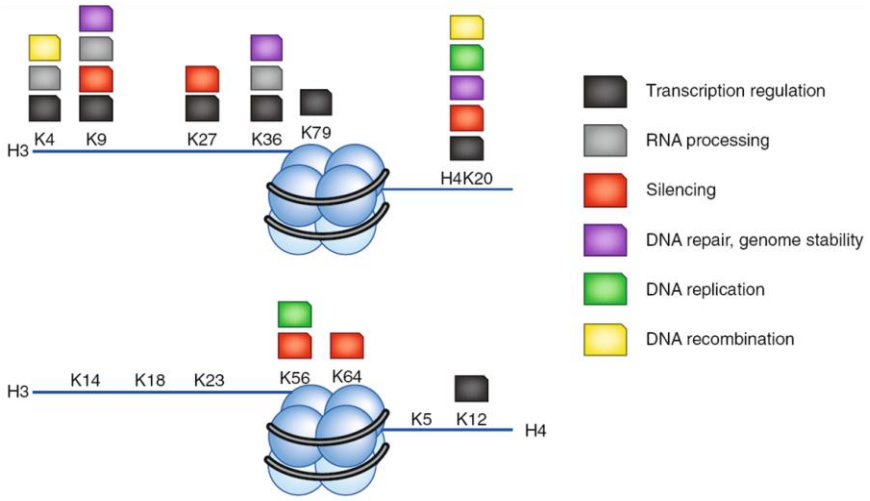


Figure 19. Principal sites of lysine methylation on mammalian histones and their associated chromatin functions. Canonical (top) and non-canonical (bottom) lysine methylation marks on histone 3 (H3) and 4 (H4) and the biological functions in which they participate (color code, right). Extracted from (129)

Site & type	Enzyme	Site & type	Enzyme
Histone H3		*Histone H4*	
R2me2a	PRMT6	R3me2a	PRMT1
R2me2s	PRMT5 PRMT7	R3me2a	PRMT6
R8me2a	PRMT2	R3me2s	PRMT5
R8me2s	PRMT5	R3me2s	PRMT7
R17me2a	CARM1	R17me1/me2	PRMT7
R26me2a	CARM1	R19me1/me2	PRMT7
R42me2a	CARM1 PRMT6	R23me1/me2	Unknown
R63me1	Unknown	R35me1	Unknown
R83me1/me2	Unknown	R55me1	Unknown
R128me1	Unknown	R67me1	Unknown
R134me1	Unknown	R92me1	Unknown
H2A		*H2B*	
R3me2a	PRMT6	R29me1	PRMT7
R3me2s	PRMT5	R31me1	PRMT7
R3me2s	PRMT7	R33me1	PRMT7
R11me1	PRMT1	R72me1	Unknown
R29me2a	PRMT6	R79me1	Unknown
R42me1	Unknown	R99me1	Unknown
R71me1	Unknown		
R88me1	Unknown		
R96me1	Unknown		

Table 3. Arginine methylation sites in histone tails and globular domains and methyltransferases (writers) that have been reported to act on them. Extracted from (130)

Writers, erasers and readers of histone methylation

Histone lysine methylation

Most lysine methyltransferases (writers) contain a conserved Su(Var)3-9, Enhancer-of-zeste and Trithorax **(SET) domain**, with the exception of H3K79-specific

methyltransferase **Dot1/Dot1L**, which lacks a SET domain. While the SET domain targets lysines in histone tails, Dot1/Dot1L is the only enzyme known to methylate a lysine residue (H3K79) in a globular domain. On the other hand, lysine demethylases (erasers) catalyze demethylation of substrates and can be divided into two families: 1) flavin-dependent KDM1, also termed **LSD amine oxidases**, including LSD1 and LSD2 subtypes, and 2) iron-dependent **Jumonji C (JMJC)-domain-containing dioxygenases**. Regarding lysine methylation readers, these proteins contain methyl-lysine-binding motifs, including chromo, tudor, PWWP, PHD, WD40, BAH, ADD, ankyrin repeat, MBT and zn-CW domains, and can distinguish target methyl-lysines depending on their methylation state (me1, me2, me3) and neighboring amino acid sequence (131–134).

Histone arginine methylation

Equivalently, the role of writers in histone arginine methylation is performed by protein arginine N-methyltransferases (PRMTs). PRMTs are classified as type I, II, III or IV enzymes. Type I and II are involved in gene transcription regulation via histone methylation, as well as other cellular functions via non-histone protein methylation. In contrast to lysine methylation, which is widely accepted as being a dynamic process, the reversibility of arginine methylation by arginine demethylases is not so clear. Indeed, deaminase enzymes peptidyl arginine deaminase 4 (PAD4) and Jumonji domain-containing protein 6 (JMJD6) are the only two arginine demethylases identified and their specificity is under investigation (133,135). Methyl-arginines are mainly bound by proteins (readers) containing Tudor family binding domains, although they can also be bound by some specific PHD and WD40 domains. The main Tudor domains that bind methyl-arginine motifs are SMN, SPF30, TDRD3, SND1, TDRD1, TDRKH, TDRD6 and TDRD9 (130). **Figure 20** summarizes the principal methylation sites on H3 and H4 tails, as well as writers and erasers for each specific residue.

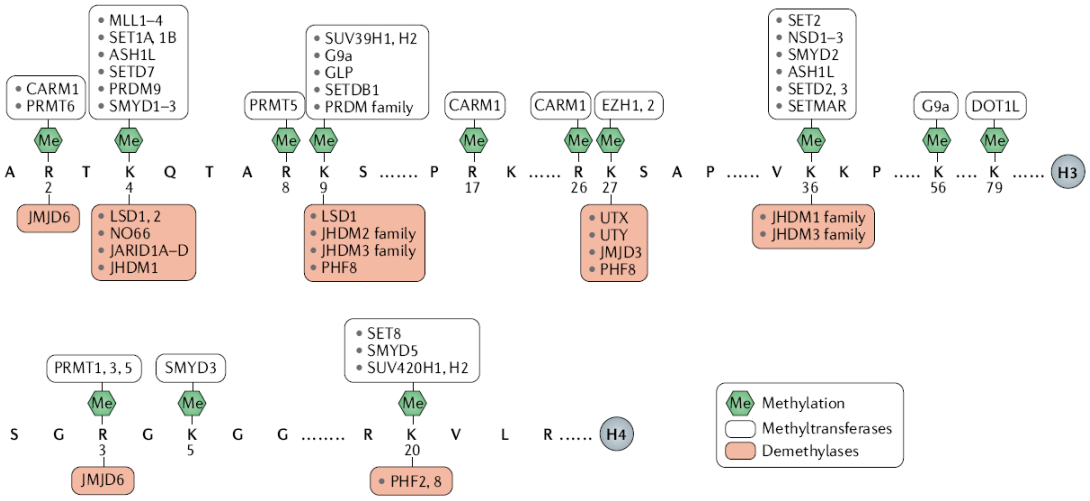


Figure 20. Principal methylation sites within the tails of histone 3 (H3) and 4 (H4), with writers (methyltransferases, above) and erasers (demethylases, below) indicated for each site. Extracted from (136)

Histone methylation and carcinogenesis

Importantly, a plethora of cancer types present deregulation of histone methyltransferases that generate epigenetic marks associated with either activation or inhibition of transcription (**Table 4**). Since tumorigenesis involves hyperactivation of oncogenes and silencing of tumor suppressors, a possible mechanism for this is that enzymes that add marks associated with gene activation (such as MLLs, NSD2, Smyd2 or Smyd3) may specifically target oncogenes, whereas those adding repressive marks (such as EZH2, G9a) would target tumor suppressor genes. In addition, as shown in **Table 5**, dysregulation of demethylases is also observed in many cancer types. The aberrations of histone methyltransferases and demethylases that have been identified in cancer are mutations, loci amplifications and gene fusions, as well as downregulation and overexpression. In some cases, downregulation is due to promoter methylation, but in other cases the molecular mechanism that leads to upregulation or downregulation of the enzyme remains unknown (137).

Non-histone proteins methylation and carcinogenesis

In addition to histone methylation-mediated modulation of chromatin structure, histone methyltransferases and demethylases also regulate gene expression by targeting non-histone protein substrates, including transcription factors, the activity of which can be modulated by methylation (**Table 6**). Many of these transcription factors have direct roles in proliferation control, thereby representing another layer of cancer development regulation (137).

Histone methyltransferase	Histone target	Examples of cancer types in which it is dysregulated	References
MLL / KMT2A MLL2 / KMT2D MLL3 / KMT2C	H3K4me1/2/3 H3K4me1/2/3 H3K4me1/2/3	Leukemia, large intestine, lung, endometrial, breast, bladder and brain cancers	(138)
EZH2 / KMT6	H3K27me2/3	Prostate, breast and ovarian cancers	(139)
NSD1 / KMT3B NSD2 / MMSET NSD3	H3K36me1/2/3 H3K36me1/2/3 H3K36me1/2	Leukemia, myeloma and prostate cancers	(140)
SETD2 / KMT3A	H3K36me1/2/3	Leukemia, glioma, melanoma, renal cell carcinoma, bladder, breast, colorectal, stomach, lung and uterine cancers	(141)
Suv39H2 / KMT1B	H3K9me2/3	Leukemia, lymphoma, hepatocellular carcinoma, gastric carcinoma, colorectal, lung, breast and ovarian cancers	(142)
GLP1 / KMT1D G9a / KMT1C	H3K9me1/2 H3K9me1/2	Myeloma, hepatocellular carcinoma, lung and brain cancers	(143,144)
SETDB1 / KMT1E	H3K9me2/3	Hepatocellular carcinoma, melanoma, colorectal, lung, ovarian and breast cancers	(145,146)
SET8 / PR-SET7 / KMT5A	H4K20me1	Leukemia, hepatocellular carcinoma, bladder, lung, breast, prostate and pancreatic cancers	(147)
DOT1L / KMT4	H3K79me1/2/3	Leukemia, gastric and colorectal cancers	(148–150)
PRDM1		Lymphoma, lung and colorectal cancers	(151–153)
PRDM2 / KMT8	H3K9me1/2/3	Hepatocellular carcinoma, leukemia, lymphoma, melanoma, neuroblastoma, colorectal, ovarian and breast cancers	(154)
PRDM5		Hepatocellular carcinoma, leukemia, glioma, myeloma and gastric cancers	(155–158)
PRDM12		Leukemia	(159)
PRDM14		Leukemia, breast and pancreatic cancers	(160)

Table 4. A list of histone methyltransferases, their targets (if known) and some examples of cancer types in which their dysregulation is observed. Modified and updated from (137)

Histone demethylase	Histone target	Examples of cancer types in which it is dysregulated	References
LSD1 / KDM1A	H3K4me1/2 H3K9me1/2	Leukemia, lymphoma, myeloma, hepatocellular carcinoma, prostate and breast cancers	(161,162)
JHDM1A / KDM2A	H3K36me2	Prostate and breast cancers	(163)
JHDM1B / KDM2B	H3K36me1/2 H3K4me3	Leukemia, breast, cervical, ovarian, pancreatic, gastric, lung, bladder and prostate cancers	(164)
JMJD1A / KDM3A	H3K9me1/2	Hepatocellular carcinoma, neuroblastoma, myeloma, colorectal, prostate, breast, lung, bladder, pancreatic and ovarian cancers	(165)
JMJD2B / KDM4B	H3K9me2/3 H3K36me2/3	Hepatocellular carcinoma, leukemia, myeloma, neuroblastoma, prostate, colorectal, gastric, bladder, lung, breast and ovarian cancers	(166)
JMJD2C / KDM4C	H3K9me2/3 H3K36me2/3	Leukemia, lymphoma, prostate, colorectal, gastric and breast cancers	(167)
JARID1A / KDM5A	H3K4me2/3	Leukemia, breast and pancreatic cancers	(168)
JARID1B / KDM5B	H3K4me1/2/3	Hepatocellular carcinoma, melanoma, prostate and ovarian cancers	(168)
JARID1C / KDM5C	H3K4me2/3	Breast and prostate cancers	(168)
UTX / KDM6A	H3K27me2/3	Leukemia, colorectal, pancreatic, breast and bladder cancers	(169)
JMJD3 / KDM6B	H3K27me2/3	Leukemia, lymphoma, glioma, melanoma, lung, pancreatic, breast and prostate cancers	(170)

Table 5. A list of histone demethylases, their targets and some examples of cancer types in which their dysregulation is observed. Modified and updated from (137)

Methyltransferase	Demethylase	Substrate	Effect	References
Set9 / KMT7		TAF10	Pol-II association, transcription activation	(171)
	LSD1 / KDM1A	p53	Protein stabilization	(172)
Set9 / KMT7	LSD1 / KDM1A	NFκB	Protein stabilitzation by demethylation	(173,174)

Set9 / KMT7	LSD1 / KDM1A	Stat3	Repression of transcriptional activity by methylation	(175)
Set9 / KMT7	LSD1 / KDM1A	E2F1	Protein degradation by methylation	(176,177)
Set9 / KMT7	LSD1 / KDM1A	MYPT1	Protein destabilitzation by demethylation and cell cycle progression by affecting Rb1 and E2F activity	(178)
Set9 / KMT7	LSD1 / KDM1A	DNMT1	Protein stabilitzation by demethylation	(179)
Set9 / KMT7	LSD1 / KDM1A	HIF-1α	Protein stabilitzation by demethylation	(180)
Set9 / KMT7		Foxo3	Repression of transcriptional activity	(181)
PR-SET7/ KMT5A		p53	Repression of transcriptional activity	(182)
SMYD2 / KMT3C	LSD1 / KDM1A	p53	Repression of transcriptional activity by methylation	(183)
SetD6		NFκB	GLP interaction	(184)
SETD1A / KMT2F		Hsp70	Nuclear localization of Hsp70	(185)
EZH2 / KMT6		RoRα	Protein degradation	(186)
NSD1 / KMT3B	JHDM1A /KDM2A	NFκB	Increased transcriptional activity by methylation	(187)
G9a / KMT1C		MyoD	Repression of transcriptional activity	(188,189)
G9a / KMT1C		C/EBPβ	Repression of transcriptional activity	(190,191)
G9a / KMT1C		G9a	Interaction with HP1	(192)
G9a / KMT1C		p53	Inactivation of p53	(193,194)
Suv39H1 /KMT1A		PC2	Non-coding RNA interaction, subnuclear repositioning in three-dimensional structures	(195)

(Previous page) Table 6. Methyltransferases and demethylases involved in cancer and their non-histone targets, most of which are transcription factors. Modified and updated from (137), (196) and (197)

SMYD family proteins

The present work is focused on the study of the role of the SMYD5 protein in the liver and intestine. SMYD5 belongs to the SET and Myeloid, Nervy, and DEAF-1 (MYND) domain-containing (SMYD) family of proteins. The SMYD family comprises five members; SMYD1, SMYD2, SMYD3, SMYD4 and SMYD5 (198).

The SMYD family of proteins is a special class of lysine methyltransferases the catalytic SET domain (which often co-exists with post-SET and SET-I domains, both present in this family) of which is split by a MYND domain. The SET domain is known to have lysine-specific methyltransferase activity, while the MYND domain contains a zinc-finger motif that has been shown to bind proline-rich regions and mediate protein-protein interactions. Although the MYND domain is highly positively charged, which in some proteins contributes to DNA binding, only SMYD3 has been shown to directly bind DNA *in vitro* (199). SET domains often exist with post-SET, SET-I and pre-SET domains, which also aid in cofactor and substrate binding as well as protein stability, although SMYD proteins do not contain the pre-SET domain. In addition, a C-terminal domain (CTD) is present in SMYD1-4, but absent in SMYD5. The CTD is structurally similar to tetratricopeptide repeats (TPR) and is an important motif for binding of co-chaperones with heat shock protein-90 (Hsp90). Co-immunoprecipitation assays demonstrated that SMYD2 and SMYD3 interact with Hsp90 *in vitro*, and mass spectrometry and *in vitro* methylation assays defined K209 and K615 as Hsp90 residues methylated by SMYD2. Consistent with the lack of a TPR domain, SMYD5 did not interact with Hsp90. Proteomic analysis described the interactome of SMYD2, SMYD3 and SMYD5, although SMYD5 predicted interaction networks have not been verified experimentally (200).

SMYD family members are known to methylate histone residues associated with both transcriptional gene activation and repression, as well as non-histone proteins that regulate various cellular processes. SMYD1 has been described to methylate H3K4 and proteins such as skNAC and TRB3. SMYD2 also methylates H3K4, as well as H3K36 and a plethora of non-histone substrates, including Hsp90α, p53, RB, PARP1, MAPKAPK3, PTEN, ERα, BTF3, PDAP1, AHNAK, AHNAK2, MAPT, CCAR2, eEF2, NCOA3, STUB1, UTP14A, Six1, Six2, SIN3B and DHX15. SMYD3 has methylase activity in both H3 (H3K4) and H4 (H4K5 and H4K20), and in MAP3K2 and VEGF receptor 1 proteins. SMYD4 histone and non-histone targets are unknown and SMYD5 has been shown to methylate H4K20 (See “The SMYD5 protein” section). With respect to their role in gene expression, SMYD1, SMYD2 and SMYD3 can activate or repress gene expression depending on the biological context, while, to date, studies on SMYD4

and SMYD5 concluded that these enzymes are only capable of repressing gene expression. While SMYD1 and SMYD4 activity seems to inhibit growth, SMYD2 and SMYD3 can either promote or inhibit cell growth. The effect of SMYD5 on the regulation of cellular growth is unclear (198).

The Smyd gene family is present in all eukaryotes, since the *Saccharomyces cerevisiae* (yeast) SET5 and SET6 proteins also contain a putative zinc finger resembling the MYND-type zinc finger domain. Calpena *et al*. defined Smyd5 as the most evolucionary parsimonious (with the fewest evolutionary changes) member of the family, since all metazoan species included in their study had one Smyd5 gene (201). Indeed, all SMYD proteins maintain significant homology from fish to humans (198,202). Linear representation of the structural domains of the SMYD proteins and their percentage of homology between selected species used in biomedical research are depicted in **Figure 21**.

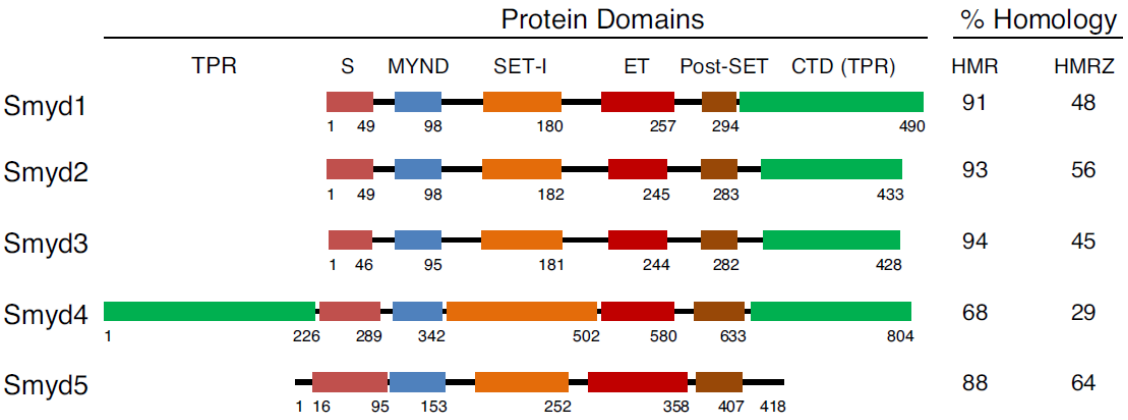
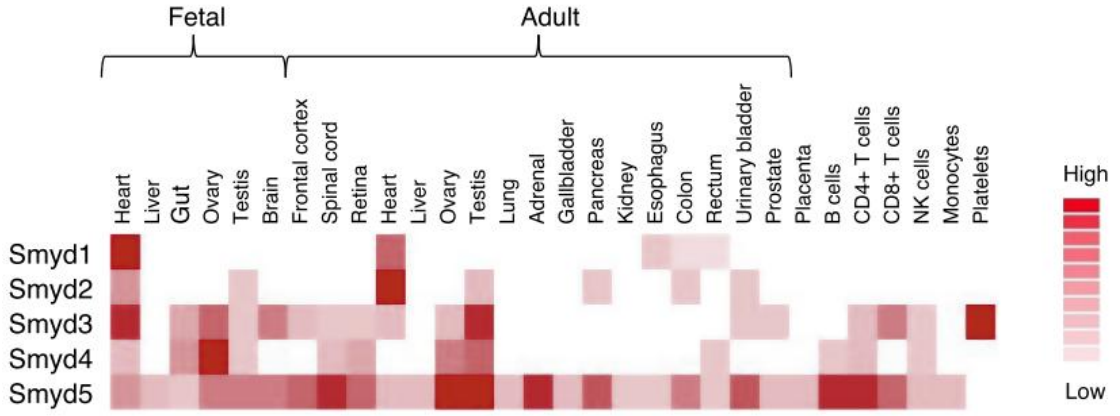


Figure 21. Linear representation of the structural domains of the SMYD proteins. The split Su(Var)3-9 Enhancer-of-zeste and Trithorax (SET; "S" and "ET", red), Myeloid Nervy and DEAF-1 (MYND) (blue), SET-I (orange), Post-SET (brown) and C-terminal domain (CTD) or tetratricopeptide repeats (TPR) (green) domains are indicated. The numbers represent the final residue number in each domain. Percent (%) homology is reported as a comparison for human, mouse, and rat (HMR) sequences as well as a comparison for human, mouse, rat, and zebrafish (HMRZ). Extracted from (198)



(Previous page) Figure 22. Expression of Smyd proteins in human tissue and cells. The relative expression levels of SMYD1, SMYD2, SMYD3, SMYD4 and SMYD5 across different tissues are shown as measured by mass spectrometry by (203). Each protein is normalized individually using the scale bar (right) indicating the degree of expression (low to high). Extracted from (198)

Importantly, SMYD proteins are expressed in a wide range of healthy fetal and adult human tissues and cells (**Figure 22**) and cancerous tissues (**Figure 23**). Remarkably, SMYD5 expression is considerably low in healthy adult liver and colon (**Figure 22**). In any case, SMYD proteins-mediated methylation of different targets is known to regulate tumorigenesis (204).

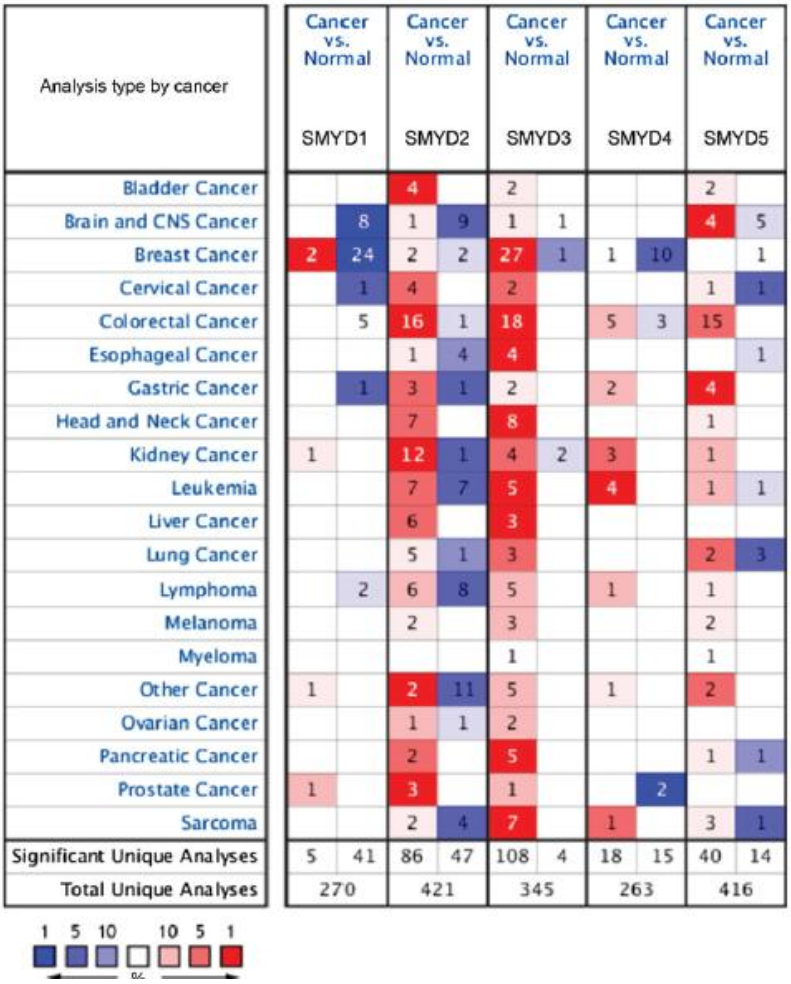


Figure 23. Expression levels of the SMYD family in different cancer types. Upregulation (red) and downregulation (blue) of the different members of the SMYD family of proteins in various different cancer types. All the alterations in expression identified were considered statistically significant ($P<0.05$). Cell color indicates the gene rank percentile. Image generated using ONCOMINE. Abbreviations: CNS, central nervous system. Extracted from (205).

Role of SMYD proteins in cancer

Many studies suggest that SMYD proteins function in cancer development. While mainly oncogenic roles have been proposed for SMYD2 and SMYD3, SMYD1 and SMYD4 may act as tumor suppressors. SMYD1 expression can be repressed by the mitogenic protein hepatoma-derived growth factor (HDGF) in G-7 myoblast cells (206).

Since HDGF is overexpressed in a number of types of human cancer, including HCC (207), SMYD1 interaction with HDGF suggests that SMYD1 may function as a tumor suppressor (205). Regarding SMYD2, Huang *et al.* demonstrated that SMYD2 methylates p53 on Lys 370, and this methylation repressed p53-mediated transcriptional regulation, providing a mechanism for SMYD2 as an oncogene (183). SMYD2 also methylates the tumor suppressor RB at lysine 860 (208) and PTEN at lysine 313 *in vitro* and *in vivo*. Knockdown of SMYD2 suppressed the cell growth of breast cancer cells through negative regulation of PTEN tumor suppressor activity (209). The CCAAT-enhancer-binding protein alpha (C/EBPα) 30 kDa isoform has been described as a negative regulator of SMYD2. Consistent with that notion, downregulation of C/EBPα accompanies SMYD2 upregulation in HCC samples (210). Concerning SMYD3, previous work by our group showed that *Smyd3* expression in mice is required for chemically-induced liver and colon cancer, since organ-specific KO mice are protected against the development of these pathologies. Nevertheless, SMYD3 is not sufficient for liver cancer development since spontaneous liver tumor formation was not detected in a transgenic murine model in which SMYD3 was constitutively overexpressed in hepatocytes. In these organs SMYD3 functions in the nucleus as a transcriptional potentiator of multiple cancer-promoter genes by invading active chromatin domains via association with H3K4me3 and RNA Pol-II. Remarkably, this association is mediated by its TPR domain, which is absent in Smyd5. Furthermore, *SMYD3* expression correlates with poor clinical prognosis of human HCC (211). In contrast to other SMYD family members, SMYD4 has been proposed as a potential tumor suppressor gene involved in breast tumorigenesis. Disruption of one allele of the *Smyd4* gene led to increased proliferation of a nontumorigenic mouse mammary epithelial cell line NOG8, which was inhibited by re-expression of *Smyd4 in vitro* and in xenograft experiments in nude mice, an immuno-compromised mouse strain (212). Furthermore, these results are consistent with the low *Smyd4* expression described in breast cancer patient samples compared to normal tissues (205,212).

The SMYD5 protein

SMYD5, the main focus of the present study, is a novel, recently identified, SMYD protein family member. Its gene is located in chromosome 6 in mice (*Mus musculus*) and chromosome 2 in humans (*Homo sapiens*) and the protein mass is 47 kDa. In 2012, Stender *et al.* published the first article describing a mechanism of action for SMYD5 (213). They proposed a role for SMYD5 in the negative regulation of inflammatory response genes (summarized in **Figure 24**). According to their data SMYD5 is involved in the active basal repression of Toll-like receptor (TLR)4-responsive promoters in macrophages. SMYD5 is associated with the nuclear receptor corepressor (NCoR) complex and is recruited to TLR4-responsive genes where it establishes histone 4 lysine 20 trimethylation (H4K20me3), a mark for transcriptional repression generally associated with heterochromatin. The study also described a mechanism for the activation of TLR4-dependent genes involving PHD finger protein 2 (PHF2) demethylation of H4K20me3. Of note, the authors' conclusions were based on biochemical assays and *in vitro* studies using primary murine macrophages and cell lines such as human embryonic kidney HEK293 cells (213). Therefore, no data regarding the *in vivo* role of SMYD5 was provided. We also note that mass spectrometry analyses of purified NCoR complexes by different investigators failed to identify SMYD5 as a *bona fide* subunit of NCoR complex. Consequently, we assume that the recruitment of SMYD5 in the TLR4-responsive genes is likely due to indirect interaction with either the NCoR complex or other transcription factors.

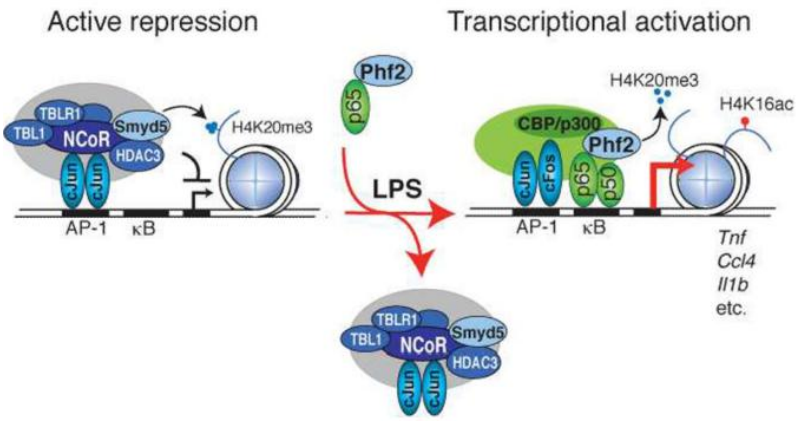


Figure 24. Role of histone 4 lysine 20 trimethylation (H4K20me3) in regulation of Toll-like receptor (TLR)4-responsive genes. Scheme of an integrated model for SMYD5 and PHF2 regulation of inflammatory gene promoters proposed by (213). Extracted from (213).

In 2016, an *in vivo* study in zebrafish was published (214). The authors selected zebrafish as a model because all five members of the Smyd family have been identified in this organism, being Smyd5 one of the most conserved members with around 70%

identity and 90% similarity to human homolog protein (215). Quantitative real time polymerase chain reaction (qPCR) and whole-mount *in situ* hybridization experiments performed in whole zebrafish embryos concluded that *smyd5* expression peaked at early developmental stages (3 hours post fertilization) and decreased in later stages. However, *smyd5* mRNA was present in all the adult tissues examined, including liver and gut, albeit at low levels. Smyd5 expression was abolished using two different approaches (morpholino oligonucleotides and CRISPR/Cas9 [which stands for clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9] targeted mutation) and neither gross morphological phenotype nor heart or skeletal muscle abnormalities were observed in the *smyd5* knockdown embryos. The authors claimed that Smyd5 is crucial for both primitive and definitive hematopoiesis, but this statement was only based on the enhanced expression of primitive and definitive hematopoietic markers in *smyd5* knockdown embryos and the restoration of normal marker levels by injection of *smyd5* mRNA. Importantly, nearly all the embryos analyzed displayed the same changes in hematopoietic markers (214). Nevertheless, despite the high degree of consistency of results, the biological implications of the deregulated expression of these markers remain to be clarified. Furthermore, the study does not relate the phenotype to Smyd5 enzymatic activity.

The role of SMYD5 and H4K20me3 in embryonic stem cells (ESCs) self-renewal and differentiation has been explored in two studies published by the same research group (216,217). In the first study (216) the authors performed RNA sequencing (RNA-seq) and observed high *Smyd5* expression in the murine ESC ES-R1 cell line and subsequent downregulation upon differentiation. Knockdown of *Smyd5* with a lentivirus encoding short hairpin RNA (shRNA) in this murine ESC cell line lead to normal ESC colony morphology loss as well as a remarkable reduction of alkaline phosphatase staining, a marker of undifferentiated ESCs. Restoration of normal SMYD5 levels by overexpression of an shRNA-resistant version of the protein reversed the phenotype. RNA-seq revealed global gene expression changes in *Smyd5*-depleted ESC colonies compared to wild-type (WT), including downregulation of the expression of pluripotency regulators such as *Pou5f* (Oct4), *Nanog* and *Tbx3*. Moreover, since many upregulated genes are also enriched in committed lineages, the authors hypothesized a role for SMYD5 in repressing expression of lineage-specific genes in ESCs. Indeed, embryoid body induction and teratoma formation assays confirmed that ESCs differentiation is altered in the absence of SMYD5. Biotin-mediated ChIP-seq and FLAG ChIP-seq revealed that this abnormal differentiation was accompanied by deregulated gene expression and decreased global H4K20me3 levels, that were restored upon

overexpression of an shRNA-resistant version of SMYD5. Moreover, an important percentage of SMYD5-occupied regions contained H4K20me3, H3K9me3 (described as a repression mark) and repetitive DNA elements from the LINE/LTR repeat subfamilies. Furthermore, their results demonstrate that shSmyd5 ESCs display increased expression of repetitive DNA sequences, decreased levels of H4K20me3 and H3K9me3/2, and abolished HIP1α and G9a binding to DNA. The authors also proved that HIP1α and G9a proteins interact with SMYD5 in the nucleus of ESCs. Further analysis of the role of DNA repetitive elements suggested a model in which repetitive DNA elements recruit SMYD5 to the vicinity of differentiation genes in order to maintain them silenced. Overall, this study describes a role of SMYD5 in regulating ESC maintenance by silencing lineage-specific genes (216).

