PhD degree in Systems Medicine (curriculum in Molecular Oncology)

European School of Molecular Medicine (SEMM),

University of Milan and University of Naples "Federico II"

Settore disciplinare: BIO/11

Uncovering epigenetic vulnerabilities in intestinal cancer development

Cecilia Toscani

IEO, Milan

Matricola n. R11468

Supervisor: Prof. Diego Pasini

IEO, Milan

Anno accademico 2018-2019

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

5-methylcytosine (5mC)

Adenomatous polyposis coli (APC)

Bone morphogenic protein (BMP)

Breast metastasis suppressor 1 (BRMS1)

Chromatin immunoprecipitation (ChIP)

Class II transactivator (CIITA)

Colorectal cancer (CRC)

Cytosine-guanine dinucleotides (CpG)

DNA methyltransferases (DNMTs)

Embryonic stem cells (ESC)

Epidermal growth factor (EGF)

Esophageal squamous cell carcinoma (ESCC)

Green fluorescent protein (GFP)

Histone acetyltransferases (HATs)

Histone deacetylases (HDACs)

Histone demethylases (HDMs)

Histone lysine methyltransferases (HKMTs)

Histone methyltransferases (HMTs)

Histone H2A lysine K119 mono ubiquitin (H2AK119ub)

Histone H3 lysine K4 monomethylated (H3K4me1)

Histone H3 lysine K4 trimethylated (H3K4me3)

Histone H3 lysine K27 monomethylated (H3K27me1)

Histone H3 lysine K27 trimethylated (H3K27me3)

Histone H3 lysine K27 acetylated (H3K27ac)

Intestinal stem cells (ISCs)

Mouse embryonic fibroblast (MEF)

Polycomb Repressive Complex 1 (PRC1)

Polycomb Repressive Complex 2 (PRC2)

Post translational modification (PTM)

Protein arginine methyltransferase (PRMTs)

Real time quantitative PCR (qRT-PCR)

RNA interference screening (RNAi)

RNA sequencing (RNA-seq)

RT-qPCR Quantitative Reverse transcription PCR

Ten-eleven translocation (TET)

S-adenosylmethionine (SAM)

Short hairpin RNA (shRNA)

Valproic acid (VPA)

WD-repeat protein-5 (WDR5)

Western blot (WB)

Wingless integration site (WNT)

Wild-type (WT)

FIGURE INDEX

INTRODUCTION

MATHERIALS AND METHODS

RESULTS

ABSTRACT

Colorectal cancer (CRC) arises from a multi-step process leading to the progressive accumulation of genetic and epigenetic mutations, thus causing deregulation in homeostasis followed by neoplastic transformation.Epigenetic and genetic alterations are able to induce a constitutive activation of the WNT signaling pathway, whose aberrant activity converges into deregulation of proliferation, differentiation and cell death pathways. The most common causes of WNT pathway hyper-activation are APC loss of function, or β-catenin constitutive activation mutations. Despite this knowledge of aberrant WNT activity, upstream interference with this signaling pathway induces adverse effects due to high cross-talk with other pathways, highlighting a need to find alternative ways to indirectly target the effectors of this pathway.

Deregulated pathways in CRC provoke aberrant signaling that converges into the nucleus where transcription and chromatin-remodeling factors cooperate to maintain or modify the identity of a cell. In recent years, several studies have been focused on epigenetic players, which act by depositing specific and reversible post-translational modifications. For this reason they are being recognized as promising new targets for the development of cancer therapeutic strategies. In this context, my project takes advantage of murine 3D intestinal organoid cultures, carrying oncogenic deregulations of the WNT pathway, as a platform for pooled and arrayed RNA interference screens, to identify novel regulators that control the nuclear/transcriptional aspect of this oncogenic pathway. I also implemented the validation of selected targets in human metastatic colorectal cancer organoids to highlight their clinical relevance. Moreover, since several chromatin modifier inhibitors have been already developed, the findings of this project should prompt the development of new molecules for CRC treatment to target the novel sensitivities I identified. Finally, this project generated important technical knowledge through this pioneering approach that will open up the possibility of performing functional screens in other tissues from which organoid cultures have already been established.

INTRODUCTION

1. Genetic events in colorectal cancers

Colorectal cancer (CRC) is the second cause of tumor-related death in industrialized countries[1]. It develops as a consequence of a combination of genetic, demographic, behavioral and environmental factors.

The risk of developing this cancer is related to bad nutritional habits, alcohol intake, smoking, intestinal inflammatory disease, polyps, genetic factors, and aging[2].

The clinical goal to treat this disease is to reduce the mortality rate with early diagnosis and treatment. The stage in which the cancer is detected determines the prognosis, survival, and treatment of the patient[2].

At a phenotypic level CRC arise in most of the cases from dysplastic adenomatous polyps. This confers a selective growth advantage to the colonic epithelial cell and drives the transformation from normal epithelium to adenomatous polyp to invasive colorectal cancer[3].

At a molecular level, the same signaling pathways that sustain and constrain intestinal stem cells (ISCs) activity at the bottom of intestinal crypts are always affected by mutations in CRC patients. Such tumor, that usually is sporadic, arises from a multistep accumulation of somatic genetic mutations process involving the inactivation of a variety of genes that suppress tumors and repair DNA and the simultaneous activation of oncogenes. It indeed reported that CRC presents WNT constitutive activation by frequent APC inactivating mutations in the ~80% of the patients[4], or less frequently by ß-Catenin stabilizing mutations (~5%); EGF constitutive activation via activating RAS mutations (~30%-50%)[5]; TGFß/BMP signaling inactivation via frequent SMAD4 or TGFß receptor loss of function (~40-50%)[6].

The loss of APC induces hyperproliferation and expansion of undifferentiated progenitor cells, resulting in the formation of a small adenoma[7]. Subsequently KRAS and BRAF activations induce a second round of expansion, marking a switch that leads to progression into large adenoma. Mutations of PIK3CA, SMAD4 and TP53 finally generate a malignant

phenotype that has potential for invasion and metastasis[8]. It is important to mention that CRC cells can revert to functionally normal cells, upon restoration of APC function, restoring the homeostasis n the intestinal crypts[9].

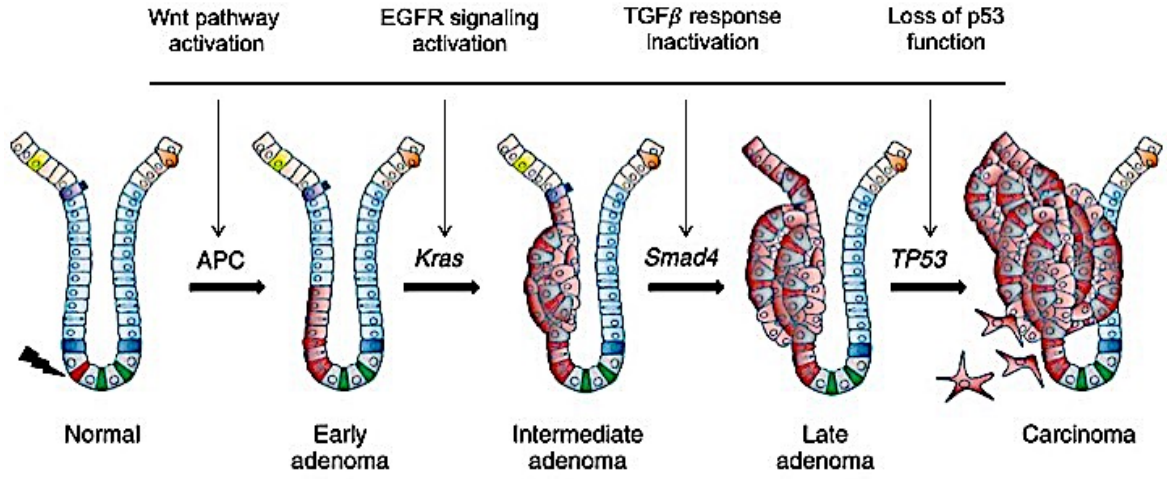

Fig. 1 Model of intestinal cancer initiation and progression. Loss of APC leads to WNT pathway constitutive activation, resulting in increased proliferation. Subsequent KRAS activation promotes the transition from a dysplastic crypt to an early adenoma. The gradual accumulation of additional oncogenic hits drives the progression from a benign adenoma to an invasive carcinoma. (Image from Pamela Rizk, Nick Barker 2012)

2. The intestinal compartment

2.1 Intestinal epithelium

The mammalian small intestine exerts diverse functions, such as nutrient absorption of food arriving from the stomach, or the formation of a protective barrier against pathogens in the intestinal lumen. Such functions are performed by a columnar epithelium structured into invaginations, termed crypts of Lieberkuhn, interspersed with protrusions, called villi, whose role is to expand the absorptive surface area of the gut[10]. The intestinal epithelium undergo complete turnover in 4-5 days, a mechanism by which is able to deal with the harsh intestinal environment[11],

The intestinal epithelium is a rapidly self-renewing tissue whose tissue architecture and regeneration is sustained by highly proliferating intestinal stem cells (ISCs) that reside near the base of the crypts[12]. The progeny of ISCs, termed transient amplifying (TA) cells, is expanded through a limited number of rounds of mitosis while it migrates upwards along

the crypt axis. Close to the intestinal lumen, TA cells undergo cell cycle arrest, terminal differentiation and apoptosis at the tip of the villi[12].

Due to the fast renewal time of the intestinal epithelium, it has been proposed that only mutations occurring in ISCs can be fixed in the tissue to give rise to tumor, suggesting that ISCs represent the *bona fide* cells of origin for CRC. Their identity is maintained by well-established signaling pathways among which Wnt, BMP and EGF signaling integrate with each other to maintain intestinal homeostasis[13]. WNT is the central signaling sustaining self-renewal and proliferation of ISCs. Epidermal growth factor (EGF) is the key growth factor regulating cell survival that exerts mitogenic effects on stem and TA cells. BMP belongs to the transforming growth factor B (TGF-B) family, prevents crypt formation and ISC self-renewal and favors the maturation of the secretory lineage.

2.2 Wnt Signaling Pathway

The Wnt signaling pathway is an ancient and highly evolutionarily conserved pathway that regulates crucial processes such as cell fate determination, cell migration, cell polarity, neural patterning and organogenesis during embryonic development.

Wnts are secreted glycoproteins that compose a family of 19 proteins in humans involved in diverse layers of signaling regulation, function and biological output.

To date, two signaling branches downstream of the Fz receptor have been identified. These comprise a canonical or Wnt/β-catenin dependent pathway and a non-canonical or β-catenin-independent pathway. This second branch can be further divided into the Planar Cell Polarity and the Wnt/Ca2+ pathways[14].

2.2.1 The canonical Wnt Pathway

The canonical pathway was first identified and through genetic screens carried in Drosophila and further studies in the fly, worm, frog, fish and mouse have led to the identification of a basic molecular signaling framework. The main feature of the canonical Wnt pathway is the accumulation and the translocation of the adherent junction associated-protein β-catenin into the nucleus. Without Wnt signaling, cytoplasmic β-catenin is degraded by a β-catenin destruction complex, which includes Axin, adenomatosis polyposis coli (APC), protein phosphatase 2A (PP2A), glycogen synthase kinase 3 (GSK3) and casein kinase 1α (CK1α)[15]. Casein Kinase and GSK3 can phosphorylate β-catenin, which is targeted for ubiquitination and the subsequent proteolytic destruction by the proteosomal machinery[16]. Binding of Wnt to its receptor complex, composed of the Fz and the LRP5/6, triggers a series of events that mediates the disruption of the APC/Axin/GSK3 complex, required for the targeted destruction of β-catenin[17]. The binding of Wnt to the Fz/ LRP5/6 complex induces the membrane translocation of a key negative regulator of signaling Axin, which binds to a conserved sequence in the cytoplasmic tail of LRP5/6[18,19].

Upon membrane translocation of Axin, its binding to LRP5/6 is catalyzed by the phosphorylation of LRP5/6, mediated by either CK1γ or GSK3.

It has been proposed that the binding of Axin to remove its negative activity on canonical Wnt signaling thus leading to the activation of the phosphoprotein Dsh, also known to bind to Axin and Fz[20]. Dsh is thus phosphorylated by several kinases, leading to a regulation of both its subcellular localization and its ability to interact with effectors of Wnt signaling[21].

Once Dsh is activated, it inhibits the activity of the GSK3 enzyme, and activates series of different processes leading to the prevention of degradation of β-catenin and its consequent stabilization and accumulation in the cytoplasm[22]. Stabilized β-catenin then translocates into the nucleus, acting as a transcriptional co-activator, interacting with T cell factor (TCF)/ lymphocyte enhancer factor (LEF) to induce specific gene transcription[23].

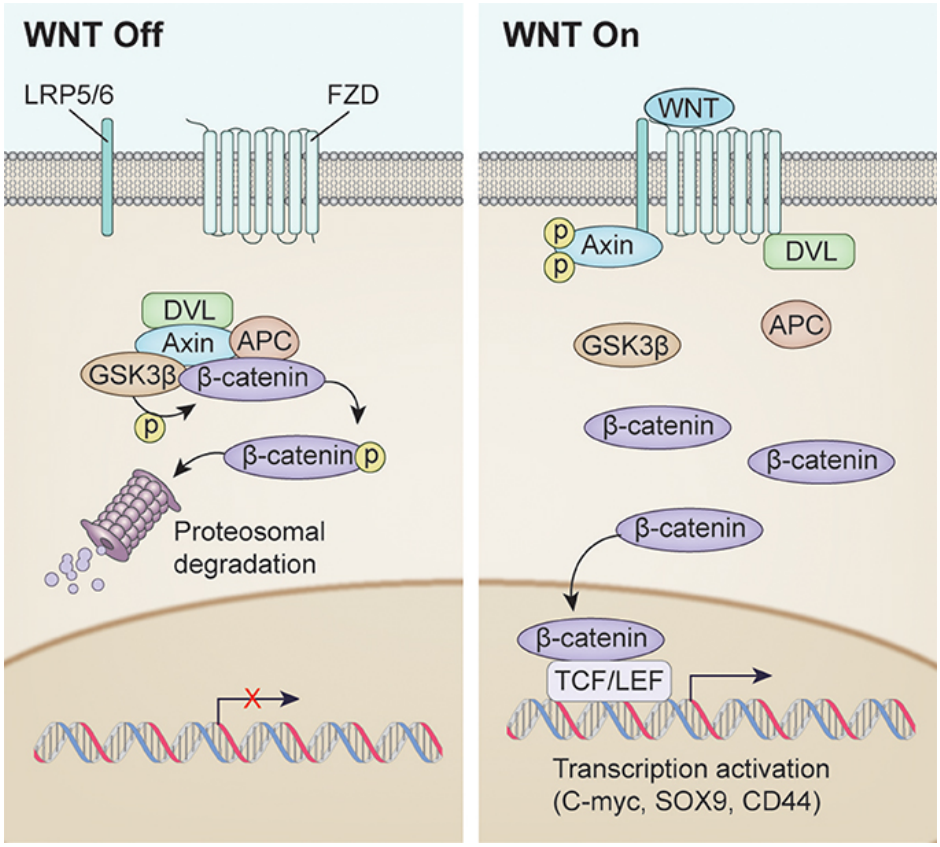

Fig. 2 Overview of WNT signaling pathway activation. Left: WNT signaling pathway inactivation. In absence of WNT ligand, phosphorylation of β-catenin resulted in Dishevelled (DVL), Axin, APC, and GSK3β complex resulting in proteosomal degradation. Right: canonical WNT signaling activation after WNT ligand binding. Unphosphorylated β-catenin enters the nucleus to drive transcription. (image from McCord et al. Front. Cell. Neurosci. 2017)

2.3 Wnt Signaling and cancer

Wnt signaling alterations are frequently found in colorectal cancer. APC gene can present germ-line heterozygous mutations that lead to a hereditary cancer syndrome defined as familiar adenomatous polyposis (FAP), inherited in an autosomal dominant manner[24]. The second allele is frequently lost in individual cells, which is translated for patients in an uncontrolled growth of hundreds to thousands adenomatous polyps in the colon during the second and third decade of life. Such polyps frequently develop into colorectal carcinoma by the age of 40–50 years. About 10% of FAP patients present a milder course of disease with a significant diminution in number of adenomas and a later onset of disease. This variant is termed attenuated FAP (AFAP)[25]. The majority of the cases of sporadic colorectal

cancer result from loss of both APC alleles[26], leading to accumulation and stabilization of β-catenin that converge in the nucleus forming a complex with the intestinal TCF family member Tcf4[27]. Less frequently colorectal cancers present a mutation in Axin2[28], or mutations leading to a stable form of β-catenin[29]. Some patients present also fusions between VTl1A and Tcf7l2, the gene encoding Tcf4[30].