In their second study (217), the authors extended the embryoid body assays from 14 to 21 days and observed the formation of transformed cells during differentiation in the absence of SMYD5. These shSmyd5 cells were capable of proliferating in suspension as well as a monolayer or as clusters of cells for more than 2 months *in vitro*, and displayed a high proliferative rate and proteolytic activity. Depletion of SMYD5 in human H1 embryonic stem cells (hESCs) led to a similar transformed cell phenotype. Xenograft assays consisting of the injection of murine shSmyd5 transformed cells into immunodeficient mice led to the formation of tumors *in vivo*. Thus, these cells can be considered cancer cells, bearing chromosomal aberrations and copy number alterations. Whole genome DNA sequencing revealed that copy number alterations in shSmyd5 cancer cells are associated with diminished H4K20me3/H3K9me3 and are enriched with LINE and LTR repetitive elements. Furthermore, the shSmyd5 cancer cell gene expression signature was obtained by RNA-seq. Kaplan-Meier analysis found that lung adenocarcinoma, colon adenocarcinoma and invasive breast carcinoma patients with low expression of shSmyd5 cancer cell–repressed genes had a decreased rate of survival when compared to patients with a high expression profile of shSmyd5 cancer cell-repressed genes, suggesting that a shSmyd5 cancer cell signature predicts patient survival and loss of SMYD5 is associated with cancer formation. Moreover, depletion of SMYD5 led to partial malignant transformation of the nontumorigenic epithelial cell line MCF10A and increased upregulation of a subset of already-overexpressed genes in human HCT-116 colon cancer, A549 lung cancer and MCF7 breast cancer cell lines. Notably, xenograft assays performed by injecting SMYD5-depleted human HCT-116 colon cancer and A549 lung cancer cell lines into immunodeficient mice led to increased tumor growth *in vivo*, compared to the injection

of the same cancer cell lines expressing endogenous SMYD5 (217). These data implies that SMYD5 may behave as a tumor suppressor in this biological context.

Two recently-published integrative bioinformatics analyses suggested a role for *SMYD5* expression levels in the diagnosis and prognosis of breast (205) and gastric (218) cancer using patient samples, supporting the notion that SMYD5 may function in cancer development and emphazising the clinical relevance of SMYD5. On one hand, Song *et al.* explored publicly available databases (ONCOMINE and The Cancer Genome Atlas data [TCGA]) and identified no significant differences in *SMYD5* expression between breast cancer and normal tissues. *SMYD5* mRNA was increased in specific breast cancer subtypes, such as HER2-positive and triple-negative breast cancer. Furthermore, decreased *SMYD5* mRNA was associated with good relapse-free survival in patients with breast cancer (205). On the other hand, Meng *et al.* used datasets from public databases (ONCOMINE, TCGA and Gene Expression Omnibus [GEO]) to define the role of histone modification-associated genes in gastric cancer. *SMYD5* was upregulated in gastric cancer compared to normal tissues in all three datasets. Aberrant expression of *SMYD5* was more obvious in gastric intestinal-type adenocarcinoma than in other subtypes. According to Kaplan-Meier plotter analysis, patients with high *SMYD5* expression displayed a poorer overall survival than those with low expression. Nonetheless, no significant associations were observed between *SMYD5* and a number of clinicopathological characteristics including tumor size and clinical scores (218). These studies suggest a role for SMYD5 in promoting carcinogenesis, as has been proposed for SMYD2 and SMYD3 by our group (211) and others (see “Role of SMYD proteins in cancer” section).

In contrast, genetic screening in mice defined SMYD5 as a suppressor of metastatic activation (219). The process of colonization of distant organs by cancer cells involves adaptation to each specific microenvironment, survival during dormancy and subsequent reactivation and outgrowth of these cancer cells. Only a small percentage of cancer cells acquire the genetic and epigenetic alterations necessary to complete each step. Gao *et al.* used a mouse xenograft model in which breast cancer cells were injected intravenously to syngeneic mice (immuno-competent animals from the same genetic background as the breast cancer cells donors) in order to mimic metastatic dormancy and reactivation of breast cancer to produce lung metastases. Loss-of-function shRNA library screen recovered shRNA targeting *Smyd5*, suggesting that *Smyd5* inactivation is required to exit dormancy. Ectopic overexpression of *Smyd5* was achieved by introduction of vectors encoding the gene in mammary tumor cells and injection of these cells into mice revealed that expression of *Smyd5* completely

suppresses the ability of these cells to metastasize to the lung. Although *Smyd5* was found to be upregulated in apocrine tumors (apocrine breast carcinoma) compared with tumors of a different origin (basal-like or luminal-like subtypes of invasive breast carcinoma), the results of this study strongly support the hypothesis that *Smyd5* mediates tumor dormancy (219).

In addition to all the cancer-related data, there is a study indicating the possible role of Smyd5 in intestine stem cell maintenance (220). Fevr *et al.* reported that intestinal stem cells are induced to terminally differentiate upon deletion of β-catenin in mice, resulting in a complete block of intestinal homeostasis and fatal loss of intestinal function. To characterize the expression changes regulated by the Wnt/β-catenin pathway, the authors isolated mRNA from crypt cells from WT and β-catenin KO mouse intestinal epithelium and performed microarrays. Then, in order to obtain an indication of which of the genes regulated by the Wnt/ β-catenin pathway could be involved in stem cell maintainance, they generated a list of downregulated genes in β-catenin KO crypts which overlap with genes characteristic for other stem cell populations and intestinal tumors. *Smyd5*, as well as *Smyd2*, were included in the list, meaning that these genes could be involved in stem cell mainteinance (220).

Furthermore, Nestorov *et al.* analyzed mRNA expression of different genes by single-cell qPCR at early stages of development (221). *Smyd5* expression was ubiquitous, meaning that it was expressed in the maternal cells as well as the zygotic cells, during preimplantation development. Later, when the blastomer was formed, *Smyd5* positively correlated with expression of a set of pluripotency genes in the inner cell mass (a mass of cells inside the primordial embryo that will form the entire mouse) but not with expression of a set of genes overexpressed in the trophectoderm, that will give rise to the placenta (221,222), suggesting that it might be a stemness-related gene.

Aim of the study

Based on the functions of the SMYD family, the involvement of methyltransferases in many biological processes, including cancer, and the scarce publications studying SMYD5, we hypothesized that SMYD5 could play a role in promoting carcinogenesis and/or stem cell maintainance *in vivo*.

Hence, the main purpose of this study was to investigate the role of SMYD5 in liver cancer in mice. Specifics aims included the following:

1) To identify SMYD5-mediated methylation targets *in vitro*.
2) To analyze SMYD5 expression levels in human cancer samples and their correlation with clinical parameters.
3) To generate SMYD5 KO murine models.
4) To determine the role of SMYD5 in chemically induced *in vivo* models of liver and colon carcinogenesis using SMYD5 KO mice.
5) To characterize liver cancer in liver-specific SMYD5 KO murine models at the histological and gene expression levels.

Materials and Methods

Mice and *in vivo* treatments

Mice carrying the *Smyd5* targeted allele (*Smyd5*$_{tm1a(EUCOMM)Wtsi}$) in C57BL/6N background (generated by the European Conditional Mouse Mutagenesis Program and supplied by Wellcome Trust Sanger Institute) were crossed with Flipase-expressing mice (003946, Jackson Laboratories), Alb-Cre mice (223–225), Alfp-Cre mice (226) and Villin-Cre mice (227), or MX-Cre mice (002527, Jackson Laboratories) to obtain appropriate *Smyd5* KO models (see Results: "Generation of SMYD5 knockout mouse models" section).

Mice were maintained in group cages in a temperature-controlled, pathogen-free facility on a 12/12-hour light/dark cycle and fed with a standard chow diet (Altromin 1324; 19% protein, 5% fat) and water *ad libitum*. The following primer sets were used for *Smyd5* KO mice genotying:

1) To detect the WT allele:

SMYD5_F 5'-GGTCTCATGGGGAACTGAGG-3'
SMYD5_R 5'-GCTTTCAGCCAAGCCAAGTC-3'

2) To detect the transgenic cassette (could be combined with (1) to detect heterozygotes):

SMYD5_F 5'-GGTCTCATGGGGAACTGAGG-3'
CAS_R1_Term 5'-TCGTGGTATCGTTATGCGCC-3'

3) To detect the transgenic cassette after FRT site recombination:

SMYD5_F 5'-GGTCTCATGGGGAACTGAGG-3'
SMYDR1 5'-GGCGAGCTCAGACCATAACT-3'

4) To detect the WT allele and the transgenic cassette before LoxP recombination (this PCR is unable to distinguish the presence or absence of FRT sites; for this purpose there is [3]):

SMYD51 5'-GGGCTCTGGCTACCTTTGTG- 3'
SMYDR2 5'-GTTTCAGTCGGGATGTCTAAG-3'

To induce liver tumors with DEN, we treated mice with a single intraperitoneal injection of 25 mg/kg DEN (N0756; Sigma) at postnatal day 14 (P14) and euthanized them at 8

or 10 months of age (tumors are macroscopically visible in WT C57BL/6 livers at least 7.5 months after DEN treatment).

Treatments for colitis-associated colon cancer were as follows. Two-month-old mice were treated with 30 mg/kg DMH (D161802, Sigma) by intraperitoneal injection. One week after injection animals were given 2% (w/v) DSS (42867, Sigma) in drinking water for 6 days followed by pure drinking water (without any additional substance) for 14 days. The 6-day DSS cycle was repeated two more times with 14-days interval of pure drinking water, after which the mice were euthanized for analysis of tumor formation.

Short-term CCl_4 treatment for damage-induced liver regeneration consisted of a single intraperitoneal injection of 1 ml/kg CCl_4 (289116, Sigma) diluted in corn oil. Mice were euthanized 24, 48, 72 hours or 7 days after treatment.

All animal experiments were approved by the Prefecture of Attica and the Institutional Review Board of the Biomedical Sciences Research Center “Alexander Fleming” and were performed in accordance with the respective national and European Union regulations. All the experiments were performed in randomly chosen age-matched male mice. No blinding was used in this study.

RNA and DNA analysis

Genomic DNA isolation from mouse tail

A small piece of tail was cut at postnatal day 10-15 (P10-P15) mice. The tissue was placed in a tube containing 600 μl TNES buffer (50 mM Tris pH 8, 400 mM NaCl, 100 mM EDTA, 0.5% SDS) and 10 μl of freshly added (just before use) Proteinase K (10 mg/ml). The solution was incubated overnight (O/N) at 55°C to allow proteins to be digested by the enzyme. Then 167 μl 6M NaCl were added, and the sample was shaken for 15 sec and centrifuged at maximum speed for 5 min. The supernatant was transferred to a clean eppendorf tube. DNA was precipitated with 95% ethanol and centrifuged. The pellet was washed with 70% ethanol and resuspended in 100 μl ddH_2O. 50-100 ng genomic DNA were used to perfom PCR for genotyping (See below).

Total RNA isolation from mouse liver

Total RNA from mouse liver was isolated according to the manufacturer’s instructions for TRIzol (TRI Reagent® Solution AM9738, ThermoFisher), a widely used product for isolation and the first purification step. After liver resection from mouse, we washed the organ with 1x PBS and placed a piece of the liver (usually an entire middle lobe) in 4 mL TRIzol and homogenized it with a Polytron homogenizer. Chloroform was added at

a proportion of 200 µl per 1000 µl TRIzol, and the sample was shaken manually and vortexed for 10 sec. After a 3 min-incubation at room temperature (RT), the sample was centrifuged for 15 min at 4ºC at 12000 g, and the supernatant was transferred to a clean eppendorf tube to which 100% ethanol and 1/10 volume 3 M CH3COONa pH 5.8 were added in order to precipitate RNA. After manual shaking the sample was placed at -20ºC at least O/N or -80ºC for long-term storage. Then we centrifuged the eppendorf tube at 4ºC at 14000 rpm for 30 min, we removed the supernatant and rinsed the pellet with 70% ethanol. The pellet was then washed once more, centrifuged again and, after supernatant removal, the pellet was dried in a speedvac centrifuge to completely eliminate the ethanol. The pellet was resuspended in 200-300 µl double distilled water (ddH_2O) and RNA was quantified by photometry (NanoDrop™ Spectrophotometer) at 260 nm and RNA quality control was performed by running a few microliters on a 2% agarose gel.

Total RNA Purification with DNAase and reverse transcription

Forty µg of total RNA were incubated at 37ºC for 30 min with 1 µl DNaseI RNase-free (10 u/µl) (04716728001, Roche) diluted in the corresponding buffer supplied by the company, and 0.5 µL RNaseOUT (RNAse inhibitor) (10777-019, Invitrogen) per 50µL reaction. After incubation, we added the appropiate volume of ddH2O to lose as little sample as possible and then washed the sample firstly with phenol: chloroform: isoamyl alcohol (volume ratio 25: 24: 1) and secondly with chloroform: isoamyl alcohol (24: 1). Afterwards, we added 2.5x volumes 100% etanol and 1/10 volume sodium acetate (3 M CH3COONa pH 5.2) to the supernatant and induced RNA precipitation by storing the sample at -80ºC for 1 hour or at -20ºC O/N. Then the sample was centrifuged for 30 min at 4ºC at 14000 rpm, and the pellet was rinsed with 70% ethanol and resuspended in ddH2O (~ 50-100 µL).

We used 1 µg of this DNAseI-treated RNA for reverse transcription reaction, DNaseI-treated as previously described. We added 1.2 µL OligoDT random primers (100 ng/µl) (79237, Qiagen), 2 µL 10 mM dNTPs (4 dNTPs mixture, 10297-018, Invitrogen), and ddH2O until achieving a final volume of 19 µL. The mixture was incubated for 5 min at 65ºC and cooled for 5 min RT. Finally, we added 1 µL 5x Moloney Murine Leukemia Virus (M-MLV) Reverse Transcriptase (M-MLV RT) (200 u/µl) (28025-013, Invitrogen) and 0.1 µL RNaseOUT, and this was incubated at 37ºC for 1 hour to allow the enzymatic reaction to take place, and we stored the eppendorf tube at -20ºC. The final product is the so-called complementary DNA (cDNA).

Polymerase Chain Reaction (PCR)

In this study we used PCR technology to identify genes (genotyping PCR) or quantify RNA molecules (Reverse Transcriptase reaction followed by qPCR).

Regarding genotyping, we used a PCR thermocycler (Biorad) and followed this program: 1) DNA denaturation at 95ºC, 2) primer annealing at 54-60ºC, depending on the specific set of primers used, and 3) extension by Taq polymerase at 72ºC. The program was repeated for 30-40 cycles. Reactions were performed with 50-200 ng DNA at a final volume of 50 µL with the following composition: 5 µL 10x polymerase buffer (100 mM Tris pH 8.5, 500 mM KCl, 1% Triton X-100, 15 mM $MgCl_2$), 2 µL 5´ primer 10 µM, 2 µL 3´ primer 10 µM, 1 µL dNTPs mix (10 mM), 1 µL Taq DNA polymerase (5 u/µL) (Recombinant E.coli, 203-1, Minotech), 3 µL $MgCl_2$ (25 mM), 8 µL Betaine (1 M) (61962, Sigma), and ddH2O until 50 µL final volume. For 20 µL final volume the composition was as follows: 2 µL 10x polymerase buffer (100 mM Tris pH 8.5, 500 mM KCl, 1% Triton-X100, 15 mM $MgCl_2$), 0.5 µL 5´ primer 10 µM, 0.5 µL 3´ primer 10 µM, 0.5 µL dNTPs mix (10 mM), 0.4 µL Taq DNA polymerase (5 u/µL), 2.8 µL $MgCl_2$ (25 mM), 4 µL Betaine (1 M), and ddH2O until achieving a final volume of 20 µL. The primers used for *Smyd5*-deficient mice genotyping are listed in the section "Mice and *in vivo* treatments".

Quantitative Real Time Polymerase Chain Reaction (qPCR)

The qPCR program is the same as described above, but qPCR allows product quantification in each cycle. For qPCR we used the StepOne™ Real-Time PCR System and LightCycler® 96 Instrument thermocyclers. The reaction composition for 5 ng cDNA was as follows: 2 µL 10x polymerase buffer (100 mM Tris pH 8.5, 500 mM KCl, 1% Triton X-100), 0.3 µM 5´ primer, 0.3 µM 3´ primer, 100 µM 4 dNTPs, 1 µL Taq DNA polymerase (5 u/µL), 200 mM Betaine, 5 mM $MgCl_2$ and water until 20 µl. We also used the commercial ready-to-use mixture FastStart Universal SYBR Green Master (Rox) (04913914001, Roche), following the manufacturer's instructions, as an alternative to the previously described reaction composition. We used the commercial mixture, primers and cDNA to perform qPCR reactions at a final volume of 20 µL.

To study the differential expression of genes in a variety of conditions (including untreated and treated WT and *Smyd5*-deficient mice), we generated cDNA from RNA from isolated liver, colon, gallbladder or other organs from adult mice, as well as murine embryo liver, and then performed qPCR, by following the previously described protocols. We used 5 ng cDNA for each reaction, in 20 µL final volume, and the following qPCR primers to target the murine gene:

Sequences of qPCR primers used in this study

Smyd5 (exon 2)	NM_144918.2	5' GACAGCATCAAGGGAAAGGG 3'
		5' AAGTGCATTCCAGAGGAACT 3'
Smyd5 (exon 4-6)	NM_144918.2	5' GGCACCCCCTCAATAAGCTG 3'
		5' ACCCAGTGGTCCTTGTCCTT 3'
Rplp1	NM_018853.3	5' CTTCCGAGGAAGCTAAGGCCGCGTT 3'
		5' AGCAGTTTAGTCAAAAAGACCGAAG 3'
Gapdh	NM_008084	5' CCA ATG TGT CCG TCG TGG ATC T 3'
		5' GTT GAA GTC GCA GGA GAC AAC C 3'
Acly	NM_001199296.1	5' GGC TTC ATT GGG CAC TAC CTT G 3'
		5' GTA GGG CTC CTG GCT CAG TTA CA 3'
Cd36	NM_001159558.1	5' CAA AGT TGC CAT AAT TGA GTC C 3'
		5' CGA ACA CAG CGT AGA TAG ACC T 3'
Cyp8B1	NM_010012.3	5' GTA CGC TTC CTC TAT CGC CTG A 3'
		5' TGG AGG GAT GGC GTC TTA TG 3'
Fasn	NM_007988.3	5' GGA CTT GGG TGC TGA CTA CAA C 3'
		5' CCT CCC GTA CAC TCA CTC GT 3'
Apob	NM_009693.2	5' CCTTAGAAGCCTTGGGCACA 3'
		5' GGATTCGAGCACAGATGACC 3'
Cyclin A2	NM_009828.2	5' CAGCCTGCAAACTGTAAGGT 3'
		5' CAATGACTCAGGCCAGCTCT 3'
Cyclin B1	NM_172301.3	5' AGCAAATATGAGGAGATGTACC 3'
		5' CGACTTTAGATGCTCTACGGA 3'
Cyclin D1	NM_007631.2	5' GCC GAG AAG TTG TGC ATC TAC 3'
		5' GGA GAG GAA GTG TTC GAT GAA 3'

Cyclin E1	NM_007633.2	5' GTTCCAAGCTCAAGCACTTCC 3' 5' CACTCGGAGGAGGAGAAATCC 3'
Cyclin E2	NM_001037134.2	5' CTGCTGCCGCCTTATGTCAT 3' 5' TACACACTGGTGACAGCTGC 3'
c-Myc	NM_001177352.1	5' CAGCGACTCTGAAGAAGAGCA 3' 5' GACCTCTTGGCAGGGGTTTG 3'
Ctnnb1	NM_007614	5' TCT TGG ACT GGA CAT TGG TGC 3' 5' GGC CAC CCA TCT CAT GCT C 3'
Jak1	NM_146145.2	5' TTTAGTCCCATGGCCTTGTTCC 3' 5' CAGTTGGGTGGACATGGCAG 3'
Jak2	NM_001048177.2	5' GCAACGGAAGATTGCCAAGG 3' 5' CCGGATTTGATCCACCCGAA 3'
Snai1	NM_011427.2	5' CCACACTGGTGAGAAGCCATTC 3' 5' GACATGCGGGAGAAGGTTCG 3'
Snai2	NM_011415.2	5' CAAGGCTTTCTCCAGACCCT 3' 5' GCCCTCAGGTTTGATCTGTCT 3'
Twist1	NM_011658.2	5' GCTACGCCTTCTCCGTCTG 3' 5' AATGACATCTAGGTCTCCGGC 3'
Pcna	NM_011045.2	5' CACAAAAGCCACTCCACTGT 3' 5' TGCCTAAGATGCTTCCTCATC 3'
Timp1	NM_001044384.1	5' TTTCTCATCACGGGCCGC 3' 5' GAGCAGGGCTCAGAGTACG 3'
Sox4	NM_009238.2	5' TTTCAGCTCCTCATCGGCG 3' 5' CCAGGTTAGAGATGCTGGACTC 3'

Mmp3	NM_010809.2	5' CAATCCATGGAGCCAGGATTT 3'
		5' CTGCGAAGATCCACTGAAGAAG 3'
Mmp7	NM_010810.4	5' GTGAGGACGCAGGAGTGAA 3'
		5' CCGGGAACAGAAGAGTGAC 3'
Mmp8	NM_008611.4	5' TGGACCCAGGTTACCCCAA 3'
		5' ATTGTGGTCCACTGAAGAAGAG 3'
Mmp14	NM_008608.3	5' GCTCTCTTCTGGATGCCCAAT 3'
		5' CACTGCCCATGAATGACCCC 3'
Mmp23	NM_011985.3	5' GGGAAAGTATACTGGTACAAGG 3'
		5' CAGGAGTAGGTGCTGAGAA 3'
Cd45 (Ptprc)	NM_001111316.2	5' CCAGCAGACAGGGTTGTTCT 3'
		5' CGGGATAGATGCTGGCGATG 3'
F4/80 (Emr1)	NM_010130.4	5' CCATCCACTTCCAAGATGGGTTA 3'
		5' TGCCATCAACTCATGATACCCT 3'

In vitro methylation assay

In vitro methylation reactions were performed at a final volume of 20 µL with the following composition: 2 µl of 10x methylation buffer (500 mM Tris pH 9.1, 20 mM DTT and 40 mM EDTA, ddH_2O), 5 µl of the substrate (histone 4 [H4], histone 3 [H3], histone 2A [H2A], histone 2B [H2B], histone 1 [H1], a mixture of histones, P53 or E2F1; all were recombinant proteins, purified by previous laboratory members), 1 µl of human SMYD5 recombinant protein (0.3µg/µl) (31409, Active Motif), 2 µl of ^{3}H-labeled SAM and ddH_2O. The reactions were incubated at 30ºC for 1 hour shaking (300rpm) and terminated by the addition of 7.5µl 4x SDS Loading Buffer (200 mM Tris pH 6.8, 8% SDS, 40% Glycerol, 400 mM DTT, 0.4% bromophenol blue, ddH_2O). We loaded 10 µl from each reaction and in 12% agarose gel and performed electrophoresis at 120 Volt (Mini-PROTEAN® Tetra Handcast Systems, Biorad). Then the gel was stained for 20 min RT in Coomassie blue solution (1 L Coomassie blue solution: 2.5 g Coomassie bromophenol blue, 500 ml Methanol, 100 ml Acetic acid, 400 ml H_2O), washed with Destaining solution (1 L Destaining solution: 250 ml Methanol, 100 ml Acetic acid, 650

ml H_2O) to eliminate excess of Coomassie blue, and incubated for 1 hour at RT with gentle agitation in Amplify solution (Amplify Fluorographic Reagent, NAMP100, Amersham). The gel was dried at 80ºC for 1 hour prior to radioactive signal detection (KODAK X-OMAT 1000 machine).

Antibodies

The antibodies used in this study for immunostaining and/or Western Blot were from the following companies: Santa Cruz Biotechnology, anti-TfIIb (sc-225), anti-Gapdh (sc-32233), anti-HNF4α (sc-8987); Cell Signaling Technology, anti-phospho histone H2A.X (#9718); Abcam, anti-Ki67 (ab15580); Biolegend, anti-CD45 (#103101); AbDSerotec, anti-F4/80 (MCA497); BD Transduction Laboratories, anti-E-Cadherin (#610181) and DAKO, anti-CK19 (#A0575). Our Smyd5 antibody was raised in-house (in collaboration with Professor George Chalepakis research group, Department of Biology, University of Crete) in New Zealand White female rabbits, which were injected with full-length recombinant SMYD5 protein purified under native conditions.

Protein extracts and analysis

Whole Cell Extracts (WCE) preparation from murine liver

Freshly dissected liver pieces (usually half of the left lobe) were rinsed in 1x PBS and supplemented with 10 volumes of cold modified RIPA Buffer (50 mM Tris pH 7.5, 1% NP40, 0.25% Na-Deoxycholate, 150 mM NaCl, 1mM EDTA pH 8, 10% Glycerol, 1mM PMSF and 2 μg/ml Aprotinin) supplemented with 1 commercial protease inhibitor cocktail tablet (Complete mini EDTA-free, 11836170001, Roche) per 10 mL RIPA buffer. After homogenization with a Polytron homogenizer for 10 sec, the extracted tissue was incubated at 4ºC for 20 min with constant agitation. After centrifugation at 4ºC at 14000 rpm for 20 min extracted proteins were recovered and stored at -80ºC.

Whole Cell Extracts (WCE) preparation from cell culture

Cells cultured in dishes 10 cm in diameter were rinsed 3 times with 10 mL 1x PBS, collected by scraping in 1 ml 1x PBS in an eppendorf tube and centrifuged at 1500 rpm for 5 min at RT. After discarding the supernatant, the pellet was resuspended in modified RIPA buffer (50 mM Tris pH 7.5, 1% NP40, 0.25% Na-Deoxycholate, 150 mM NaCl, 1 mM EDTA pH 8, 10% Glycerol, 1 mM PMSF and 2 μg/ml Aprotinin) supplemented with 1 commercial protease inhibitor cocktail tablet (Complete mini EDTA-free, 11836170001, Roche) per 10 mL RIPA buffer. Optionally, cells could be

disrupted by sonication using a Bioruptor (3 cycles of 30 sec each, separated by 20-sec intervals to avoid sample heating) and then centrifuged at 14000 rpm for 20 min at 4ºC. The supernatant was collected in a clean eppendorf and stored at -80º C.

Nuclear Extracts (NE) preparation from murine liver

Liver tissue was washed with 1x PBS and the right and left lobe were minced into small pieces in 10 ml Sucrose buffer A (0.32 M sucrose, 15 mM Hepes pH 7.9, 60 mM KCl, 2 mM EDTA, 0.5 mM EGTA, 0.5 % BSA, 0.5 mM spermidine, 0.15 mM spermine, 0.5 mM DTT, 0.5 mM PMSF, 0.5 mM Aprotinin). After 10 strokes of homogenization with glass Teflon pestle, the sample was passed through a plastic mesh to remove agglomerates and extracellular material, followed by a second homogenization with Teflon pestle. The isolated nuclei were layered over equal volume of Sucrose buffer 2 (30% sucrose, 15 mM Hepes pH 7.9, 60 mM KCL, 2 mM EDTA, 0.5 mM EGTA, 0.5 mM spermidine, 0.15 mM spermine, 0.5 mM DTT, 0.5 mM PMSF, 0.5 mM Aprotinin) and centrifuged at 4ºC at 3000 rpm for 15 min. Pelleted nuclei were washed with 1x PBS, resuspended in equal volume of Buffer A (25 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.1% NP40, 1 mM DTT, 0.5 mM PMSF, 10 µg/mL Aprotinin) and transferred to a clean eppendorf tube containing equal volume of NLB800 buffer (25 mM Hepes pH 7.9, 10% Glycerol, 0.8 M KCl, 0.1% NP40, 0.2 mM EDTA pH 8.0, 1 mM DTT, 0.5 mM PMSF, 10 µg/mL Aprotinin). The sample was incubated at 4ºC for 20 min with constant agitation and then centrifuged at 14000 rpm for 20 min at 4ºC. The supernatant containing all nuclear proteins was transferred to a clean eppendorf tube and stored at -80ºC.

Nuclear and cytoplasmic extracts from cell culture

Cells cultured in 150 mm diameter dishes were rinsed 3 times with 15 mL 1x PBS, collected by scraping in 15-20 ml 1x PBS in a falcon tube and centrifuged at 2000 rpm for 5 min at 4ºC. After discarding the supernatant, the pellet was resuspended in 1 ml 1x PBS and transferred to an eppendorf tube. It was again centrifuged at 6000 rpm for 2 min at 4ºC. The supernatant was discarded and a double volume of Buffer A (25 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.1% NP40, 1 mM DTT, 0.5 mM PMSF, 10 µg/ml Aprotinin) was added to the pellet. After resuspension, 10 µl of the solution were mixed with 10 µl Trypan blue solution (T8154, Sigma) to observe isolated blue-stained nuclei under the microscope. We centrifuged the sample at 6000 rpm for 2 min at 4ºC. Firstly, we transferred the supernatant (cytoplasmic extract) to a clean eppendorf tube, added 2x SDS Loading Buffer (L.B.: 100 mM Tris pH=7.5, 4% SDS, 20% Glycerol, 200 mM DTT, 0.2% bromophenol blue) and placed it at -80ºC for long-term storage.

Secondly, in parallel, we added 1 ml Nuc Buffer (15 mM Hepes pH 7.5, 60 mM KCL, 15 mM NaCl, 0.34 mM Sucrose, 0.15 mM mercaptoethanol, 0.5 mM spermidine, 0.15 mM spermine) to the pellet, mixed manually and centrifuged at 6000 rpm for 2 min at 4 ºC. After discarding the supernatant, we added 2x SDS Loading Buffer (L.B.: 100 mM Tris pH=7.5, 4% SDS, 20% Glycerol, 200 mM DTT, 0.2% bromophenol blue) to the pellet (nuclear extract). If necessary, DNA was eliminated by sonicating the samples in a Bioruptor (3 cycles of 30 sec each, separated by 20-sec intervals to avoid sample heating) before storage at -80º C.

Western Blot

Protein extracts were quantified with Bradford reagent (1976, Bio-rad Protein assay, Biorad) and the desired quantity of protein (usually 40 ng protein per well) was mixed with equal volume of 2x SDS Loading Buffer (L.B.: 100 mM Tris pH=7.5, 4% SDS, 20% Glycerol, 200 mM DTT, 0.2% bromophenol blue). The samples were denatured by boiling at 100ºC for 5 min, vortexed vigorously and centrifuged to avoid losing part of the sample. Then the samples were separated by SDS-PolyAcrylamide Gel Electrophoresis (SDS-PAGE) for approximately 90 min at 210 Volt (Mini-PROTEAN® Tetra Handcast Systems, Biorad) and transferred to a nitrocellulose membrane (Biorad). The membranes were blocked by incubation in TBS-T buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 0.1% Tween 20) supplemented with 5% nonfat dry milk or 5% w/v BSA for 1 hour RT. The membranes were washed 3 times with 1x PBS, incubated for 1 hour RT or O/N at 4ºC with primary antibody diluted in 1% BSA w/v in 1x TBS-T, washed 3 times and incubated for 1 hour RT with horseradish-peroxidase (HRP)-conjugated secondary antibody diluted in 1% BSA w/v in 1x TBS-T. After 3 washing steps with excess 1x TBS-T buffer, the membranes were developed using ECL chemiluminescent kit (Amersham) and exposed to X-ray films (KODAK X-OMAT 1000 machine).

Histological analysis

For histological analysis, tissues were embedded in optimal cutting temperature compound (OCT), frozen in liquid nitrogen and stored at -80ºC. Alternatively, tissues were fixed with 4% formaldehyde in 1x PBS O/N, washed 3 times with 1x PBS, dehydrated and embedded in paraffin in the Biomedical Sciences Research Center “Alexander Fleming” Histology facility, in order to make paraffin blocks.

B-galactosidase staining

B-galactosidase activity was performed in 20 μm-thick cryosections (OCT embedded-tissue). After fixation in 0.5% glutaraldehyde for 20 min, the sections were washed with 1x PBS supplemented with 2 mM $MgCl_2$ and stained for 3 hours with 1 mg/ml X-gal (7240-90-6, Roche) diluted in X-gal staining solution (0.1 M sodium phosphate buffer pH 7.3 [containing equal molarity of Na_2HPO_4 and NaH_2PO_4], 2 mM $MgCl_2$, 0.01% Na deoxycholate, 0.02% NP-40, 5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$). Sections were counterstained with eosin and observed with light microscopy.

Hematoxilin-Eosin staining

5 μm-paraffin sections were cut and dried for 2 hours at 37ºC. The slides were immersed in the following containers for the indicated periods of time: Xylene 1, 2 and 3, for 2.5 min each, 100% EtOH for 2 min, 96% EtOH for 2min, 96% EtOH for 2 min, 70% EtOH for 2 min, 50% EtOH for 2 min, H_2O 1 and 2 for 2 min each, Hematoxilin for 2.5 min, washes with H_2O, Acid/70 % EtOH for 1 sec, washes with H_2O, Scot's Salt for 3 min, washes with H_2O, 50% EtOH for 30 sec, 70% EtOH for 30 sec, Eosin for 1.5 min, washes with H_2O, 50% EtOH for 30 sec, 70% EtOH for 30 sec, 96% EtOH for 30 sec, 96% EtOH for 30 sec. Slides were mounted with DPX medium (06522, Sigma-Aldrich) and a coverslip. When DPX medium had solidified, the slides were ready for long-term storage or observation with light microscopy.

Immunofluorescence in cryosections

OCT blocks containing tissues were transferred from -80ºC to the cryotome (-20ºC), in which 5 μm cryosections were cut and placed in positively-charged surface slides for better adhesion of tissues. After 5 min of incubation at RT, the slides were fixed with 4% formaldehyde diluted in 1x PBS for 10 min RT, washed 3 times with 1x PBS and permeabilized with 0.1-0.2% v/v Triton X-100 diluted in 1x PBS for 10 min RT. The slides were washed with 1x PBS, blocked with Blocking buffer (5% BSA, 0.1% Triton, in 1x PBS) for 30 min RT and incubated with primary antibody diluted in Blocking buffer (1:50-1:500 dilution) O/N at 4ºC in a humidity chamber. After 3 washes with 0.1% Triton diluted in 1x PBS, the slides were incubated with fluorochrome-conjugated secondary antibody (1:500 dilution, Molecular Probes) and DAPI (1:2000 dilution, 4083, Cell signaling), both diluted in Blocking buffer for 1 hour at RT. Then slides were washed 3 times with 1x PBS and mounted with Mowiol mounting medium (81381, Sigma-Aldrich) and a coverslip, and stored at 4ºC protected from light.

Reactive Oxygen Species (ROS) measurement

Accumulation of ROS was measured by staining 5 μm frozen liver sections (cryosections) with 2 mM of 5-(and-6)-chloromethyl-20,70-dichlorodihydrofluorescein diacetate, acetyl ester (CM-H2DCFDA) for 30 min at 37ºC, according to the manufacturer's instructions (C6827, Invitrogen).