2.4 Three-dimensional intestinal organoid cultures

In the recent years, a great effort was put to develop *in vitro* culture methods of adult stem cells in order to better understand tissue homeostasis and stem cell biology. This involves *in vitro* systems to cultivate and differentiate ISCs in three-dimensional structures[31], named organoids, which closely recapitulate the physiology, functions and cell type content of the *in vivo* intestinal epithelium. Such system has revolutionized *in vitro* studies of the small intestine epithelium.

Organoids can be obtained from purified crypts embedded in a laminin/collagen-rich matrix mimicking the basal lamina and grown in a medium containing factors that preserve the stem cell niche. Such organoid culture medium contains several niche factors: EGF, Noggin, and R-spondin. These growth factors were defined taking advantage from several observations reported in literature. First, the discovery of Wnt signaling as an essential cellular signaling pathway for stem cell maintenance *in vivo*[32], and the observation that R-spondins, Wnt agonists binding to LGR5, are mitogens causing stem cell hyperplasia. Furthermore, a study reported EGF as another potent mitogen. Finally Noggin was discovered to be crucial for the maintenance of stem cell niche[33-35]. In addition to these essential growth factors, it is possible to use in particular circumstances Rho kinase (ROCK) inhibitor to inhibit anoikis[36]. These crypts embedded in Matrigel and cultured with such growth factors can form self-organized and self-renewed organoids structures closely recapitulating the native intestinal epithelium. Organoids form as cysts and differentiate through bud formation in mini-guts composed by a central lumen flanked by a highly polarized villus epithelium and crypt-like

protrusions. The stem cells reside at the bottom of these crypts and gradually give rise to all the other cell types that form the intestinal epithelium *in vivo*. All major oncogenic mutations that drive CRC development are sufficient to relief organoids growth from the stimuli present in the culture media leading to an indefinite expansion of organoids as symmetric round cysts that mimics in vivo tumor growth[37].

This makes organoid structures a powerful tool to investigate the molecular events that governs both normal homeostasis and cancer development[38].

2.4.1 Tumoroids

In vitro tumor models failed to mimic the heterogeneity of human cancers, leading to a significant lack in understanding the tumor pathogenesis process and consequently in the therapeutic approaches. The rise of 3D organoid cultures represents a new frontier not only for the modeling of human cancers, but also for the development of new strategies of drug treatment, thus being recognized as robust preclinical models[39].

Intestinal cancer organoids present diverse behaviors in comparison with their wild-type counterpart, showing different levels of dependency for the growth factors commonly utilized in the *in vitro* culture, according to their mutational landscape of origin[40,41].

More in details, in physiologic conditions murine and human intestinal stem cells can be cultured with a medium containing the stem-cell-niche factors WNT, R-spondin, EGF and Noggin, but when precise mutation as APC, P53, KRAS and SMAD4 occurs, organoids acquire the capability to grow independently of such growth factors[42].

Phenotypic and genotypic profiling of cancer organoids showed a high degree of similarity to the tumor *in vivo*. Furthermore, it has been reported that orthotopic transplantation of colorectal cancer organoids is able to mimic more accurately tumorigenesis and liver metastasis formation in the native colon environment[43].

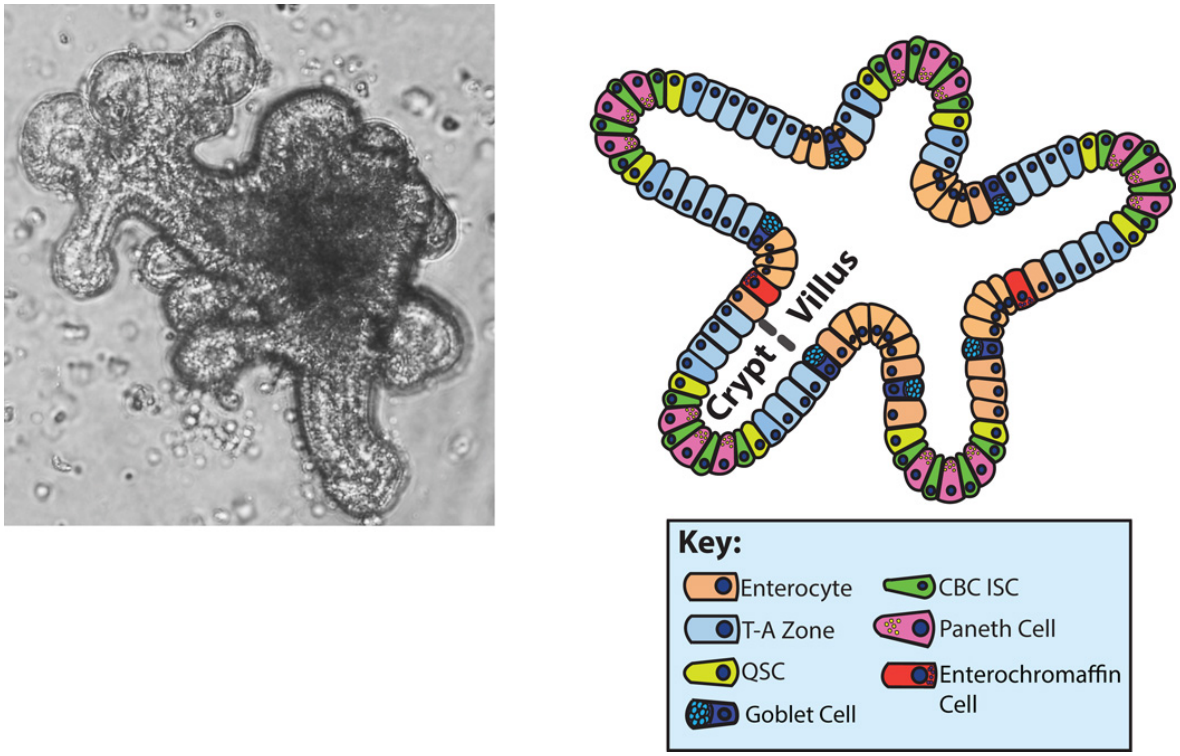

Fig. 3 Organoids resemble in vivo intestinal epithelium. Left: murine small intestinal organoid. Right: illustration of typical intestinal organoid structure and distribution of intestinal epithelial cell types.(from Wallach et al. J.Pediatr. Gastroenterol Nutr 2018)

3. Functional genetic screens

Functional genomics takes advantage of the information generated from genomes, aiming to address on a genomic scale the relationship between genotype and phenotype, The basic aim of a genomic screen is to modify gene function and to study the genetic perturbation so generated to dissect what happens to the resulting phenotype. Possible outputs of a genetic screen could be modifications of cell viability or proliferation. A classic genetic screen follow the phenotype-to-genotype strategy, starting with the alteration of the expression of multiple, then clones displaying the phenotype of interest selected, and the responsible genetic change identified. Such technique provided valuable insights in the oncology field, unraveling new putative diagnostic markers and therapeutic targets, or identifying mechanisms of drug resistance and thereby delivering strategies to combat resistance. A further application of such loss-of-function screens is the identification of drug targets whose inactivation is effective in a specific context. This approach, defined as synthetic lethality

screen enables the identification of drugs that act only in cancer cells having a cancer-specific mutation and the discovery of potent combination therapies[44].

This technique can be performed both *in vivo,* by means of whole living organisms, or *in vitro,* by introducing genetic material in cells to induce genetic alterations causing changes in the morphological or biochemical phenotype.

3.1 RNA interference screens

RNA interference (RNAi) is a powerful tool to silence gene expression, facilitating the assessment of genes function. Stable suppression of specific genes mediated by a collection of thousands of short-hairpin RNAs (shRNAs) allows the simultaneous screening of several different genes in order to identify which one acts in specific cellular pathways[45]. The procedure usually relies on a lentiviral library encoding a huge population of shRNA sequences packaged in viral particles. To avoid off-target effects, it is necessary to use several different shRNAs to target efficiently each gene.

Such shRNA sequences are inserted in a backbone background characterized by the presence of both constitutive promoters for shRNA expression and selection cassette that enable the selection of infected cells by antibiotic resistance or by fluorescence visualization through a reporter. Commonly, libraries also contain unambiguous barcodes to enable identification of each specific shRNA through High-Throughput sequencing technique. To assess which specific shRNA caused a precise phenotype, it is important to ensure that the majority of the cells are only infected with one copy of shRNA. This can be achieved by infecting the cells with a low multiplicity of infection (MOI), a parameter that determine the number of viral integration per each cell, that has to be accurately calculated[46].

The relative abundance of a specific shRNA can be determined through PCR amplification from genomic DNA followed by Next Generation Sequencing. To assess without any risk of off-target effect the function of the gene studied, multiple shRNAs targeting such gene have to be identified in the screen[44].

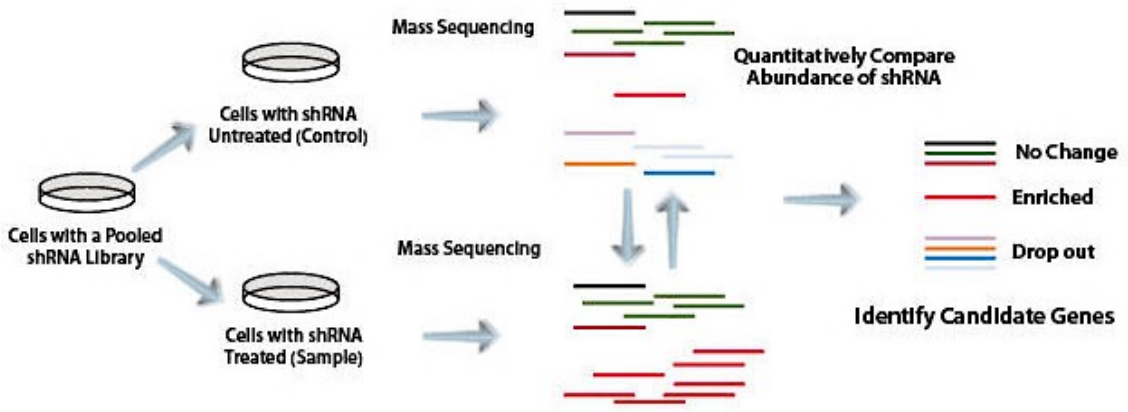

Fig.4 Grahical representation of RNAi screening technique. Cells of interest are first infected with a lentiviral pooled library. Control sample and experimental sample are collected at the time-points established for the procedure, then genomic DNA is extracted and barcode amplified by PCR. HT-sequencing analysis allow the quantitation of the abundance of each barcode in the two samples, identifying genes that has to be object of further analyses. (from sigmaaldrich.com)

.

4.Overview of chromatin features and functions

4.1 Chromatin structure

The eukaryotic genome is compacted through several hierarchical levels of organization to fit within the nucleus. To achieve this, genomic DNA is associated with specialized proteins to form the chromatin fibers. The basic unit of the chromatin, characterized for the first time by Kornberg in 1974 through x-ray diffraction pattern from nuclei of cells, is the nucleosome[47]. It is constituted by 147 base pairs of DNA wrapped around a proteic core formed by an octamer of two copies each of the four histones proteins H2A, H2B, H3 and H4.

As early event, two copies of histones H3 and H4 form a tetramer. This successively associates with two dimers of H2A and H2B forming the complete octamer. The core histones make contact with the DNA through the phosphodiester backbone. No specific contacts between the DNA bases and the histones are reported, explaining how the nucleosomes can interact with the DNA in a sequence-independent fashion[48]. A specific region of DNA, the linker region, separates two adjacent nucleosomes. Unlike core histone proteins building a nucleosomal core on which DNA is wrapped, histone H1 binds the linker DNA in close proximity to the sites of DNA entry and exit to the nucleosome core. Histone

H1 association has a role in the establishment and maintenance of the assembly of three-dimensional chromatin structures, leading to an increased degree of chromatin compaction[49]. The distance in terms of number of nucleotides between two adjacent nucleosomes is variable according to the organism of interest and is known as the nucleosome repeat length[50]. Histone proteins possess biochemical properties that allow the formation of a very stable structure. More in details, the octamer core is abundant in amino acids with basic lateral chain that make them positively charged. These lateral chains mediate the establishment of ionic bonds with the negatively charge of the phosphate group of DNA establishing a tight interaction with it.

Importantly, nucleosomes are the main characters in determining DNA accessibility. The characteristics of each nucleosome relies on the DNA sequence incorporated, and this can have functional significance for specific gene promoters[51]. Nucleosomes can be more or less compacted, thus determining the folding of chromatin and consequently influencing the rate of gene transcription. Indeed, it is possible to define two configurations of chromatin according to their compaction status and transcriptional activity: euchromatin and heterochromatin. Euchromatin is characterized by a more relaxed compaction rate, thus being considered more accessible to transcription factors, while heterochromatin is tightly compacted and it is related to stable transcriptional repression. However, this organization is reversible and dynamic. In particular facultative heterochromatin is able to undergo conformational changes through the action of particular proteins, namely chromatin remodelers. These proteins mediate the change of the folding rate after specific stimuli sensed by the cell, allowing changes in gene expression by a switch from heterochromatin to euchromatin.

There are also regions of constitutive heterochromatin, such as pericentromeric or telomeric chromatin, characterized by a poor transcriptional activity and the presence of repeated sequences. This structure is physically separated from euchromatin and facultative heterochromatin in the cell nucleus[52].

4.2 The histone structure

The interactions between each of the histone chains of the histone were clarified by the 2.8 Å resolution structure of the nucleosome core particle[53]. Histones H2A, H2B, H3 and H4 are composed by two common regions, denominated histone fold and histone tails.

All the histones composing the core contain the histone fold domain that is composed by three □□helices connected by two loops, which allow heterodimeric interactions between core histones known as the handshake motif. These regions bind 2.5 turns of DNA, which is wrapped around them along their long axes generating a 140° bend.

Complementary internal packing causes the formation of only H3–H4 and H2A–H2B heterodimeric pairs instead of homodimers or other heterodimers. *In vitro* experiments demonstrated that if a physiologically relevant ionic strength is applied the only oligomeric aggregates of histones formed are the H3–H4 tetramer and the H2A–H2B dimer in solutions, and the entire histone octamer in the presence of DNA or high salt concentration.