Image acquisition and quantification

Fluorescence and bright-field images were acquired using the Zeiss Axioscope 2 Plus Microscope (Carl Zeiss). Stained nuclei quantification was performed semi-automatically, aided by Image J software (National Institutes of Health).

Biochemical analysis

Blood was collected from the heart with the anti-coagulant heparin and centrifuged at 1000 g for 10 min. Freshly isolated supernatant serum fractions were used for measuring alanine aminotransferase (ALAT) activity using the ALAT assay kit (1 2701, Diasys), according to the manufacturer instructions.

Statistical analysis

All statistical analyses were performed with GraphPad Prism graphing and statistics software, version 5.0. The significance level threshold for all tests was 0.05.

Other techniques performed by our collaborators

The Cancer Genome Atlas (TCGA) data analysis was done by Antonis Giakountis.

Villae and crypt isolation and subsequent RNA-seq of each fraction were done by Antonis Giakountis, Vassiliki Zarkou and Pantelis Hatzis.

RNA-Seq was done in the Biomedical Sciences Research Center "Alexander Fleming" Genomics Facility by Vaggelis Harokopos.

RNA-Seq data analysis was done by Panagiotis Moulos.

Results

Smyd5 is expressed in many wild-type C57BL/6 adult male mouse tissues in homeostasis and upon different insults

As mentioned above (see Introduction), data regarding SMYD5 biological function are scarce. Thus, first we wanted to determine *Smyd5* mRNA and protein levels in a range of WT C57BL/6 adult male mouse tissues in homeostasis and upon different insults, by qPCR and Western blot, respectively. Unfortunately we failed to find a reliable antibody against SMYD5 that could be used for immunostaining. Regarding Western blot, we had some difficulties with liver cell lines, especially when using murine tissue samples (see below and Discussion).

For that reason we restricted our analysis to mRNA levels in mice. As shown is **Figure 25A**, many WT adult tissues (postnatal day 60 [P60]), including the liver and the colon, expressed *Smyd5* at different levels, similarly to what has been shown for WT neonatal (postnatal day 4 [P4]) tissues (228). For this analysis we used the gene encoding the ribosomal protein lateral stalk subunit P1 (*Rplp1*) as a housekeeping gene since we observed very low variation of its expression among tissues. Conversely, expression of the genes encoding the cytoskeleton component *B-actin*, the component of the purine salvage pathway hypoxanthine phosphoribosyltransferase 1 (*Hprt)*, ribosomal protein L13a (*Rpl13A*) and glycolysis pathway enzymes phosphoglycerate kinase 1 (*Pgk1*) and glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) was very different among tissues (data not shown). The variability in the expression levels of these genes was not due to different amounts of mRNA extracted but to intrinsic properties of the tissues, and therefore, these genes cannot be used as reference genes for value normalization.

We also aimed to explore *Smyd5* expression during murine liver development due to the well-accepted concept that tumorigenesis partially recapitulates embryogenesis (229,230). In line with this idea, genes upregulated during liver development and silenced in the adult tissue may be reactivated in liver cancer (e.g. Insulin-like growth factor 2 mRNA binding protein 3 [IGF2BP3] (231) or Yes-associated protein 1 [YAP1] (232)), conferring survival advantage to the malignant cells (233). Indeed, *Smyd5* was significantly overexpressed at embryonic day 15.5 (E15.5) compared to both embryonic day 18.5 (E18.5) and postnatal day 60 (P60) (**Figure 25B**). Since the presence of

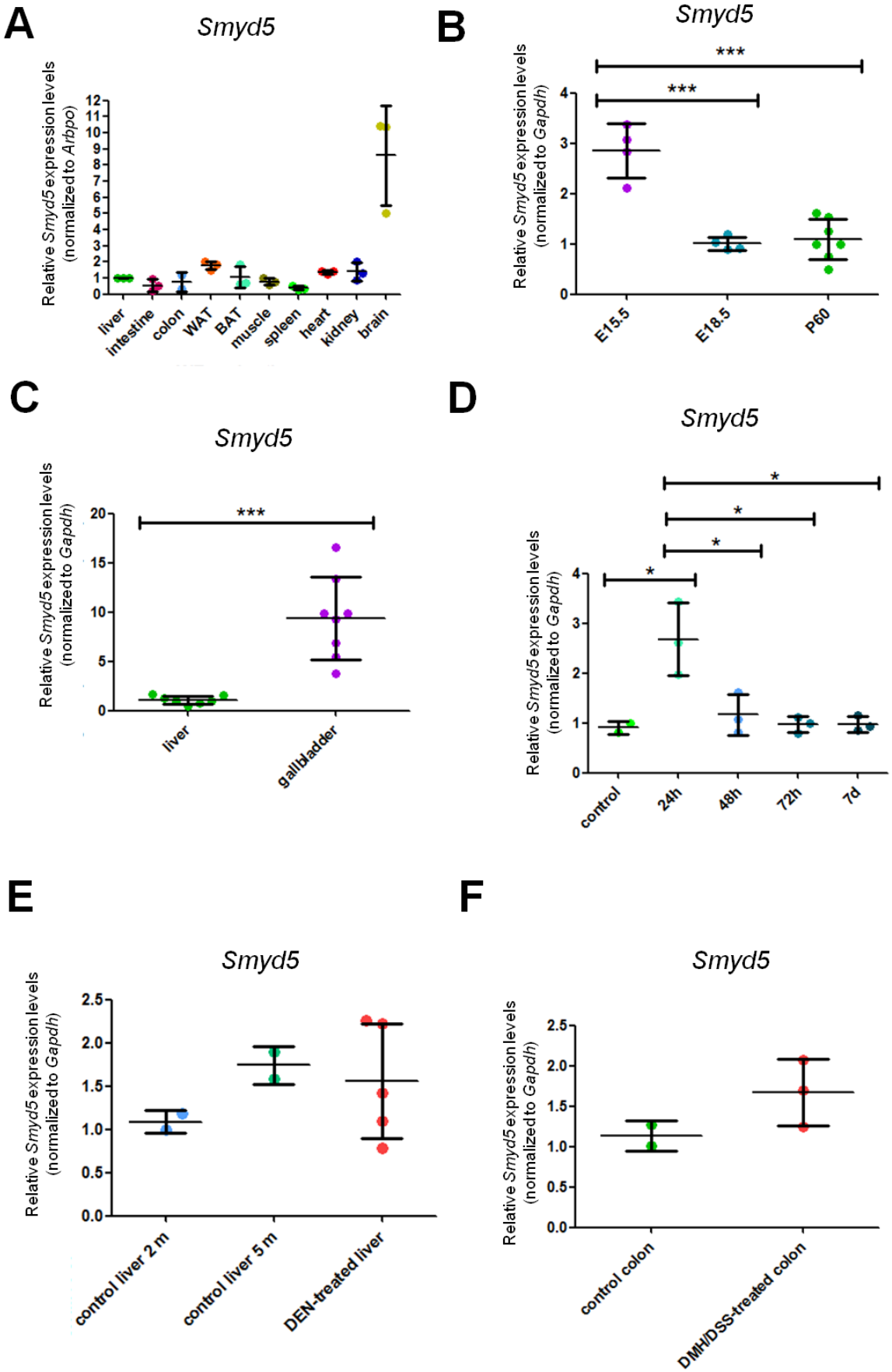


Figure 25. *Smyd5* expression in different wild-type (WT) murine organs and conditions. (**A**) *Smyd5* expression in many WT C57BL/6 adult male mouse tissues (N = 3 per tissue) in homeostasis. (**B**) *Smyd5*

expression in WT liver at embryonic day 15.5 (E15.5) (N = 4), embryonic day 18.5 (E18.5) (N = 4) and postnatal day 60 (P60) (N = 7). (**C**) *Smyd5* expression in WT adult liver (N = 7) and gallbladder (N = 8 pools of 4 gallbladders). (**D**) *Smyd5* expression in untreated healthy WT liver (control) (N = 2) and livers from mice sacrificed 24 hours (24h) (N = 3), 48 hours (48h) (N = 3), 72 hours (72h) (N = 3) and 7 days (7d) (N = 3) after a single carbon tetrachloride (CCl_4) injection. (**E**) *Smyd5* expression in untreated healthy WT liver from 2-month-old mice (control liver 2m) (N = 2), untreated healthy WT liver from 5-month-old mice (control liver 5m) (N = 2) and diethylnitrosamine (DEN)-treated liver from 8.5-month-old mice (DEN-treated liver) (N = 5). (**F**) *Smyd5* expression in untreated healthy WT colon from 5-month-old mice (control colon) (N = 2) and 1,2-dimethylhydrazine/dextran sodium sulfate (DMH/DSS)-treated liver from 5-month-old mice (DMH/DSS-treated colon) (N = 5). All quantitative real time PCR (qPCR) analyses were performed with a primer set targeting *Smyd5* exon 4 to 6 (see below and Methods). Bars represent mean mRNA levels normalized to *Rplp1* or *Gapdh* mRNA and standard deviation (±SD) from each group. T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05; ***P value ≤ 0.001).

common hepatocyte and biliary epithelium progenitor populations in non-damaged murine gallbladders has been reported (234), we isolated whole mouse gallbladders and observed an increase in *Smyd5* expression in pools of four gallbladders compared to whole liver tissue (**Figure 25C**), reinforcing the idea that SMYD5 function in the liver could be related to stemness and/or cancer.

To investigate the role of *Smyd5* in liver injury, which can eventually lead to cancer, we injected carbon tetrachloride (CCl_4) intraperitoneally in WT mice. In the liver, CCl_4 is metabolized by cytochrome P-450, leading to lipid peroxidation and, consequentely, injury to the cell membrane. This provokes an increase in cellular permeability that promotes the leakage of enzymes and the disruption of calcium homeostasis, inducing calcium-dependent degradative enzymes that give rise to hepatic damage (235). *Smyd5* expression was enhanced 24 hours after a single CCl_4 injection (**Figure 25D**), when hepatic cell death (necrosis and apoptosis) normally occurs and before the peak of damage (between 48 hours and 72 hours) (previous data by our research group). *Smyd5* expression returns to basal levels 48-72 hours after the injection and remains stable until day 7, when the liver is fully recovered from the damage. Remarkably, Huch *et al.* identified a LGR5-positive stem cell population that arises after liver injury by CCl_4 (236). These results raised the possibility that SMYD5 may characterize this stem cell population or another one, but we did not perform further experiments to explore this possibility because we focused on the role of SMYD5 in cancer.

To determine a possible role in liver and colon cancer, we reviewed *Smyd5* expression in RNA-seq data from cancer murine experimental setups previously published by our research group. We were not able to detect changes in *Smyd5* expression in these datasets. *Smyd5* was not differentially expressed in liver-specific PR-SET7 KO that

developed cancer compared to WT liver samples (fold change = 1.2, P value = 0.66) (91). Similarly, its expression did not vary in DEN-treated WT cancerous livers when compared to untreated WT controls (fold change = 1.06, P value = 0.09), or in the same DEN-treated samples when compared to DEN-treated liver-specific SMYD3 KO, a model that attenuates cancer development (fold change = 2.18, P value = 1.13) (211). Likewise, RNA-seq from DMH/DSS-treated WT colon did not display *Smyd5* overexpression in comparison to untreated WT colon samples (fold change = 0.69, P value = 0.44) (211). Nevertheless, we decided to assess *Smyd5* expression in these DEN-treated WT livers and DMH/DSS-treated colons by qPCR. We observed a trend of increasing *Smyd5* expression in both types of cancer compared to the corresponding untreated control group (**Figure 25E** and **25F**). This trend was not statistically significant most probably due to the intrinsic variability of cancer but pointed to an involvement of SMYD5 in carcinogenesis.

SMYD5 preferentially methylates histones 2A and 2B *in vitro*

In order to reproduce the results by Stender *et al.*, which describe SMYD5 as a histone methyltransferase that catalyzes H4K20me3, and to identify new histone and non-histone targets we performed an *in vitro* methylation assay by incubating SMYD5 enzyme with ^{3}H-labeled SAM and H4, H3, H2A, H2B, H1, a mixture of histones, or proteins, such as P53 and E2F1, involved in cancer progression (**Figure 26A**). We identified H2A and H2B as previously unknown SMYD5 targets *in vitro,* but we failed to observe methylation signal in H4 (**Figure 26B**). No signal is detected in the histone mix, because, based on Coomassie staining, each histone is at a lower concentration, probably below the detection limit.

We studied subcellular localization of SMYD5 in a murine hepatoma cell line, Hepa 1-6, and mouse liver. Due to suboptimal antibody performance, interpretation of the results was not straightforward. Nevertheless, Western blot analysis revealed that SMYD5 is mostly located in the cytoplasm of Hepa1-6 cells and WT murine liver cells (**Figure 27A** and **27B**). Since nuclear extracts were pure and had lower protein content than cytoplasmic or whole cells extracts, SMYD5 protein levels in liver extracts from untreated mice could be below the detection limit using our antibody (see Discussion).

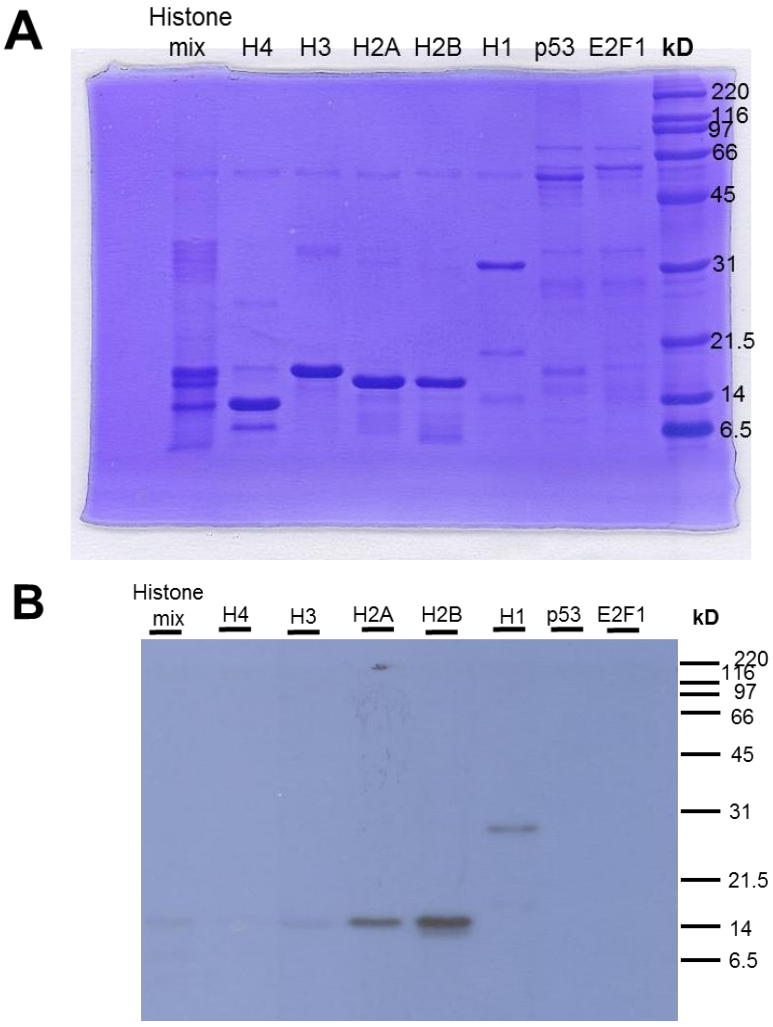


Figure 26. SMYD5 methyltransferase activity *in vitro*. Purified human baculovirus-expressed SMYD5 enzyme was incubated with ^{3}H-labeled SAM and histones or proteins involved in cancer progression to determine SMYD5 methylation substrates *in vitro*. (**A**) Coomassie Blue-stained gel. (**B**) Autoradiogram in which bands correspond to ^{3}H-labeled SAM signal.

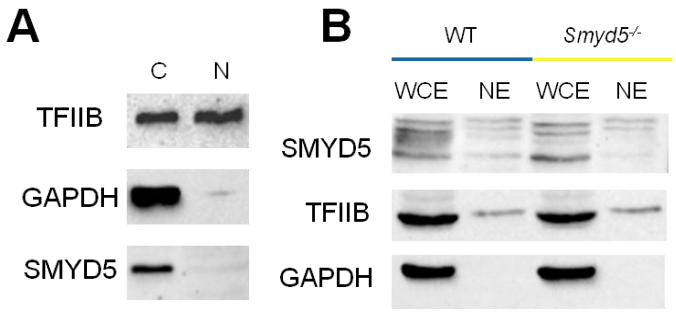


Figure 27. SMYD5 subcellular localization in murine liver. (**A**) Cytoplasmic (C) and nuclear (N) protein extracts of murine hepatoma cell line Hepa1-6 analyzed by Western blot. (**B**) Whole cell extracts (WCE) and nuclear extracts (NE) from wild-type (WT) and *Smyd5* full knockout (KO) (*Smyd5*$^{-/-}$) mouse livers. TFIIB and GAPDH were used as controls. TFIIB is a transcription factor located in the nucleus and GADPH is a cytoplasmic enzyme. As expected, cytoplasmic extracts are contaminated with nuclear content but nuclear extracts are pure and have less protein content.

Smyd5 is localized in the liver and in the intestinal crypts where stem cells are located

As mentioned above, none of the antibodies tested against SMYD5 were capable of specifically detecting the protein by immunostaining. Therefore, we performed β-galactosidase enzymatic reaction-based staining as antibody-independent evidence that can provide information about SMYD5 cellular distribution in WT adult mouse liver and intestine. Bacterial gene LacZ encodes for β-galactosidase enzyme. B-galactosidase can hydrolize the colorless artificial analog of lactose X-gal, which, after a series of reactions, yields 5,5'-dibromo-4,4'-dichloro-indigo, a blue-colored insoluble product (237) (**Figure 28A**). This blue precipitate can be observed under the microscope. The bacterial LacZ gene is regulated by the SMYD5 promoter in our SMYD5 transgenic mouse carrying the transgene that can give rise to the KO mouse when exon 2 is excised. Remarkably, although all SMYD5 exons were present, mice carrying the complete transgene (as depicted in **Figure 28A**) were expected to present a variable percentage of decrease in *Smyd5* expression in all tissues due to a splice acceptor site that can lead to alternative splicing. However, these mice did not present any spontaneous phenotype in homeostasis (see below). Thus, β-galactosidase staining recapitulates *Smyd5* expression. After β-galactosidase staining, we observed blue color in the liver and the intestine, accounting for the presence of the SMYD5 protein. In the liver β-galactosidase staining did not show an obvious specific pattern due to technical limitations. In the intestine β-galactosidase staining appeared in the bottom of the intestinal crypts, suggesting that SMYD5 is located mainly in the intestinal stem cell compartment in WT mice (**Figure 28B**). Unfortunately, we were not able to perform co-staining with antibodies or mRNA probes against intestinal stem cell markers such as the Wnt-targeted gene *Lgr5,* achaete-scute family bHLH transcription factor 2 (*Ascl2*), olfactomedin 4 (*Olfm4*) or TNF receptor superfamily member 19 (*Tnfrsf19*) (238) due to technical limitations. Nevertheless, our data support the hypothesis that SMYD5 could be involved in stemness maintenance.

To obtain further indirect evidence of a role of SMYD5 in stem cells, we analyzed public mRNA expression datasets. We assessed *Smyd5* expression in published mRNA microarray data from isolated *Lgr5*-expressing cells (intestinal stem cells) and its daughter cells (differentiated cells) (238). Unexpectedly, *Smyd5* was not significantly overexpressed in these intestinal stem cells compared to differentiated cells, while expression of well-established intestinal stem cell markers such as *Lgr5, Ascl2*, *Olfm4* or *Tnfrsf19* was significantly increased (238). Moreover, we examined the RNA-seq

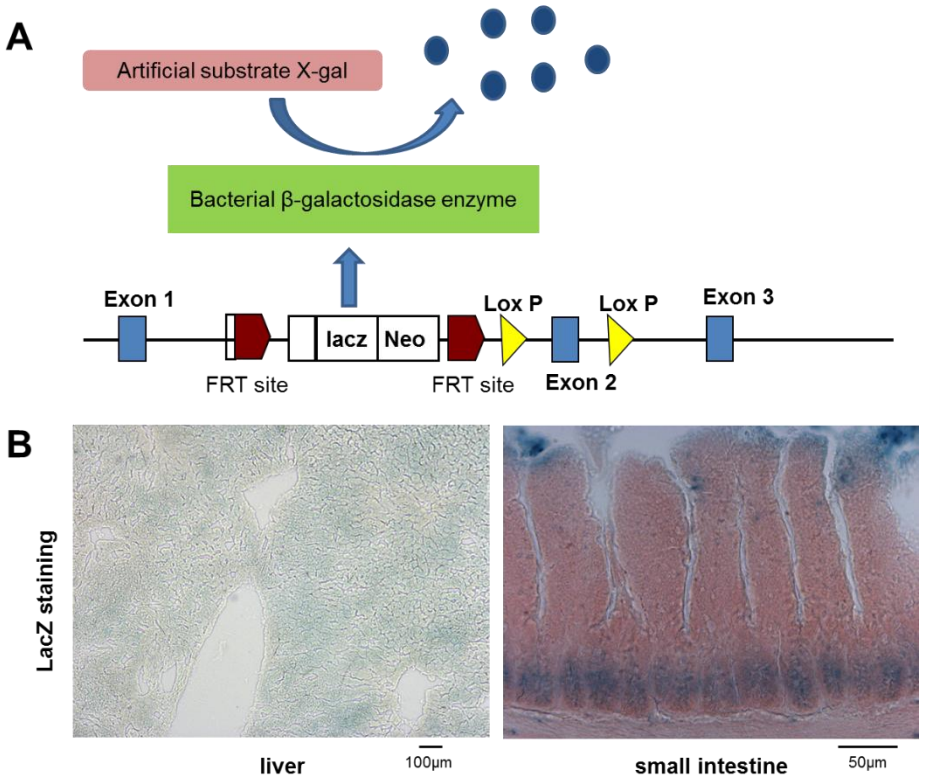
dataset published by Wang *et al.* They used lineage tracing with Wnt-responsive gene *Axin2* in mice to identify a population of proliferating and self-renewing cells adjacent to the central vein in the liver lobule (239). RNA-seq showed no differential expression of *Smyd5* in *Axin2*-positive and *Axin2*-negative populations.

We also analyzed unpublished data from the Pantelis Hatzis research group. RNA-seq data from isolated murine intestinal crypts and villi confirmed *Smyd5* overexpression in the crypts (14.65-fold increase) relative to the villi fraction.

The above data raise the possibility that SMYD5 may mark adult stem/progenitor cells.

Clinicopathological analysis of human liver, colon and lung cancer samples points to a possible role of SMYD5 in carcinogenesis

Since carcinogenesis requires cell renewal and other SMYD family members have been involved in cancer, we decided to analyze *SMYD5* gene expression in human data available in the TCGA project (in collaboration with Antonis Giakountis). We found significant overexpression of *SMYD5* in liver (**Figure 29A**), colon (**Figure 30A**) and lung (**Figure 31A**) cancer. For further analysis, values were obtained by median dichotomization and tumors were stratified in "*SMYD5* Low" and "*SMYD5* High" according to their expression in the three cancer types (**Figure 29B**, **30B** and **31B**).



(Previous page) Figure 28. SMYD5 cellular distribution in adult mouse liver and small intestine. (**A**) Our transgenic mice contain the LacZ gene encoding for β-galactosidase enzyme under the control of *Smyd5* promoter. β-galactosidase can catalyze the conversion of X-gal, which, after a series of reactions, yields the blue-colored product 5,5'-dibromo-4,4'-dichloro-indigo. (**B**) β-galactosidase staining was performed in the liver and the small intestine (ileum) of mice carrying the transgene described in (**A**). Presence of blue color accounts for the presence of the SMYD5 protein.

Kaplan-Meier analysis comparing liver cancer patients with high (Q3 group) and low (Q1 group) *SMYD5* expression revealed that high *SMYD5* expression in liver tumors decreases the probability of patient survival after diagnosis (**Figure 29C**). The probability of surviving as a "tumor-free" patient for the corresponding timeframe after chemotherapy and/or surgery (**Figure 29D**), the probability of surviving and remaining in the same stage of the disease (**Figure 29E**) and the probability of remaining in the group of patients classified as G1 or G2 grade and not progressing to G3-G4 (more advanced stages) (**Figure 29F**) were significantly lower in the "*SMYD5* High" group compared to the "*SMYD5* Low" group.

Analysis of *SMYD5* expression in colon adenocarcinoma (**Figure 30**) and lung adenocarcinoma (**Figure 31**) was also conducted using TCGA datasets. Kaplan-Meier analysis comparing colon cancer patients with high (Q3 group) and low (Q1 group) *SMYD5* expression failed to detect differences in overall survival (**Figure 30C**) and in the probability of surviving as a "tumor-free" patient after chemotherapy for a given time frame (**Figure 30D**). Conversely, the probability of progressing to high-grade (G3/G4) tumors was enhanced in the "*SMYD5* High" compared to the "*SMYD5* Low" group (**Figure 30E**). Contrary to what has been described in a murine model that mimics lung metastases of breast cancer (219), the probability of developing metastasis in other organs apart from the primary colon tumor was increased in "*SMYD5* High" compared to "*SMYD5* Low" patients (**Figure 30F**).

Negative results were obtained for lung adenocarcinoma (**Figure 31**), despite *SMYD5* overexpression in tumor samples (**Figure 31A** and **31B**). Kaplan-Meier analysis comparing lung cancer patients with high (Q3 group) and low (Q1 group) *SMYD5* expression did not detect differences in overall survival (**Figure 31C**), in the probability to survive as a "tumor-free" patient after chemotherapy (**Figure 31D**), in the probability of progressing to high-grade (G3/G4) tumors (**Figure 31E**) and in the probability of developing metastasis in other organs apart from the primary lung tumor (**Figure 31F**).

In conclusion, *SMYD5* is overexpressed in human liver, colon and lung cancer and correlates with bad prognostic factors in liver and colon cancer, but not in lung cancer.

Generation of SMYD5 knockout mouse models

Once we assessed clinical relevance of *SMYD5* for liver and colon cancer prognosis, we aimed to generate *Smyd5*-deficient mice as an *in vivo* model to study the role of SMYD5 in the carcinogenesis in these organs. We generated *Smyd5*-deficient mice which are devoid of *Smyd5* mRNA and protein in either all the tissues (*Smyd5*^{flox}/MX-Cre) or in specific cell types; specifically in hepatocytes in the adult liver (*Smyd5*^{flox}/Alb-Cre), in both hepatocytes and biliary epithelium cells since embryonic day 15 (E15) (*Smyd5*^{flox}/Alfp-Cre) or in intestinal epithelium cells (*Smyd5*^{flox}/Villin-Cre) (**Figure 32A**).

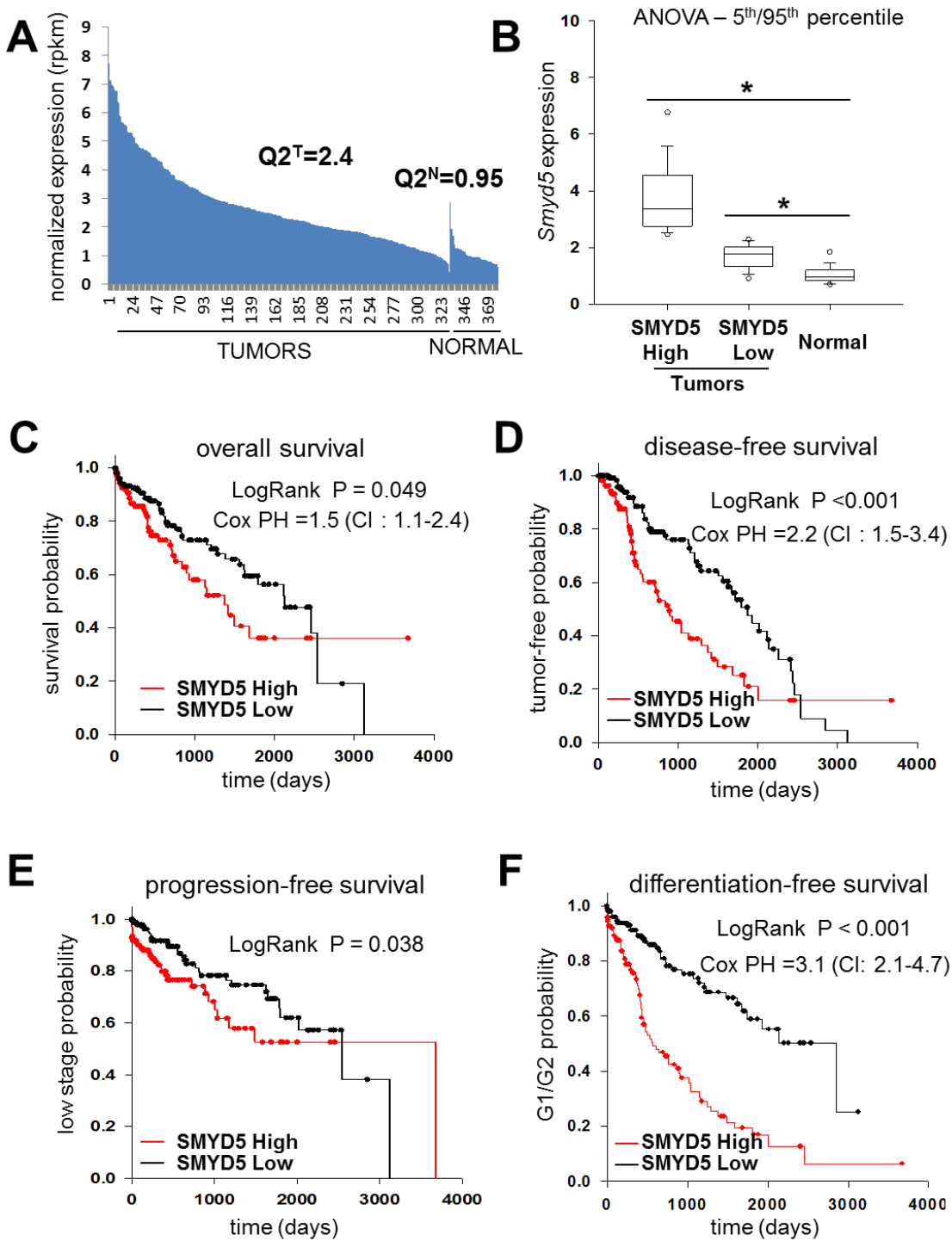
Mice carrying *Smyd5* targeted allele in C57BL/6N background contain a splice acceptor site in the region flanked by FRT sites (each site usually has a length of 34 bp) that can lead to alternative splicing and a variable percentage of decrease in *Smyd5* expression in all tissues. Nevertheless, to assure that our mice completely lacked *Smyd5* mRNA and protein, we deleted the *Smyd5* exon 2 by two different strategies (**Figure 32B**).

First, to generate conditional *Smyd5* KO mice, the original mice carrying the *Smyd5* targeted allele were crossed with *Flipase*-expressing mice in order to achieve delection of the region between FRT sites by flipase enzyme. We selected the progeny that carried the targeted allele in which the region flanked by FRT sites had been recombined but did not carry the *Flipase* enzyme. Then we crossed them with mice expressing the appropiate Cre recombinases (capable of deleting the sequence flanked by specific 34 bp sequences called LoxP sites); either Alb-Cre, Alfp-Cre or Villin-Cre.

Second, to generate standard *Smyd5* KO mice, we directly crossed commercial mice carrying the original *Smyd5* targeted allele with mice expressing MX-Cre recombinase.

(Next page) Figure 29. *SMYD5* expression in human liver cancer samples deposited in "The Cancer Genome Atlas" (TCGA) database. (**A**) Relative expression of *SMYD5* in liver tumor (N=323) and normal (N=46) biopsies. mRNA levels retrieved from RNA-seq data are expressed as reads per kilobase of exons per mbp (rpkm). (**B**) For further analysis patients were stratified into two groups with "High" (Q3) and "Low" (Q1) *SMYD5* expression. *P value ≤ 0.05 by ANOVA. (**C, D, E, F**) Kaplan-Meier analysis comparing liver cancer patients with high (Q3 group) and low (Q1 group) *SMYD5* expression are shown. Panel (**C**) depicts differences in overall survival probability between the two groups. Panel (**D**) shows differences in the probability of surviving as "tumor-free" patient for the corresponding timeframe after chemotherapy and/or surgery between the two groups. Panel (**E**) depicts differences in the probability of remaining in the same stage of the disease between the two groups. Panel (**F**) shows differences in the probability of remaining in the group of patients classified as G1 or G2 grade and not progressing to G3-G4 between the two groups. LogRank P < 0.05 means statistical significance. For clinical traits with statistically significant differences

between the two groups (*SMYD5* High and *SMYD5* Low), Proportional Hazard (Cox PH) shows the X-fold increase in the probability for the event to happen in one group compared to another, with a specific confidence interval (CI).



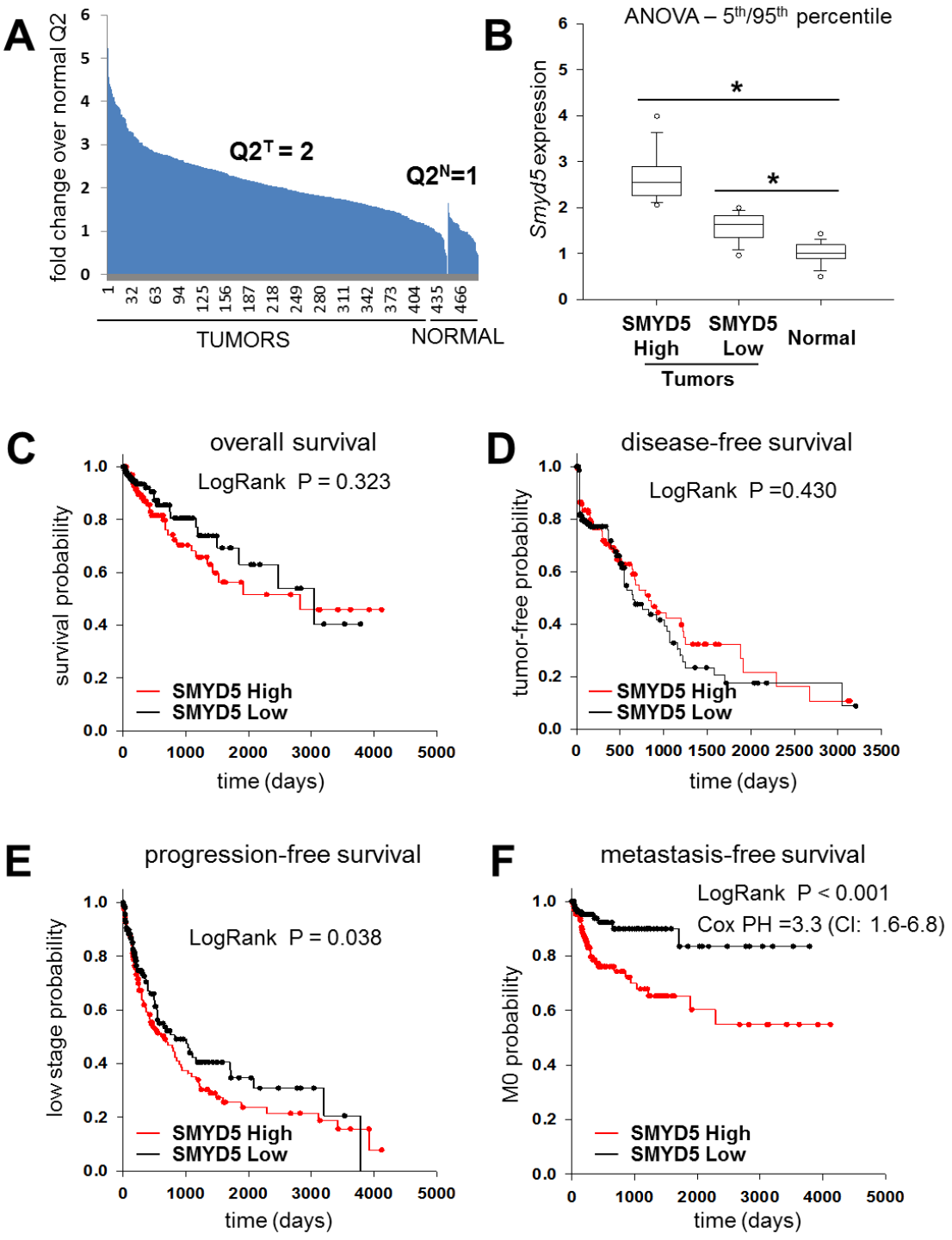


Figure 30. *SMYD5* expression in human colon cancer samples deposited in "The Cancer Genome Atlas" (TCGA) database. (**A**) Relative expression of *SMYD5* in liver tumor (N=404) and normal (N=62) biopsies. mRNA levels retrieved from RNA-seq data are expressed as fold change over normal. (**B**) For further analysis patients were stratified into two groups with "High" (Q3) and "Low" (Q1) *SMYD5* expression. *P value ≤ 0.05 by ANOVA. (**C, D, E, F**) Kaplan-Meier analysis comparing colon cancer patients with high (Q3 group) and low (Q1 group) *SMYD5* expression are shown. Panel (**C**) depicts no differences in overall survival probability between the two groups. Panel (**D**) shows no differences in the probability of surviving as "tumor-free" patient for the corresponding timeframe after chemotherapy and/or surgery between the two groups. Panel (**E**) depicts differences in the probability of remaining in the same stage of the disease between the two groups. Panel (**F**) shows differences in the probability of developing

metastasis in other organs apart from the primary colon tumor between the two groups. LogRank P < 0.05 means statistical significance. For clinical traits with statistically significant differences between the two groups (*SMYD5* High and *SMYD5* Low), Proportional Hazard (Cox PH) shows the X-fold increase in the probability for the event to happen in one group compared to another, with a specific confidence interval (CI).

All *Smyd5* KO mice (**Figure 32A**) were born at expected Mendelian ratios, developed normally, were fertile and lived for at least 1.5 years. The KO mice did not display any visible phenotype in the different organs. This was partly expected because previous studies by our research group (211) and others (240) revealed that mice lacking *Smyd2* and *Smyd3* mRNA and protein did not present any macroscopic or histological alteration in any organ examined, including liver and intestine.

Liver-specific *Smyd5* knockout mice do not present any spontaneous macroscopic or histological phenotype at 2 months of age

To ensure that *Smyd5* exon 2 was removed by recombination in our KO models, but also that the whole *Smyd5* mRNA was not produced, we designed two primer sets: one pair targeting exon 2 and one pair targeting exons 4 to 6 (**Figure 33A**). We used both primer sets to perform qPCR in livers from 2-month-old WT and *Smyd5*^{flox}/Alb-Cre KO mice. As expected, *Smyd5* exon 2 was deleted in KO livers, and the deletion led to a decrease in *Smyd5* mRNA (**Figure 33B**). The same results were obtained by Western blot (**Figure 33C**). Similar results were obtained for *Smyd5*^{flox}/Alfp-Cre (data not shown), the full KO (*Smyd5*^{flox}/MX-Cre) (data not shown) and *Smyd5*^{flox}/Villin-Cre mice (see below). Hematoxylin and eosin (H&E) staining indicated physiological liver histology in *Smyd5*^{flox}/Alb-Cre KO mice under homeostasis (**Figure 33D**). Preliminary insights into the regulation of liver metabolic pathways in *Smyd5*^{flox}/Alb-Cre KO mice did not show any statistically significant differences when compared with WT (**Figure 34**).

In summary, *Smyd5* mRNA and protein were present in WT but absent in KO tissues (liver or intestine), meaning that our KO mice were verified KO and could be used in this study. We could not detect any macroscopic or histological alterations in these animal models under physiological conditions, suggesting that *Smyd5* is dispensable for normal liver and intestine development.

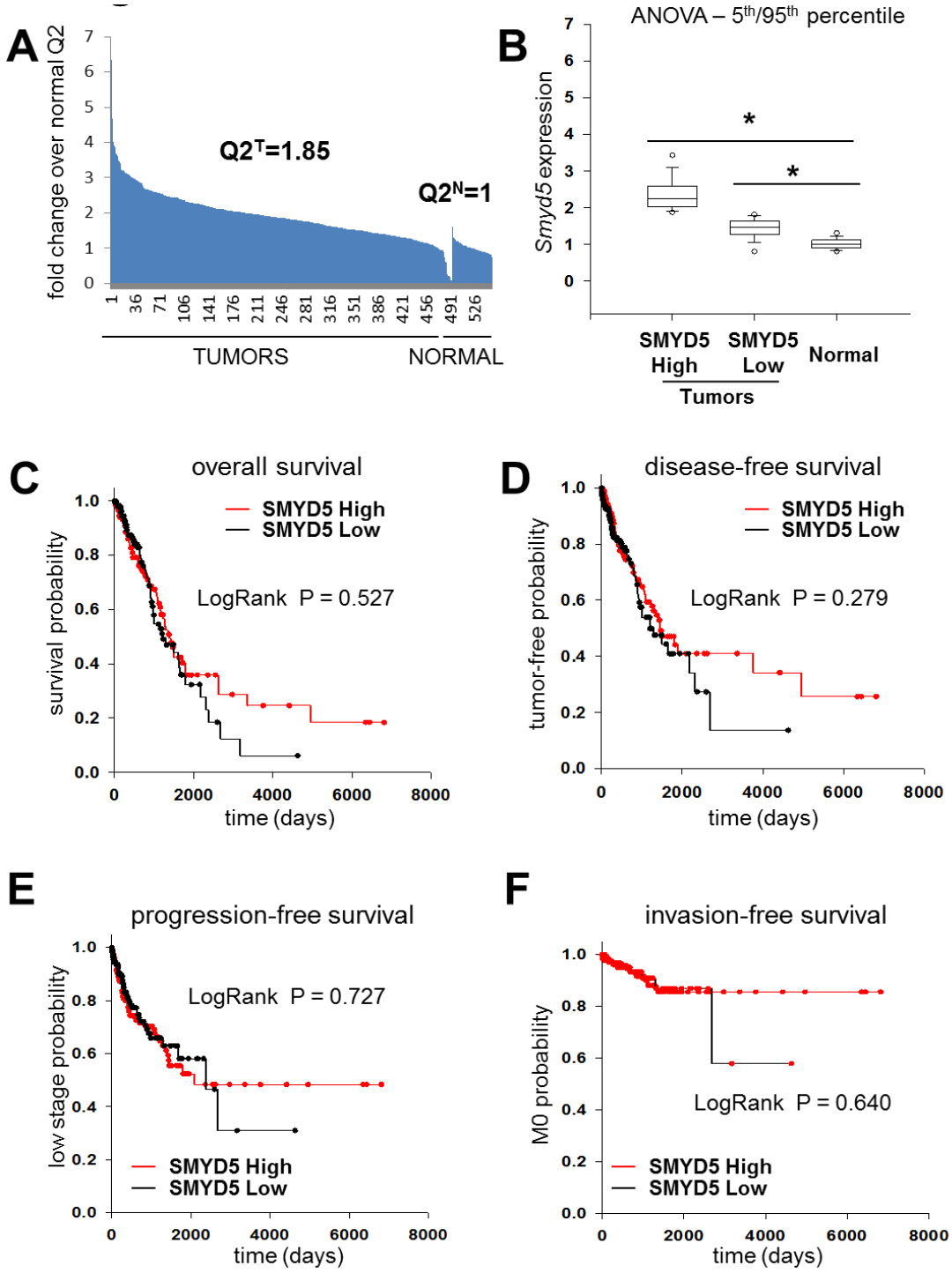


Figure 31. *SMYD5* expression in human lung cancer samples deposited in "The Cancer Genome Atlas" (TCGA) database. (**A**) Relative expression of *SMYD5* in lung tumor (N=456) and normal (N=70) biopsies. mRNA levels retrieved from RNA-seq data are expressed as fold change over normal. (**B**) For further analysis patients were stratified into two groups with "High" (Q3) and "Low" (Q1) *SMYD5* expression. *P value ≤ 0.05 by ANOVA. (**C, D, E, F**) Kaplan-Meier analysis comparing lung cancer patients with high (Q3 group) and low (Q1 group) *SMYD5* expression are shown. Panel (**C**) depicts no differences in overall survival probability between the two groups. Panel (**D**) shows no differences in the probability of surviving as "tumor-free" patient for the corresponding timeframe after chemotherapy and/or surgery between the two groups. Panel (**E**) depicts no differences in the probability of remaining in the same stage

of the disease between the two groups. Panel (**F**) shows no differences in the probability of developing metastasis in other organs apart from the primary lung tumor between the two groups. LogRank P < 0.05 means statistical significance. For clinical traits with statistically significant differences between the two groups (*SMYD5* High and *SMYD5* Low), Proportional Hazard (Cox PH) shows the X-fold increase in the probability for the event to happen in one group compared to another, with a specific confidence interval (CI).

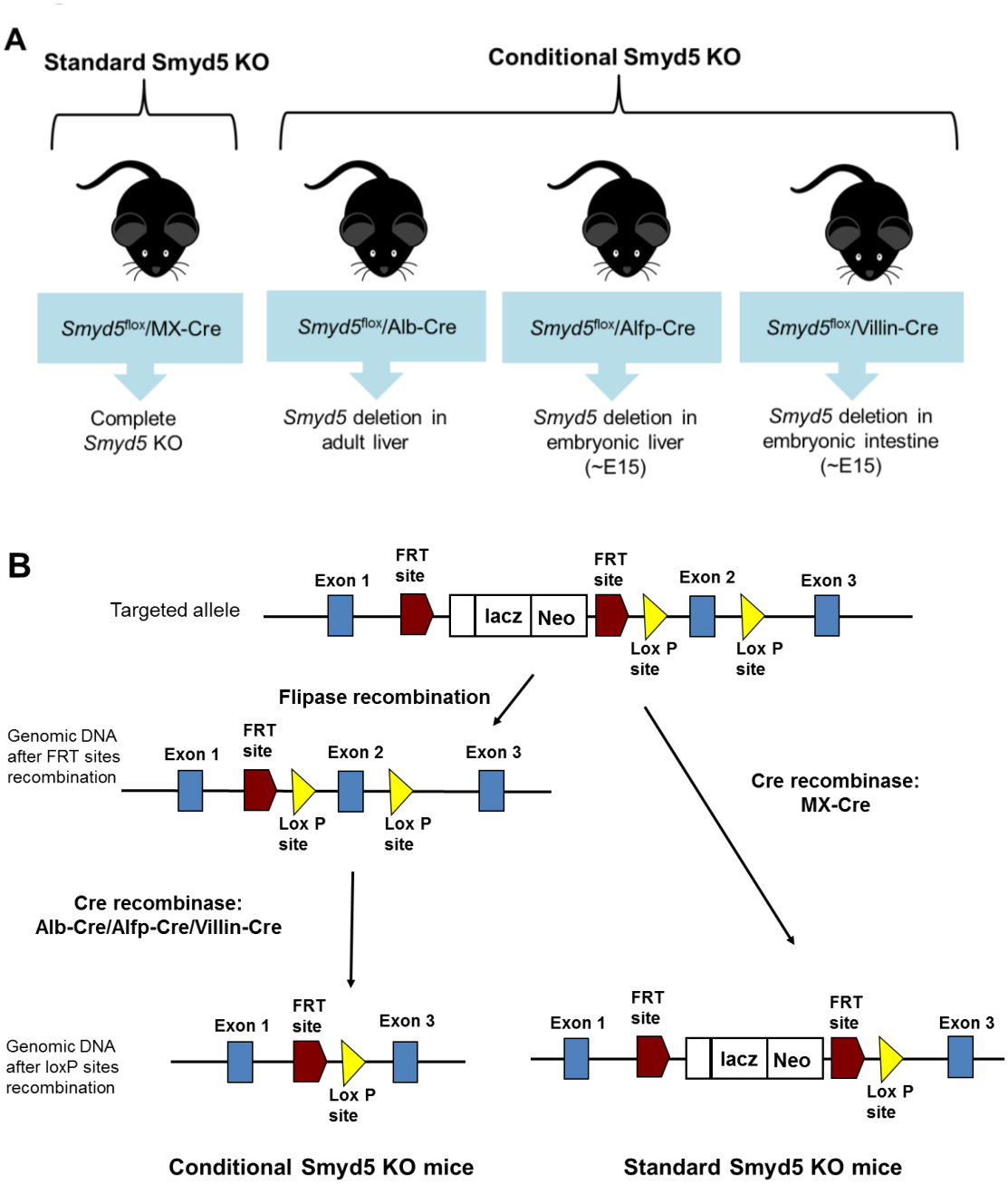


Figure 32. Generation of Smyd5 knockout (KO) mouse models. (**A**) Schematic representation of standard and conditional tissue-specific *Smyd5*-KO mouse models. (**B**) Schematic representation of the process followed in order to obtain standard and conditional tissue-specific *Smyd5*-KO mouse models and the transgenic DNA structure in each step.

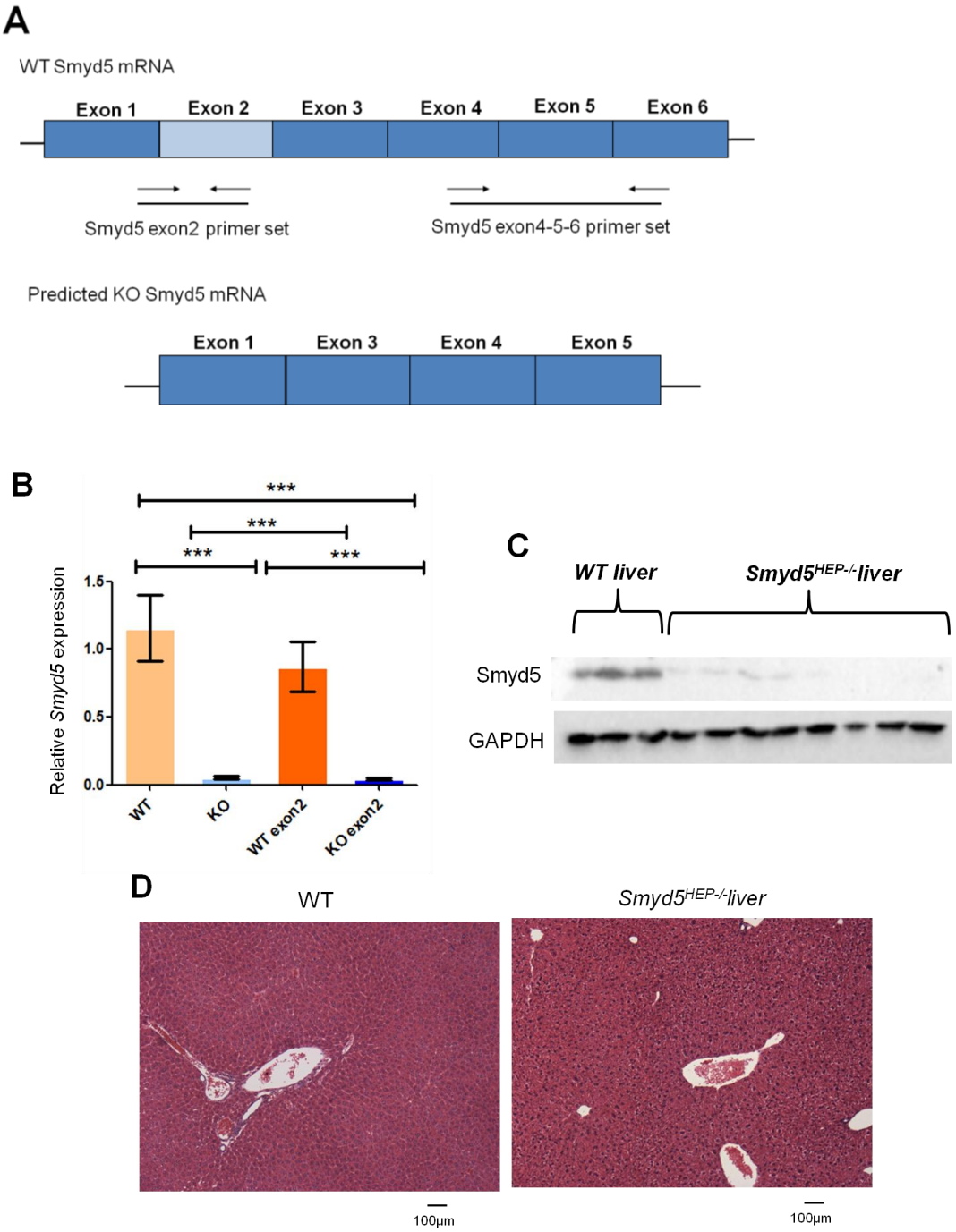


Figure 33. Histological analysis of untreated *Smyd5*^{flox}/Alb-Cre knockout (KO) livers at 2 months of age. (**A**) Primer set localization in *Smyd5* mRNA. (**B**) Relative *Smyd5* expression analysis by quantitative real time PCR (qPCR) in wild-type (WT) and *Smyd5*^{flox}/Alb-Cre KO livers using one primer set targeting exon 4 to 6 ("WT", "KO" samples) and one primer set targeting exon 2 ("WT exon 2", "KO exon 2"). Values are from at least N = 3 mice per group and average values and standard deviation (±SD) are shown. Values were normalized to *Gapdh* mRNA levels and compared to the mean value of the WT group. T-test was performed in all cases, comparing groups one by one (***P value ≤ 0.001). (**C**) SMYD5 protein levels assessed by Western blot in WT and *Smyd5*^{flox}/Alb-Cre KO (*Smyd5*$^{HEP-/-}$liver) livers. (**D**) Hematoxylin and eosin (H&E) staining in WT and *Smyd5*^{flox}/Alb-Cre KO livers under homeostasis.

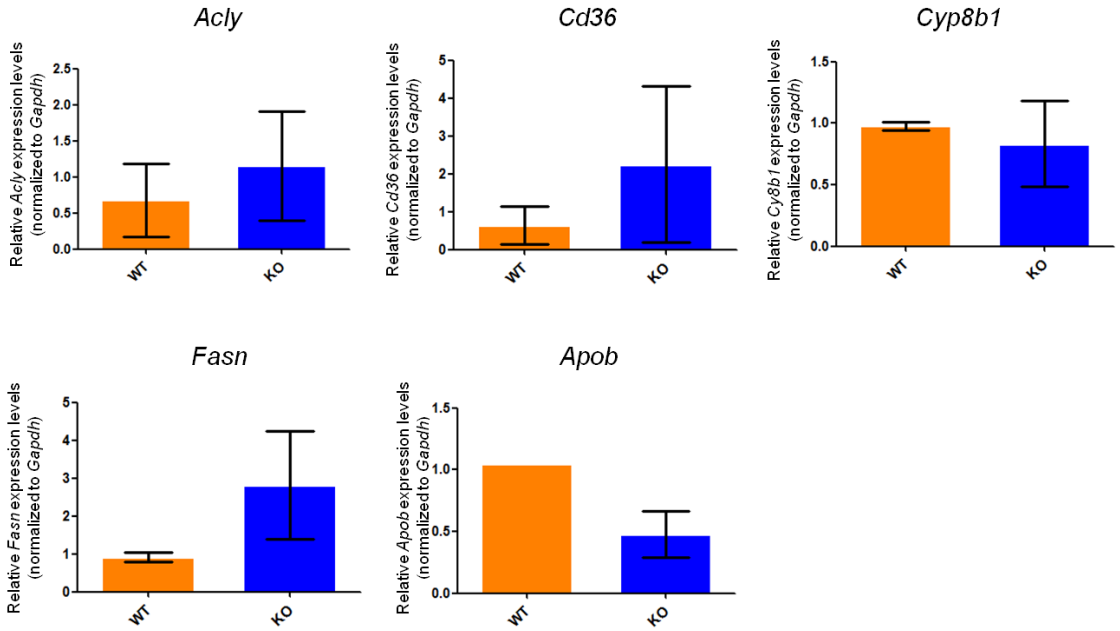


Figure 34. Quantitative real time PCR (qPCR) analysis of ATP-citrate synthase (*Acly*), *Cd36*, Cytochrome P450 family 8 subfamily b polypeptide 1 (*Cyp8b1*), Fatty acid synthase (*Fasn*) and Apolipoprotein B (*Apob*) mRNA levels in the livers of 2-month-old untreated wild-type (WT) and *Smyd5*^{flox}/Alb-Cre (KO) mice. All analyses were performed with a primer set targeting *Smyd5* exons 4 to 6 (see Methods). Bars represent mean mRNA levels normalized to *Gapdh* mRNA and compared to WT, and standard deviation (±SD) from each group (at least N = 2 individual mice). T-test was performed in all cases, comparing groups one by one.

Smyd5 liver-specific knockout shows delayed carcinogenesis after DEN treatment

To investigate a potential causative relationship between *Smyd5* overexpression and liver cancer we analyzed chemically induced tumor formation in WT and liver specific *Smyd5*-deficient mice. We performed the analysis in *Smyd5*^{flox}/Alfp-Cre KO, with similar results being obtained for *Smyd5*^{flox}/Alb-Cre (data not shown) and the full KO (*Smyd5*^{flox}/MX-Cre) (data not shown). We intraperitoneally injected 25 mg DEN per mouse kg to postnatal day 14 (P14)-litters and sacrificed the animals 7.5 or 9.5 months after DEN treatment (**Figure 35A**). Macroscopic examination revealed a dramatic reduction of the number and size of tumor foci in liver specific *Smyd5*-deficient livers compared with WT littermates, although most were not statistically significant, probably due to high variability and low sample number. The differences in tumor formation capacity were more striking 7.5 months after DEN treatment (**Figure 35B**, **36A** and **36B**). There was a trend indicating that hepatomegaly, a sign of liver cancer, was attenuated in liver-specific *Smyd5*-deficient livers compared with WT littermates (**Figure 35B**, **37A** and **37B**). Since the delayed carcinogenesis phenotype was more obvious 7.5 months after DEN treatment, further analyses were performed at this time point.

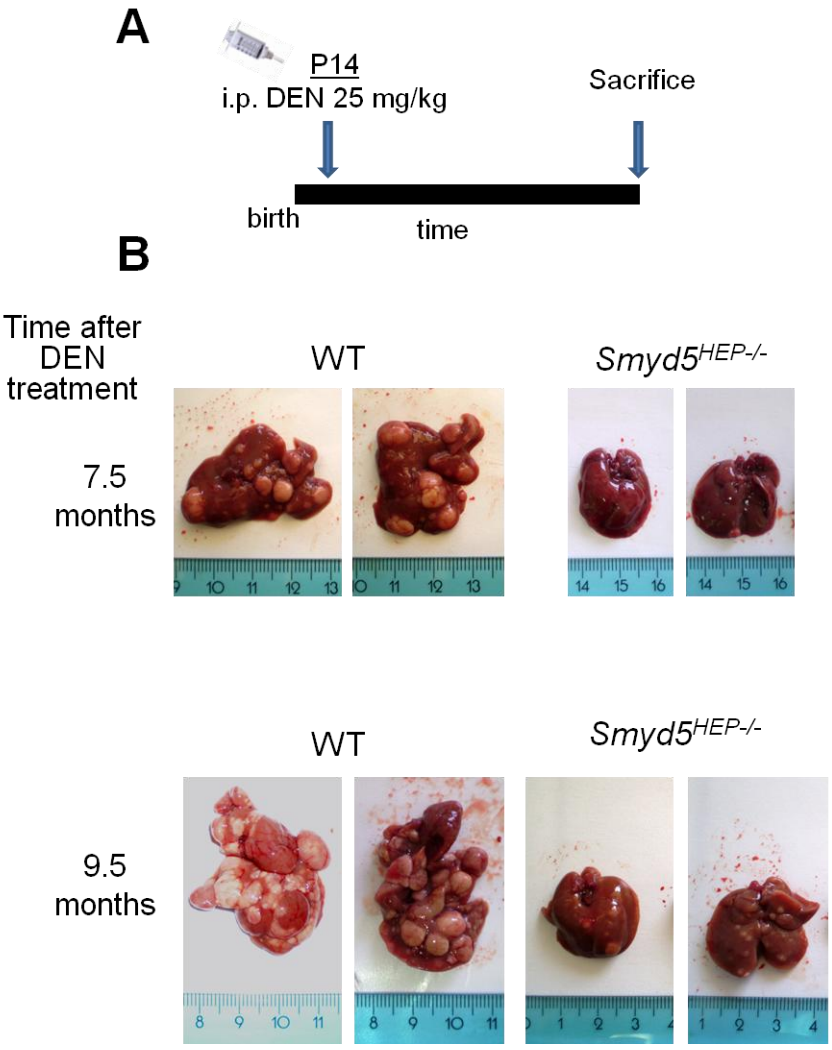


Figure 35. Macroscopic analysis of diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO livers. (**A**) Mice were subjected to a single intraperitoneal (i.p.) injection of 25 mg DEN per mouse kg at postnatal day 14 (P14) and sacrificed 7.5 or 9.5 months after DEN treatment. (**B**) Macroscopic appearance of livers in 8- and 10-month-old DEN-treated and untreated WT and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice. Size in centimeters.

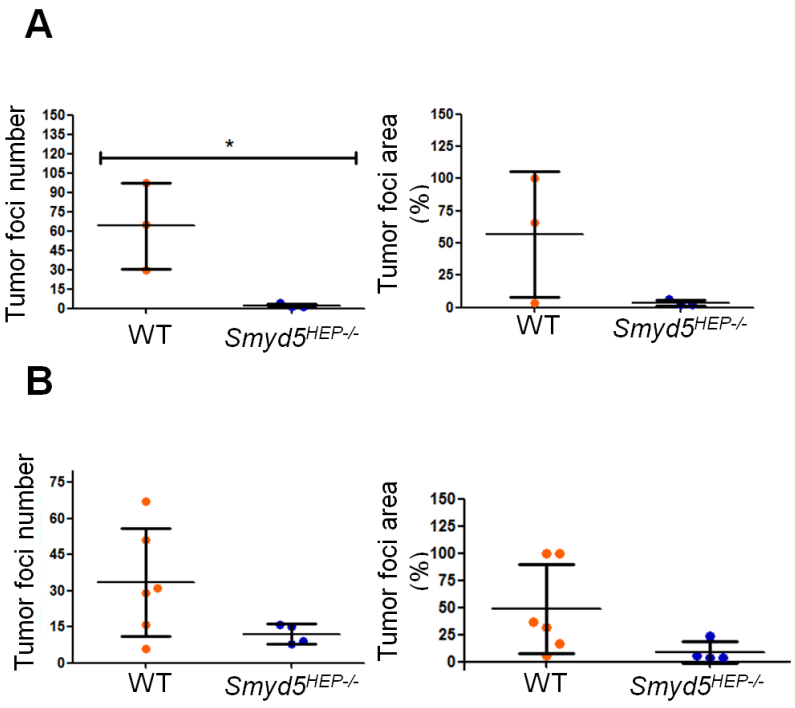


Figure 36. Quantification of tumor foci number and area in diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) livers. Quantifications (**A**) 7.5 and (**B**) 9.5 months after DEN treatment. Average values and standard deviation (±SD) were from at least N = 3 mice. T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05).

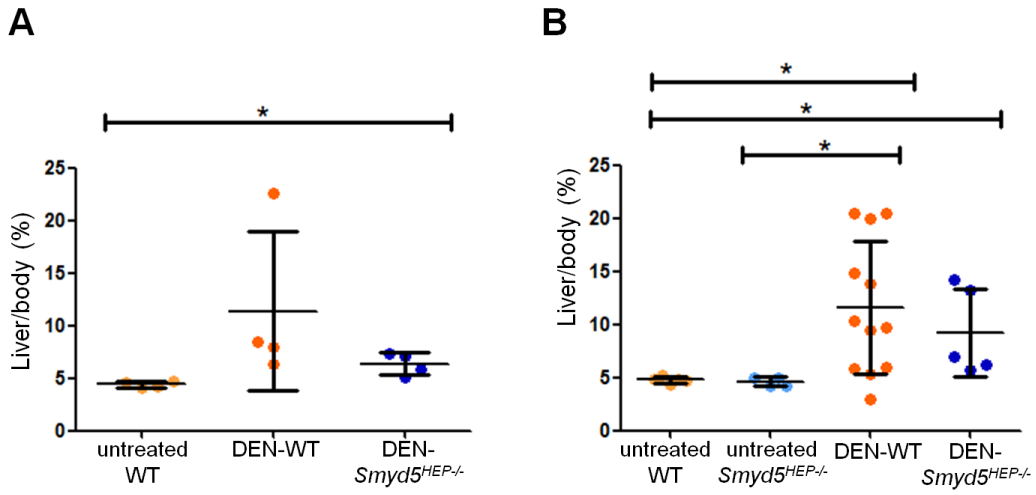


Figure 37. Percentage of liver mass respect to body mass in untreated and diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) livers. Quantifications (**A**) 7.5 and (**B**) 9.5 months after DEN treatment. Average values and standard deviation (±SD) were from at least N = 4 mice. T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05).

Only a few focal nodular hyperplastic areas were detected in histological sections of livers from liver-specific *Smyd5*-deficient mice. In contrast, the livers of DEN-treated WT mice contained several large cancerous foci with features of HCC, such as obliterated portal tracts, irregular trabeculae, frequent mitotic figures and cells displaying nuclear atypia, cytoplasmic clearance, or increased eosinophilic inclusions (**Figure 38**).

In healthy conditions the liver is a slow-cycling organ, however, cancer is associated with a high proliferation rate. In agreement with the highly reduced cancerous areas in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO livers, we detected a significantly decreased number of proliferating cells by Ki67 staining (**Figure 39A** and **39B**), concomitant to a dramatic reduction of E-cadherin staining intensity in the tumor areas of livers of carcinogen-treated WT but not in *Smyd5*^{flox}/Alfp-Cre KO (**Figure 39C**). Loss of E-cadherin expression in parallel with disruption of cell polarity or cell-cell contact and degradation of extracellular matrix are the hallmarks of EMT, a process involved in cancer initiation and metastatic spread (241).

Moreover, we measured alanine aminotransferase (ALAT) serum levels as an indicator of liver injury in untreated and DEN-treated WT, liver-specific *Smyd5*^{flox}/Alfp-Cre KO and full KO (*Smyd5*^{flox}/MX-Cre). ALAT levels were higher in DEN-treated WT mice serum compared to all the other groups (**Figure 40**).

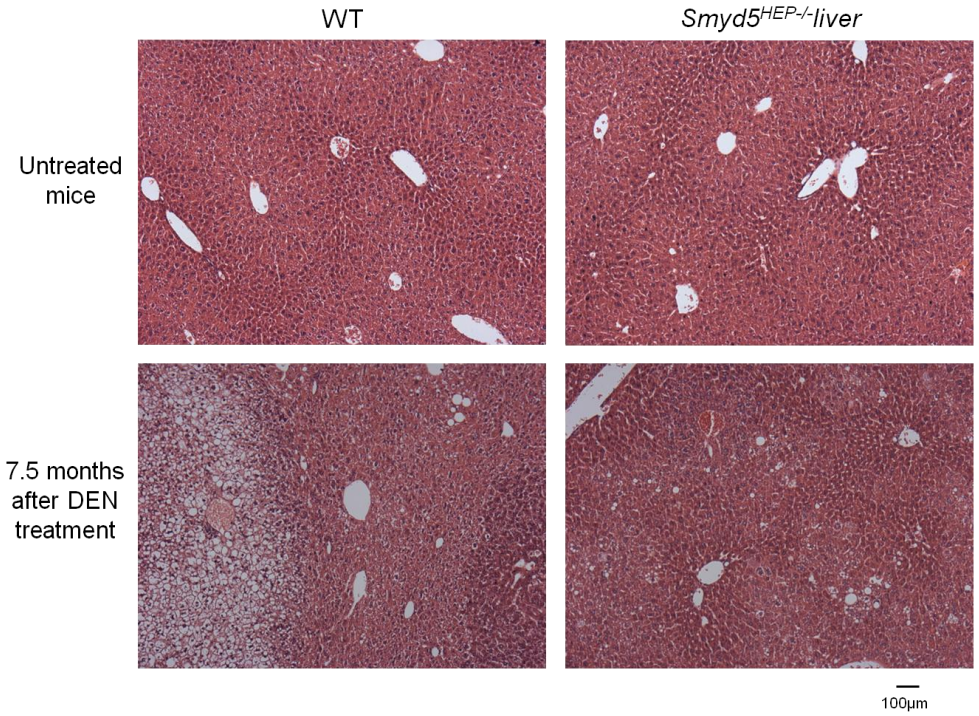


Figure 38. Representative hematoxylin and eosin (H&E) staining of formaldehyde-fixed paraffin-embedded liver sections of 8-month-old diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice.