Furthermore, each core histone contains also a tail. These tails are regions prevalently amino-terminal to the histone fold domain with the exception of H2A, whose tails are extended to both sides of his fold region. These regions are highly basic and contribute to the formation of the chromatin fiber through interactions with neighboring nucleosomes and in addiction they have a role in supporting the establishment of higher order structures by facilitating nucleosome assembly or disassembly[54]. The N-tails protrude as disordered structures from the center of the nucleosome and are known to interact with both the nucleosomal DNA and the DNA linker.

The core histone tails play instead important roles in nucleosome stability. It is reported that the function of each histone could be redundant in maintaining intact cellular chromatin. In particular *in vivo* studies have demonstrated that the removal of a single N-tail of each histone does not show any effect on cell viability, while the combined deletion of two N-terminals tails from the histone pairs H2A/H2B, H3/H4 or H2A/H4 compromises cell survival[55-57].

4.3 Chromatin remodeling

The formation and maintenance of specific chromatin states is tightly regulated during development and during all DNA related processes that require rapid rearrangements of chromatin architecture. This plasticity allows access of compacted genomic DNA to the regulatory transcription machinery proteins, and thereby controls gene expression. Such chromatin dynamicity is achieved through the activity of different remodeling complexes. Among them a crucial role is exerted by adenosine triphosphate (ATP)-dependent chromatin remodeling complexes. These are large complexes, highly conserved within eukaryotes, which use energy from ATP hydrolysis to slide, eject and restructure nucleosomes. In particular chromatin-remodeling enzymes catalyze a broad range of chromatin transformations that include sliding the histone octamer across the DNA, changing the conformation of nucleosome DNA and changing the composition of the histone octamer, leading to a fine modulation of chromatin structure and functionality[58,59]. In general, nucleosome remodeling results in increased or reduced accessibility of genomic loci thus leading to transcriptional activation or repression[60,61].

Chromatin remodeling complexes are all characterized in terms of structure by the presence of an ATPase subunit homologous to ATP-binding helicase of the DEAD/H box-containing family[62]. These ATPases have been grouped into 4 families: SWI/SNF, ISWI, IN080 and CHD[63]. The additional presence of multiple accessory scaffold subunits contributes to a combinatorial assembly of these complexes that allows the diversification of their biochemical properties and functions. These non-catalytic subunits are required for the targeting and regulation of distinct nucleosome positioning activities of remodeling complexes and thereby determine the gene expression program and the cell fate. The different remodeling enzymes recognize different histone modifications, DNA structures/sequences and RNA signals that target them to specific genomic loci recognized by dedicated protein domains, termed reader domains. The accessibility of the DNA through the activity of chromatin remodelers is precisely tuned and therefore the malfunction of these

enzymes trigger cells to undergo tumorigenesis, with this class of enzymes being frequently and specifically mutated in a wide variety of cancers[64,65].

4.4 Roles of the major histone and DNA modifications

In addition to DNA sequence, which influences both nucleosome positioning and unwrapping characteristics, nucleosome stability and plasticity are further modulated by chemically altering the histones themselves, which changes the energy landscape of histone–DNA interactions and therefore increases the dynamic range of DNA accessibility. These chemical changes can be in the form of post-translational modifications (PTMs) that can be added and removed enzymatically in a reversible fashion.

The first experiment that demonstrated that histone proteins are subject of PTMs at their N-terminal tails date back to early sixties[66]. Nowadays, about 100 different modifications have been identified and described. These modifications act influencing the chromatin architecture, as demonstrated by the X-ray structure of the nucleosome[53]. Besides governing nucleosome interactions, histone PTMs were shown to be fundamental in recruiting specific proteins, such as remodeling enzymes, to reposition nucleosomes throughout the genome.

The most functionally characterized histone PTMs are acetylation and methylation of lysine residues, or the phosphorylation of threonine, tyrosine and serine. Recently more modification were appreciated, such as crotonylation, succinylation, and malonylation.

The majority of PTMs occurs at the N- and C-terminal tail domains that protrude from the nucleosome center, but a significant fraction of modification takes place also in the globular domain of the histones, which regulates histone-histone and histone-DNA interactions[67]. A detailed list of known histone PTMs is shown in the Table 1.

Recently, the establishment and improvement of techniques like mass spectrometry (MS) and chromatin immunoprecipitation (ChIP) assay coupled to high throughput DNA sequencing technology (ChIP-seq) permitted to intensively investigate the nature and the patterns of deposition of several histone PTMs in the genome of different organisms, and

helped to unravel and dissect their functions, allowing specific histone modifications with gene transcription status or to specific genomic features, including promoters, transcribed regions, enhancers and insulators. The combinatorial nature of different histone PTMs mediates specific function on chromatin. Indeed, through the acquisition or loss of specific chemical properties, PTMs can influence nucleosome architecture. Moreover, particular histone PTMs combinations can drive the selective recruitment or exclusion of effector proteins. The histone PTMs can function both sinergically and antagonistically, and can be deposited asymmetrically within the nucleosome adding further layers of complexity and plasticity of their functions[68].

4.4.1 DNA modifications

Additionally to DNA sequence, an in cooperation with histone PTMs, DNA modifications drive and regulate the association and the downstream functions of factors that bind DNA without changing its sequence. Eukaryotic DNA can be methylated at 5 position of cytosine (5mC) occurring on cytosines that precede a guanine nucleotide; DNA methylation is heritable and has been correlated with gene silencing either through direct impairment of interaction between DNA and transcription factors[69], or acting as docking sites for protein bearing methyl binding domain (MBD) such as chromatin remodelers and transcriptional co-repressor complexes[70]. However, regions with high density of CpG di-nucleotides, called CpG islands (CpGi), are present in the genome and are generally maintained free of DNA methylation: about 60-70% of promoter regions contains a CpGi, whose methylation would result in stable transcriptional repression[71]; this is highlighted in some types of human cancer, where specific promoters of tumor suppressor genes get hyper-methylated at CpGi leading to their silencing. Cytosine methylation is catalyzed by a family of enzymes called *de novo* methyltransferases (DNMTs) that transfer a methyl group from *S*-adenyl methionine (SAM) to the fifth carbon of a cytosine residue to form 5mC: specific DNMTs introduce cytosine methylation at previously unmethylated CpG sites, whereas others are responsible

to copy the DNA methylation pattern from the parental DNA strand onto the newly synthesized daughter strand maintaining such modification[72]. It was recently demonstrated that the ten-eleven translocation (TET) family of enzymes catalyze iterative 5mc oxidation to 5hmC[73], 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC)[74]. These two cytosine modifications are substrate for thymidine DNA glycosilase (TDG)-mediated base excision repair, thus suggesting an active role for TET proteins in the DNA demethylation process independently from replication process[75]. The establishment, maintenance and removal of DNA methylation have been shown to be influenced by the recognition of specific patterns of histone PTMs[70]. Overall, histone and DNA modifications function in a cooperative manner to rapidly change the chromatin organization and facilitate the recruitment of effector proteins throughout the genome.

Histone modification	Modified aminoacid	Modified residues
Phosphorilation	**Serine, thyrosine, threonine**	H1S171, H1S172, H1S17, H1S186, H1S188, H1S26, H1T10, H1T137, H1T145 H1T153, H1T154, H1T17,H1T30, H2AS137 H2AS139, H2AS1, H2AT120ph, H2AY142, H2BS14, H3S10, H3S28, H3S31, H3S6,H3T11, H3T3, H3T45, H3T6, H3Y41, H4S1
Methylation	**Lysine, arginine**	H1K186, H1K25, H2BK5, H3K27,H3K36, H3K4, H3K79, H3K9, H4K20; H2AR3, H3R17, H3R26, H3R2, H3R8, H4R3,
Acetylation	**Lysine**	H1K25, H2AK5, H2AK9, H2BK120, H2BK12, H2BK15, H2BK16, H2BK20, H2BK46, H2BK5, H3K14,H3K18, H3K23, H3K27, H3K36, H3K4, H3K56, H3K9, H4K12, H4K16, H4K5, H4K8, H4K91
O-linked glycosilation	**Serine, threonine**	H3S10, H4S37, H2AT101, H2B36, H2BS91, H2BS112, H2BS123
ADP-ribosylation	**Arginine, glutamate**	H1E2, H1R33, H2AK13, H2BK30, H3K27a H3K37, H4K16
Deimination (citrullination)	**Arginine**	H2AR3, H3R17, H3R26, H3R2, H3R8, H4R3
Isomerization	**Proline**	H3P30, H3P38
Ubiquitylation	**Lysine**	H2AK119 H2AK121, H2BK120, H4K91
SUMOylation	**Lysine**	H2AK126, H2B6,7
Biotinilation	**Lysine**	H2AK125, H2AK127, H2AK129, H2AK13,H 2AK9, H3K18, H3K9, H4K12, H4K8
Clipping	**Lysine**	H3K20

Table 1 Summary of histone PTMs.

4.4.2 Histone acetylation

Since the discovery of the basic principles of chromatin organization, the question of how such extensive packaging can be compatible with reactions that involve reading the DNA has stimulated extensive research. It soon became evident that many features of chromatin structure could be explained by interactions between nucleosomal histones and DNA, neighboring nucleosomes and non-histone proteins. Most of these interactions involve the N-terminal tails of the core histones, which protrude from the nucleosome center, and are modified in several ways. Of these PTM modifications, histone acetylation was the first described[76]. Such modification causes the neutralization of the positive charge of lysines, weakening the interactions between histone and DNA and thus leading to an enhanced accessibility of DNA to the transcription machinery. Indeed, acetylated histones are found to be enriched at actively transcribed regions. Interestingly, several studies have demonstrated that is the charge neutralization, rather than the acetylation of specific lysines, to influence the transcriptional status of modified gene loci[77]. It has been showed that histone acetylation is also important in the DNA replication process and is required for replication origin firing[78]. Lysine acetylation is the enzymatic product of histone acetyltransferases (HATs). These enzymes show low substrate specificity and can acetylate both cytoplasmic free and nucleosomal histones. HATs are present at sites of active transcription to facilitate polymerase transit exerting their action weakening DNA–histones interactions[79]. This modification can be removed by the action of histone deacetylase enzymes (HDACs), which restore the positive charge of lysines leading to transcriptional repression. A further layer of regulation is constituted by the fact that acetylated lysines can be recognized by specific domains defined as bromo-domains that are present in several proteins with chromatin remodeling activity, thus influencing the chromatin structure[80]. Among the acetylated lysine residues of histone H3, the lysine 27 (H3K27ac) plays a particularly interesting role. High throughput techniques revealed that H3K27ac was enriched at promoter regions of transcriptionally active genes[81], including Polycomb target sites. Such modification, which

is catalyzed by both p300 and CREB binding protein (CBP), have been shown to prevent H3K27me3 at Polycomb target genes, revealing dynamic and complementary temporal deposition profiles during embryogenesis of H3K27ac and H3K27me3[82]. H3K27ac has also important function in the regulation of enhancer element. Indeed it discriminates active (H3K27ac positive) from poised (H3K27ac negative) enhancers on the basis of expression of proximal genes[83].

4.4.3 Histone methylation

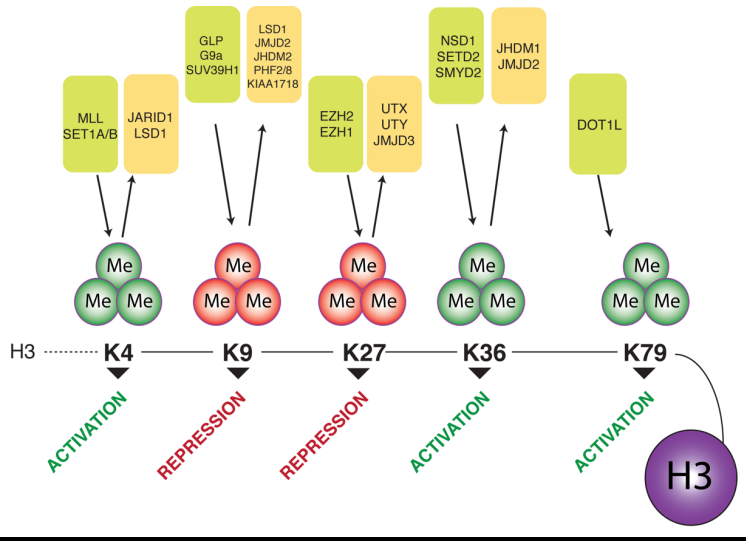

Fig.5 Histone H3 methylation and its related enzymes. In the image the depicted lysine residues are known methylation sites on histone H3. In red are represented the methylation states involved in transcriptional repression, while in green methylation states correlate with transcriptional activation and elongation. The enzymes in light green can modify the lysine by depositing the mark on it, while the enzymes represented in orange remove such marks.

Methylation of histones can occur on both arginine and lysine residues and, contrarily to acetylation, it does not lead to charge changes. Such modification can be catalyzed in different manners, thus providing several combinations and functional outcomes: arginines can be mono- or symmetrically and asymmetrically dimethylated[84] differently from lysines that instead can be mono, di, or trimethylated. It has been demonstrated that the effects on transcriptional regulation and nucleosome dynamics seem to be exerted almost completely by lysines methylation rather than methylated arginines[77]. Methylation of histones can either increase or decrease transcription of genes, depending on which amino acids in the histones are methylated, and how many methyl groups are attached. This modification is indeed associated to transcriptional repression if deposited on H3K9 or H3K27, while it determines transcriptional activation when involves H3K36 or H3K4. The enzymes responsible for these PTMs are the histone lysine methyltransferases (HKMTs).

Among the most relevant lysine methylations H3K4me1/me3 and H3K27me play a particularly interesting function in regulating gene transcription. The COMPASS complexes, which are the mammalian orthologues of trithorax proteins in Drosophila,

regulate all H3K4 methylations and it exerts anti-repressor functions antagonizing the repressive activity of Polycomb complexes[85]. *In vitro* experiments demonstrated that H3K4me3 deposition impairs the ability of the Polycomb Repressive complex 2 (PRC2) complex to deposit H3K27me3[86]. H3K4me3 is enriched at many promoters of eukaryotic genes, in particular it strongly marks the actively transcribed ones[87], thus positively regulating transcription process through the recruitment of nucleosome remodeling enzymes and histone acetyltransferases[88,89]. Moreover, H3K4me3-marked promoters are also enriched for RNA polymerase II[90]. In mESC, a subset of promoters bearing high CpG content is both marked by H3K4me3 and H3K27me3. These promoters have been termed "bivalent chromatin domains" showing features of both active and repressive chromatin. The expression of such genes is low and is poised for rapid change in gene expression during development[91]. These genes are mostly repressed in pluripotent cells, but can be rapidly induced or stably inactivated, depending on the developmental stages[91]. Monomethylated H3K4 was identified as marker of enhancer regulatory elements in human cells in association with other PTMs[92]. According to the literature, the signature of enhancers' PTMs might indicate general genome accessibility or chromatin dynamics at these sites.