Decreased ALAT serum levels indicated attenuated liver damage in DEN-treated liver-specific *Smyd5*^{flox}/Alfp-Cre KO compared to DEN-treated WT. Thus, we stained the liver sections with an antibody against γH2AX, a histone modification that indicates DNA damage. We observed reduced DNA damage in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO in comparison to DEN-treated WT (**Figure 41A**). Due to metabolic and signaling aberrations, cancer cells exhibit elevated ROS levels that can lead to DNA damage (242,243). We detected a decrease in ROS accumulation in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO compared to DEN-treated WT mice, by staining the liver sections with the H_2O_2-sensitive fluorescent dye CM-H_2DCFDA (**Figure 41B**).

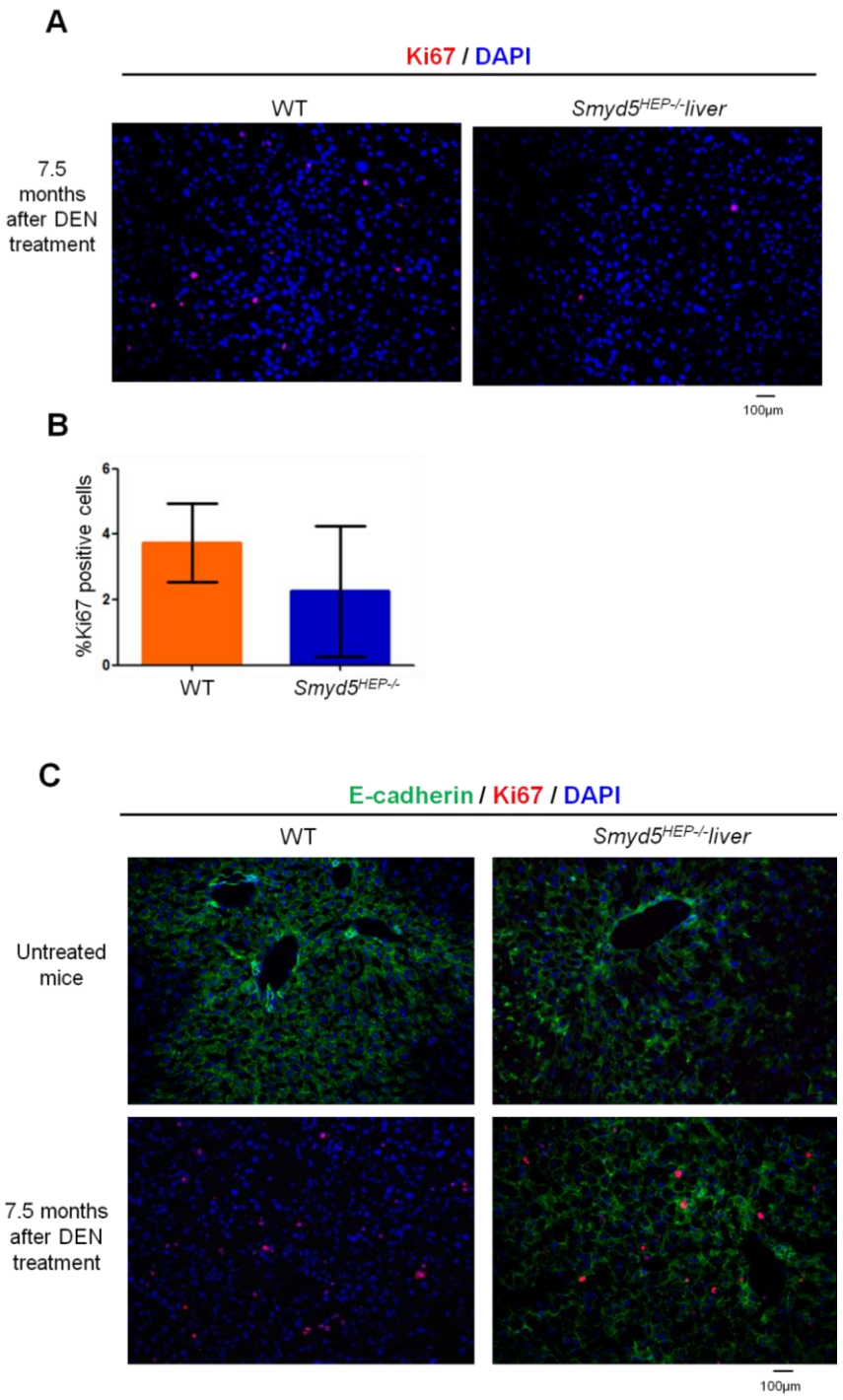


Figure 39. Cell proliferation after diethylnitrosamine (DEN) treatment. (**A**) Representative immunohistological staining of frozen liver sections from 8-month-old DEN-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with Ki67 antibody. Nuclei stained with DAPI. (**B**) Quantification of Ki67-positive cells in 8-month-old DEN-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice. Bars represent average percentages and standard deviation (±SD) of cells positively stained for Ki67 over all cells (estimated by DAPI staining) examined in 10 high-power fields (HPF) of sections from three different animals. (**C**) Representative immunohistological staining of frozen liver sections from 8-month-old untreated and DEN-treated WT and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with E-cadherin and Ki67 antibodies. Nuclei stained with DAPI.

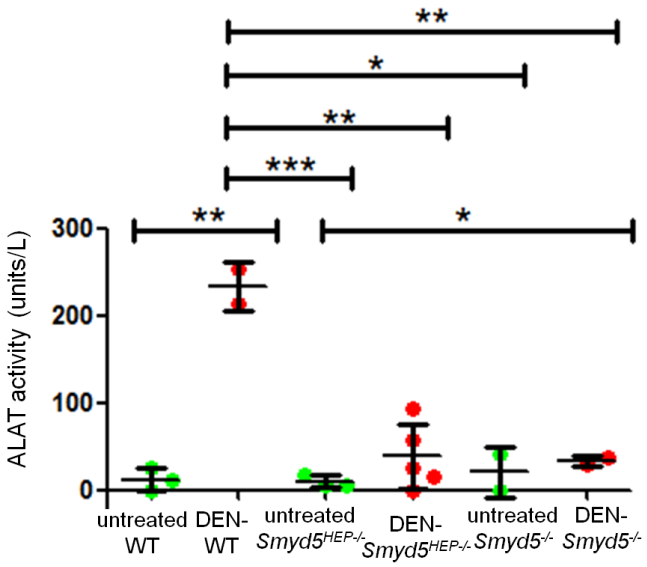


Figure 40. Alanine aminotransferase (ALAT) quantification in serum from untreated and diethylnitrosamine (DEN)-treated wild-type (WT), liver specific *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) and full KO (*Smyd5*$^{-/-}$) mice. Units per liter (units/L). Average values and standard deviation (±SD) were from at least N = 2 mice. T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05; **P value ≤ 0.01; ***P value ≤ 0.001).

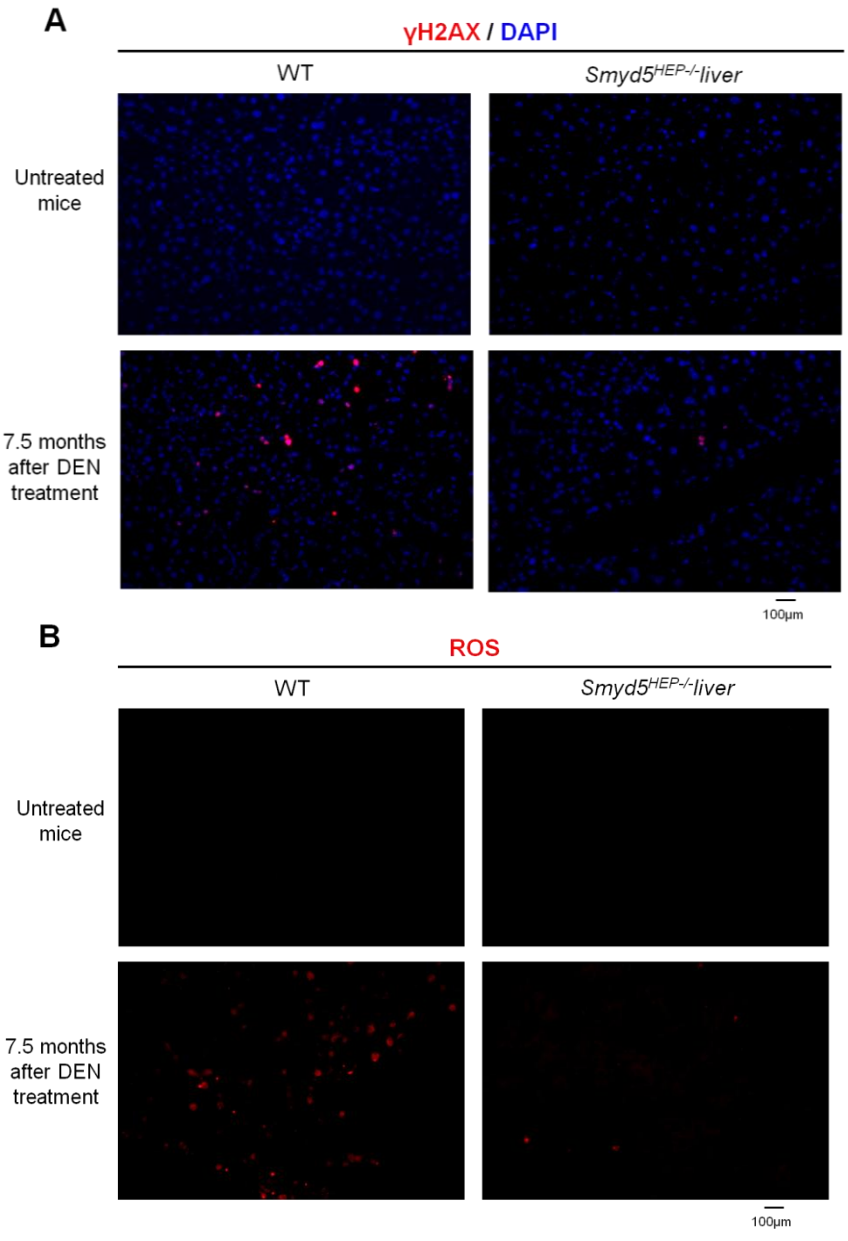


Figure 41. DNA damage after diethylnitrosamine (DEN) treatment. (**A**) Representative immunohistological staining of frozen liver sections from 8-month-old untreated and DEN-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with γH2AX antibody. Nuclei stained with DAPI. (**B**) Analysis of reactive oxygen species (ROS) accumulation in frozen liver sections from 8-month-old untreated and DEN-treated WT and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice by staining with the H_2O_2-sensitive fluorescent dye CM-H_2DCFDA.

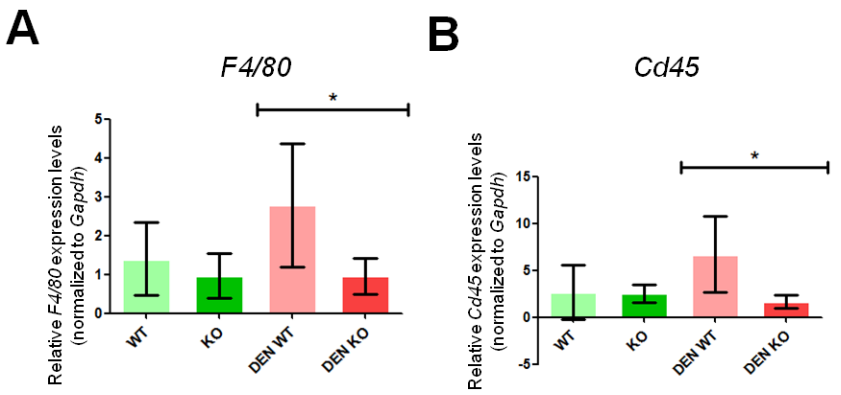


Figure 42. Quantitative real time PCR (qPCR) analysis of *F4/80* and *Cd45* mRNA levels in the livers of 8-month-old untreated and DEN-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre (KO) mice. Bars represent mean mRNA levels normalized to *Gapdh* mRNA and compared to WT and standard deviation (±SD) from each group (at least N = 3 individual mice). T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05).

Analysis of macrophage (F4/80) and pan-leukocyte (CD45) marker mRNA levels revealed a decrease in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO livers compared to DEN-treated WT (**Figure 42**). Contrary to the mRNA results, staining of F4/80 and CD45 proteins revealed a similar degree of inflammatory cell infiltration in the livers of both DEN-treated WT and *Smyd5*^{flox}/Alfp-Cre KO mice (**Figure 43A** and **43B**). Moreover, we stained liver sections with antibodies against the hepatocyte transcription factor HNF4α and the biliary epithelium cytokeratin CK19, which are markers of hepatocyte and biliary epithelium cell identity, respectively. Cell identity was not lost in any of the conditions studied (**Figure 44A** and **44B**), meaning that these chemically-induced tumors are not highly dedifferentiated. We concluded that the highly attenuated tumor formation in the livers of *Smyd5*^{flox}/Alfp-Cre KO mice can be due to an effect of SMYD5 on DNA damage and/or proliferation, but it is not associated with acute or chronic inflammation or abolishment of cell dedifferentiation processes.

RNA-seq reveals major changes in diethylnitrosamine-treated liver-specific knockout compared with diethylnitrosamine-treated wild-type mice

We next sought to elucidate what global changes in the mRNA expression landscape are concomitant to delayed carcinogenesis in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO livers (in collaboration with Panagiotis Moulos). First, we performed RNA-seq in livers from the four experimental conditions; untreated WT, untreated *Smyd5*^{flox}/Alfp-Cre KO, DEN-treated WT and DEN-treated *Smyd5*^{flox}/Alfp-Cre KO mice. RNA-seq data were used to

calculate pairwise correlations and perform multidimensional scaling. As shown in **Figure 45**, variations between WT, *Smyd5*^{flox}/Alfp-Cre KO, and DEN-treated *Smyd5*^{flox}/Alfp-Cre KO samples were lower than the degree of variation between any of samples and those from DEN-treated WT livers, indicating that DEN-WT was the most different condition regarding mRNA expression. Hierarchical clustering analysis of differentially expressed genes in DEN-treated WT and *Smyd5*^{flox}/Alfp-Cre KO mice revealed a highly distinctive pattern (**Figure 46**). Mean-difference plots comparing different conditions confirmed that subsets of genes were upregulated and downregulated in all comparisons and a high number of genes were upregulated in

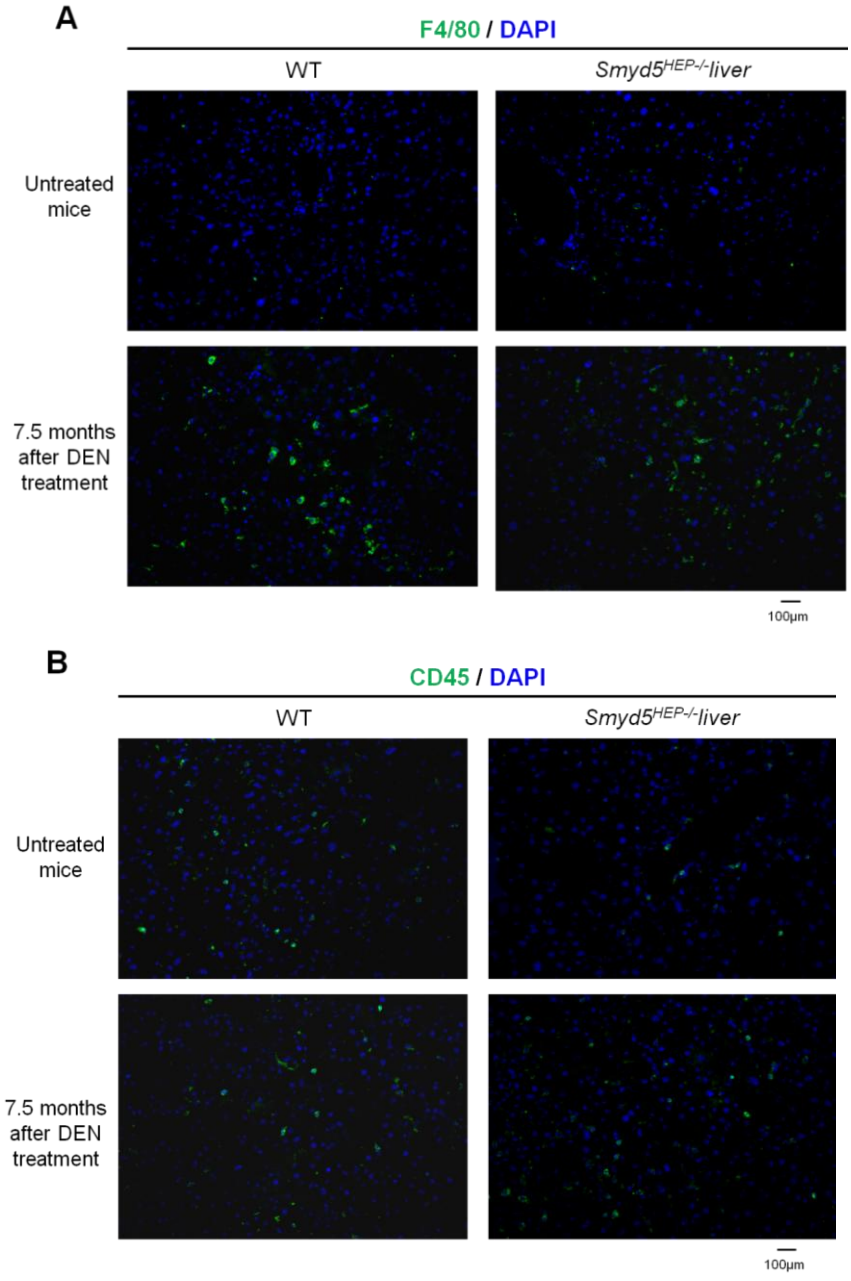


Figure 43. Immune cell infiltration after diethylnitrosamine (DEN) treatment. (**A**) Representative immunohistological staining of frozen liver sections from 8-month-old untreated and DEN-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with F4/80 antibody. Nuclei stained with DAPI. (**B**) Representative immunohistological staining of frozen liver sections from 8-month-old untreated and DEN-treated WT and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with CD45 antibody. Nuclei stained with DAPI.

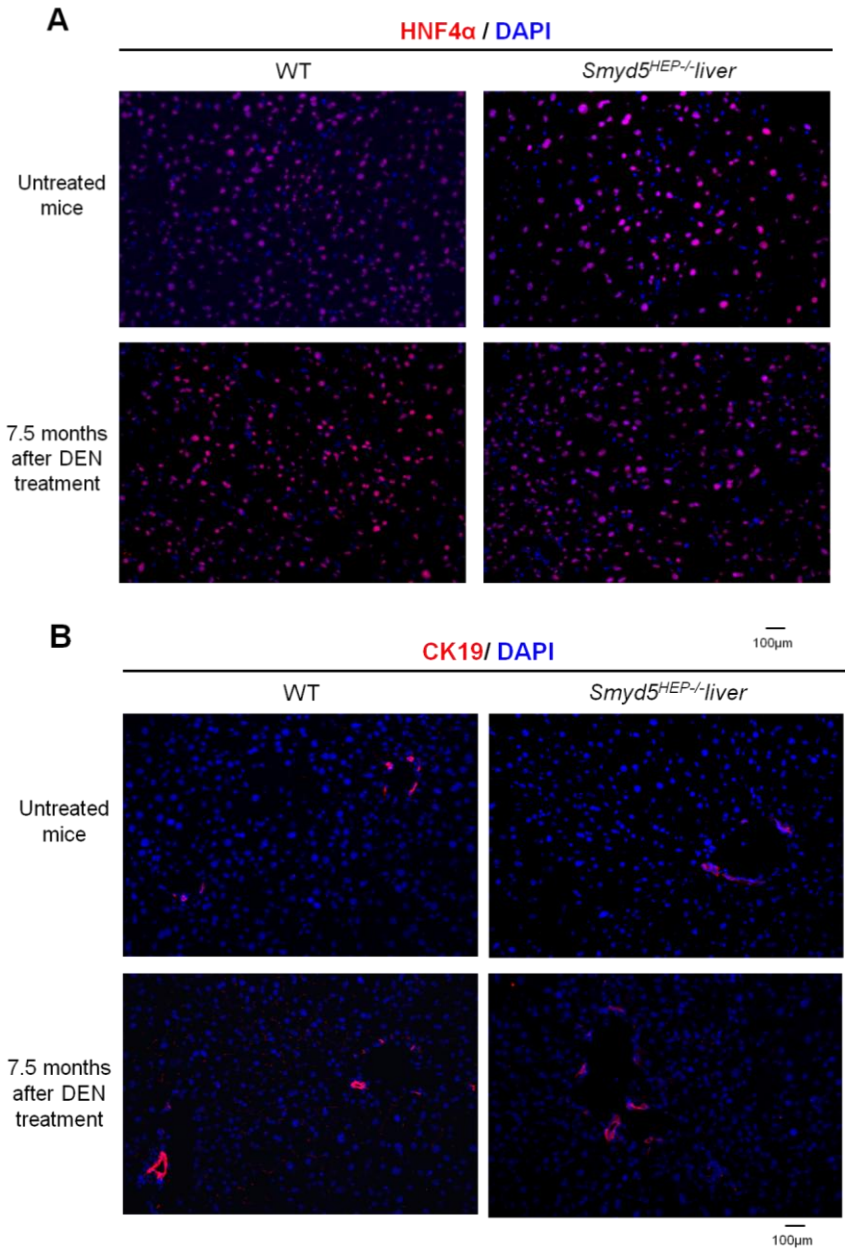


Figure 44. Cell identity after diethylnitrosamine (DEN) treatment. (**A**) Representative immunohistological staining of frozen liver sections from 8-month-old untreated and DEN-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with HNF4α antibody. Nuclei stained with DAPI. (**B**) Representative immunohistological staining of frozen liver sections from 8-month-old untreated and DEN-treated WT and *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) mice with CK19 antibody. Nuclei stained with DAPI.

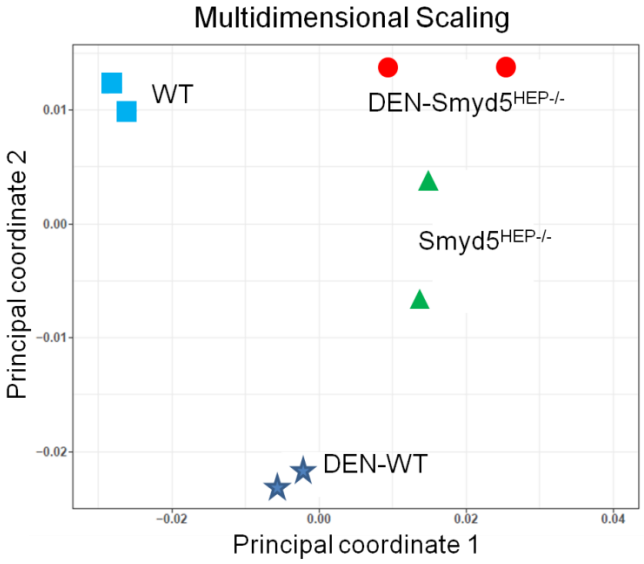


Figure 45. Multidimensional scaling (principal coordinate analysis) plots of the mRNA expression levels in the livers of 8-month-old untreated wild-type (WT), *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) and diethylnitrosamine (DEN)-treated wild-type (DEN-WT) or DEN-treated *Smyd5*^{flox}/Alfp-Cre KO (DEN-*Smyd5*$^{HEP-/-}$) mice.

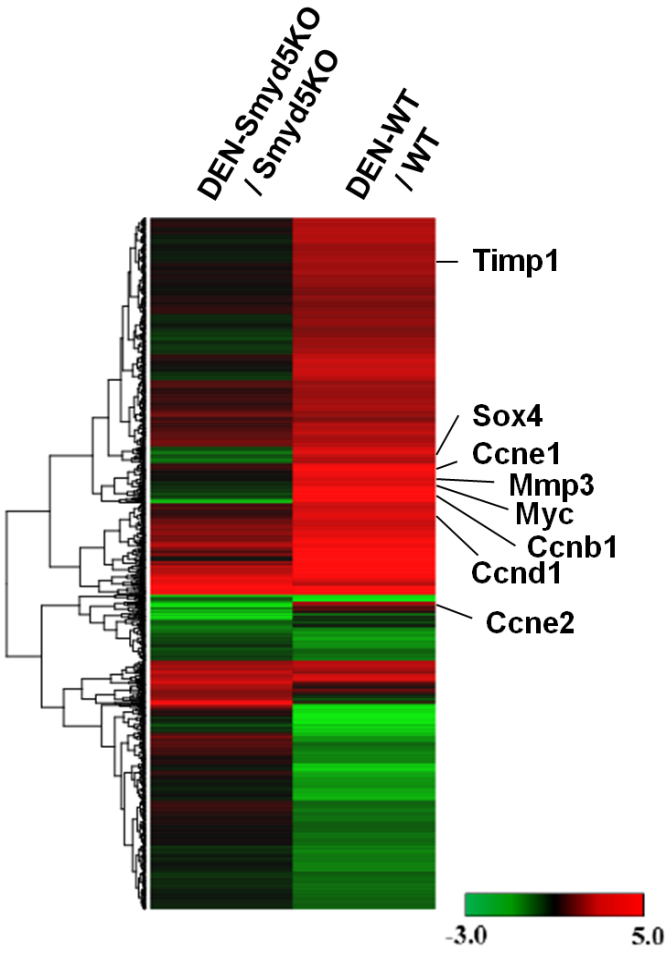


Figure 46. Hierarchical clustering analysis and corresponding heatmap of differentially expressed genes in the livers of diethylnitrosamine (DEN)-treated versus untreated 8-month-old wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre KO (Smyd5KO) mice. The color scale bar indicates the fold change ranges for the comparisons (ratios) indicated on the top of the heatmap. Representative genes the differential expression of which was validated by quantitative real time PCR (qPCR) are indicated on the right; Tissue inhibitor of metalloproteinase 1 (*Timp1*), SRY (sex determining region Y)-box 4 (*Sox4*), Cyclin E1 (*Ccne1*), Matrix metallopeptidase 3 (*Mmp3*), Myelocytomatosis oncogene (*Myc*), Cyclin B1 (*Ccnb1*), Cyclin D1 (*Ccnd1*), and Cyclin E2 (*Ccne2*).

cancerous DEN-treated WT compared to untreated WT (**Figure 47**). Differential expression of genes in which dysregulated activity has been involved in carcinogenesis was validated by qPCR, confirming its downregulation in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO livers compared with DEN-treated WT (**Figure 46** and **48**). Moreover, we observed downregulation of other genes related to carcinogenesis in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO livers compared with DEN-treated WT (**Figure 49**). In addition, *Smyd5* expression was increased in these DEN-treated WT samples compared with untreated WT (**Figure 49**). *Jak1* and *Jak2* (genes encoding for tyrosine kinases known to be involved in malignant processes (244)) mRNA levels were also upregulated in DEN-treated WT compared to both untreated and DEN-treated *Smyd5*^{flox}/Alfp-Cre KO (**Figure 50**).

In conclusion, bioinformatic analysis of RNA-seq data and subsequent qPCR validation confirmed the macroscopic and histological phenotype of delayed carcinogenesis in DEN-treated *Smyd5*^{flox}/Alfp-Cre KO livers when compared to DEN-treated WT.

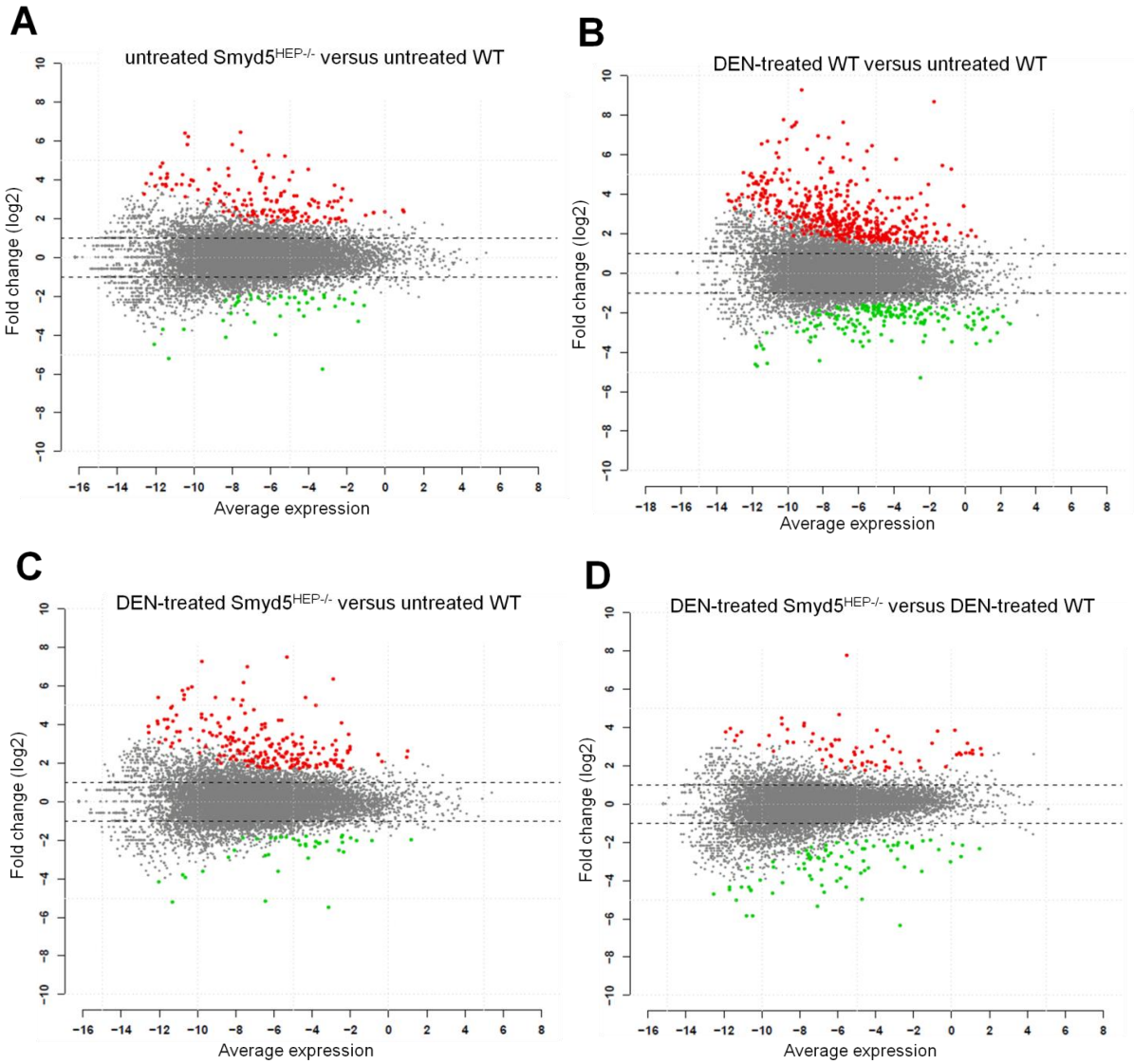
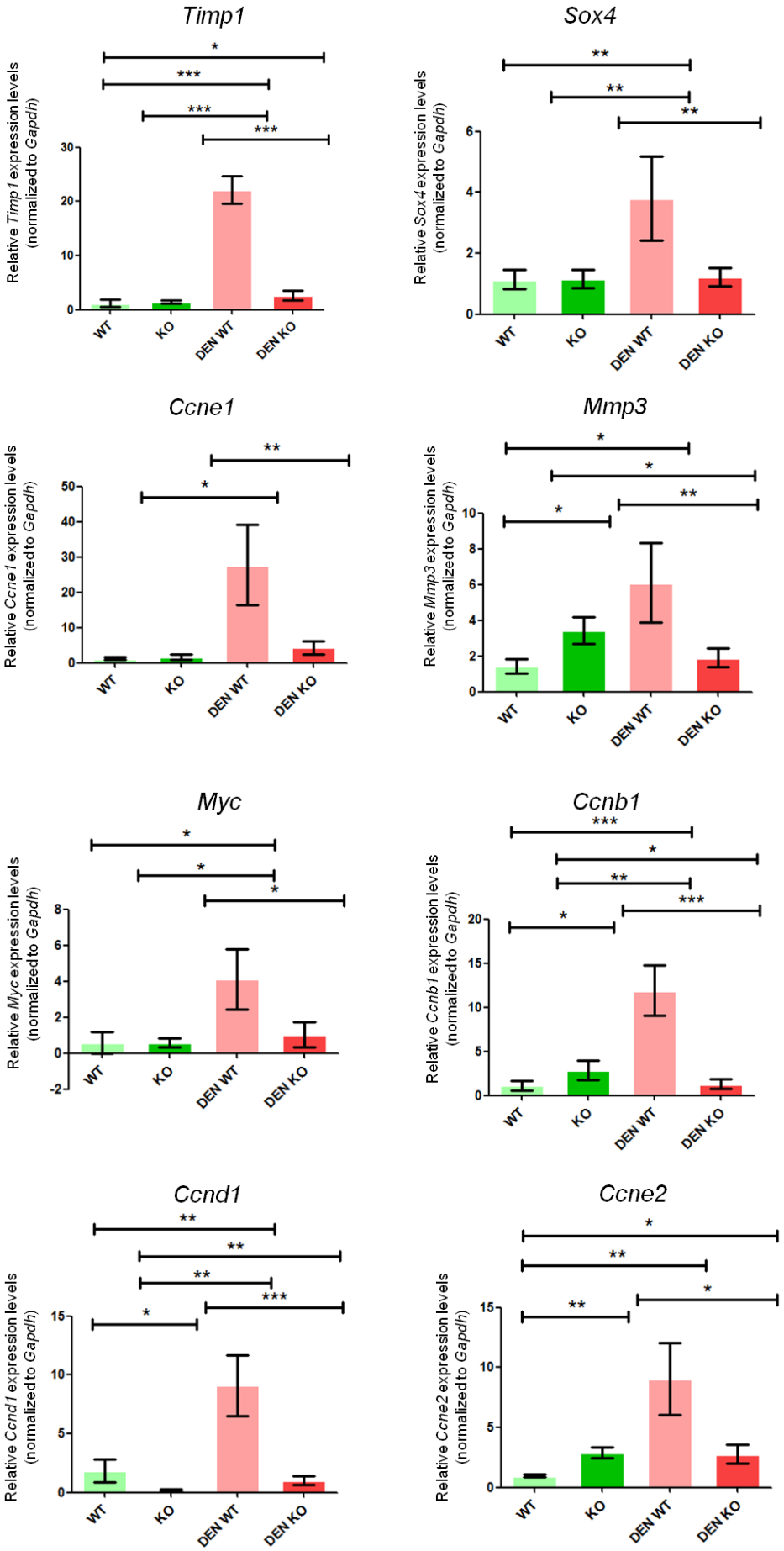
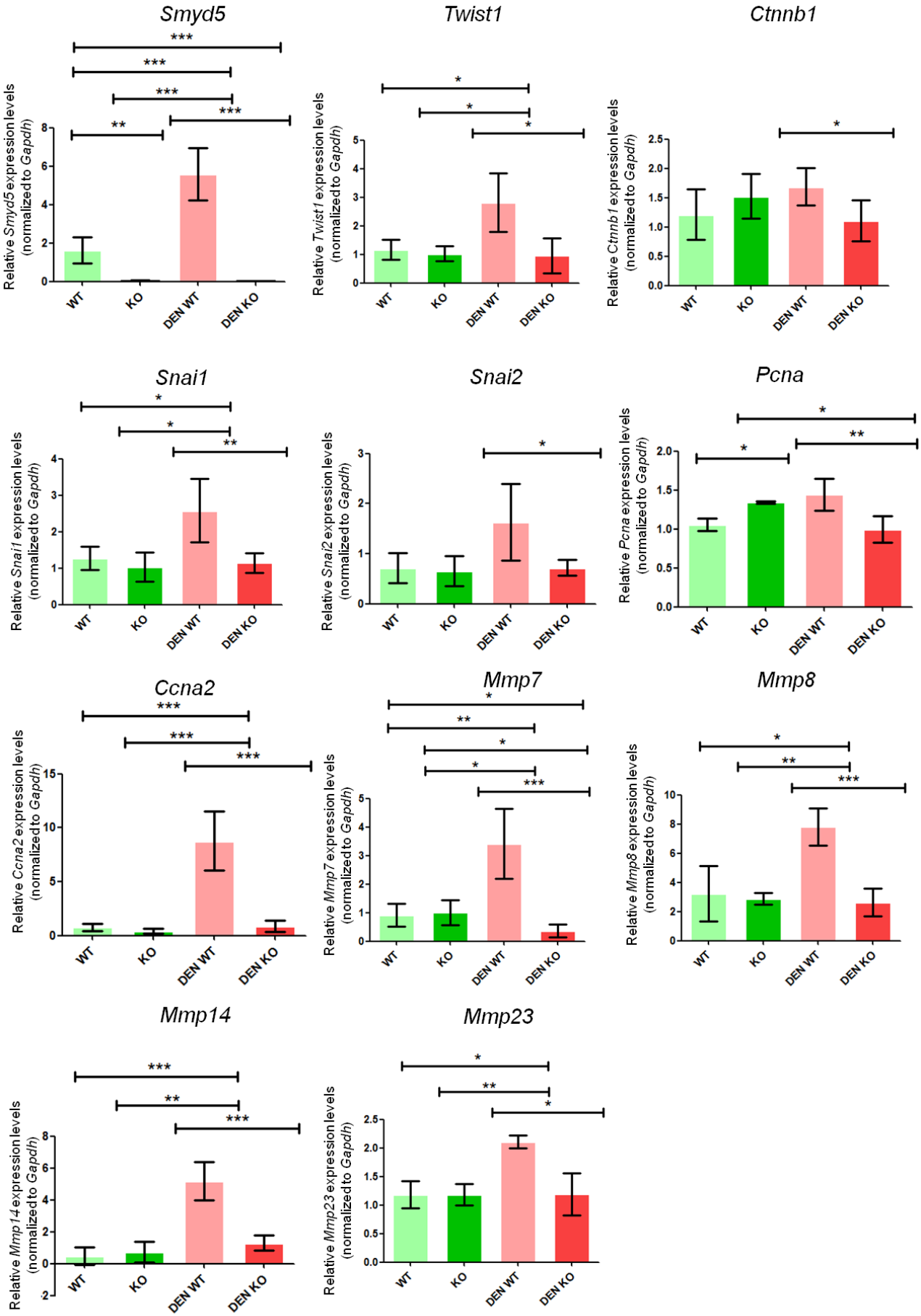


Figure 47. Mean-difference plots corresponding to genes differentially expressed in (**A**) untreated *Smyd5*^{flox}/Alfp-Cre KO (*Smyd5*$^{HEP-/-}$) versus untreated wild-type (WT), (**B**) diethylnitrosamine (DEN)-treated WT versus untreated WT, (**C**) DEN-treated *Smyd5*$^{HEP-/-}$ versus untreated WT and (**D**) DEN-treated *Smyd5*$^{HEP-/-}$ versus DEN-treated WT. Differences (fold change) between the two variables are plotted against the averages of the two. Significantly upregulated genes (False Discovery Rate [FDR] < 0.05) are depicted in red, significantly downregulated genes (FDR < 0.05) are depicted in green, and other genes are depicted in grey. Dashed lines represent upper and lower fold change thresholds.

(Next page) Figure 48. Quantitative real time PCR (qPCR) analysis of Tissue inhibitor of metalloproteinase 1 (*Timp1*), SRY (sex determining region Y)-box 4 (*Sox4*), Cyclin E1 (*Ccne1*), Matrix metallopeptidase 3 (*Mmp3*), Myelocytomatosis oncogene (*Myc*), Cyclin B1 (*Ccnb1*), Cyclin D1 (*Ccnd1*), and Cyclin E2 (*Ccne2*) mRNA levels in the livers of 8-month-old untreated and diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre (KO) mice. Bars represent mean mRNA levels normalized to *Gapdh* mRNA and compared to WT and standard deviation (±SD) from each group (at least N = 3 individual mice). T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05; **P value ≤ 0.01; ***P value ≤ 0.001).


Timp1
Relative Timp1 expression levels (normalized to Gapdh)
WT
KO
DEN WT
DEN KO
Sox4
Relative Sox4 expression levels (normalized to Gapdh)
Ccne1
Relative Ccne1 expression levels (normalized to Gapdh)
Mmp3
Relative Mmp3 expression levels (normalized to Gapdh)
Myc
Relative Myc expression levels (normalized to Gapdh)
Ccnb1
Relative Ccnb1 expression levels (normalized to Gapdh)
Ccnd1
Relative Ccnd1 expression levels (normalized to Gapdh)
Ccne2
Relative Ccne2 expression levels (normalized to Gapdh)


Smyd5
Relative Smyd5 expression levels (normalized to Gapdh)
WT
KO
DEN WT
DEN KO
Twist1
Relative Twist1 expression levels (normalized to Gapdh)
Ctnnb1
Relative Ctnnb1 expression levels (normalized to Gapdh)
Snai1
Relative Snai1 expression levels (normalized to Gapdh)
Snai2
Relative Snai2 expression levels (normalized to Gapdh)
Pcna
Relative Pcna expression levels (normalized to Gapdh)
Ccna2
Relative Ccna2 expression levels (normalized to Gapdh)
Mmp7
Relative Mmp7 expression levels (normalized to Gapdh)
Mmp8
Relative Mmp8 expression levels (normalized to Gapdh)
Mmp14
Relative Mmp14 expression levels (normalized to Gapdh)
Mmp23
Relative Mmp23 expression levels (normalized to Gapdh)

(Previous page) Figure 49. Quantitative real time PCR (qPCR) analysis of *Smyd5 (*with the primer set targeting exon 4 to 6), Twist basic helix-loop-helix transcription factor 1 (*Twist1*), Catenin (cadherin associated protein) beta 1 (*Ctnnb1*), Snail family zinc finger 1 (*Snai1*) and 2 (*Snai2*), Proliferating cell nuclear antigen (*Pcna*), Cyclin A2 (*Ccna2*) and Matrix metalloproteases (*Mmp 7, 8, 14, 23*) mRNA levels in the livers of 8-month-old untreated and diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre (KO) mice. Bars represent mean mRNA levels normalized to *Gapdh* mRNA and compared to WT and standard deviation (±SD) in each group (at least N = 3 individual mice). T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05; **P value ≤ 0.01; ***P value ≤ 0.001).

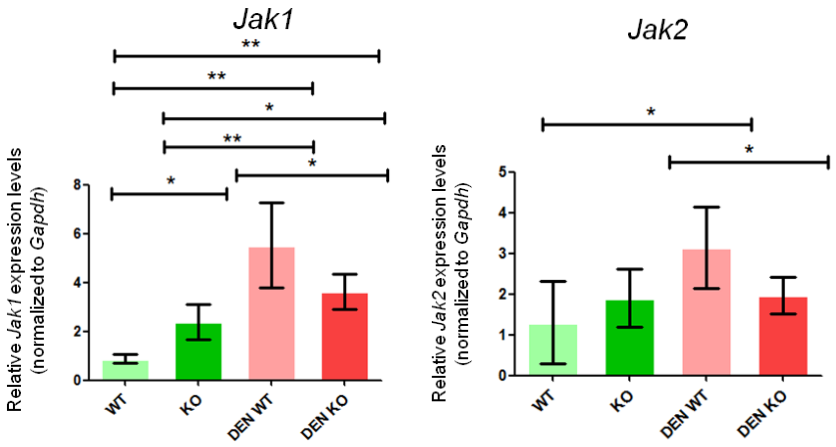


Figure 50. Quantitative real time PCR (qPCR) analysis of Janus kinase 1 (*Jak1*) and 2 (*Jak2*) mRNA levels in the livers of 8-month-old untreated and diethylnitrosamine (DEN)-treated wild-type (WT) and *Smyd5*^{flox}/Alfp-Cre (KO) mice. Bars represent mean mRNA levels normalized to *Gapdh* mRNA and compared to WT and standard deviation (±SD) in each group (at least N = 3 individual mice). T-test was performed in all cases, comparing groups one by one (*P value ≤ 0.05; **P value ≤ 0.01).

Smyd5 intestine-specific knockout mice do not show differences in carcinogenesis after 1,2-dimethylhydrazine / dextran sodium sulfate treatment compared to wild-type

Analysis of *SMYD5* expression in colon adenocarcinoma from the TCGA database showed that the probability of progressing to high-grade (G3/G4) colon tumors and the probability of developing metastasis was lower in patients with low *SMYD5* expression compared to patients with high *SMYD5* expression (**Figure 30E** and **30F**). These results prompted us to investigate potential causative relationship between *SMYD5* overexpression and colon cancer. Following the same rationale as that which we previously applied to the study of SMYD5 in liver cancer, WT and intestine-specific *Smyd5*^{flox}/Villin-Cre KO mice were subjected to 1,2-dimethylhydrazine / dextran sodium sulfate (DMH/DSS) treatment to chemically induce colon cancer. The protocol starts

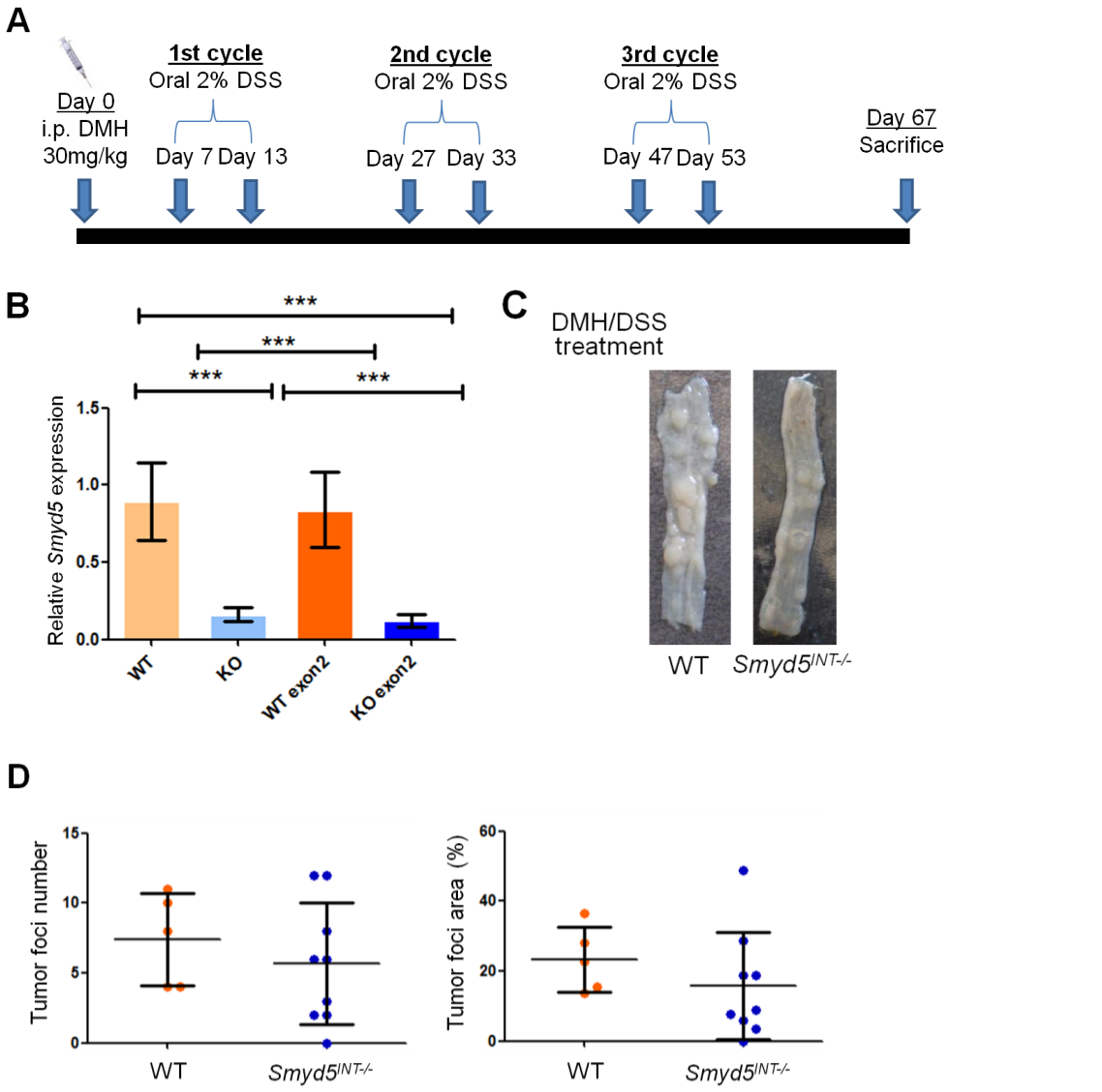


Figure 51. Analysis of 1,2-dimethylhydrazine/ dextran sodium sulfate (DMH/DSS)-treated wild-type (WT) and *Smyd5*^{flox}/Villin-Cre KO livers. (**A**) Two-month-old mice were subjected to a single intraperitoneal (i.p.) injection of 30 mg DMH per mouse kg**.** One week after injection, the animals were given 2% (w/v) DSS in drinking water for 6 days followed by pure drinking water (without any additional substance) for 14 days. The 6-day DSS cycle was repeated two more times with a 14-day interval of pure drinking water, after which the mice were sacrificed for the analysis of colon tumor formation. (**B**) Relative *Smyd5* expression levels were analyzed by quantitative real time PCR (qPCR) in WT and *Smyd5*^{flox}/Villin-Cre KO intestines using one primer set targeting exon 4 to 6 ("WT", "KO" samples) and one primer set targeting exon 2 ("WT exon 2", "KO exon 2") (see **Figure 33A**). Bars represent mean mRNA levels normalized to *Gapdh* mRNA and compared to WT. Average values and standard deviation (±SD) were from at least N = 3 mice per group. T-test was performed in all cases, comparing groups one by one (***P value ≤ 0.001). (**C**) Macroscopic appearance of colons in DMH/DSS-treated WT and *Smyd5*^{flox}/Villin-Cre KO mice (*Smyd5*$^{INT-/-}$). (**D**) Quantification of tumor foci number and area in DMH/DSS-treated WT and *Smyd5*$^{INT-/-}$. Average values and standard deviation (±SD) were from at least N = 5 mice. T-test was performed in all cases, comparing groups one by one.

with a single intraperitoneal injection of DMH, followed by repetitive cycles of exposure to DSS in drinking water alternated with periods of pure drinking water (106,245) (**Figure 51A**). Unfortunately we failed to observe obvious differences in tumor formation capacity between colons in which *Smyd5* was expressed (WT) and colons devoid of *Smyd5* expression (*Smyd5*^{flox}/Villin-Cre KO), probably due to high intragroupal variability (**Figure 51B** and **51C**). Quantification of tumor foci number and area confirmed the high variability between individual mice. Although there may be a trend, differences between WT and intestine-specific *Smyd5*^{flox}/Villin-Cre KO colons were not significant (**Figure 51D**). Similar results were obtained for DMH/DSS-treated *Smyd5* full KO (*Smyd5*^{flox}/MX-Cre) colons (data not shown).

In conclusion, our data from human samples and murine models points to an important role for SMYD5 in liver cancer.

Discussion

Summary of our results

This study sheds light on the function of the SET and MYND domain-containing protein SMYD5 in liver carcinogenesis in mice. Here we have shown that SMYD5 methyltransferase acts on H2A and H2B *in vitro* (**Figure 26**). Analysis of *SMYD5* expression in TCGA datasets of HCC revealed a significant positive correlation between *SMYD5* expression and the incidence of HCC. In addition, *SMYD5* expression positively correlated with clinicopathological parameters of bad prognosis (**Figure 29**). To further investigate the mechanism by which SMYD5 could be involved in cancer formation we have generated *Smyd5*-deficient mice, lacking *Smyd5* mRNA and protein either in all tissues (*Smyd5*^{flox}/MX-Cre) or specifically in the liver (*Smyd5*^{flox}/Alfp-Cre and *Smyd5*^{flox}/Alb-Cre) (**Figure 32**). None of these mice displayed any obvious abnormal phenotype, neither during development nor in adulthood (**Figure 33**). Nevertheless, we found that *Smyd5* expression is required for the development of chemically-induced liver cancer (**Figure 35** and **36**), and we focused on the analysis of the delayed liver carcinogenesis phenotype observed in our *Smyd5*-deficient mice at the histological and gene expression level (**Figures 36** to **50**). Additionaly, a similar study was conducted for colon cancer. We analyzed TCGA datasets of colon cancer and chemically-induced colon cancer in intestine-specific *Smyd5*-deficient mice (*Smyd5*^{flox}/Villin-Cre). In this case TCGA analysis was not so promising (**Figure 30**), and no obvious attenuated tumor formation phenotype was observed in *Smyd5*^{flox}/Villin-Cre (**Figure 51**). In summary, we discovered SMYD5 methyltransferase activity in previously unknown histone substrates *in vitro*. We also describe an *in vivo* role for SMYD5 in liver cancer progression in mice, which is in accordance with patient data. Further studies are needed to clarify the mechanism by which SMYD5-mediated H2B and/or H2A methylation regulates cancer development.

Histone targets of SMYD5 and their role in regulation of gene expression and carciogenesis

At the beggining of this study little was known about SMYD5. The first mechanistic study (213) described SMYD5 as a histone methyltransferase capable of depositing H4K20me3 marks at the promoters, by its association with NCoR corepressor complexes. H4K20me3 is a repression checkpoint that restricts expression TLR4 target genes in macrophages (213). Stender *et al.* performed *in vitro* H3 and H4 methylation assays and concluded that bacteria-purified GST-SMYD5 fusion protein exhibits methyltransferase activity in recombinant H4, but not H3. The main histone methylation

sites in H4 are H4K20 and H4R3, and SET domains are lysine-specific. Accordingly, after incubation of bacteria-expressed GST-SMYD5 with H4, methylation-specific antibodies detected H4K20me3 as a reaction product, but not H4K20me2. FLAG-SMYD5 produced in mammalian cells also had the ability to deposit H4K20me3 marks in unmodified H4, but not chemically-generated H4K20me3. Conversely, FLAG-SMYD5 could not generate H3K4me3, H3K9me3, H3K27me3, or H3K36me3 marks. Importantly, mutations in the SET catalytic domain of SMYD5 abolished its methyltransferase activity (213).

Therefore, first of all we wanted to identify SMYD5-mediated methylation targets *in vitro*, attempting to confirm the above findings. Unexpectedly, our *in vitro* methylation assay (**Figure 26**) clearly concludes that baculovirus-purified SMYD5 methylates H2A and H2B, but not H4.

In summary, despite the mounting data indicating that SMYD5 is capable of generating H4K20me3 epigenetic marks (213), we could not reproduce or confirm this result. Instead, we identified H2A and H2B as SMYD5-mediated methylation histone targets (**Figure 26**). At this point we do not know which lysine residue(s) in H2A or H2B is methylated. Possible candidates are H2BK5, H2BK37 and H2BK120. Some of these residues are known to be subjected to other types of modifications.

Indeed, H2A and H2B are well established substrates for ubiquitination. In a context-dependent manner, both writers and erasers of H2A and H2B ubiquitination (H2Aub and H2Bub) can act as tumor suppressors or oncogenes (246). For example, E3 ubiquitin ligase RNF20 typically catalyzes H2BK120 monoubiquitination (H2BK120ub1), resulting in tumor growth suppression, while E3 ubiquitin ligase TRIM37 can behave as an oncogene by mediating monoubiquitination of H2AK119 (H2AK119ub1), which facilitates silencing of tumor suppressors and other genes and promotes cancer development (246–248). USP22 deubiquitinase typically promotes tumor growth by acting on both H2A and H2B, although a role for USP22 as a tumor suppressor is emerging (249). H2A and H2B are also subjected to acetylation. In fact, acetylation of H2BK5 (H2BK5ac), in a specific set of genes may be involved in the maintainance of the epithelial phenotype, thereby controlling EMT programme activation (250). As it is typical for histone PTMs, PTMs in H2A or H2B participate in crosstalk with other chromatin marks. For example, H2Bub is required for H3K4me3 and H3K79me3, which are gene activation marks (246,251,252). In addition, PTMs of histone variants can also regulate carcinogenesis; H2A.Z primarily harbors oncogenic properties, while H2A.X and macroH2A generally act as tumor suppressors (253).

Finally, mutations occuring in critical regions of different histones (known as oncohistone mutations), including H2A and H2B, are found in many human cancers and can have an impact on the disease (254). Therefore, although many studies have focused on PTMs of H3 and H4, the biological relevance of PTMs of H2A and H2B is undoubted.

Subcellular localization of SMYD5: concerns about existing antibodies and implications for our project

A major consideration is the subcellular localization of the SMYD5 protein. The results of our *in vitro* methylation assay would logically lead to the assumption that SMYD5 is localized in the nucleus. However, SMYD family members can be found in the nucleus or the cytoplasm, suggesting that they might play different roles by targeting different substrates depending on their subcellular localization; these targets are typically histone substrates in the nucleus and non-histone proteins in the cytoplasm. For example, SMYD2 dimethylates H3K36 in the nucleus and mediates inhibitory monomethylation of p53 at K370 and pRb at K860 in the cytoplasm, among many other targets in both subcellular compartments (255). Of note, SMYD proteins lack a nuclear localization signal, which suggests that they may bind to other factors in order to be transported to the nucleus (256). Importantly, our results indicate that SMYD5 is abundant in the cytoplasm of murine healthy liver cells and the murine hepatoma cell line Hepa1-6, whereas its presence in the nucleus is not clear (**Figure 27**).

Taking all these data into account, we hypothesize that, similarly to other members of the SMYD protein family, SMYD5 may be preferentially localized in the cytoplasm, but can be translocated to the nucleus upon different stimuli, for example during cell cycle entrance, to act as a lysine methyltransferase *in vivo*. This hypothesis is based on our results (**Figure 26** and **27**) as well as compelling data regarding other SMYD proteins (see above), but has an important limitation: we failed to find specific reliable antibodies for the SMYD5 protein that perform well in immunoblotting and staining experiments. Accordingly, to date there is no published immunoblotting or staining experiment using an antibody against endogenous SMYD5. For example, one study performed ChIP against FLAG-tagged SMYD5 in embryonic stem cells but not against the endogenous protein (216).

Nevertheless, we made an important effort to optimize conditions to detect the SMYD5 protein by immunoblotting and staining. We managed to detect SMYD5 by Western blot with our Smyd5 antibody raised in-house by collaborators using rabbits as hosts (**Figure 27** and **33C**, see Materials and Methods). However, interpretation of the results

when using cell lines is complex and even more difficult when liver samples are used instead, due to the presence of unspecific bands that can vary among tissues (data not shown) and differences in signal intensity among antibody aliquots. Regarding immunostaining, we could not detect SMYD5 with any of the antibodies tested, including our in-house antibody, or with many anti-SMYD5 antibodies raised in rabbit (ab81419, Abcam), goat (sc-82738, Santa Cruz) or mouse (OAAH00025, Aviva Systems Biology; orb48125, Biorbyt). None of the listed antibodies worked properly for staining, since they produced no signal, very faint signal or non-specific staining. Staining was considered non-specific when WT and KO tissue showed the same staining pattern. The above mentioned considerations should be taken into account when designing new experiments to further explore the biological role of SMYD5 methyltransferase.

Indeed, one of the experiments planned in this study was ChIP-seq with SMYD5 antibody, which could not be optimized in murine whole liver samples due to technical limitations (lack of specific antibodies for SMYD5) and/or biological limitations (the potentially dynamic subcellular localization of SMYD5, regulated by currently known stimuli). Lack of ChIP-seq data hinders the assessment of SMYD5 activity in histones and its biological relevance for cancer. After addressing these constraints, future studies need to focus on the link between SMYD5 methyltransferase activity in H2A and/or H2B and liver cancer development. Nevertheless, the existence of other nuclear or cytoplasmic targets for SMYD5 that may play a role in carcinogenesis cannot be ruled out.

SMYD5 and other SMYD family members in human and murine liver cancer

In parallel to the study of SMYD5 enzymatic activity, we analyzed *SMYD5* expression in TCGA datasets of HCC. We found that *SMYD5* expression positively correlates with the incidence of HCC and with clinicopathological parameters of bad prognosis in HCC (**Figure 29**). These results emphasize the clinical revelance of SMYD5 for cancer and therefore serve as a rationale for the translational potential of this study into clinical practice.

However, the robust correlations of *SMYD5* mRNA expression and the clinicopathological features of HCC do not necessarily mean that SMYD5 is a driver of cancer, since correlation does not necessarily imply causality. Thus, in this study we demonstrated that SMYD5 is a driver of liver carcinogenesis by means of our KO murine models. We concluded that absence of SMYD5 delays DEN-induced

carcinogenesis in mice, without completely abolishing it. Nevertheless, the phenotype of our DEN-treated KO mice is not as striking as expected from the TGCA results, since SMYD5 may promote liver cancer development, but it is not fully required for cancer initiation. A possible explanation for this could be the partial functional redundancy of SMYD5 with other members of the SMYD protein family.

For years, members of our research group, including myself, have been maintaining colonies of SMYD3 full KO, SMYD2 full KO and SMYD2 liver-specific KO (*Smyd2*^{flox}/Alfp-Cre) mice. We concluded that all these genetically-modified mouse strains are viable and healthy, meaning that these genes are non-essential in mice. Bagislar *et al.* bred SMYD2 KO mice, crossed them with SMYD3 KO mice to generate the double KO, and reached the same conclusions regarding viability and health (240). Collectively, observations by our research group and others concluded that SMYD2, SMYD3 and SMYD5 are dispensable for murine development and survival, suggesting a degree of functional redundancy that could be proven by generating and studying a triple KO murine model.

In addition, a study on SMYD3 full KO mice by our research group (211) used an experimental set up similar to ours and reached similar conclusions; SMYD3 full KO mice show delayed carcinogenesis in liver after DEN treatment. In contrast to our results, SMYD3 full KO mice also show delayed colon carcinogenesis after DMH/DSS treatment (211). In any case, the absence of SMYD3 and SMYD5 protein does not completely abolish the development of cancer but has a protective effect.

SMYD5 in different cell types and cancer stages

The concept of preponderance of different “forces” pulling towards opposite outputs (promotion or inhibition of cancer development) is much obvious when these “forces” are cellular types. The results of this study suggest that all SMYD5 KO models develop liver cancer at a slower rate than WT mice. Thus, these models provide no clue about the impact of SMYD5 depletion in different cell types, since the final output regarding cancer development is similar in all of them. In contrast to our results, a study using IKKβ KO mice (257) showed that, upon DEN treatment, hepatocyte-specific IKKβ KO (*IKKβ*^{flox}/Alb-Cre) mice develop more aggressive liver cancer than WT mice, while decreased carcinogenesis was observed in mice lacking IKKβ in both hepatocytes and hematopoietic-derived Kupffer cells, thereby emphasizing the opposite role of IKKβ in immune cell types (Kupffer cells) and parenchymal cells (hepatocytes) in this experimental set up (257). Similarly, JNK deficiency in both parenchymal and non-parenchymal cells protects against hepatitis and liver

cancer development are suppressed in myeloid cell-specific JNK KO mice, indicating that myeloid cells are responsible for the phenotype (258–260).

As mentioned above, we did not observe any differences in the cancer phenotype of SMYD5 full KO and cell-specific KO mice and, therefore, we did not generate immune cell types-specific SMYD5 KO mice that could allow a description of the role of SMYD5 in each cell type.

Cancer development involves interactions between a variety of cell types that are dynamic in time. In the present study mice were sacrified at time points when cancer is well established in WT mice, meaning that we do not have information about the process of cancer development, but, according to the multi-step nature of carcinogenesis, we assume that liver cancer development occurs at a slower rate in our KO models. Importantly, we can affirm that SMYD5 is not crucial for cancer initiation since our KO developed cancer, but we have no other information about the time point at which SMYD5 is capable of controlling cancer development in our experimental model. Molecular studies such as RNA-seq at initial stages of cancer development are needed to solve this issue.

The possibility that the absence of SMYD5 in different cell types and/or at different stages of carcinogenesis (initiation, progression, metastasis) could play distinct and even antagonistic roles cannot be entirely excluded, although the final output of the hypothetical interactions of all these cell types is a similar degree of carcinogenesis in each SMYD5 KO model.

SMYD5 in cancer and healthy liver stem cells

Indeed, cancer biology cannot be understood without considering the crosstalk between many cell types, including tumor subpopulations (see Introduction: “The relevance of cancer stem cells and tumor microenvironment” section). Cancer stem cells are considered the most clinically relevant tumor population, because their stem-like properties make them key drivers of cancer development and metastasis, as well as tumor recurrence after treatment. Previously published data (216,220), as well as our own, indicate that SMYD5 might play a role in maintaining stemness.

SMYD5 in colon cancer and metastasis

SMYD5 mRNA levels do not affect the overall survival and disease-free survival in colon cancer patients (**Figure 30C** and **30D**), which is in accordance with our results in SMYD5 KO mice (**Figure 51**). Nevertheless, we found an increased probability of developing metastasis in other organs in patients with high SMYD5 expression (**Figure**

30F). DMH/DSS-treated mice do not spontaneously develop metastatic lessions. Indeed, most chemically-induced cancer murine models do not recapitulate this tumoral feature, which accounts for most of the cancer-related deaths in human patients. Thus, metastatic models usually require genetically-modified mice and/or tumor transplantation (261). The research group led by Eduard Batlle has recently created a colon cancer metastasis murine model that faithfully mimmicks human colon cancer. They generated a quadruple mutant mouse bearing the four central human colon cancer genetic alterations (inactivation of *Apc*, *Trp53* and *Tgfbr2*, and activating mutation of *Kras*), cultured intestinal organoids derived from these mice and injected the organoids into different tissues of immunocompetent mice to study metastasis in each site (262). We did not further explore the role of SMYD5 in distant metastasis, but the appearance of these new murine models raises the possibility of inhibiting *Smyd5* expression in these organoids using the CRISPR/Cas9 methodology and then studying the subsequent alteration of the metastatic phenotype.

Our study within the context of previous publications: discussing apparently contradictory results

Although our data clearly points to a role of SMYD5 as an oncogene, a few previously mentioned publications could seem to point in another direction. For example, the study by Kidder *et al.* provides *in vitro* and *in vivo* evidence for a role of SMYD5 as a tumor suppressor (217) (see Introduction: “The SMYD5 protein” section). This idea contradicts our results in HCC patients samples, in which *SMYD5* is overexpressed compared to normal livers, and this overexpression is associated with bad prognosis (**Figure 29**). It also contradicts our results in DEN-treated *Smyd5* KO mice, in which there is a delay in the development of liver cancer compared to WT mice (**Figure 35**). However, their findings are not related to the liver. In contrast, our data were obtained in liver cancer experimental setups and refer to liver cell types (hepatocytes and biliary epithelial cells).

Moreover, the experiments performed by Gao *et al.* in a mouse model of breast cancer that produces lung metastases indicate that *Smyd5* mediates tumor dormancy (219) (see Introduction: “The SMYD5 protein” section). Nevertheless, their results and our results can be brought together by taking into account that carcinogenesis and metastasis are distinct processes, with each being comprised by differentiated steps and involving different molecular mechanisms.

SMYD5 as a hypothetical target for cancer therapy

Taken together, our data indicate that Smyd5 acts as an oncogen in HCC but is dispensable for murine organism homeostasis. Further studies are needed; first, to increase sample size and verify the reproducibility of our experimental results and, second, to evaluate their clinical translation, i.e. whether these results can be extrapolated to human patients. If future studies confirm the clinical relevance of our discoveries, SMYD5 could be a promising target for cancer therapy.

On one hand, SMYD2 and SMYD3 inhibitors have been designed and tested *in vitro* and *in vivo*, although none has been entered clinical trials so far (263). Indeed, despite the amount of literature involving SMYD2 and SMYD3 in cancer progression, there is some controversy about their role in autonomous cancer cell proliferation (264). This raises the importance of a more thorough study to clarify the function of SMYD proteins and define whether SMYD2/3/5 inhibitors are effective options for human cancer treatment.

On the other hand, other epigenetic drugs, including inhibitors of DNA methylation and histone acetylation as well as microRNA(miRNA)-based therapies, have already been tested in liver cancer-related pre-clinical and clinical trials (265), although more research is needed. Drugs targeting epigenetic alterations are a potential new therapeutic tool for HCC, alone or in combination with other therapies (266,267).

Conclusions and future directions

In summary, we discovered SMYD5 methyltransferase activity in histone substrates *in vitro* and its *in vivo* role in carcinogenesis in mice. Moreover, our *Smyd5* KO mouse model liver cancer phenotype is in accordance with data obtained in patients. The mechanism by which SMYD5-mediated H2B and/or H2A methylation regulates cancer development is currently being investigated.