Another histone PTM correlating with active gene transcription is H3K36me3, a modification that is found enriched at gene bodies. Its deposition by the KMT Setd2 is mediated by its interaction with elongating RNAPII[93,94]. This interaction is conserved in different organisms from yeast to human[95,96]and it is regulated by the phosphorylation of Serine 2 (pS2) in the RNAPII C-terminal Domain (CTD). The role of H3K36me3 in the transcriptional elongation is still not fully understood. In yeast, H3K36me3 recruits the Rpd3 deacetylase complex to deacetylate newly incorporated histones[97-99]. Alternatively, H3K36me3 could regulate pre-mRNA splicing since it was shown to be preferentially accumulated on exons of actively transcribed genes and to associate with spliceosome components[100,101].

A histone modification correlated instead with transcriptional repression in euchromatin is H3K9 mono- or dimethylated, often found along with H3K27 mono- and dimethylation. In addiction, focal enrichments of H3K9me3 are a characteristic marker of constitutive heterochromatin, composed by regions that totally lack H3K4me but in which the monomethylated form of the H3K27 is present. This fact suggests that the selective combination of distinct H3-K9 and H3-K27 methylation states may be able to index different chromatin regions[102].

4.5 Epigenetic tools: the Writers, the Readers and the Erasers

Readers, writers and erasers are enzymes responsible for the regulation of gene expression patterns and have a role in determining the cellular identity without changing the genetic information encoded in the nucleotide sequence. Such regulation is dynamic, reversible, but when is deregulated also contribute to several human disease outcomes.

4.5.1 Writers

DNA and histones proteins can be modified through the addition of different chemical groups through the action of several specific enzymes, which govern the gene expression pattern of each specific cell by regulating transcription process.

The most well characterized epigenetic writers include DNA methyltransferases, histone lysine methyltransferases, protein arginine methyltransferases and histone acetyltransferases.

DNA methyltransferases (DNMTs) are responsible for the transfer of a methyl group from the universal methyl donor, *S*-adenosyl-L-methionine (SAM), to the 5-position of cytosine residues in DNA, and are essential for mammalian development[103]

Among mammals, this class of proteins is conserved and is composed by four members, including DNMT1, DNMT3A, DNMT3B and DNMT3L. DNMT3L, differently from the other DNMTs, does not possess any inherent enzymatic activity DNMT3L but it is able to

physically associate with the active *de novo* DNA methyltransferases, DNMT3A and DNMT3B, stimulating their catalytic activities in a cell culture system[104].

DNMT1 is the major enzyme involved in maintenance of the pattern of DNA methylation after DNA replication. It has been demonstrated that this protein is located at the replication fork, where it could directly modify nascent DNA immediately after replication. Moreover DNMT1 prefers hemimethylated DNA to unmethylated DNA as a substrate up to 40-fold *in vitro*[105]. *In vivo*, it is supposed to be responsible for copying the existing methylation pattern of the DNA and is therefore defined as maintenance methyltransferase. Structurally, it has an extended N-terminal regulatory domain sequence forming a platform for the binding of proteins involved in chromatin condensation and gene regulation, a replication foci targeting domain, and a C-terminal catalytic domain[106].

DNMT3A and DNMT3B enzymes have been shown to be essential for the establishment of *de novo* DNA methylation during gametogenesis and early embryogenesis[107].

Indeed, such enzymes are highly expressed in developing embryos while their levels are much lower in most adult tissues[108]. Although these two enzymes are generally considered *de novo* DNA methyltrasferases, they are required together with DNMT1 for the accurate maintenance of the methylation at particular repeated regions in mESC as well as at single copy genes[109]. Interestingly, DNMT3A and DNMT3B were not able to *de novo* methylate DNA efficiently in the absence of DNMT1[109], which is in agreement with other data showing that DNMT1 is also involved in *de novo* methylation[110]. Overall, although DNMT1 seems to be more important for methylation maintenance, whereas DNMT3A and DNMT3B are generally involved in *de novo* methylation, both cooperate to establish and maintain the methylation patterns of the cell.

The enzymes responsible for the histone methylation are the histone methyltransferases (HMTs), which acts using S-adenylmethionine to add methyl groups to their specific substrate. Such enzymes are divided in two classes, the histone lysines methyltransferases (HKMT) and the histone arginine methyltransferases (HRMT).

The HKMT can be divided in two different subgroups based on the features of their catalytic domain. The first class endows a SET domain, a highly evolutionary conserved sequence motif of 130 amino acid initially identified in the position-effect-variegation (PEV) suppressor gene *Su(Var)3-9*, in the Polycomb group (Pc-G) gene *Enhancer of Zeste* (E[z]) and in the activating Thritorax group (Trx-G) gene *Thritorax*[111]. Such domain acts transferring the methyl group from S-Adenosyl Methionine (SAM) to the ε-amino group of the substrate lysine. The second subgroup consists in one single protein, Dot1L, which does not possess a SET domain[112]. Overall HKMTs class of enzymes retain high substrate specificity and some of them are even specific for a given methylation state. In particular, it has been shown that an aromatic residue within the SET catalytic pocket domain (a tyrosine or a phenylalanine) is crucial to control the state of methylation[113].

Regarding the arginine methyltransferases, PRMTs are ubiquitously expressed and have been found in many organisms; supporting the fact that arginine methylation is evolutionarily conserved. This family of AdoMet-dependent methyltransferases harbors a set of four conserved signature amino acid sequence motifs (I, post-I, II, and III), and a THW loop. Motifs I, post-I, and the THW loop form part of the AdoMet-binding pocket[114].

To date, eleven mammalian PRMTs have been identified. These methyltrasferases are classified into two subgroups, type I and type II, on the basis of the modification introduced. Although both the subgroups are able to catalyze the monomethylation, they differ in the final type of arginine modification. Type I enzymes, constituted by PRMT1, PRMT3, PRMT4, PRMT6 and PRMT8 are able to catalyze the formation of the asymmetrical dimethylarginine (ADMA), while the Type II enzymes, that include PRMT5, PRMT7 and PRMT9, are able to transfer methyl groups to the ω-nitrogen of the arginine residue, resulting in symmetrical dimethylarginine (SDMA)[115,116]. Finally, PRMT2, PRMT10 and PRMT11 do to possess enzymatic activity[117].

Histone acetyltransferase (HATs) enzymes catalyze the transfer of an acetyl group from acetyl-CoA to ε-amino group of conserved lysine aminoacids on histones[118]. Such enzymes

directly link chromatin modification to gene activation, resulting in a dispersed structure of chromatin, which becomes accessible by transcriptional factors. HATs are enzymes known to acetylate also a variety of non-histone substrates, thus the HATs are generally identified as lysine acetyltransferases[79]. Based on their cellular localization, HATs can be classified into two subfamilies: type A and type B HATs. The type A HATs are commonly found in the nucleus and likely catalyze the processes related to transcription[119]. This subfamily can be further divided into five subgroups: the GNAT family members, including PCAF, Gnc5 and ELP3, The CBP/p300 family, the MYST family composed by Tip60, MOZ, MORF, HBO1 and HMOF[120], the transcriptional factor related HAT family comprising TAF1 and TIFIIIC90, and finally the family of steroid receptor co-activators, such as p600, SRC1, CLOCK and AIB1/ACTR/SCR3 etc.[121,122].

Type B HATs are instead localized in the cytoplasm and it has been demonstrated that their function is to acetylate the newly synthesized histones. For example, HAT1 is one of type B HAT members and functions in DNA repair and histone deposition[123].

4.5.2 Readers

Over the past 10 years, it has been discovered that multiple families of conserved domains are able to recognize modified histones. These proteins can achieve this function using specialized domains that allow surveying the epigenetic landscape and docking at specific regions within the genome. The binding modules within these proteins, termed chromatin readers, recognize and bind different covalent modifications of the nucleosome and assemble functional complexes onto specific loci to facilitate DNA-templated processes.

Specific residues within the binding pocket of the reader domain confer an affinity for specific modification states, whereas residues outside the binding pocket contribute in the determination of the histone-sequence specificity. These two combinatorial features ensure that proteins with similar binding domains dock at different modified residues, or at the same amino acid displaying a different modification state[124,125].

Histone methylation readers constitute a heterogeneous class of proteins, each harboring specialized domains that mediate the recognition of these signals. Two residues, lysine and arginine can be methylated, and each of them has three possible methylation states.

Readers of methylated lysines are the most deeply studied group and include several families including ADD (ATRX-DNMT3-DNMT3L), ankyrin, bromo-adjacent homology (BAH), chromobarrel, chromodomain, double chromodomain (DCD), MBT (malignant brain tumor), PHD (plant homeodomain), PWWP (Pro-Trp-Trp-Pro), tandem Tudor domain (TTD), Tudor, WD40 and the zinc finger CW (zf-CW).

Such methyl-lysine specific readers are able to bind this PTM through an aromatic cage, and the specificity for a particular methylated lysine is imparted by interaction with surrounding residues. Some histone readers are highly specific; others are selective for only a certain methylation state or otherwise bind in a promiscuous fashion.

The structurally related chromodomain, chromobarrel, MBT, PWWP, Tudor and TTD modules possess a characteristic β-barrel topology and comprise the Royal superfamily. One example of methyl-lysine readers is represented by chromodomains, which usually prefer tri-methylated lysine, although some have been shown to bind also di-methylated species. They are predominantly involved in chromatin remodeling and gene transcription, in particular transcriptional repression. The chromodomain is found in several nuclear proteins, including heterochromatin protein 1 (HP1), where it recognizes histone tails with specifically methylated lysines[126]. Chromodomain can be contained also in a small number of other proteins that are involved in chromatin remodeling and gene transcriptional silencing.

Histone acetylation readers are proteins involved in the recognition of acetylated lysines; this family of proteins includes bromodomain (BRD), double PHD finger and Yeats domains. The Bromodomain and Extra-terminal domain (BET) is a proteins family including BRD2, BRD3, BRD4 and BRDT, which are encoded by paralogous genes. These proteins

acts as regulators of gene transcription through epigenetic interactions between bromodomains and acetylated histones during differentiation and proliferation processes[127]. PHD finger domains are structurally conserved modules involved in several biological roles, controlling gene expression through recruitment of complexes of chromatin regulators or transcription factors[128].

4.5.3 Erasers

The main feature of epigenetic modifications is their reversibility. This property is sustained by the removal of epigenetic marks mediated by a group of enzymes known as erasers. Depending on the stimuli sensed by the cells, erasers can counteract the activity of the writers by removing the tags and this results in gene expression modulation.

The most characterized enzymes involved in removing epigenetic marks are histone demethylases and deacetylases whereas DNA methylation is removed by DNA demethylases.

TET proteins are large multi-domain enzymes. This family of proteins contain a core catalytic region in the C-terminus that is composed by three functional modules: a conserved double-stranded β-helix (DSBH) domain, a cysteine-rich domain, and binding sites for the cofactors Fe(II) and 2- oxoglutarate (2-OG) that together form. In addition to their catalytic domain, TET1 and TET3 have an N-terminal CXXC zinc finger domain that can bind DNA[129,130].

The deposition and removal of DNA methylation requires the action of several proteins. During S-phase DNMT1, the maintenance DNA methyltransferase, copies the DNA methylation pattern of the template strand onto the newly synthesized strand, thus re-establishing the symmetry on hemi-methylated DNA. Once methylated, the cytosine potentially can be substrate of a stepwise TET-mediated oxidation process forming 5hmC, 5fC, and 5caC. It has been proposed that the conversion of 5mC and its oxidized derivatives back to the unmodified state can occur by either “passive” or “active” demethylation.

"Passive" DNA demethylation refers to the failure to maintain DNA methylation patterns across cell divisions resulting in replication-dependent dilution of 5mC, whereas "active" DNA demethylation consists in an enzymatic process in which 5mC bases are replaced by unmodified cytosines in a replication-independent fashion[131].

Histone methylation can occur both on arginine and lysine residues. Arginine residue demethylation occurs via peptidyl arginine deiminase 4 (PADI4), which opposes methylation by converting arginine to citrulline[132]. However such modification is not considered a proper demethylation, since conversion from arginine to citrulline results in loss of a methyl group without leaving a free arginine. Lysine residues can be mono, di- or tri-methylated, and the methyl groups can be removed by histone demethylases. Two distinct classes specifically mediate histone demethylation: the amino oxidase homolog lysine demthylase 1 (KDM1) family and the Jumonji C (JmjC) domain-containing histone demethylases.

The first group of proteins includes two members, the histone lysine demethylase 1A (LSD1/KDM1A) and LSD2/ KDM1B.

LSD1 is a highly conserved flavin-containing amino oxidase homolog that is able to remove in a specific manner the mono- and di-methylated lysine at lysine 4 or lysine 9 of H3 via oxidation resulting in the formation of formaldehyde[133]. This protein catalyzes the demethylation at H3K4me1 and H3K4me2 through the interaction between its tower domain and CoREST, leading to transcriptional inactivation[134]. However, it reported that when it is complexed with androgen receptor, LSD1 demethylates H3K9me1 and H3K9me2, resulting in transcriptional activation[135]. This enzyme is unable to demethylate tri-methylated lysines[135]. Non-histone proteins, such as p53, are also substrates for LSD1.

Like LSD1, LSD2 is another FAD-dependent amino oxidase homolog that specifically targets H3K4me1 and H3K4me2[136].

The second class of the histone demethylases is the Jumonji C (JmJC) domain-containing demethylases, which comprises nearly 20 lysine specific demethylases, capable of removal

of trimethylations[137]. Such JmjC KDMs comprise different subfamilies including KDM2, KDM3, KDM4, KDM5, KDM6 and others.

The KDM2/FBXL subfamily consists of 2 members: KDM2A, which specifically demethylate H3K36me2, and KDM2B that demethylates both H3K4me3 and H3K36me2 [138]. KDM3A, KDM3B and JMJD1C make up the KDM3/JMJD1C subfamily. KDM3A and KDM3B are specific for removing H3K9me1 and H3K9me2, whereas JMJD1C has no reported histone demethylase activity. Specific demethylation of H3K9me2, H3K9me3, H3K36me2 and H3K36me3 is catalyzed by the KDM4 subfamily. KDM5 subfamily has catalytic activity at H3K4me2 and H3K4me3, whereas KDM6A/UTX and KDM6B/JMJD3 act on H3K27me3, a mark of gene repression.

Histone deacetylases comprise 18 human grouped into four classes, which can be further subdivided into two categories: classical (classes I, II and IV) and sirtuins (class III). Classical HDACs depends on zinc ions (Zn^{2+}), whereas the sirtuins utilize nicotinamide adenine dinucleotide (NAD^{+}), a phosphate linked dinucleotide coenzyme.

Class I HDACs are located both in the nucleus and in the cytoplasm or specialized cellular organelles. This class comprise HDAC1, HDAC2 and HDAC3, HDAC8 negatively regulate transcription by being recruited to DNA as a corepressor[139,140]. Like HDAC1 and HDAC2, represses transcription, binds to and is recruited by transcription factors, and is expressed in many different cell types.

Class II HDACs show cytoplasmic localization, suggesting a major cytoplasmic functional role.

Class IV HDAC is composed by the sole member HDAC11, which regulates the protein stability of DNA replication factor CDT1 and the expression of interleukin 10[141].

Class III HDACs include five human sirtuins (SIRT1, SIRT2, SIRT3, SIRT4, SIRT5), which are characterized by two enzymatic activities: mono-ADP-ribosyltransferase and histone deacetylase. SIRT5 further possesses an additional protein lysine desuccinylase and demalonylase activity *in vitro*[142]. Regarding their localizations, SIRT1 and SIRT2 are both

in the nucleus and cytoplasm, SIRT3 is found in the nucleus and mitochondria, SIRT4 and SIRT5 are specifically in the mitochondria, SIRT6 is only nuclear, and SIRT7 in exclusive for the nucleolus. Like the Class I, II, and IV HDACs, sirtuins shows non-histone substrates.