Indeed, other member of the Ioannis Talianidis research group have been breeding all these SMYD5 KO strains with the aim of performing further analyses of the delayed cancer phenotype, such as chromatin immunoprecipitation (ChIP) sequencing (ChIP-seq) or xenografts assays with SMYD5 KO cells.

In any case, SMYD5 methyltransferase activity seems to be relevant in cancer and is worthy of further exploration.

Dissemination of this research

This work was presented in the following scientific meetings:

- Marie Curie NR-NET First Annual Meeting that took place in Debrecen, Hungary, from November 16 to 19, 2014. The title of the oral presentation was “Characterization of Smyd5 (SET and MYND domain-containing protein 5) function in the liver and intestine”.

- Marie Curie NR-NET Mid-Trem Review meeting that took place in Lausanne, Switzerland, from June 26 to 29, 2015. The title of the oral presentation was “Analysis of transcription factor network of lipid metabolism regulators”.

- Spetses Summer School on Nuclear Receptor Signaling in Physiology and Disease (FEBS Advanced Lecture Course) that took place in the island of Spetses, Greece, from August 23 to 28, 2015. The title of the poster was “Characterization of Smyd5 (SET and MYND domain-containing protein 5) function in the liver and intestine”.

- Marie Curie NR-NET Annual Meeting that took place in Stockholm, Sweden, from September 26 to 27, 2016. The title of the oral presentation was “Characterization of Smyd5 (SET and MYND domain-containing protein 5) role in the liver”.

- Spetses Summer School on Nuclear Receptors and Epigenomic Mechanisms in Human Disease and Ageing (FEBS Advanced Lecture Course) that took place in the island of Spetses, Greece, from August 27 to September 1, 2017. The title of the poster was “Smyd5 is required for the development of hepatocellular carcinoma”.

- Marie Curie NR-NET Final Annual Meeting that took place in the island of Spetses, Greece, from September 1 to 3, 2017. The title of the oral presentation was “Smyd5 is required for the development of hepatocellular carcinoma”.

The peer-reviewed publication related to this work is “Endothelial dysfunction markers predict short-term mortality in patients with severe alcoholic hepatitis. Delia Blaya, Teresa Rubio-Tomás, Daniel Rodrigo-Torres, JuanJosé Lozano, Mar Coll, Josepmaria Argemi *et al.* Hepatology International, 2021.”
This article was accepted in Hepatology International last 20th February 2021. Here I reproduce the letter of acceptance:

On behalf of the Editorial Board of Hepatology International, we are delighted to share with you that your manuscript has been accepted for publication. It is scientific works like this, which add value to the existing literature.

We are grateful to you for choosing our journal for this publication. We do hope that you would continue to enrich the scientific community through submission of your novel work to Hepatology International
in the future also.
With kind regards,

Shiv Kumar Sarin, M.D.
Editor-in-Chief
Hepatology International

The DOI assigned to this article is 10.1007/s12072-021-10165-y

The pre-printed proofs are already available (see attached document).

References

1. Vander AJ, Sherman J, Luciano DS. Human physiology: The mechanism of body function. 8th ed. Widmaier EP, Raff H, Strang H, editors. Ballots and Bibles. Boston: McGraw-Hill; 2001.
2. Rogers AB, Dintzis RZ. Liver and Gallbladder. In: Comparative Anatomy and Histology. Elsevier Inc.; 2012. p. 193–201.
3. Molina DK, Dimaio VJM. Normal organ weights in men: Part II-the brain, lungs, liver, spleen, and kidneys. Am J Forensic Med Pathol. 2012 Dec;33(4):368–72.
4. Kruepunga N, Hakvoort TBM, Hikspoors JPJM, Köhler SE, Lamers WH. Anatomy of rodent and human livers: What are the differences? Vol. 1865, Biochimica et Biophysica Acta - Molecular Basis of Disease. Elsevier B.V.; 2019. p. 869–78.
5. Gordillo M, Evans T, Gouon-Evans V. Orchestrating liver development. Vol. 142, Development (Cambridge). Company of Biologists Ltd; 2015. p. 2094–108.
6. Terai M, Mastrangleo MJ, Sato T. Immunological aspect of the liver and metastatic uveal melanoma. J Cancer Metastasis Treat. 2017 Oct 31;3(10):231.
7. Bogdanos DP, Gao B, Gershwin ME. Liver immunology. Comperative Physiol. 2013;3(2):567–98.
8. Racanelli V, Rehermann B. The liver as an immunological organ. Hepatology. 2006;43(2 SUPPL. 1).
9. Häussinger D, Kordes C. Space of Disse: A stem cell niche in the liver. Biol Chem [Internet]. 2019 Dec 18 [cited 2020 Oct 29];401(1):81–95. Available from: https://doi.org/10.1515/hsz-2019-0283
10. M Zorn A. Liver Development. In: The Stem Cell Research Comunity, editor. StemBook. Cambridge (MA): Harvard Stem Cell Institute; 2008.
11. Øie CI, Mönkemöller V, Hübner W, Schüttpelz M, Mao H, Ahluwalia BS, et al. New ways of looking at very small holes - using optical nanoscopy to visualize liver sinusoidal endothelial cell fenestrations. Nanophotonics. 2018;7(3):575–96.
12. Birnbaum E. Surgical anatomy of the colon, rectum, and anus. In: Ratto C, Parello A, Donisi L, Litta F, editors. Colon, rectum and anus: anatomic, physiologic and diagnostic bases for disease management. Springer. 2017.
13. Sorenson RL, Brelje Clark T. Atlas of Human Histology. A Guide to Microscopic Structure of Cells, Tissues and Organs [Internet]. third edit. 2014. 367 p. Available from: http://www.histologyguide.com/about-us/atlas-of-human-histology.html
14. American Cancer Society. Colorectal Cancer Stages [Internet]. American Cancer Society. 2020 [cited 2020 Jan 13]. p. 1–6. Available from: https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/staged.html
15. Clevers H, Batlle E. SnapShot: the intestinal crypt. Cell [Internet]. 2013 [cited 2019 Apr 27];152(5):1198-1198.e2. Available from: http://dx.doi.org/10.1016/j.cell.2013.02.030
16. Hong SN, Dunn JCY, Stelzner M, Martín MG. Concise review: the potential use of intestinal stem cells to treat patients with intestinal failure. Stem Cells Transl Med. 2017;6(2):666–76.
17. The National Institutes of Health. National Cancer Institute. NCI Dictionaries [Internet]. NCI Dictionary of Cancer Terms. 2019 [cited 2020 Jan 15]. Available from: https://www.cancer.gov/publications/dictionaries
18. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell [Internet]. 2011 [cited 2019 Jun 25];144(5):646–74. Available from: https://www.cell.com/action/showPdf?pii=S0092-8674%2811%2900127-9
19. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, et al. Toward understanding and exploiting tumor heterogeneity. Nat Med [Internet]. 2015 [cited 2019 Jul 7];21(8):846–53. Available from:

http://www.nature.com/reprints/index.html.
20. Stratton MR. Exploring the genomes of cancer cells: progress and promise. Science (80-) [Internet]. 2011 Mar 25 [cited 2019 Jul 7];331(6024):1553–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21436442
21. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57–70.
22. Feitelson MA, Arzumanyan A, Kulathinal RJ, Blain SW, Holcombe RF, Mahajna J, et al. Sustained proliferation in cancer: mechanisms and novel therapeutic targets. Semin Cancer Biol. 2015 Dec 1;35:S25–54.
23. Huang Y-H, Hu J, Chen F, Lecomte N, Basnet H, David CJ, et al. ID1 mediates escape from TGF-β tumor suppression in pancreatic cancer. Cancer Discov [Internet]. 2020 [cited 2020 Jan 16];10(1):142–57. Available from: http://cancerdiscovery.aacrjournals.org/
24. Yang Y, Klionsky DJ. Autophagy and disease: unanswered questions. Cell Death Differ [Internet]. 2020; Available from: http://dx.doi.org/10.1038/s41418-019-0480-9
25. Levy JMM, Towers CG, Thorburn A. Targeting autophagy in cancer. Nat Rev Cancer [Internet]. 2017;17(9):528–42. Available from: http://dx.doi.org/10.1038/nrc.2017.53
26. Tang D, Kang R, Berghe T Vanden, Vandenabeele P, Kroemer G. The molecular machinery of regulated cell death [Internet]. Vol. 29, Cell Research. 2019 [cited 2020 Jan 15]. p. 347–64. Available from: https://doi.org/10.1038/s41422-019-0164-5
27. Najafov A, Chen H, Yuan J. Necroptosis and Cancer. Trends in Cancer. 2017;3(4):294–301.
28. Jiao D, Cai Z, Choksi S, Ma D, Choe M, Kwon HJ, et al. Necroptosis of tumor cells leads to tumor necrosis and promotes tumor metastasis. Cell Res. 2018 Aug 1;28(8):868–70.
29. Xia X, Wang X, Cheng Z, Qin W, Lei L, Jiang J, et al. The role of pyroptosis in cancer: pro-cancer or pro-“host”? Cell Death Dis [Internet]. 2019 [cited 2020 Jan 17];10(9). Available from: https://doi.org/10.1038/s41419-019-1883-8
30. Kang R, Kroemer G, Tang D. The tumor suppressor protein p53 and the ferroptosis network. Free Radic Biol Med. 2019 Mar 1;133:162–8.
31. Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, et al. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. J Hematol Oncol [Internet]. 2019 [cited 2020 Jan 17];12(1). Available from: https://doi.org/10.1186/s13045-019-0720-y
32. Faget D V., Ren Q, Stewart SA. Unmasking senescence: context-dependent effects of SASP in cancer. Nat Rev Cancer [Internet]. 2019;19(8):439–53. Available from: http://dx.doi.org/10.1038/s41568-019-0156-2
33. Ségal-Bendirdjian E, Geli V. Non-canonical roles of telomerase: unraveling the imbroglio. Vol. 7, Frontiers in Cell and Developmental Biology. Frontiers Media S.A.; 2019.
34. De Palma M, Biziato D, Petrova T V. Microenvironmental regulation of tumour angiogenesis. Nat Rev Cancer [Internet]. 2017 [cited 2020 Jan 17];17(8):457–74. Available from: www.nature.com/nrc
35. Peinado H, Zhang H, Matei IR, Costa-Silva B, Hoshino A, Rodrigues G, et al. Pre-metastatic niches: Organ-specific homes for metastases. Nat Rev Cancer [Internet]. 2017 [cited 2020 Jan 17];17(5):302–17. Available from: www.nature.com/nrc
36. Dongre A, Weinberg RA. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. Nat Rev Mol Cell Biol [Internet]. 2019;20(2):69–84. Available from: http://dx.doi.org/10.1038/s41580-018-0080-4
37. Friedl P, Locker J, Sahai E, Segall JE. Classifying collective cancer cell invasion. Nat Cell Biol. 2012;14(8):777–83.

38. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability an evolving hallmark of cancer [Internet]. Vol. 11, Nature Reviews Molecular Cell Biology. 2010 [cited 2020 Jan 16]. p. 220–8. Available from: www.nature.com/reviews/molcellbio
39. Duijf PHG, Nanayakkara D, Nones K, Srihari S, Kalimutho M, Khanna KK. Mechanisms of genomic instability in breast cancer. Trends Mol Med [Internet]. 2019 [cited 2020 Jan 16];25(7):595–611. Available from: https://doi.org/10.1016/j.molmed.2019.04.004
40. Zhang Q, Zhu B, Li Y. Resolution of cancer-promoting inflammation: A new approach for anticancer therapy. Front Immunol. 2017 Feb 1;8.
41. Hua Y, Bergers G. Tumors vs. chronic wounds: an immune cell's perspective. Front Immunol. 2019 Sep 12;10.
42. Pribluda A, Elyada E, Wiener Z, Hamza H, Goldstein RE, Biton M, et al. A Senescence-inflammatory switch from cancer-inhibitory to cancer-promoting mechanism. Cancer Cell [Internet]. 2013 [cited 2020 Jan 16];24(2):242–56. Available from: http://dx.doi.org/10.1016/j.ccr.2013.06.005
43. Lecot P, Alimirah F, Desprez PY, Campisi J, Wiley C. Context-dependent effects of cellular senescence in cancer development. Br J Cancer. 2016;114(11):1180–4.
44. Coppé JP, Desprez PY, Krtolica A, Campisi J. The senescence-associated secretory phenotype: The dark side of tumor suppression [Internet]. Vol. 5, Annual Review of Pathology: Mechanisms of Disease. Annual Reviews ; 2010 [cited 2020 Jul 31]. p. 99–118. Available from: www.annualreviews.org
45. Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. Science (80-). 2018 Mar 23;359(6382):1350–5.
46. O'Donnell JS, Teng MWL, Smyth MJ. Cancer immunoediting and resistance to T cell-based immunotherapy. Nat Rev Clin Oncol [Internet]. 2019 [cited 2020 Jan 17];16(3):151–67. Available from: www.nature.com/nrclinonc
47. Courtnay R, Ngo DC, Malik N, Ververis K, Tortorella SM, Karagiannis TC. Cancer metabolism and the Warburg effect: the role of HIF-1 and PI3K. Mol Biol Rep. 2015;42(4):841–51.
48. Liberti M V, Locasale JW. The Warburg effect: how does it benefit cancer cells? Trends Biochem Sci. 2016;41(3):211–8.
49. Vaupel P, Schmidberger H, Mayer A. The Warburg effect: essential part of metabolic reprogramming and central contributor to cancer progression. Int J Radiat Biol [Internet]. 2019 [cited 2020 Jan 17];95(7):912–9. Available from: https://www.tandfonline.com/action/journalInformation?journalCode=irab20
50. Balkwill FR, Capasso M, Hagemann T. The tumor microenvironment at a glance. J Cell Sci. 2012;125(23):5591–6.
51. Maman S, Witz IP. A history of exploring cancer in context. Nat Rev Cancer [Internet]. 2018 [cited 2020 Jan 17];18(6):359–76. Available from: www.nature.com/nrc
52. Sun HR, Wang S, Yan SC, Zhang Y, Nelson PJ, Jia HL, et al. Therapeutic strategies targeting cancer stem cells and their microenvironment. Front Oncol. 2019;9(OCT).
53. Junttila MR, De Sauvage FJ. Influence of tumour microenvironment heterogeneity on therapeutic response. Nature. 2013;501(7467):346–54.
54. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin [Internet]. 2018 Nov [cited 2019 Jun 25];68(6):394–424. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30207593
55. Cancer R. Global cancer observatory [Internet]. Vol. 593, Malaysia Cancer Statistics. 2019 [cited 2019 Jun 29]. p. 1–2. Available from: http://gco.iarc.fr/databases.php

56. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet [Internet]. 2018 Mar 31 [cited 2019 Jun 29];391(10127):1301–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29307467
57. Bertuccio P, Turati F, Carioli G, Rodriguez T, La Vecchia C, Malvezzi M, et al. Global trends and predictions in hepatocellular carcinoma mortality. J Hepatol [Internet]. 2017 Aug 1 [cited 2019 Jun 29];67(2):302–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28336466
58. Degasperi E, Colombo M. Distinctive features of hepatocellular carcinoma in non-alcoholic fatty liver disease. Lancet Gastroenterol Hepatol [Internet]. 2016 [cited 2019 Jun 29];1:156–64. Available from: www.thelancet.com/gastrohep
59. Nault J-C, Datta S, Imbeaud S, Franconi A, Mallet M, Couchy G, et al. Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. Nat Genet [Internet]. 2015 Oct 24 [cited 2019 Jul 7];47(10):1187–93. Available from: http://www.nature.com/articles/ng.3389
60. Villanueva A. Hepatocellular Carcinoma. N Engl J Med [Internet]. 2019;380(15):1450–62. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30970190
61. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer. 2006;6(9):674–87.
62. Schlageter M, Terracciano LM, Angelo SD', Sorrentino P. Histopathology of hepatocellular carcinoma. World J Gastroenterol [Internet]. 2014 [cited 2019 Jun 29];20(43):15955–64. Available from: http://www.wjgnet.com/esps/HelpDesk:http://www.wjgnet.com/esps/helpdesk.aspx:15955-15964Availablefrom:URL:http://www.wjgnet.com/1007-9327/full/v20/i43/15955.htmDOI:http://dx.doi.org/10.3748/wjg.v20.i43.15955
63. Schulze K, Nault JC, Villanueva A. Genetic profiling of hepatocellular carcinoma using next-generation sequencing. J Hepatol [Internet]. 2016;65(5):1031–42. Available from: http://dx.doi.org/10.1016/j.jhep.2016.05.035
64. Shafizadeh N, Kakar S. Diagnosis of well-differentiated hepatocellular lesions: role of immunohistochemistry and other ancillary techniques. Adv Anat Pathol [Internet]. 2011 Nov [cited 2019 Jun 30];18(6):438–45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21993269
65. McQueen CA. Comprehensive toxicology. Elsevier Science; 2010.
66. Marquardt JU, Thorgeirsson SS. SnapShot: Hepatocellular Carcinoma. Cancer Cell [Internet]. 2014 [cited 2019 Apr 27];25:550-550.e1. Available from: http://dx.doi.org/10.1016/j.ccr.2014.04.002
67. Pinyol R, Montal R, Takayama T, Chau G-Y, Mazzaferro V, Roayaie S, et al. Molecular predictors of recurrence prevention with sorafenib as adjuvant therapy in hepatocellular carcinoma: Biomarker study of the STORM phase III trial. J Hepatol [Internet]. 2017 Jan 1 [cited 2019 Jun 29];66(1):S12–3. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168827817302878
68. Llovet J, Brú C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. Semin Liver Dis [Internet]. 1999 Mar 17 [cited 2019 Jun 30];19(03):329–38. Available from: http://www.thieme-connect.de/DOI/DOI?10.1055/s-2007-1007122
69. Nault JC, Calderaro J, Di Tommaso L, Balabaud C, Zafrani ES, Bioulac-Sage P, 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 [Internet]. 2014 Dec 27 [cited 2019 Jul 7];60(6):1983–92. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/hep.27372
70. Fujimoto A, Totoki Y, Abe T, Boroevich KA, Hosoda F, Nguyen HH, et al. Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. Nat Genet [Internet]. 2012 Jul 27 [cited 2019 Jul 17];44(7):760–4. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/22634756
71. Guichard C, Amaddeo G, Imbeaud S, Ladeiro Y, Pelletier L, Maad I Ben, et al. Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. Nat Genet [Internet]. 2012 Jun 6 [cited 2019 Jul 17];44(6):694–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22561517
72. Huang J, Deng Q, Wang Q, Li K-Y, Dai J-H, Li N, et al. Exome sequencing of hepatitis B virus–associated hepatocellular carcinoma. Nat Genet [Internet]. 2012 Oct 26 [cited 2019 Jul 17];44(10):1117–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22922871
73. Kan Z, Xu J, Hauptschein R, Rejto PA, Fernandez J, Wang K, et al. Whole-genome sequencing identifies recurrent mutations in hepatocellular carcinoma. Genome Res. 2013;23(9):1422–33.
74. Totoki Y, Tatsuno K, Covington KR, Ueda H, Creighton CJ, Kato M, et al. Trans-ancestry mutational landscape of hepatocellular carcinoma genomes. Nat Genet [Internet]. 2014 Dec 2 [cited 2019 Jul 7];46(12):1267–73. Available from: http://www.nature.com/articles/ng.3126
75. Ahn SM, Jang SJ, Shim JH, Kim D, Hong SM, Sung CO, et al. Genomic portrait of resectable hepatocellular carcinomas: Implications of RB1 and FGF19 aberrations for patient stratification. Hepatology. 2014;60(6):1972–82.
76. Schulze K, Imbeaud S, Letouzé E, Alexandrov LB, Calderaro J, Rebouissou S, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. Nat Genet [Internet]. 2015 May 30 [cited 2019 Jul 7];47(5):505–11. Available from: http://www.nature.com/articles/ng.3252
77. Fujimoto A, Furuta M, Totoki Y, Tsunoda T, Kato M, Shiraishi Y, et al. Whole-genome mutational landscape and characterization of noncoding and structural mutations in liver cancer. Nat Genet [Internet]. 2016 May 11 [cited 2019 Jul 17];48(5):500–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27064257
78. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin A V., et al. Signatures of mutational processes in human cancer. Nature [Internet]. 2013 Aug 14 [cited 2019 Jul 7];500(7463):415–21. Available from: http://www.nature.com/articles/nature12477
79. Humphries C. Sex differences : Luck of the chromosomes. Nature. 2014 Dec 4;516(729):S10–1.
80. Manieri E, Herrera-Melle L, Mora A, Tomás-Loba A, Leiva-Vega L, Fernández DI, et al. Adiponectin accounts for gender differences in hepatocellular carcinoma incidence. J Exp Med. 2019 May 1;216(5):1108–19.
81. Llovet JM, Ricci S, Mazzaferro V, Hilgard P, Gane E, Blanc J-F, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med [Internet]. 2008 Jul 24 [cited 2019 Jul 7];359(4):378–90. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa0708857
82. Cheng A-L, Kang Y-K, Chen Z, Tsao C-J, Qin S, Kim JS, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomised, double-blind, placebo-controlled trial. Lancet Oncol [Internet]. 2009 Jan 1 [cited 2019 Jul 7];10(1):25–34. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19095497
83. Kudo M, Finn RS, Qin S, Han K-H, Ikeda K, Piscaglia F, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet [Internet]. 2018 Mar 24 [cited 2019 Jul 7];391(10126):1163–73. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29433850
84. Bruix J, Qin S, Merle P, Granito A, Huang Y-H, Bodoky G, et al. Regorafenib for patients with hepatocellular carcinoma who progressed on sorafenib treatment (RESORCE): a randomised, double-blind, placebo-controlled, phase 3 trial.

Lancet [Internet]. 2017 Jan 7 [cited 2019 Jul 7];389(10064):56–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27932229
85. Abou-Alfa GK, Meyer T, Cheng A-L, El-Khoueiry AB, Rimassa L, Ryoo B-Y, et al. Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. N Engl J Med [Internet]. 2018 Jul 5 [cited 2019 Jul 7];379(1):54–63. Available from: http://www.nejm.org/doi/10.1056/NEJMoa1717002
86. Zhu AX, Kang Y-K, Yen C-J, Finn RS, Galle PR, Llovet JM, et al. Ramucirumab after sorafenib in patients with advanced hepatocellular carcinoma and increased α-fetoprotein concentrations (REACH-2): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol [Internet]. 2019 Feb 1 [cited 2019 Jul 7];20(2):282–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30665869
87. Dhanasekaran R, Bandoh S, Roberts LR. Molecular pathogenesis of hepatocellular carcinoma and impact of therapeutic advances. F1000Research [Internet]. 2016 [cited 2019 Jul 20];12(5). Available from: https://doi.org/10.12688/f1000research.6946.1
88. Wai-Hung Ho D, Cheuk-Lam Lo R, Chan L-K, Oi-Lin Ng I, L Ng IO. Molecular pathogenesis of hepatocellular carcinoma. Liver Cancer [Internet]. 2016 [cited 2019 Jul 20];5(4):290–302. Available from: www.karger.com/lic
89. Oishi N, Wei Wang X. Novel therapeutic strategies for targeting liver cancer stem cells. Int J Biol Sci [Internet]. 2011 [cited 2019 Apr 27];7(5):517–35. Available from: http://www.biolsci.org517
90. Wang K, Sun D. Cancer stem cells of hepatocellular carcinoma. Oncotarget. 2018 May 1;9(33):23306–14.
91. Nikolaou KC, Moulos P, Chalepakis G, Hatzis P, Oda H, Reinberg D, et al. Spontaneous development of hepatocellular carcinoma with cancer stem cell properties in PR - SET 7-deficient livers. EMBO J. 2015;34(4):430–47.
92. He L, Tian D-A, Li P-Y, He X-X. Mouse models of liver cancer: Progress and recommendations. Oncotarget [Internet]. 2015 [cited 2019 Jun 30];6(27). Available from: www.impactjournals.com/oncotarget
93. Bakiri L, Wagner EF. Mouse models for liver cancer. Mol Oncol [Internet]. 2013 [cited 2019 Jun 30];206–23. Available from: http://dx.doi.org/10.1016/j.molonc.2013.01.005
94. Verna L, Whysner J, Williams GM. N-Nitrosodiethylamine mechanistic data and risk assessment: Bioactivation, DNA-adduct formation, mutagenicity, and tumor initiation. Vol. 71, Pharmacology and Therapeutics. Elsevier Inc.; 1996. p. 57–81.
95. Tolba R, Kraus T, Liedtke C, Schwarz M, Weiskirchen R. Diethylnitrosamine (DEN)-induced carcinogenic liver injury in mice. Lab Anim. 2015;49:59–69.
96. Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. Gut. 2017;66(4):683–91.
97. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet [Internet]. 2019;394(10207):1467–80. Available from: http://dx.doi.org/10.1016/S0140-6736(19)32319-0
98. Zeuner A, Todaro M, Stassi G, De Maria R. Colorectal cancer stem cells: From the crypt to the clinic. Cell Stem Cell [Internet]. 2014 [cited 2020 Jan 19];15(6):692–705. Available from: http://dx.doi.org/10.1016/j.stem.2014.11.012
99. Markowitz SD, Bertagnolli MM. Molecular basis of colorectal cancer. N Engl J Med. 2009;361(25):2449–60.
100. Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. Nat Rev Dis Prim [Internet]. 2015;1:1–25. Available from: http://dx.doi.org/10.1038/nrdp.2015.65
101. East JE, Atkin WS, Bateman AC, Clark SK, Dolwani S, Ket SN, et al. British Society of Gastroenterology position statement on serrated polyps in the colon

and rectum. Gut [Internet]. 2017 [cited 2020 Jan 19];66(7):1181–96. Available from: http://gut.bmj.com/
102. Curtin K, Slattery ML, Samowitz WS. CpG island methylation in colorectal cancer: past, present and future. Patholog Res Int. 2011;2011.
103. Rosenberg DW, Giardina C, Tanaka T. Mouse models for the study of colon carcinogenesis. Carcinogenesis. 2009;30(2):183–96.
104. Tanaka T. Development of an inflammation-associated colorectal cancer model and its application for research on carcinogenesis and chemoprevention. Int J Inflam. 2012;2012.
105. Chassaing B, Aitken JD, Malleshappa M, Vijay-Kumar M. Dextran sulfate sodium (DSS)-induced colitis in mice. Curr Protoc Immunol [Internet]. 2014 [cited 2019 Apr 27];104. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3980572/pdf/nihms-566982.pdf
106. Wang JG, Wang DF, Lv BJ, Si JM. A novel mouse model for colitis-associated colon carcinogenesis induced by 1,2-dimethylhydrazine and dextran sulfate sodium. World J Gastroenterol. 2004;10(20):2958–62.
107. Sweatt JD, Nestler EJ, Meaney MJ, Akbarian S. An overview of the molecular basis of epigenetics. In: Epigenetic regulation in the nervous system [Internet]. First Edit. Elsevier Inc.; 2013. p. 3–33. Available from: http://dx.doi.org/10.1016/B978-0-12-391494-1.00001-X
108. Llinàs-Arias P, Esteller M. Epigenetic inactivation of tumour suppressor coding and non-coding genes in human cancer: an update. Open Biol [Internet]. 2017 [cited 2020 Jan 22];7(9). Available from: http://dx.doi.org/10.1098/rsob.170152
109. Nebbioso A, Tambaro FP, Dell'Aversana C, Altucci L. Cancer epigenetics: Moving forward. PLoS Genet. 2018;14(6):1–25.
110. Berdasco M, Esteller M. Clinical epigenetics: seizing opportunities for translation. Nat Rev Genet [Internet]. 2019 [cited 2020 Jan 24];20(2):109–27. Available from: www.nature.com/nrg
111. Chen Z, Li S, Subramaniam S, Shyy JY-J, Chien S. Epigenetic regulation: a new frontier for biomedical engineers. Annu Rev Biomed Eng. 2017;19(1):195–219.
112. Harshman SW, Young NL, Parthun MR, Freitas MA. H1 histones: current perspectives and challenges. Nucleic Acids Res. 2013;41(21):9593–609.
113. McGinty RK, Tan S. Nucleosome structure and function. Chem Rev. 2015;115(6):2255–73.
114. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693–705.
115. Dawson MA, Kouzarides T, Huntly BJP. Targeting epigenetic readers in cancer. N Engl J Med. 2012;367(7):647–57.
116. Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell. 2011;146(6):1016–28.
117. Wan J, Liu H, Chu J, Zhang H. Functions and mechanisms of lysine crotonylation. J Cell Mol Med. 2019;23(11):7163–9.
118. Weinert BT, Schölz C, Wagner SA, Iesmantavicius V, Su D, Daniel JA, et al. Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. Cell Rep [Internet]. 2013 [cited 2020 Jan 23];4(4):842–51. Available from: http://dx.
119. Xie Z, Dai J, Dai L, Tan M, Cheng Z, Wu Y, et al. Lysine succinylation and lysine malonylation in histones. Mol Cell Proteomics [Internet]. 2012 [cited 2020 Jan 23];11(5):100–7. Available from: http://www.mcponline.org
120. Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, et al. Lysine propionylation and butyrylation are novel post-translational modifications in histones. Mol Cell Proteomics [Internet]. 2007 [cited 2020 Jan 23];6(5):812–9. Available from: http://www.mcponline.org
121. Jones PA, Issa JPJ, Baylin S. Targeting the cancer epigenome for therapy. Nat

Rev Genet [Internet]. 2016;17(10):630–41. Available from: http://dx.doi.org/10.1038/nrg.2016.93
122. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21(3):381–95.
123. Keller C, Kulasegaran-Shylini R, Shimada Y, Hotz HR, Bühler M. Noncoding RNAs prevent spreading of a repressive histone mark. Nat Struct Mol Biol. 2013;20(8):994–1000.
124. Pennacchio LA, Bickmore W, Dean A, Nobrega MA, Bejerano G. Enhancers: Five essential questions. Nat Rev Genet [Internet]. 2013 [cited 2020 Jan 23];14(4):288–95. Available from: http://genes.uchicago.edu/contents/faculty/nobrega-marcelo.html
125. Liu F. Enhancer-derived RNA: A Primer. Genomics, Proteomics Bioinforma [Internet]. 2017 Jun 1 [cited 2020 Nov 1];15(3):196–200. Available from: http://dx.doi.org/10.1016/j.gpb.2016.12.006
126. Greer EL, Shi Y. Histone methylation: A dynamic mark in health, disease and inheritance. Nat Rev Genet. 2012;13(5):343–57.
127. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007;129(4):823–37.
128. Hon G, Wang W, Ren B. Discovery and annotation of functional chromatin signatures in the human genome. PLoS Comput Biol. 2009;5(11).
129. Husmann D, Gozani O. Histone lysine methyltransferases in biology and disease. Nat Struct Mol Biol [Internet]. 2019;26(10):880–9. Available from: http://dx.doi.org/10.1038/s41594-019-0298-7
130. Gayatri S, Bedford MT. Readers of histone methylarginine marks. Biochim Biophys Acta - Gene Regul Mech [Internet]. 2014 [cited 2020 Jan 24];1839(8):702–10. Available from: http://dx.doi.org/10.1016/j.bbagrm.2014.02.015
131. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. Exp Mol Med. 2017;49(4).
132. Tropberger P, Schneider R. Going global: Novel histone modifications in the globular domain of H3. Epigenetics [Internet]. 2010 [cited 2020 Jan 23];5(2):112–7. Available from: https://doi.org/10.4161/epi.5.2.11075
133. Zhang J, Jing L, Li M, He L, Guo Z. Regulation of histone arginine methylation/demethylation by methylase and demethylase. Mol Med Rep. 2019;19(5):3963–71.
134. Kooistra SM, Helin K. Post-translational modifications: Molecular mechanisms and potential functions of histone demethylases. Nat Rev Mol Cell Biol [Internet]. 2012 [cited 2020 Jan 23];13(5):297–311. Available from: www.nature.com/reviews/molcellbio
135. Blanc RS, Richard S. Arginine methylation: The coming of age. Mol Cell [Internet]. 2017 [cited 2020 Jan 24];65(1):8–24. Available from: http://dx.
136. Michalak EM, Burr ML, Bannister AJ, Dawson MA. The roles of DNA, RNA and histone methylation in ageing and cancer. Nat Rev Mol Cell Biol [Internet]. 2019;20(10):573–89. Available from: http://dx.doi.org/10.1038/s41580-019-0143-1
137. Sarris M, Nikolaou K, Talianidis I. Context-specific regulation of cancer epigenomes by histone and transcription factor methylation. Oncogene [Internet]. 2014 [cited 2020 Jan 21];33(10):1207–17. Available from: www.nature.com/onc
138. Rao RC, Dou Y. Hijacked in cancer: The KMT2 (MLL) family of methyltransferases. Nat Rev Cancer [Internet]. 2015 May 22 [cited 2020 Dec 26];15(6):334–46. Available from: /pmc/articles/PMC4493861/?report=abstract
139. Gan L, Yang Y, Li Q, Feng Y, Liu T, Guo W. Epigenetic regulation of cancer progression by EZH2: From biological insights to therapeutic potential [Internet].