4.6 Epigenetic drugs

Epigenetic modifications regulate gene expression, mediating activation or silencing of genes transcription. Deregulation of such processes has been linked to several diseases, such as cardiovascular diseases, neurological disorders and cancer development.

It has been reported that several epigenetic players results expressed in an aberrant fashion or mutated in several different neoplastic diseases, thus indicating a crucial role of epigenetics mechanisms deregulation in cancer onset and suggesting its valuable therapeutic potential. More in details the reversible nature of the epigenetic changes that occur in cancer has led to the possibility of epigenetic therapy as a treatment option, aiming to restore the normal epigenomic landscape.

Currently, several classes of epigenetic drugs have been developed. This includes DNA methylation inhibiting drugs, which are nucleoside-like compounds acting through an irreversible trapping of DNMT1[143], or in alternative antisense oligonucleotides through the binding with the 3' untranslated region of DNMT1 prevents its transcription[144]. Bromodomain inhibitors act effectively in the downregulation of c-Myc, often related to cancer, by disrupting the interaction between the BET proteins and acetylated histones[145,146]. Other inhibitors are histone methyltransferase inhibitors, such as Enhancer of Zeste homologue 2 (EZH2) inhibitors, targeting the catalytic subunit of Polycomb Repressive Complex 2 (PRC2). Such protein results particularly attractive from a clinical point of view due to its involvement in sustaining several tumor types, both solid and liquid, through point mutations leading to enhanced activity and overexpression[147,148].

Another promising example of epigenetic drug is represented by HDAC inhibitors, which act modifying the balance between histone deacetylases (HDAC) and histone

acetyltransferases (HAT). HDAC inhibitors exert their action on several processes such as cell cycle arrest, differentiation and cell death, or also reducing angiogenesis or modulating immune response. Mechanisms of anticancer effects of HDAC inhibitors vary according to the tissue in exam. Nowadays HDAC inhibitors resulted promising anti-cancer drugs particularly in combination with other anti-cancer drugs and/or radiotherapy[149].

AIM OF THE PROJECT

Colorectal cancer (CRC) is the second cause of cancer-related death in Western countries. In the last decades the implementation of the standard treatment for advanced CRC has been systematically improved, however life expectancy still continues to be limited for patients affected by metastatic CRC. The driving force of tumorigenesis comes from genetic and epigenetic alterations that deregulate proliferation, differentiation and cell death pathways. Molecularly, colorectal cancers present frequently WNT constitutive activation by APC inactivating mutations or by ß-Catenin stabilizing mutations.

Since inhibition of upstream signaling of such pathway shows side effects due to a high level of cross talk within the cell, an attractive way to target such pathway could be represented by the modulation of the signaling that converge into the nucleus where transcription and chromatin-remodeling factors cooperate to establish new transcription programs.

Moreover, since one of the features of the chromatin modifies is their reversibility, there is a deep interest the development of several inhibitors.

In this context, the aim of this project is to identify and characterize novel epigenetic players essential for the oncogenic activation of the WNT pathway by coupling β-catenin stabilized organoids to unbiased functional screens. Once selected, candidate targets will be first validated and then characterized to provide insights on their role in tumor growth.

MATHERIALS AND METHODS

1. Synthetic lethality screen

1.1 Custom Pooled shRNA lentiviral Library

The pooled library used in the present project is composed by two modules: mEPI1 and mEPI2, each containing 1200 shRNAs, 10 shRNA per target gene. Each shRNA carries an 18 base pairs barcode that allows unambiguous identification of each sequence with high-throughput sequencing. The pooled lentiviral library vector is cloned into pRSI17-U6-(sh)-UbiC-TagGFP2-2A-Puro. This vector allows visualization of transduced cells via GFP expression and retains the possibility to select infected cells with Puromycin resistance.

The library contains some internal controls: the Luciferase as neutral control, and some essential genes as Rpl30, Psma1 and Pcna as positive controls.

1.2 Virus production of shRNA library and titration

To produce shRNA library lentiviral particles, HEK 293T cells were seeded 24 hours prior to transfection at a density of $1.1x10^7$ cells in 150 mm plates. Cells were then transfected with Lipofectamin reagent using 6 μg of plasmid shRNA library, 7 μg of VSV-G and 24 μg of PAX.2 packaging vector per dish. After 24, 48 and 72h the supernatant was harvested and filtered in 0.45 μm filters. Virus was concentrated 100X via ultracentrifuge to achieve the desired multiplicity of infection (MOI), due to the high difficulty of transduction in 3D organoids. Lentiviral titer was determined by transducing NIH-3T3 with serial dilution of virus. The efficiency of infection was checked by FACS analysis via GFP visualization.

Lentiviral titer estimation was obtained using the titer formula:

Viral particles/ml=(N cells at transduction)*(% infected cells) *(virus dilution factor used)(viral stock concentration)

1.3 Extraction of genomic DNA

To extract genomic DNA, cells were resuspended in Buffer P1 supplemented with RNase

A (Qiagen). After the addiction of 0.5% SDS the solution was incubated 5 minutes at room temperature and then homogenized using a syringe. Phenol/Chloroform was successively added and, after vortexing, sample was spinned down for 45 minutes at 20°C. Upper phase was transferred to a new tube and Chloroform was added in a volume 1:1 before spinning for 30 minutes. Na-Acetate 375 mM and Isopropanol in a volume 1:1 were added, and then the sample was spinned down to precipitate genomic DNA. After a wash in ethanol 70% pellet was air-dried and resuspended in water.

1.4 Amplification of shRNA-associated barcodes

Identification of barcodes requires the amplification of the barcode specific portion of the integrated lentiviral construct from the total genomic DNA. The pooled barcodes were amplified with two subsequent rounds of PCR taking advantage of Taq Titanium DNA polymerase (Clontech-Takara). The resulting DNA libraries were then quantified using a high sensitivity DNA Chip on Bioanalyzer instrument (Agilent) and used for cluster generation and sequencing using the HiSeq 2000 platform (Illumina) following the protocol of the manufacturer to quantify each barcode. Among the primers listed below it is present also the primer used to engineer the library to make it compatible with Novaseq 5000 platform.

List of primers used:

Primer 1st PCR	AGTTCAGACGTGTGCTCTTCCGATCTCGGATTCGCACCAGCACGCTA	**Forward1**
	AGTAGCGTGAAGAGCAGAGAA	**Reverse1**
	CAAGCAGAAGACGGCATACGAGATATTGGCGAATTCGCACCAGCACGCTACG	**Forward2**
Primer 2nd PCR	AATGATACGGCGACCACCGAGAGCACCGACAACAACGCAGA	**Reverse 2 Hiseq**
	AATGATACGGCGACCACCGAGATCTACACGAGCACCGACAACAACGCAGA	**Reverse 2 Novaseq**

2. Screen statistical analyses

Each screen replicate was aligned to a barcode reference file using Bowtie v2, ensuring so that the library complexity was represented in the sample. Alignment was executed accepting up to three mismatches within the barcodes. The resulting aligned barcodes were then counted in each screen and its corresponding reference.

The counts of the barcodes were then analyzed taking advantage of three different methods:

2.1 Robust Z-score

The number of reads for each barcode in the different replicates was normalized for the total number of reads:

$$frequency = \frac{n\,reads\;barcode}{reads\;tot}$$

and the base 2 logarithm of the fold change was derived as follow:

$$log_2\left(\frac{barcode\;frequency\;in\;reference}{barcode\;frequency\;in\;screening}\right)$$

The mean of the Log_2FC frequency was calculated among the biological replicates and the robust Z-score was finally obtained. The robust Z-score is the number of standard deviations from the mean and provides explicit information on the strength of each barcode relative to the rest of the sample distribution.

Genes selected to be validated with this method were those with at least 7 barcodes under the median of the total distribution of the barcodes.

2.2 Strictly Standardized Mean Difference (SSMD)

The mean of the barcode counts for the reference and the screen in the different biological replicates was calculated, and after this all the resulted values were grouped in order to have a single value for each gene.

The value consists in the ratio between the difference of the means and the standard deviation of the difference between two populations: the reference and the screen, and is a measure of effect size.

$$SSMD = \frac{(\mu_{ref} - \mu_{scr})}{\sqrt{(\sigma^2_{ref} + (\sigma^2_{scr})}}$$

This method provides a meaningful an interpretable criterion to classify the size of barcode depletion as described: $|ssmd| \geq 5$ for extremely strong, $5 > |ssmd| \geq 3$ for very strong, $3 > |ssmd| \geq 2$ for strong, $2 > |ssmd| \geq 1.645$ for fairly strong, $1.645 > |ssmd| \geq 1.28$ for moderate, $1.28 > |ssmd| \geq 1$ for fairly moderate, $1 > |ssmd| \geq 0.75$ for fairly weak, $0.75 > |ssmd| \geq 0.5$ for weak, $0.5 > |ssmd| \geq 0.25$ for very weak, and $|ssmd| \leq 0.25$ for extremely weak effects[150].

2.3 edgeR

edgeR[151] is a Bioconductor package for the analysis of gene expression with a dedicated pipeline for shRNA screening. A negative binomial model is used to test the abundance of counts in reference and screen and a p-value is returned for each barcode. To obtain a gene-by-gene ranking, the 10 barcodes are summarized according to their effect strength and p-value through two different competitive gene set enrichment analysis tools:

2.3.1 ROAST

The majority of gene set analysis tools is based on permutation to assess if a gene is significantly depleted or enriched, assuming independence of the barcodes across genes. ROAST uses rotation instead, which is a type of Monte Carlo simulated permutation, under an underline multivariate normal model that takes into account the cross-gene correlation structure[152]. A further advantage of this method is that the number of rotations does not depend on sample size, thus providing useful results even for experiments with minimal replication.

2.3.2 Camera

Similarly to ROAST, the Correlation Adjusted Mean Rank gene set test assumes interdependence of the gene expression but infers the variability directly from the data and abandons the normality assumption. The test statistics is calculated adjusting for the covariance structure of the entire set. This method controls the type I error rate correctly regardless of inter-gene correlations, earning excellent power of detection of genuine differences[153].

2.4 Mageck

In this method read counts from the reference and the experimental sample are first median-normalized to adjust for the effect of library sizes and read count distributions. Then, the variance of read counts is estimated and a negative binomial (NB) model is used to test whether the abundance of the barcodes varies significantly between the reference and the experimental sample. Next, barcodes were ranked according to their *P*-values. The threshold utilized to define a hit was *P*-values<0,1. The *P*-values were calculated from the NB model, and to identify depleted barcodes a modified robust ranking aggregation (RRA) algorithm named α-RRA.

More in details, α-RRA algorithm assumes that if a gene has no effect on selection, then the shRNAs targeting it should be uniformly distributed across the ranked list the total shRNAs. Genes are so ranked comparing the skew in rankings to the uniform null model, thus genes whose shRNA rankings are consistently higher than expected are prioritized[154].

3. Ethic statements

Experiments on mice were carried out with the approval of the authorized animal facility meeting the ethical standards of the Italian law (D.L.vo 116/92 and following additions) and the European directive 2010/63.

4.Mice models and treatment

Cre-Lox is a technology of recombination used in this project to create deletions in some specific sites in the DNA. The system is composed by a bacteriophage enzyme, Cre recombinase (causes recombination) that allows the recombination of a pair of short target sequences present in the two strands of the DNA called the LoxP sites (locus of crossover in phage P1)[155].

A method commonly used to control the genetic alteration in specific districts of the body involves the use of tissue-specific promoters to drive Cre expression in certain tissues.

In this project the mice used carry a Cre fused to a modified fragment of the estrogen receptor ($CreER^{T2}$) whose function is to trap Cre outside of the nuclear compartment. In presence of a stimulus given by an estrogen receptor antagonist (Tamoxifen), Cre migrates toward the nucleus where it mediates recombination of the target region of interest. Tamoxifen shows a limited affinity for the estrogen receptor, but is metabolized in the liver by various cytochrome P450 enzymes to more active metabolites as 4-hydroxitamoxifen (4-OHT) or N-desmethyl-4-hydroxytamoxifen[156]. Transcription unit was under the control of the Lgr5 promoter in order to induce the removal of the targeted genes in the intestinal stem cells (ISCs). The strains used carry a Cre-inducible constitutively active form of β-Catenin[157] indicated as Lgr5-eGFP-$CreER^{T2}$-β-$Catenin^{lox(ex3)/lox(ex3)}$. Another model, the Lgr5-eGFP-$CreER^{T2}$-$Apc^{lox(ex14)/lox(ex14),}$ resulting in the deletion of *APC* in the Lgr5 stem cells was also generated. To activate Lgr5-driven Cre recombination, mice were injected with four intra-peritoneal doses of Tamoxifen (75 mg/kg in Corn Oil each day). Mice were sacrificed at day 8 after the first injection.

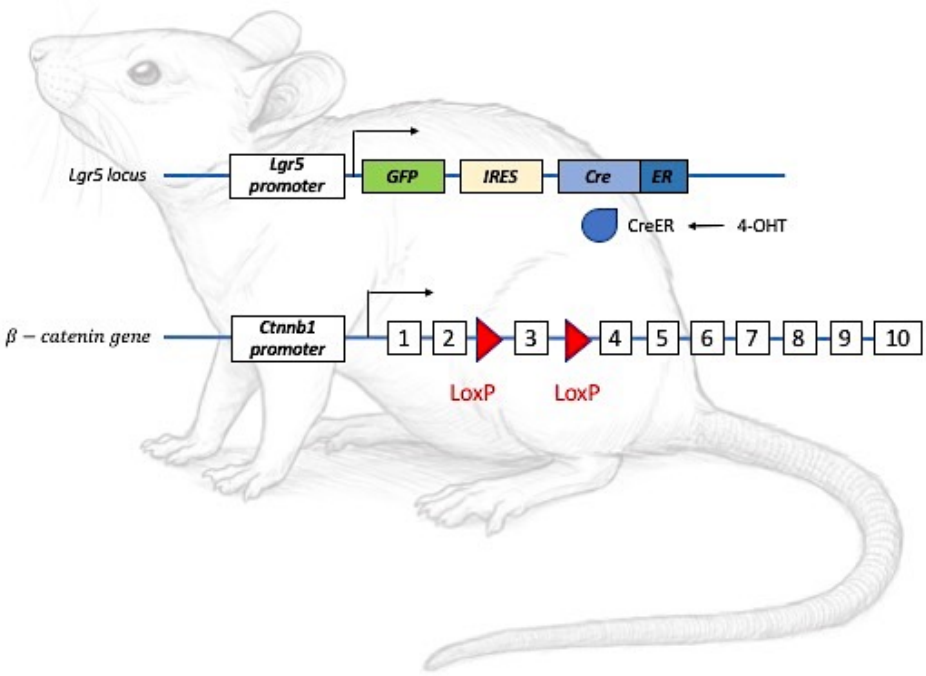

Fig.1 Lgr5GFP-CreERT2 Ctnnb1ex3/ex3 mouse model.

5.Generation of lentiviral pLKO.1 vectors

pLKO.1 lentiviral vectors were used to transduce murine and human intestinal organoids to perform efficient knock down of protein of interest.