Vol. 6, Biomarker Research. BioMed Central Ltd.; 2018 [cited 2020 Dec 26]. Available from: /pmc/articles/PMC5845366/?report=abstract
140. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol. 2012;13(2):115–26.
141. Fahey CC, Davis IJ. SETting the stage for cancer development: SETD2 and theconsequences of lost methylation. Cold Spring Harb Perspect Med [Internet]. 2017 May 1 [cited 2020 Dec 26];7(5). Available from: /pmc/articles/PMC5411680/?report=abstract
142. Li B, Zheng Y, Yang L. The oncogenic potential of SuV39H2: A comprehensive and perspective view. J Cancer [Internet]. 2019 [cited 2020 Dec 26];10(3):721–9. Available from: /pmc/articles/PMC6360419/?report=abstract
143. Shinkai Y, Tachibana M. H3K9 methyltransferase G9a and the related molecule GLP. Genes Dev [Internet]. 2011 [cited 2020 Dec 26];25(8):781–8. Available from: http://www.genesdev.org/cgi/doi/10.1101/gad.2027411.
144. Bárcena-Varela M, Caruso S, Llerena S, Álvarez-Sola G, Uriarte I, Latasa MU, et al. Dual Targeting of Histone Methyltransferase G9a and DNA-Methyltransferase 1 for the Treatment of Experimental Hepatocellular Carcinoma. Hepatology [Internet]. 2019 Feb 4 [cited 2020 Dec 26];69(2):587–603. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/hep.30168
145. Karanth AV, Maniswami RR, Prashanth S, Govindaraj H, Padmavathy R, Jegatheesan SK, et al. Emerging role of SETDB1 as a therapeutic target. Expert Opin Ther Targets [Internet]. 2017 Mar 4 [cited 2020 Dec 26];21(3):319–31. Available from: https://www.tandfonline.com/doi/full/10.1080/14728222.2017.1279604
146. Cao N, Yu Y, Zhu H, Chen M, Chen P, Zhuo M, et al. SETDB1 promotes the progression of colorectal cancer via epigenetically silencing p21 expression. Cell Death Dis [Internet]. 2020 May 1 [cited 2020 Dec 26];11(5). Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/32393761/
147. Milite C, Feoli A, Viviano M, Rescigno D, Cianciulli A, Balzano AL, et al. The emerging role of lysine methyltransferase SETD8 in human diseases [Internet]. Vol. 8, Clinical Epigenetics. Springer Verlag; 2016 [cited 2020 Dec 26]. p. 1–15. Available from: https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-016-0268-4
148. Godfrey L, Crump NT, Thorne R, Lau IJ, Repapi E, Dimou D, et al. DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. Nat Commun [Internet]. 2019 Dec 1 [cited 2020 Dec 26];10(1). Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/31243293/
149. Song Z, Wei Z, Wang Q, Zhang X, Tao X, Wu N, et al. The role of DOT1L in the proliferation and prognosis of gastric cancer. Biosci Rep [Internet]. 2020 Jan 1 [cited 2020 Dec 26];40(1). Available from: https://pubmed.ncbi.nlm.nih.gov/31939604/
150. Yang L, Lei Q, Li L, Yang J, Dong Z, Cui H. Silencing or inhibition of H3K79 methyltransferase DOT1L induces cell cycle arrest by epigenetically modulating c-Myc expression in colorectal cancer. Clin Epigenetics [Internet]. 2019 Dec 30 [cited 2020 Dec 26];11(1):1–17. Available from: https://link.springer.com/articles/10.1186/s13148-019-0778-y
151. Boi M, Zucca E, Inghirami G, Bertoni F. PRDM1/BLIMP1: A tumor suppressor gene in B and T cell lymphomas. Leuk Lymphoma [Internet]. 2015 May 1 [cited 2020 Dec 26];56(5):1223–8. Available from: http://www.tandfonline.com/doi/full/10.3109/10428194.2014.953155
152. Zhu Z, Wang H, Wei Y, Meng F, Liu Z, Zhang Z. Downregulation of PRDM1 promotes cellular invasion and lung cancer metastasis. Tumor Biol [Internet]. 2017 Apr 1 [cited 2020 Dec 26];39(4). Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/28378641/

153. Liu C, Banister CE, Weige CC, Altomare D, Richardson JH, Contreras CM, et al. PRDM1 silences stem cell-related genes and inhibits proliferation of human colon tumor organoids. Proc Natl Acad Sci U S A [Internet]. 2018 [cited 2020 Dec 26];115(22):E5066–75. Available from: www.pnas.org/cgi/doi/10.1073/pnas.1802902115
154. Sorrentino A, Rienzo M, Ciccodicola A, Casamassimi A, Abbondanza C. Human PRDM2: Structure, function and pathophysiology. Biochim Biophys Acta - Gene Regul Mech [Internet]. 2018 [cited 2020 Dec 26];1861(7):657–71. Available from: https://doi.org/10.1016/j.bbagrm.2018.06.002
155. Shu X sheng, Geng H, Li L, Ying J, Ma C, Wang Y, et al. The epigenetic modifier PRDM5 functions as a tumor suppressor through modulating WNT/β-catenin signaling and is frequently silenced in multiple tumors. PLoS One [Internet]. 2011 Nov 8 [cited 2020 Dec 26];6(11). Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/22087297/
156. Zhou P, Chen X, Li M, Sun X, Tan J, Wang X, et al. Overexpression of PRDM5 promotes acute myeloid leukemia cell proliferation and migration by activating the JNK pathway. Cancer Med [Internet]. 2019 [cited 2020 Dec 26];8(8):3905–17. Available from: http://www.abren.net/Progn
157. Wang X, Chang H, Gao G, Su B, Deng Q, Zhou H, et al. Silencing of PRDM5 increases cell proliferation and inhibits cell apoptosis in glioma. Int J Neurosci [Internet]. 2020 Mar 5 [cited 2020 Dec 26];1–10. Available from: https://www.tandfonline.com/doi/full/10.1080/00207454.2020.1733563
158. Ordoñez R, Kulis M, Russiñol N, Chapaprieta V, Carrasco-Leon A, García-Torre B, et al. Chromatin activation as a unifying principle underlying pathogenic mechanisms in multiple myeloma. Genome Res [Internet]. 2020 Oct 1 [cited 2020 Dec 26];30(9):1217–27. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/32820006/
159. Reid AG, Nacheva EP. A potential role for PRDM12 in the pathogenesis of chronic myeloid leukaemia with derivative chromosome 9 deletion [9]. Leukemia [Internet]. 2004 [cited 2020 Dec 26];18(1):178–80. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/14523459/
160. Tracey LJ, Justice MJ. Off to a Bad Start: Cancer Initiation by Pluripotency Regulator PRDM14. Trends Genet [Internet]. 2019 [cited 2020 Dec 26];35(7):489–500. Available from: http://lab.research.sickkids.ca/justice/
161. Karakaidos P, Verigos J, Magklara A. Lsd1/kdm1a, a gate-keeper of cancer stemness and a promising therapeutic target. Cancers (Basel) [Internet]. 2019 [cited 2020 Dec 27];11(12). Available from: www.mdpi.com/journal/cancers
162. Majello B, Gorini F, Saccà CD, Amente S. Expanding the role of the histone lysine-specific demethylase LSD1 in cancer. Cancers (Basel) [Internet]. 2019 Mar 1 [cited 2020 Dec 27];11(3). Available from: /pmc/articles/PMC6468368/?report=abstract
163. Rizwani W, Schaal C, Kunigal S, Coppola D, Chellappan S. Mammalian Lysine Histone Demethylase KDM2A Regulates E2F1-Mediated Gene Transcription in Breast Cancer Cells. PLoS One [Internet]. 2014 [cited 2020 Dec 27];9(7):100888. Available from: www.plosone.org
164. Yan M, Yang X, Wang H, Shao Q. The critical role of histone lysine demethylase KDM2B in cancer. Am J Transl Res [Internet]. 2018 [cited 2020 Dec 27];10(8):2222–33. Available from: www.ajtr.org/ISSN:1943-8141/AJTR0070873
165. Yoo J, Jeon YH, Cho HY, Lee SW, Kim GW, Lee DH, et al. Advances in histone demethylase KDM3A as a cancer therapeutic target. Cancers (Basel) [Internet]. 2020 [cited 2020 Dec 27];12(5). Available from: www.mdpi.com/journal/cancers
166. Wilson C, Krieg AJ. KDM4B: A nail for every hammer? Genes (Basel) [Internet]. 2019 [cited 2020 Dec 27];10(2). Available from: www.mdpi.com/journal/genes
167. Zhang C, Wang Z, Ji Q, Li Q. Histone demethylase JMJD2C: Epigenetic regulators in tumors. Oncotarget [Internet]. 2017 [cited 2020 Dec

27];8(53):91723–33. Available from: www.impactjournals.com/oncotarget/
168. Harmeyer KM, Facompre ND, Herlyn M, Basu D. JARID1 Histone Demethylases: Emerging Targets in Cancer. Trends in Cancer. 2017;3(10):713–25.
169. Wang L, Shilatifard A. UTX Mutations in Human Cancer. Cancer Cell [Internet]. 2019 [cited 2020 Dec 27];35(2):168–76. Available from: https://doi.org/10.1016/j.ccell.2019.01.001
170. Zhang X, Liu L, Yuan X, Wei Y, Wei X. JMJD3 in the regulation of human diseases. Protein Cell [Internet]. 2019 Dec 1 [cited 2020 Dec 27];10(12):864–82. Available from: https://doi.org/10.1007/s13238-019-0653-9
171. Kouskouti A, Scheer E, Staub A, Szlò Tora L, Talianidis I. Gene-Specific Modulation of TAF10 Function by SET9-Mediated Methylation methylation. Mol Cell [Internet]. 2004 [cited 2020 Dec 27];14:175–82. Available from: http://www.molecule.org/
172. Scoumanne A, Chen X. The lysine-specific demethylase 1 is required for cell proliferation in both p53-dependent and -independent manners. J Biol Chem [Internet]. 2007 May 25 [cited 2020 Dec 30];282(21):15471–5. Available from: http://www.jbc.org.sire.ub.edu/content/282/21/15471.full
173. Yang X-D, Huang B, Li M, Lamb A, Kelleher NL, Chen L-F. Negative regulation of NF-κB action by Set9-mediated lysine methylation of the RelA subunit. EMBO J [Internet]. 2009 Apr 22 [cited 2020 Dec 30];28(8):1055–66. Available from: http://emboj.embopress.org/cgi/doi/10.1038/emboj.2009.55
174. Yang X-D, Tajkhorshid E, Chen L-F. Functional Interplay between Acetylation and Methylation of the RelA Subunit of NF-κB. Mol Cell Biol [Internet]. 2010 May 1 [cited 2020 Dec 30];30(9):2170–80. Available from: http://mcb.asm.org/
175. Yang J, Huang J, Dasgupta M, Sears N, Miyagi M, Wang B, et al. Reversible methylation of promoter-bound STAT3 by histone-modifying enzymes. Proc Natl Acad Sci U S A [Internet]. 2010 Dec 14 [cited 2020 Dec 30];107(50):21499–504. Available from: https://www-pnas-org.sire.ub.edu/content/107/50/21499
176. Kontaki H, Talianidis I. Cross-talk between post-translational modifications regulate life or death decisions by E2F1 [Internet]. Vol. 9, Cell Cycle. Taylor and Francis Inc.; 2010 [cited 2020 Dec 30]. p. 3836–7. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/20890118/
177. Kontaki H, Talianidis I. Lysine Methylation Regulates E2F1-Induced Cell Death. Mol Cell [Internet]. 2010 Jul [cited 2020 Dec 30];39(1):152–60. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/20603083/
178. Cho HS, Suzuki T, Dohmae N, Hayami S, Unoki M, Yoshimatsu M, et al. Demethylation of RB regulator MYPT1 by histone demethylase LSD1 promotes cell cycle progression in cancer cells. Cancer Res [Internet]. 2011 Feb 1 [cited 2020 Dec 30];71(3):655–60. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/21115810/
179. Wang J, Hevi S, Kurash JK, Lei H, Gay F, Bajko J, et al. The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. Nat Genet [Internet]. 2009 Jan 21 [cited 2020 Dec 30];41(1):125–9. Available from: https://www-nature-com.sire.ub.edu/articles/ng.268
180. Lee JY, Park JH, Choi HJ, Won HY, Joo HS, Shin DH, et al. LSD1 demethylates HIF1α to inhibit hydroxylation and ubiquitin-mediated degradation in tumor angiogenesis. Oncogene [Internet]. 2017 Sep 28 [cited 2020 Dec 30];36(39):5512–21. Available from: www.nature.com/onc
181. Xie Q, Hao Y, Tao L, Peng S, Rao C, Chen H, et al. Lysine methylation of FOXO3 regulates oxidative stress-induced neuronal cell death. EMBO Rep. 2012;13:371–7.
182. Shi X, Kachirskaia I, Yamaguchi H, West LE, Wen H, Wang EW, et al. Modulation of p53 Function by SET8-Mediated Methylation at Lysine 382. Mol Cell. 2007 Aug 17;27(4):636–46.

183. Huang J, Perez-Burgos L, Placek BJ, Sengupta R, Richter M, Dorsey JA, et al. Repression of p53 activity by Smyd2-mediated methylation. Nature [Internet]. 2006 Nov 1 [cited 2020 Dec 30];444(7119):629–32. Available from: https://www-nature-com.sire.ub.edu/articles/nature05287
184. Levy D, Kuo AJ, Chang Y, Schaefer U, Kitson C, Cheung P, et al. Lysine methylation of the NF-κ B subunit RelA by SETD6 couples activity of the histone methyltransferase GLP at chromatin to tonic repression of NF-κ B signaling. Nat Immunol [Internet]. 2011 Jan 5 [cited 2020 Dec 31];12(1):29–36. Available from: https://www-nature-com.sire.ub.edu/articles/ni.1968
185. Cho HS, Shimazu T, Toyokawa G, Daigo Y, Maehara Y, Hayami S, et al. Enhanced HSP70Â lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B. Nat Commun [Internet]. 2012 [cited 2020 Dec 31];3. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/22990868/
186. Lee JM, Lee JS, Kim H, Kim K, Park H, Kim JY, et al. EZH2 Generates a Methyl Degron that Is Recognized by the DCAF1/DDB1/CUL4 E3 Ubiquitin Ligase Complex. Mol Cell. 2012 Nov 30;48(4):572–86.
187. Lu T, Jackson MW, Wang B, Yang M, Chance MR, Miyagi M, et al. Regulation of NF-κB by NSD1/FBXL11-dependent reversible lysine methylation of p65. Proc Natl Acad Sci U S A [Internet]. 2010 Jan 5 [cited 2020 Dec 31];107(1):46–51. Available from: https://www-pnas-org.sire.ub.edu/content/107/1/46
188. Ling BMT, Bharathy N, Chung TK, Kok WK, Li S De, Tan YH, et al. Lysine methyltransferase G9a methylates the transcription factor MyoD and regulates skeletal muscle differentiation. Proc Natl Acad Sci U S A [Internet]. 2012 Jan 17 [cited 2020 Dec 31];109(3):841–6. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/22215600/
189. Dey J, Dubuc AM, Pedro KD, Thirstrup D, Mecham B, Northcott PA, et al. MyoD is a tumor suppressor gene in medulloblastoma. Cancer Res [Internet]. 2013 Nov 15 [cited 2020 Dec 31];73(22):6828–37. Available from: /pmc/articles/PMC4014678/?report=abstract
190. Pless O, Kowenz-Leutz E, Knoblich M, Rn Lausen J, Beyermann M, Walsh MJ, et al. G9a-mediated Lysine Methylation Alters the Function of CCAAT/Enhancer-binding Protein-* □ S. 2008 [cited 2020 Dec 31]; Available from: http://www.jbc.org/
191. Tseng HH, Hwang YH, Yeh KT, Chang JG, Chen YL, Yu HS. Reduced expression of C/EBPα protein in hepatocellular carcinoma is associated with advanced tumor stage and shortened patient survival. J Cancer Res Clin Oncol [Internet]. 2009 Feb [cited 2020 Dec 31];135(2):241–7. Available from: https://link.springer.com/article/10.1007/s00432-008-0448-5
192. Sampath SC, Marazzi I, Yap KL, Sampath SC, Krutchinsky AN, Mecklenbräuker I, et al. Methylation of a Histone Mimic within the Histone Methyltransferase G9a Regulates Protein Complex Assembly. Mol Cell [Internet]. 2007 Aug 17 [cited 2020 Dec 31];27(4):596–608. Available from: https://pubmed-ncbi-nlm-nih-gov.sire.ub.edu/17707231/
193. Huang J, Dorsey J, Chuikov S, Zhang X, Jenuwein T, Reinberg D, et al. G9a and Glp Methylate Lysine 373 in the Tumor Suppressor p53 * □ S and the. J Biol Chem [Internet]. 2010 [cited 2020 Dec 31];285(13):9636–41. Available from: http://www.jbc.org
194. Zhang J, Wang Y, Shen Y, He P, Ding J, Chen Y. G9a stimulates CRC growth by inducing p53 Lys373 dimethylation-dependent activation of Plk1. Theranostics. 2018;8(10):2884–95.
195. Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, et al. NcRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. Cell. 2011 Nov 11;147(4):773–88.
196. Perillo B, Tramontano A, Pezone A, Migliaccio A. LSD1: more than

demethylation of histone lysine residues. Exp Mol Med [Internet]. 2020 Dec 14 [cited 2020 Dec 30]; Available from: http://www.nature.com/articles/s12276-020-00542-2
197. Hamamoto R, Nakamura Y. Critical roles of non-histone protein lysine methylation in human tumorigenesis. 2015 [cited 2020 Dec 31]; Available from: www.nature.com/reviews/cancer
198. Tracy CM, Warren JS, Szulik M, Wang L, Garcia J, Makaju A, et al. The Smyd family of methyltransferases: role in cardiac and skeletal muscle physiology and pathology. Curr Opin Physiol [Internet]. 2018 Feb [cited 2019 Aug 7];1:140–52. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2468867317300135
199. Xu S, Wu J, Sun B, Zhong C, Ding J. Structural and biochemical studies of human lysine methyltransferase Smyd3 reveal the important functional roles of its post-SET and TPR domains and the regulation of its activity by DNA binding. Nucleic Acids Res [Internet]. 2011 May [cited 2019 Aug 24];39(10):4438–49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21266482
200. Abu-Farha M, Lanouette S, Elisma F, Tremblay V, Butson J, Figeys D, et al. Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2. J Mol Cell Biol [Internet]. 2011 Oct 1 [cited 2019 Sep 19];3(5):301–8. Available from: https://academic.oup.com/jmcb/article-lookup/doi/10.1093/jmcb/mjr025
201. Calpena E, Palau F, Espinós C, Galindo MI. Evolutionary history of the Smyd gene family in metazoans: A framework to identify the orthologs of human Smyd genes in drosophila and other animal species. PLoS One. 2015;10(7).
202. Spellmon N, Holcomb J, Trescott L, Sirinupong N, Yang Z, Spellmon N, et al. Structure and Function of SET and MYND Domain-Containing Proteins. Int J Mol Sci [Internet]. 2015 Jan 8 [cited 2019 Apr 27];16(1):1406–28. Available from: http://www.mdpi.com/1422-0067/16/1/1406
203. Kim M-S, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, et al. A draft map of the human proteome. Nature [Internet]. 2014 May 28 [cited 2019 Aug 24];509(7502):575–81. Available from: http://www.nature.com/articles/nature13302
204. Singh PK. Histone methyl transferases: A class of epigenetic opportunities to counter uncontrolled cell proliferation. 2019 [cited 2020 Nov 2]; Available from: https://doi.org/10.1016/j.ejmech.2019.01.069
205. Song J, Liu Y, Chen Q, Yang J, Jiang Z, Zhang H, et al. Expression patterns and the prognostic value of the SMYD family members in human breast carcinoma using integrative bioinformatics analysis. Oncol Lett [Internet]. 2019 Apr [cited 2019 Sep 19];17(4):3851–61. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30930987
206. Yang J, Everett AD. Hepatoma-derived growth factor represses SET and MYND domain containing 1 gene expression through interaction with C-terminal binding protein. J Mol Biol [Internet]. 2009 Mar 6 [cited 2019 Sep 22];386(4):938–50. Available from: https://www.sciencedirect.com/science/article/pii/S0022283609000035?via%3Dihub
207. Hu T-H, Huang C-C, Liu L-F, Lin P-R, Liu S-Y, Chang H-W, et al. Expression of hepatoma-derived growth factor in hepatocellular carcinoma. Cancer [Internet]. 2003 Oct 1 [cited 2019 Sep 22];98(7):1444–56. Available from: http://doi.wiley.com/10.1002/cncr.11653
208. Saddic LA, West LE, Aslanian A, Yates JR, Rubin SM, Gozani O, et al. Methylation of the retinoblastoma tumor suppressor by SMYD2. J Biol Chem [Internet]. 2010 Nov 26 [cited 2020 Dec 30];285(48):37733–40. Available from: http://www.jbc.org.sire.ub.edu/content/285/48/37733.full
209. Nakakido M, Deng Z, Suzuki T, Dohmae N, Nakamura Y, Hamamoto R. Dysregulation of AKT pathway by SMYD2-mediated lysine methylation on

PTEN. Neoplasia [Internet]. 2015 Apr 1 [cited 2019 Sep 19];17(4):367–73. Available from: https://www.sciencedirect.com/science/article/pii/S1476558615000342?via%3Dihub
210. Wang C, Chen X, Wang Y, Gong J, Hu G. C/EBPαp30 plays transcriptional regulatory roles distinct from C/EBPαp42. Cell Res [Internet]. 2007 [cited 2019 Aug 26];17:374–83. Available from: www.nature.com/cr
211. Sarris ME, Moulos P, Haroniti A, Giakountis A, Talianidis I. Smyd3 is a transcriptional potentiator of multiple cancer-promoting genes and required for liver and colon cancer development. Cancer Cell [Internet]. 2016 Mar 14 [cited 2018 Feb 19];29(3):354–66. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26908355
212. Hu L, Zhu YT, Qi C, Zhu YJ. Identification of Smyd4 as a potential tumor suppressor gene involved in breast cancer development. Cancer Res [Internet]. 2009 [cited 2019 Sep 22];69(9):4067–72. Available from: http://cancerres.aacrjournals.org/
213. Stender JD, Pascual G, Liu W, Kaikkonen MU, Do K, Spann NJ, et al. Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. Mol Cell [Internet]. 2012 [cited 2019 Apr 27];48(1):28–38. Available from: http://dx.doi.org/10.1016/j.molcel.2012.07.020
214. Fujii T, Tsunesumi SI, Sagara H, Munakata M, Hisaki Y, Sekiya T, et al. Smyd5 plays pivotal roles in both primitive and definitive hematopoiesis during zebrafish embryogenesis. Sci Rep [Internet]. 2016 [cited 2019 Nov 10];6. Available from: www.nature.com/scientificreports
215. Sun XJ, Xu P-F, Zhou T, Hu M, Fu C-T. Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes. PLoS One [Internet]. 2008 [cited 2019 Nov 10];3(1):1499. Available from: www.plosone.org
216. Kidder BL, Hu G, Cui K, Zhao K. SMYD5 regulates H4K20me3-marked heterochromatin to safeguard ES cell self-renewal and prevent spurious differentiation. Epigenetics Chromatin. 2017;10(8).
217. Kidder BL, He R, Wangsa D, Padilla-Nash HM, Margarida Bernardo M, Sheng S, et al. SMYD5 controls heterochromatin and chromosome integrity during embryonic stem cell differentiation. Cancer Res [Internet]. 2017 [cited 2019 Nov 17];77(23):6729–45. Available from: http://cancerres.aacrjournals.org/
218. Meng X, Zhao Y, Liu J, Wang L, Dong Z, Zhang T, et al. Comprehensive analysis of histone modification-associated genes on differential gene expression and prognosis in gastric cancer. Exp Ther Med [Internet]. 2019 Sep [cited 2019 Sep 19];18(3):2219–30. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31452712
219. Gao H, Chakraborty G, Lee-Lim AP, Mavrakis KJ, Wendel HG, Giancotti FG. Forward genetic screens in mice uncover mediators and suppressors of metastatic reactivation. Proc Natl Acad Sci U S A [Internet]. 2014 [cited 2019 Nov 10];111(46):16532–7. Available from: www.pnas.org/cgi/doi/10.1073/pnas.1403234111
220. Fevr T, Robine S, Louvard D, Huelsken J. Wnt/-Catenin Is essential for intestinal homeostasis and maintenance of intestinal stem cells. Mol Cell Biol [Internet]. 2007 [cited 2019 Apr 27];27(21):7551–9. Available from: http://mcb.asm.org/
221. Nestorov P, Hotz HR, Liu Z, Peters AHFM. Dynamic expression of chromatin modifiers during developmental transitions in mouse preimplantation embryos. Sci Rep [Internet]. 2015 [cited 2019 Apr 27];5(14347). Available from: www.nature.com/scientificreports
222. Thomson JA, Odorico JS. Human embryonic stem cell and embryonic germ cell lines. Trends Biotechnol [Internet]. 2000 Feb 1 [cited 2019 Apr 27];18(2):53–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10652509

223. Postic C, Shiota M, Niswender KD, Jetton TL, Chen Y, Moates JM, et al. Dual roles for glucokinase in glucose homeostasis as determined by liver and pancreatic β cell-specific gene knock-outs using Cre recombinase. J Biol Chem. 1999;274(1):305–15.
224. Postic C, Magnuson MA. DNA excision in liver by an albumin-Cre transgene occurs progressively with age. Genesis. 2000;26(2):149–50.
225. Weisend CM, Kundert JA, Suvorova ES, Prigge JR, Schmidt EE. Cre activity in fetal albCre mouse hepatocytes: Utility for developmental studies. Genesis. 2009;47(12):789–92.
226. Kellendonk C, Opherk C, Anlag K, Schütz G, Tronche F. Hepatocyte-specific expression of Cre recombinase. Genesis. 2000;26(2):151–3.
227. El Marjou F, Janssen KP, Chang BHJ, Li M, Hindie V, Chan L, et al. Tissue-specific and inducible Cre-mediated recombination in the gut epithelium. Genesis. 2004;39(3):186–93.
228. Diehl F, Brown MA, van Amerongen MJ, Novoyatleva T, Wietelmann A, Harriss J, et al. Cardiac deletion of Smyd2 is dispensable for mouse heart development. PLoS One [Internet]. 2010 Mar 17 [cited 2019 Apr 27];5(3):e9748. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20305823
229. Monk M, Holding C. Human embryonic genes re-expressed in cancer cells. Oncogene. 2001 Dec 6;20(56):8085–91.
230. Wong DJ, Liu H, Ridky TW, Cassarino D, Segal E, Chang HY. Module map of stem cell genes guides creation of epithelial cancer stem cells. Cell Stem Cell. 2008;2:333–44.
231. Zhao W, Lu D, Liu L, Cai J, Zhou Y, Yang Y, et al. Insulin-like growth factor 2 mRNA binding protein 3 (IGF2BP3) promotes lung tumorigenesis via attenuating p53 stability. Oncotarget. 2017 Nov 1;8(55):93672–87.
232. Yu F-X, Zhao B, Guan K-L. Hippo pathway in organ size control, tissue homeostasis, and cancer. Cell. 2015;163(4):811–28.
233. Becker D, Sfakianakis I, Krupp M, Staib F, Gerhold-Ay A, Victor A, et al. Genetic signatures shared in embryonic liver development and liver cancer define prognostically relevant subgroups in HCC [Internet]. Vol. 11, Molecular Cancer. 2012 [cited 2019 Dec 25]. Available from: http://www.molecular-cancer.com/content/11/1/55
234. Lugli N, Kamileri I, Keogh A, Malinka T, Sarris ME, Talianidis I, et al. R-spondin 1 and noggin facilitate expansion of resident stem cells from non-damaged gallbladders. EMBO Rep [Internet]. 2016 May [cited 2018 Feb 19];17(5):769–79. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26993089
235. Manibusan MK, Odin M, Eastmond DA. Postulated carbon tetrachloride mode of action: a review. J Environ Sci Heal [Internet]. 2007 [cited 2019 Apr 27];25(3):185–209. Available from: http://www.tandfonline.com/doi/abs/10.1080/10590500701569398
236. Huch M, Dorrell C, Boj SF, Van Es JH, Van De Wetering M, Li VSW, et al. In vitro expansion of single Lgr5+ liver stem cells induced by Wnt-driven regeneration. Nature [Internet]. 2013 [cited 2019 Apr 27];494(7436):247–50. Available from: www.nature.com/nature.
237. Kiernan JA. Indigogenic substrates for detection and localization of enzymes. Biotech Histochem. 2007;82(2):73–103.
238. Muñoz J, Stange DE, Schepers AG, Van De Wetering M, Koo BK, Itzkovitz S, et al. The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent ′ +4′ cell markers. EMBO J [Internet]. 2012 [cited 2019 Apr 27];31(14):3079–91. Available from: www.embojournal.org
239. Wang B, Zhao L, Fish M, Logan CY, Nusse R. Self-renewing diploid Axin2 + cells fuel homeostatic renewal of the liver. Nature. 2015;524(7564):180–5.
240. Bagislar S, Sabò A, Kress TR, Doni M, Nicoli P, Campaner S, et al. Smyd2 is a Myc-regulated gene critical for MLL-AF9 induced leukemogenesis. Oncotarget

[Internet]. 2016 [cited 2019 Apr 27];7(41):66398–415. Available from: http://www.oncotarget.com/fulltext/12012
241. Puisieux A, Brabletz T, Caramel J. Oncogenic roles of EMT-inducing transcription factors. Vol. 16, Nature Cell Biology. Nature Publishing Group; 2014. p. 488–94.
242. Gorrini C, Harris IS, Mak TW. Modulation of oxidative stress as an anticancer strategy. Vol. 12, Nature Reviews Drug Discovery. 2013. p. 931–47.
243. Weinberg F, Ramnath N, Nagrath D. Reactive oxygen species in the tumor microenvironment: An overview. Cancers (Basel). 2019 Aug 1;11(8).
244. Verma A, Kambhampati S, Parmar S, Platanias LC. Jak family of kinases in cancer. Cancer Metastasis Rev. 2003;22(4):423–34.
245. Kohno H, Suzuki R, Sugie S, Tanaka T. β-Catenin mutations in a mouse model of inflammation-related colon carcinogenesis induced by 1,2-dimethylhydrazine and dextran sodium sulfate. Cancer Sci. 2005;96(2):69–76.
246. Espinosa J. Histone H2B ubiquitination: the cancer connection. Genes Dev [Internet]. 2008;22:2743–9. Available from: citeulike-article-id:3906840
247. Jing YY, Cai FF, Zhang L, Han J, Yang L, Tang F, et al. Epigenetic regulation of the Warburg effect by H2B monoubiquitination. Cell Death Differ [Internet]. 2019; Available from: http://dx.doi.org/10.1038/s41418-019-0450-2
248. Bhatnagar S, Gazin C, Chamberlain L, Ou J, Zhu X, Tushir JS, et al. TRIM37 is a new histone H2A ubiquitin ligase and breast cancer oncoprotein. Nature [Internet]. 2014;516(729):116–20. Available from: http://dx.doi.org/10.1038/nature13955
249. Jeusset LMP, McManus KJ. Ubiquitin specific peptidase 22 regulates histone H2B mono-ubiquitination and exhibits both oncogenic and tumor suppressor roles in cancer. Cancers (Basel). 2017 Dec 1;9(12).
250. Mobley RJ, Abell AN. Controlling epithelial to mesenchymal transition through acetylation of histone H2BK5. J Nat Sci. 2017;3(9):1–10.
251. Sun ZW, Allis CD. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature. 2002;418(6893):104–8.
252. Kim J, Hake SB, Roeder RG. The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. Mol Cell. 2005;20(5):759–70.
253. Corujo D, Buschbeck M. Post-translational modifications of H2A histone variants and their role in cancer. Cancers (Basel). 2018;10(3):1–25.
254. Nacev BA, Feng L, Bagert JD, Lemiesz AE, Gao JJ, Soshnev AA, et al. The expanding landscape of 'oncohistone' mutations in human cancers. Nature. 2019;567(7749):473–8.
255. Yi X, Jiang XJ, Fang ZM. Histone methyltransferase SMYD2: ubiquitous regulator of disease. Clin Epigenetics [Internet]. 2019 Dec 1 [cited 2020 Feb 3];11(1):112. Available from: https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-019-0711-4
256. Du SJ, Tan X, Zhang J. SMYD proteins: Key regulators in skeletal and cardiac muscle development and function. Anat Rec [Internet]. 2014 Sep [cited 2019 Aug 26];297(9):1650–62. Available from: http://doi.wiley.com/10.1002/ar.22972
257. Maeda S, Kamata H, Luo JL, Leffert H, Karin M. IKKβ couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. Cell. 2005;121(7):977–90.
258. Das M, Sabio G, Jiang F, Rincón M, Flavell RA, Davis RJ. Induction of hepatitis by JNK-mediated expression of TNF-α. Cell. 2009 Jan 23;136(2):249–60.
259. Das M, Garlick DS, Greiner DL, Davis RJ. The role of JNK in the development of hepatocellular carcinoma. Genes Dev. 2011 Mar 15;25(6):634–45.
260. Han MS, Barrett T, Brehm MA, Davis RJ. Inflammation mediated by JNK in myeloid cells promotes the development of hepatitis and hepatocellular

carcinoma. Cell Rep [Internet]. 2016 [cited 2020 Feb 15];15(1):19–26. Available from: http://dx.doi.org/10.1016/j.celrep.2016.03.008
261. Gomez-Cuadrado L, Tracey N, Ma R, Qian B, Brunton VG. Mouse models of metastasis: Progress and prospects. Dis Model Mech. 2017;10(9):1061–74.
262. Tauriello DV, Palomo-Ponce S, Stork D, Berenguer-Llergo A, Badia-Ramentol J, Iglesias M, et al. TGFβ drives immune evasion in genetically reconstituted colon cancer metastasis. Nature. 2018;554(7693):538–43.
263. Fabini E, Manoni E, Ferroni C, Rio A Del, Bartolini M. Small-molecule inhibitors of lysine methyltransferases SMYD2 and SMYD3: Current trends. Future Med Chem. 2019;11(8):901–21.
264. Thomenius MJ, Totman J, Harvey D, Mitchell LH, Riera T V., Cosmopoulos K, et al. Small molecule inhibitors and CRISPR/Cas9 mutagenesis demonstrate that SMYD2 and SMYD3 activity are dispensable for autonomous cancer cell proliferation. Schang LM, editor. PLoS One [Internet]. 2018 Jun 1 [cited 2020 Mar 18];13(6):e0197372. Available from: https://dx.plos.org/10.1371/journal.pone.0197372
265. Toh TB, Lim JJ, Chow EK-H. Epigenetics of hepatocellular carcinoma. Clin Transl Med [Internet]. 2019 May 6 [cited 2019 Jul 20];8(1):13. Available from: http://www.ncbi.nlm.nih.gov/pubmed/31056726
266. Bayo J, Fiore EJ, Dominguez LM, Real A, Malvicini M, Rizzo M, et al. A comprehensive study of epigenetic alterations in hepatocellular carcinoma identifies potential therapeutic targets. J Hepatol [Internet]. 2019;71(1):78–90. Available from: https://doi.org/10.1016/j.jhep.2019.03.007
267. Dai Q, Zhang C, Yuan Z, Sun Q, Jiang Y. Current discovery strategies for hepatocellular carcinoma therapeutics. Expert Opin Drug Discov [Internet]. 2020;15(2):243–58. Available from: https://doi.org/10.1080/17460441.2020.1696769