The target sequences designed with specific oligonucleotides allowing the insertion and the formation of the hairpin structure were cloned into pLKO.1 backbone previously co-digested by AgeI and EcoRI restriction enzymes by means of Quick Ligase (New England Biolabs).

6. Methods to detect gene expression

6.1 RNA extraction and retro-transcription

To assess the efficiency of the knockdown of the genes selected in the functional screen, RNA was extracted from frozen pellets of mouse intestinal organoids either with Quick-RNA MiniPrep kit (Zymo Research) or with RNeasy Plus Micro Kit (Qiagen) following the instructions provided by the manufacturer. RNA obtained was then retro-transcribed taking advantage of ImProm-II (Promega) kit.

6.2 Quantitative Real Time PCR (RT-qPCR)

The quantitative polymerase chain reaction (qPCR), also called real-time quantitative PCR (RT–qPCR) allows the detection and relative quantification of a specific DNA sequence in a sample. Unspecific fluorescent dye (SYBR green) intercalates with double-strand DNA, specifically amplified through PCR made with specific primers. For a short period of the reaction, DNA amplification is exponential; therefore it can be described by a mathematical function, allowing DNA quantification. This technique can be used both to detect the amount of a DNA sequence (such as target genes in a ChIP experiment), or the abundance of a cDNA derived from an RNA sample. All the RT-qPCRs were carried out using Fast SYBR Green (Applied Biosystem) as dye.

7.Cell Culture and manipulation

7.1 Cell lines utilized

HEK 293T cells

HEK 293T cells were grown in Dulbecco's Modified Eagle Medium (Lonza) with 4,5 g/l glucose, 10% of South American Fetal Bovine Serum (FBS, Euroclone), L-glutamine (2 mM) and 1% penicillin/streptomycin (Gibco), Cells were kept at 5% CO_2 with standard oxygen tension (21% oxygen).

E14 Mouse embryonic stem cells

Mouse embryonic stem cells (mESC) of ES-E14TG2a cell line were used in the present thesis to test the efficiency of the knockdown of all the shRNA used to perform the screen validation.

These cells, are commonly called E14, were grown on 0.1% gelatinized tissue culture dishes in Glasgow Modified Eagle Medium (GMEM) supplemented with 15% FBS, 2 mM glutamine, 100 U/ml penicillin and 0.1 mg/ml streptomycin, 0.1 mM non-essential aminoacids (Gibco), 1 mM Na-Pyruvate (Gibco), 50 µM ß-mercaptoethanol-phosphate-

buffered saline (PBS; Gibco) and Leukemia Inhibitory Factor (homemade).

NIH-3T3 cells

The NIH-3T3 cells were grown in Dulbecco's Modified Eagle Medium (Lonza) with 4,5 g/l glucose, 10% of Calf Serum (Euroclone), L-glutamine (2 mM) and 1% penicillin/streptomycin (Gibco), Cells were kept at 5% CO_2 with standard oxygen tension (21% oxygen).

HCT-116

Human colorectal cancer cell line that expresses a wild-type form of APC and a mutation on β-Catenin. Moreover they carry a mutation in KRAS. They grow in McCoy's 5A, 2mM Glutamine, 10% FBS (South American).

7.2 Virus production

Single shRNA in pLKO.1 vector were produced in HEK 293T using calcium phosphate protocol in a 100 mm dish using 10 μg of lentiviral delivery vector, 3 μg of VSV-G and 6 μg of PAX.2. When used to transduce organoids, virus was concentrated using the procedure previously described.

7.3 Growth curve assay

After puromycin selection, DLD-1 and HCT-116 cell lines infected with the shRNA of interest were plated at a density of 2500 cells per cm^2 in a 6 well multi-well. 48 hours after seeding cells were collected and manually counted once a day.

8.Intestinal stem cell culture and manipulation

8.1 Intestinal single cells isolation

Small intestines were separated from mice sacrificed at 8-12 weeks of age, soaked in PBS and cut in parts. Pieces were then opened longitudinally through scissors and the villi compartment was then removed by scraping. Intestinal crypts were finally isolated by EDTA-based Ca2+/Mg2+ chelation. In particular the pieces were incubated with PBS-EDTA 5 mM for 45 minutes on a rotating wheel at 4°C and then resuspended in 20 ml of cold 1% serum PBS. The preparation was shaken roughly and filtered in a 70 μm strainer to isolate intestinal crypts. The crypts were successively resuspended in a solution composed by Tryple Express (Gibco), DNase I (Roche) 2000 U/ml, Y-27632 (Selleckem) 10 nM and incubated 15 minutes at 37 °C. Tryple Express was inactivated with FBS (South American) and falcon was shacked roughly to separate efficiently the intestinal cells. Cells were finally filtered with a 40 μm cell strainer and live cells were manually counted using Erythrosine B.

8.2 Murine organoids culture

Intestinal organoids were cultured in Matrigel (Corning), either through embedding of single dissociated cells in sitting drop of 40 μl, or seeding of single cells on a bottom of Matrigel 1of 00 μl. Organoids were then cultured with a basal medium containing Advanced Dulbecco Modified Eagle Medium (ADMEM)/F12 (Invitrogen), 2mM Glutamine (Lonza), 10 mM Hepes, 100 U/ml penicillin and 0.1 mg/ml streptomycin, 1.25 mM N-acetyl-cysteine (Sigma-Aldrich), N2 (Invitrogen), B27 (Invitrogen), gentamycin 0,1 mg/ml (Lonza). This medium was supplemented with dedicated growth factors according to the specific type of organoid. Wild-type organoids were cultured with Recombinant Wnt3a (homemade), Recombinant R-spondin1 (homemade), 0.1 μg/ml Noggin, 0.05 μg/ml EGF, while APC KO and β-Catenin stabilized organoids were cultured inly with 0.1 μg/ml Noggin and 0.05 μg/ml EGF.

8.3 Human organoids culture

Human organoids were obtained from the bio-bank of our collaborator Andrea Bertotti in Candiolo Hospital.

They were cultured in Matrigel (Corning), either through embedding of single dissociated cells in sitting drop of 40 µl, or seeding of single cells on a bottom of Matrigel 1of 00 µl. Organoids were then cultured with the basal medium described in paragraph 8.2. The medium was supplemented with human 0.02 µg/ml EGF (Peprotech).

8.4 Transduction of single intestinal cells

Single intestinal cells were counted in order to seed 100.000 cells per each well in a 48 well plate. The proper amount of cells was resuspended in basal organoids medium added with growth factors (described in previous paragraph) supplemented with 10 mM Y27632 and 10 mM Nicotinammide (Sigma-Aldrich) with the proper amount of virus particle and plated onto a bottom of Matrigel (Corning) as described in the protocol of Onuma et al.[158]. The morning after floating dead cells were removed with the change of the medium. Puromycin selection (2 µg/ml) was carried out for 72 hours, and the evaluation of knockdown efficiency was then assessed by Real-Time qPCR.

8.5 Organoids growth curve assay

To monitor the phenotypic effect given by knockdown organoids were disrupted after puromycin selection and seeded in a proper number. Every day three pictures per well were collected and organoids area was manually measured taking advantage of ImageJ tool. This procedure was used also to investigate an effect in size after drug treatments.

8.6 Treatment with MS023

Murine organoids were treated with MS023 (Selleckem) for 72 hours 30 nM, 300 nM and 3 µM. DMSO was used as a vehicle control. Efficiency of the treatment was then evaluated

by Western Blot analysis through the evaluation of the presence of H4R3me2a modification.

8.7 Treatment with valproic acid

Human organoids were dissociated into single cells and 20.000 cells/well were embedded in Matrigel and seeded in 24 well plates.

The day after, valproic acid was added to newly formed organoids in three increasing concentrations;1 mM, 5 mM and 10 mM. To test the efficacy of the treatment, organoids were collected 24h after the addiction of the drug to the medium and Western Blot analysis was performed to investigate the acetylations levels. To evaluate the growth rate of treated organoids, pictures were taken for each condition at 24h, 48h and 72h.

8.8 Treatment with JQ1

Human organoids were dissociated into single cells and 20.000 cells/well were embedded in Matrigel and seeded in 24 well plates.

The day after, JQ1 was added to newly formed organoids in three increasing concentrations: 50 nM, 200 nM, 500 nM

To assess the efficacy of the treatment, organoids were collected 24h after the addiction of the drug to the medium and RT-qPCR analysis was performed. To evaluate the growth rate of treated organoids, pictures were taken for each condition at 24h, 48h and 72h.

9. Protein detection

9.1 Immunoblot analysis

Protein extracts were obtained after lysis in high salt conditions for 20 minutes on ice (25 mM Tris HCl pH 7.6, 300 mM NaCl, 10% glycerol, 0,25% NP40), and sonicated for 6 cycles 30 sec ON+30 sec OFF using BioRuptor® Ultrasonicator (Diagenode). Quantification was done taking advantage of Bradford Proteins Assay (Bio-Rad), at OD_{595}. Laemmli Sample Buffer (312,5 mM Tris-Hcl pH 6,8, 12,5% Glycerol, 10% SDS, 0,05% Bromophenol Blue,

5% β-mercaptoethanol) was then added to the samples in the proper quantity, before the loading into the acrylamide gel. Gel-separated proteins were then transferred to a Protran nitrocellulose membrane (Whatman) at 100 V for one hour and a half at 4 °C. The membrane was then blocked in a solution of TRIS-buffered saline (TBS: 20mM TRIS/HCl, pH 7.4, 137 mM NaCl, 2.7 mM KCl) plus 0.1% Tween (TBS-T) containing 5% low-fat dry milk or BSA. After the incubation with primary antibody the membrane was rinsed in TBS-T and then incubated with a secondary HRP (horseradish peroxidase)-conjugated antibody (BioRad) diluted in the same solution for one hour at room temperature. After several washes in TBS-T, the bound secondary antibody was revealed by ECL method (Bio-Rad).

10. RNA-sequencing

Murine organoids and HCT-116 cells were counted and immediately processed taking advantage of Smart-seq2 (Switching Mechanism at the 5' end of the RNA Transcript) protocol[159]. 2500 cells were counted and lysed in 2 μL of Lysis buffer (Tryton X-100 0,2%, RNAse inhibitor 4 U/μL). Next, 1 μL of 10 mM oligo-dT30Vn and 1 μL of 10 mM DNTPs were added to the samples that were subsequently incubated for 3 minutes at 72°C, to allow RNA unfolding and oligo annealing with the polyA end of messenger-RNA. After this step, a mix containing SuperScript III reverse transcriptase enzyme 100U/μL (Invitrogen) and its buffer, RNase Inhibitor 10U/μL, Betaine 1M, DTT 5 mM, MgCl2 6 mM and TSO 1 μM (Template Switching Oligo) was added to each sample and RT was performed following the published protocol. After the end of the RT, samples were Pre-Amplified using KAPA taq HotStart enzyme working in High Fidelity buffer. To perform the Pre-amplification, 15 μL of the preamplification mix (KAPA taq HotStart, KAPA HiFi Buffer, MgCl2 0,5 mM, dNTPs 0,3 mM, ISPCR primer 0,1 μM) were directly added to the retro-transcribed samples and were incubated in the thermal cycler. Reactions were purified with AMPure XP beads

(Agencourt AMPure XP, Beckman Coulter). Purified cDNA quality was assessed using Bioanalyzer instrument (Invitrogen).

After the evaluation of the quality of the samples, tagmentation reactions occurred. 2 ng of eache sample were tagmented in the process of library preparation. More in details, the protocol enables DNA cutting and the ligation of adapters at the same time. cDNA was tagmented with 100 ng of homemade Tn5 enzyme pre-annealed with A/B-MEDS (Mosaic End Double-Stranded) oligonucleotides in working buffer containing TAPS-NaOH pH 8.5 5mM , PEG 8000 8%for 5 minutes at 55°C. Next, the samples were purified with AMPure beads and finally PCR mix (KAPA HiFi buffer, KAPA Taq HotStart, dNTPs 0,3mM, Ad1 Primer (i5 primer for Illumina sequencing Without Barcode) 2 μM and Ad2.X Primer 2 μM (i7 primer for Illumina sequencing containing different Barcodes) were directly added to the eluted tagmented DNA and were amplified through PCR reaction. Amplified library undergo an additional round of purification with AMPure. The resulting libraries were quantified by Qubit and the quality was checked through Bioanalyzer. Libraries that showed enrichment of DNA fragments that span around 200 and 800 bp length, were then sequenced using NovaSeq 5000.

10.1 RNA-sequencing analyses

Reads were aligned to the mouse reference genome mm10 and hg38 for murine organoids and for HCT-116, respectively, using STAR[160]. PCR duplicates were removed using Samblaster[161].

Gene counts were calculated using featureCounts using RefSeq mm10 annotation in unstranded mode[162]. Two biological replicates were analyzed for the control condition and each of the two shRNA mutants. To perform the differential expression analysis the control was compared with the common regulated genes in both the two different shRNAs conditions.

Differential expression analyses were carried out using R package DESeq2 v1.20 [163] using default parameters. Genes with an absolute log2 (fold change) of 1 and false discovery rate of <0.05 were considered ad differentially expressed and used for analyses. Gene enrichment analyses on down- and up- regulated genes were performed using Enrichr online tool[164].

11.ChIP sequencing analyses

Sequenced reads were obtained from ENCODE[165] data repository and aligned to human reference genome hg38 using Bowtie v1.2.2[166] with default parameters and not allowing multimapping (-m1). PCR duplicates were removed using Samblaster, and peaks were called using MACS2 v2.1.1[167] with p-value<0.00001.

Genomic peaks annotation was performed using R package ChIPpeakAnno v3.15[168], considering the promoter region major or equal to 2.5kb around TSS (Transcription start site) and the distal intergenic region, which are all the regions that are not annotated genes in the whole genome, here considered as putative enhancer . For heatmap representation of ChIP-seq signals, bigwig files, subtracted by input signal, were generated using function *bamCompare* from deepTools 2.0[169]. To normalize for differences in sample library size, a scaling factor for each sample was calculated as (1/total mapped reads) x 1 milion, and applied during bigwig file generation with the parameter *–scaleFactors* from bamCompare.

RESULTS

1.Workflow of the project

To identify novel positive and negative regulators of the major signaling pathways involved in colorectal cancer development, I took advantage of two different approaches using a new powerful tool: intestinal organoids. In this project I performed different analyses on wild-type, APC KO and β-catenin stabilized organoids. Wild-type organoids grows as a 3D small intestinal epithelium in which self-renewing $Lgr5^{+}$ stem cells and their niche-supporting Paneth cells are located in a domain that resembles the crypt, whereas differentiated cells as enterocytes, goblet and enteroendocrine cells move upward to build a villus-like domain that lines the central lumen. Since APC is a negative regulator of the Wnt pathway, organoids APC KO will grow in the absence of R-spondin 1 as cystic spheroids in contrast to wild-type organoids that show budding formation upon differentiation.

In the first approach, I identified from bulk-RNA sequencing analyses several epigenetic players upregulated in organoids that carry a mutation that results in a stabilized form of β-catenin, while in the second I performed a functional shRNA screening on β-catenin stabilized mouse organoids taking advantage of a barcoded epigenetic library designed to target 240 genes, with 10 shRNA per each gene. In this latter approach, statistical analyses were performed to recognize depleted barcodes, thus identifying other epigenetic players potentially crucial for tumorigenesis that have been object of further investigations.

All the selected epigenetic players identified in both approaches were validated independently with dedicated analyses using specific shRNA sequences in which organoids growth was tested. The functions of the validated hits were then further characterized by ChIP-seq and RNA-seq analyses in colorectal cancer cell lines, a more feasible platform for a higher level of investigation on their functions and mechanisms of action.

Moreover, I further strengthen my results using specific inhibitory compounds. Finally, to assess the relevance of these findings in a pathological context, I performed the same shRNA

loss of function and chemical inhibition analyses in human organoids derived from metastatic colorectal cancer patients with additional characterized oncogenic mutations in these signaling pathways, confirming the therapeutic potential of the identified novel mechanisms implicated in CRC development.

2. First approach: Target selection by differential expression

2.1 RNA-seq analysis of intestinal organoids reveals epigenetic players strongly regulated in β-catenin stabilized organoids

To uncover new epigenetic players involved in the regulation of the tumorigenesis process in colorectal cancer, the first step of my project was to perform bulk-RNA sequencing analysis on wild-type and β-catenin stabilized organoids. The results were obtained from two independent biological replicates. The analysis revealed the upregulation of Wnt pathways, confirming the effectiveness of the experiment performed (Fig 1A). Then, I specifically looked for epigenetic players among the regulated genes that showed a padj<0.05 and a $\log_2$foldchange of at least 1.5 by crossing the list of the total regulated genes in the bulk RNA-sequencing with an annotated list of 337 known epigenetic players.

With these parameters, 10 epigenetic players (*Cxxc5*, *Carm1*, *Hdac7*, *Orc1*, *Dnmt3b*, *Cbx6*, *Suv39h2*, *Hmgn3*, *Prmt1*, *Phf21b*) were found upregulated in β-catenin stabilized organoids (Fig1A) and their levels of expression were validated by real time qPCR (Fig 1B).

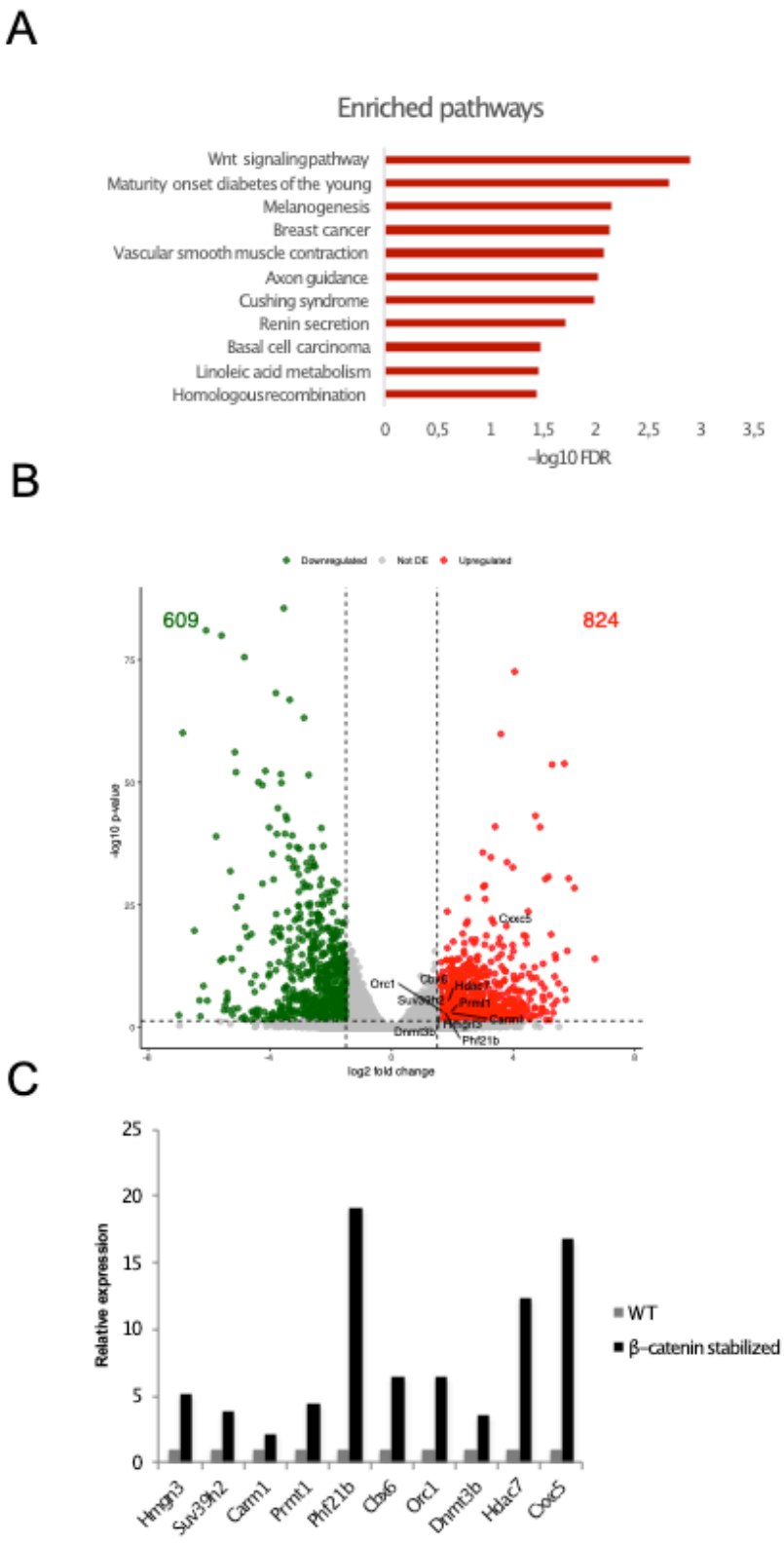

Fig. 1 Bulk RNA-seq on WT and β-catenin stabilized organoids reveals 10 epigenetic players upregulated. A) Enriched pathways detected in the bulk RNA-seq analysis. B) Volcano-plot representing upregulated and downregulated genes. Regulated epigenetic players are indicated. C) RT-qPCR analysis showing the effective upregulation of the regulated epigenetic players in β-catenin stabilized organoids.

2.2 Optimization of intestinal organoids culture manipulation by viral delivery

Since knockdown assays were the crucial technique to study the effects of the loss of the upregulated epigenetic players identified by RNA-sequencing analysis, I first needed to optimize the viral delivery in murine organoids system. The most efficient protocol of infection I set up for my entire project, consists in the incubation of dissociated single cells with the viral particles on a Matrigel bed. In this way organoids are generated from highly proliferative intestinal stem cells (ISC)[158]. Moreover, since all the organoids cells derive and are maintained exclusively by ISCs, this protocol assures the presence of the transgenes of choice in the stem cell compartment achieving their long-term stable expression (Fig. 1A).
In parallel, I chose as vector backbone a miR-E plasmid, a vector that the Zuber's laboratory has recently optimized to achieve efficient knockdown even at single insertions and that derives from the engineering of the standard miR-30 environment used in most RNA-PolII based vectors[170]. This vector also retains the system to visualize transduced cells via GFP expression and bears the puromycin resistance for antibiotic selection of infected cells. By testing the empty miR-E vector in intestinal organoids, I first established that this particular vector efficiently expresses GFP in the whole organoid, further confirming its ability to be transduced in the stem-cell compartment in my 3D culture conditions (Fig. 2B).

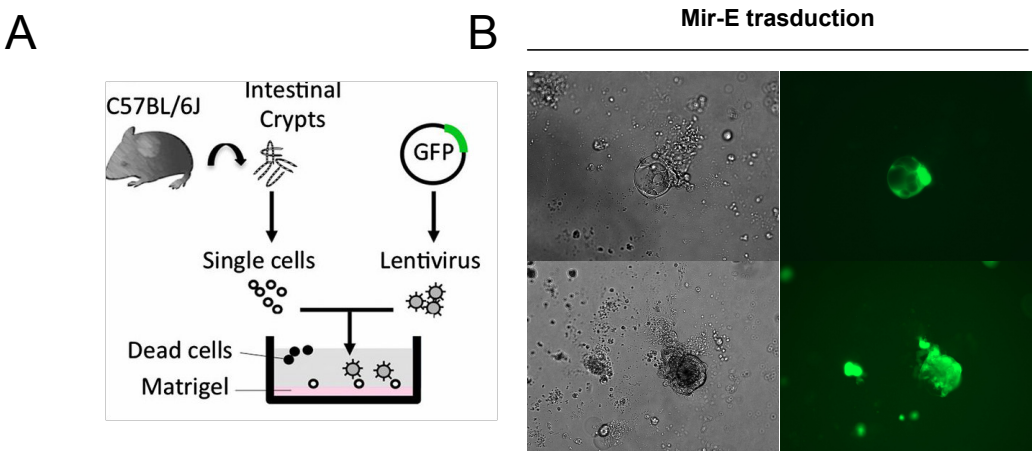

Fig. 2 Organoids transduction. A) Schema representing the infection procedure (image taken from Onuma et al.2013) B) Pictures representing transduced organoids at 48 hours after the infection event. Images were taken at 4X magnification.

2.3 Evaluation of the knockdown efficiency of regulated hits

Having assessed the feasibility of the infection procedure with Mir-E empty vector, I cloned specific sequences designed to express shRNA molecules to silence the epigenetic players I have previously selected to investigate if their depletion could exert any effect on mutated organoids. I first verified the knockdown efficiency in two different murine cell lines, either E14 mouse embryonic stem cells (mESCs) or NIH-3T3 depending on the expression levels of the genes of interest, to avoid costs and technical issues linked to the manipulation of organoids. Although several sequences were cloned and tested, I obtain efficient knockdown for only five out of the ten identified epigenetic players. I proceeded performing the knockdown assay on the remaining five (*Cxxc5*, *Carm1*, *Dnmt3b*, *Cbx6*, *Prmt1*) (Fig.3). Since this step of the project required parallel screening of such five epigenetic players, I decided to start by using only one shRNA per gene, choosing the most efficient one, to manipulate murine organoids.

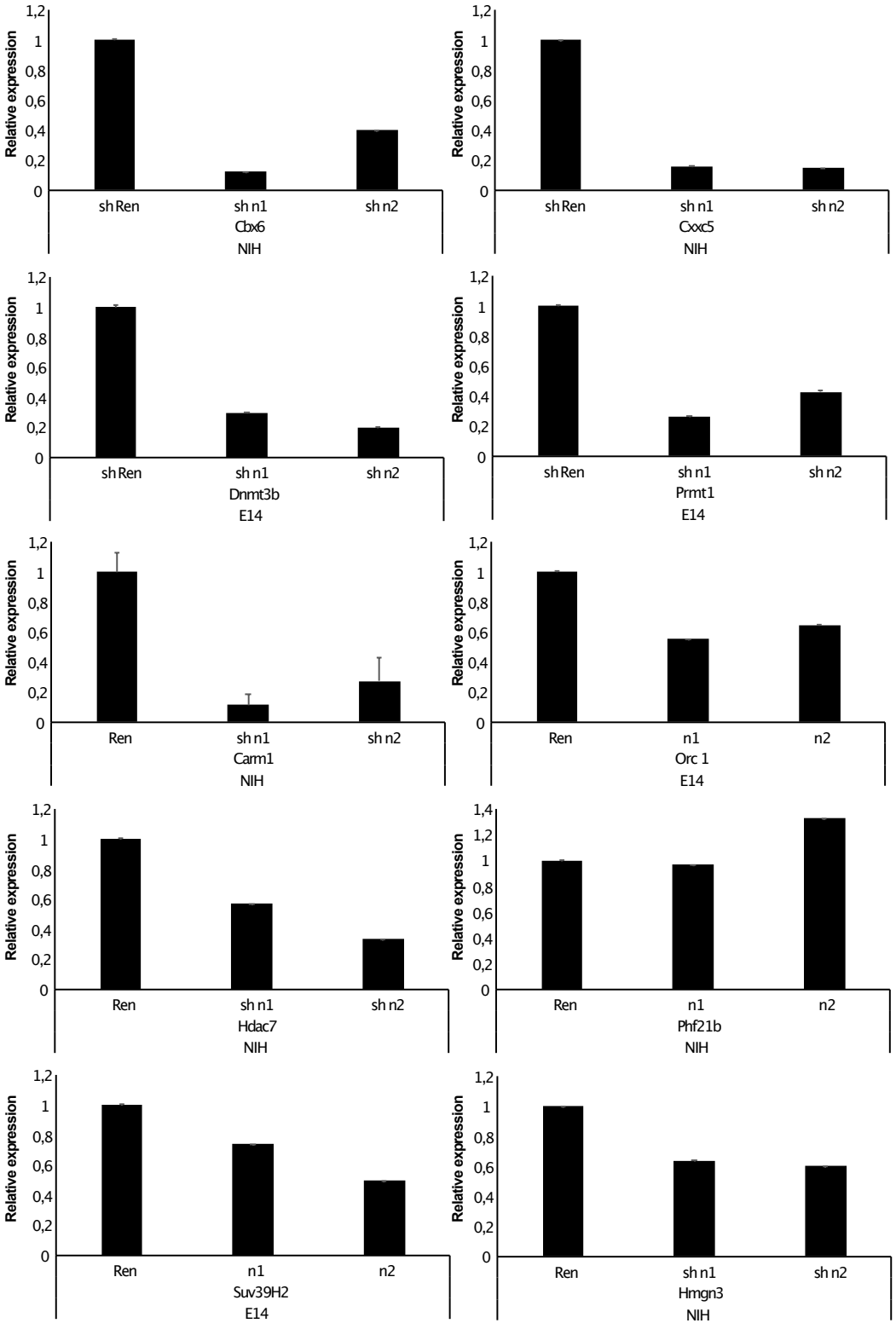

Fig. 3 Knockdown efficiency test of shRNAs on cell lines. qRT-PCR analysis of E14 and NIH-3T3 infected with dedicated shRNA. The sequence for each gene that mediated the best silencing was then used in the initial validation.

2.4 Neither Carm1 nor Cbx6 loss impact on β-catenin stabilized organoids growth

To address whether the knockdown of the selected epigenetic players was causative of changes in organoids growth, I collected transduced β-Catenin stabilized organoids after puromycin selection, dissociated them to single cells and replated the same number of living cells both of the control, an shRNA targeting the Renilla transcript not expressed in murine cells, and of the experimental samples.

I assessed the effects caused by the knockdown using the area of the reformed organoids as a parameter. To achieve this task, I acquired organoids pictures daily from day 2 to day 5 from which I measured the organoids area with ImageJ software comparing the experimental samples with the control.

As highlighted in the dot-plots representation, neither Carm1 (Fig. 4C) or Cbx6 (Fig. 4D) loss cause any change in organoids growth despite the high knockdown efficiency obtained in both replicates analyzed (Fig.4B). Indeed, silenced and control organoids showed similar size.

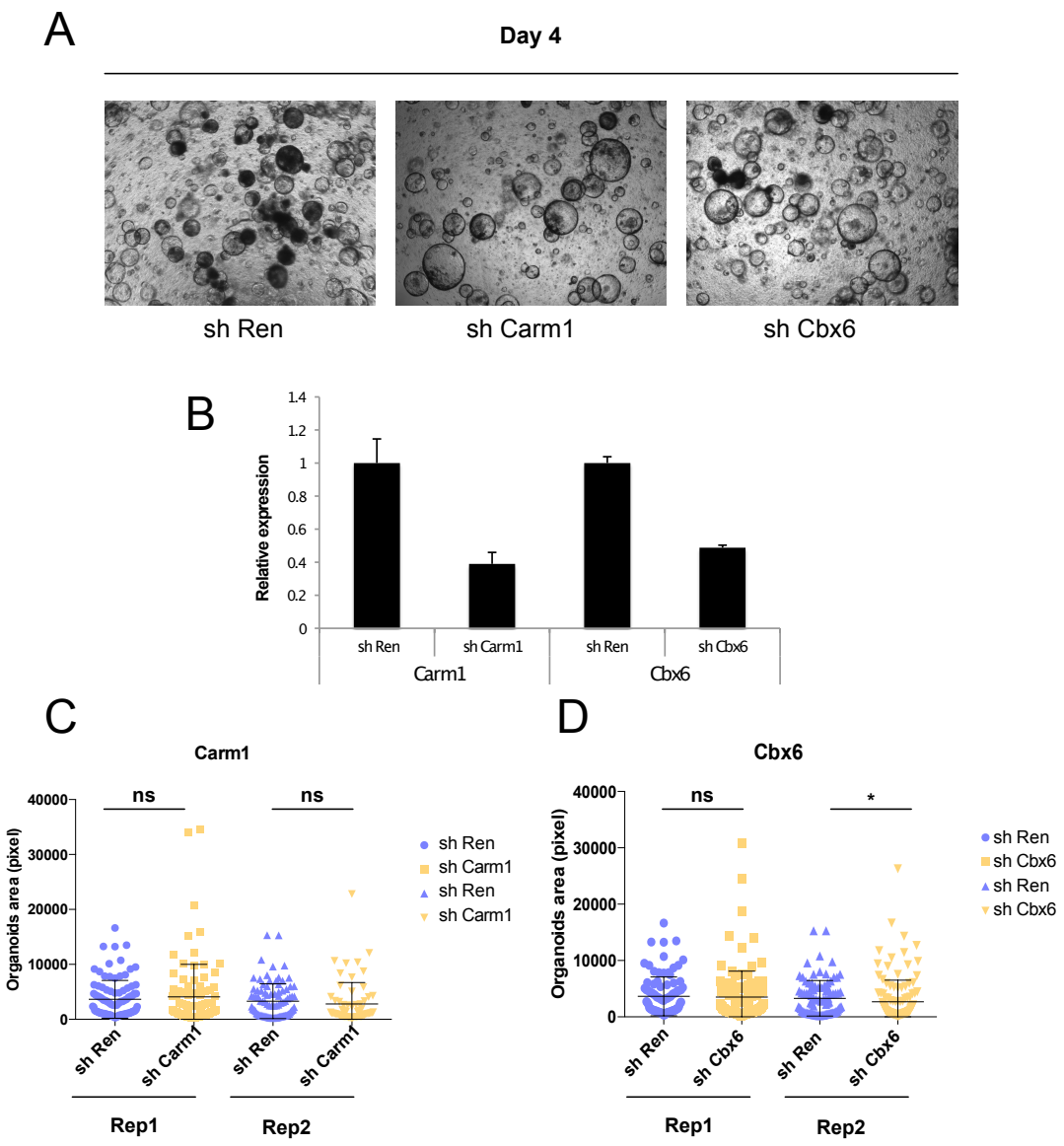

Fig. 4 Carm1 and Cbx6 knockdown do not show an effect on β-catenin stabilized organoids. A) Pictures representing organoids infected with sh Ren and shRNA targeting Carm1 (sh Carm1) or Cbx6 (sh Cbx6) at day 4 after single cells replating Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 4 days growth curve analysis. B) RT-qPCR analysis showing the knockdown efficiency mediated by the indicated shRNAs. Error bars correspond to mean± SD. C) and D) Dot-plot graphs depicting the distribution of organoids area at day 4. To assess statistically significative differences, Mann-Whitney test was applied.

2.5 Prmt1, Dnmt3b and Cxxc5 depletion cause impairment on β-catenin stabilized organoids growth

Targeting the other three epigenetic players regulated, Dnmt3b, Cxxc5 and Prmt1, I found that their depletion cause growth impairment on β-Catenin stabilized organoids (Fig. 5A). Indeed, in both replicates, upon efficient knockdown (Fig.5B), the overall organoids showed

a statistical significant reduction in size (Fig 5 C,D,E). In particular, organoids were very sensitive to the loss of Prmt1, showing a more striking phenotype.

Hence, I decided to further investigate the mechanism by which Prmt1 loss is causative of such a global growth decrease.

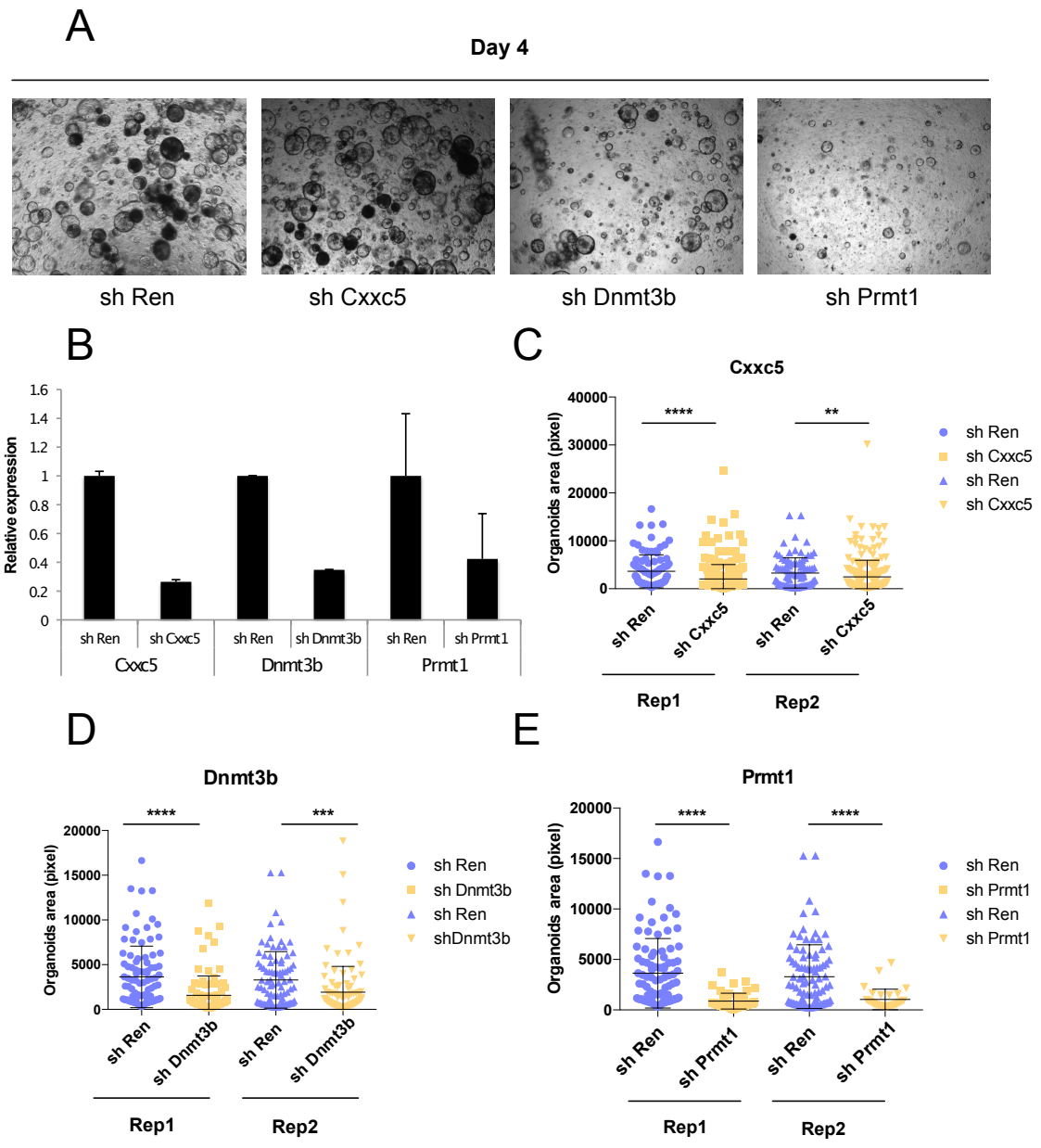

Fig. 5 Prmt1, Cxxc5 and Dnmt3b shows an impact on organoids growth. A) Pictures representing organoids infected with shRen and shRNA targeting Cxxc5 (sh Cxxc5), Dnmt3b (sh Dnmt3b) and Prmt1 (sh Prmt1) at day 4 after single cells replating. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 4 days growth curve analysis. B) RT-qPCR analysis showing the knockdown efficiency mediated by the shRNAs. Error bars correspond to mean± SD. C), D) and E) Dot-plot graphs depicting the distribution of organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

2.6 Prmt1 depletion causes severe growth impairment on APC KO organoids

Mutations in the components of the Wnt pathway are known to be the most frequently observed in colorectal carcinoma[4]. These mutations constitutively activate the Wnt signaling pathway as occurs with constitutive β-catenin activation or APC disruption. I wondered whether Prmt1 expression levels could be regulated and Prmt1 knockdown could decrease the growth of organoids with an *APC* null background.

First, I checked Prmt1 mRNA expression levels in WT and APC KO intestinal organoids and found Prmt1 up-regulation in APC null organoids (Fig 6A). Knocking down Prmt1 with two independent shRNA in APC KO intestinal organoids (Fig.6B), I observed significant growth impairment (Fig 6C,D), indicating that organoids harboring Wnt activating mutations are sensitive to Prmt1 loss.

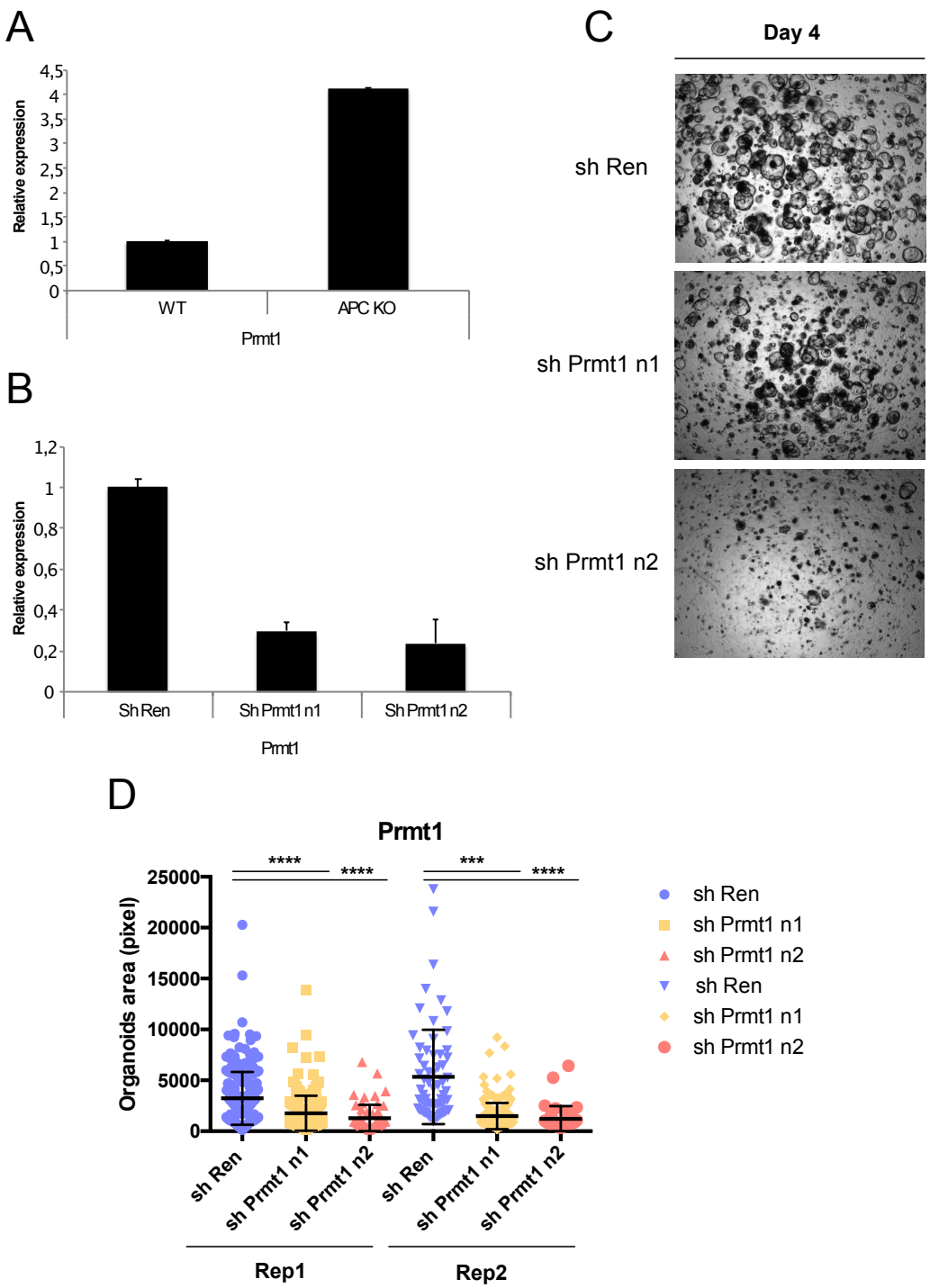

Fig. 6 Prmt1 knockdown causes a decrease in growth in APC KO organoids. A) Prmt1 expression levels in WT and APC KO organoids. B) RT-qPCR analysis showing the knockdown efficiency mediated by the shRNAs. Error bars correspond to mean± SD. C) Pictures representing organoids infected with shRen and shRNA targeting Prmt1. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 4 days growth curve analysis. D) Dot-plot depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

2.7 Prmt1 knockdown does not impair growth in WT organoids

To understand whether also wild-type organoids shared the same sensitivity to Prmt1 loss as the organoids with Wnt signaling activating mutations, I knocked down this specific arginine

methyltransferase in intestinal organoids derived from WT mice with both Prmt1 specific shRNAs. Unlike APC KO organoids, WT organoids do not grow as spheroids, but develop both crypt-like structures and villus-like domains throughout *in vitro* culture. In fact, under standard conditions, epithelial mini-guts maintain strong Wnt signals only nearby the Paneth cells, which are the only source of Wnt *in vitro,* while R-spondin-1 is ubiquitously present in the culture medium and enhances these focal Wnt signals. Paneth cells create the niche environment and secrete Wnt3a. Thus, the sharp Wnt gradient surrounding the Paneth cells induces the crypts formation[171,172].

Their resulting irregular shape made the measurement of the organoids area unreliable as a parameter for the evaluation of growth defects in WT organoid experiments. To overcome this issue, I decided to evaluate the effects of Prmt1 knockdown in this system by manually measuring the ability of WT organoids to undergo the process of differentiation, focusing my attention on the ratio of spheroids and fully differentiated organoids. As shown by the pictures in Fig. 7A, WT organoids infected with both shRNA targeting Prmt1 displayed similar size when compared to the control. Furthermore, they were able to fully differentiate, recapitulating both the crypt-like and the villus-like domains (Fig 7C). These results suggest that Prmt1 could be a promising therapeutical target, since its loss does not affect growth and homeostasis of normal tissue.

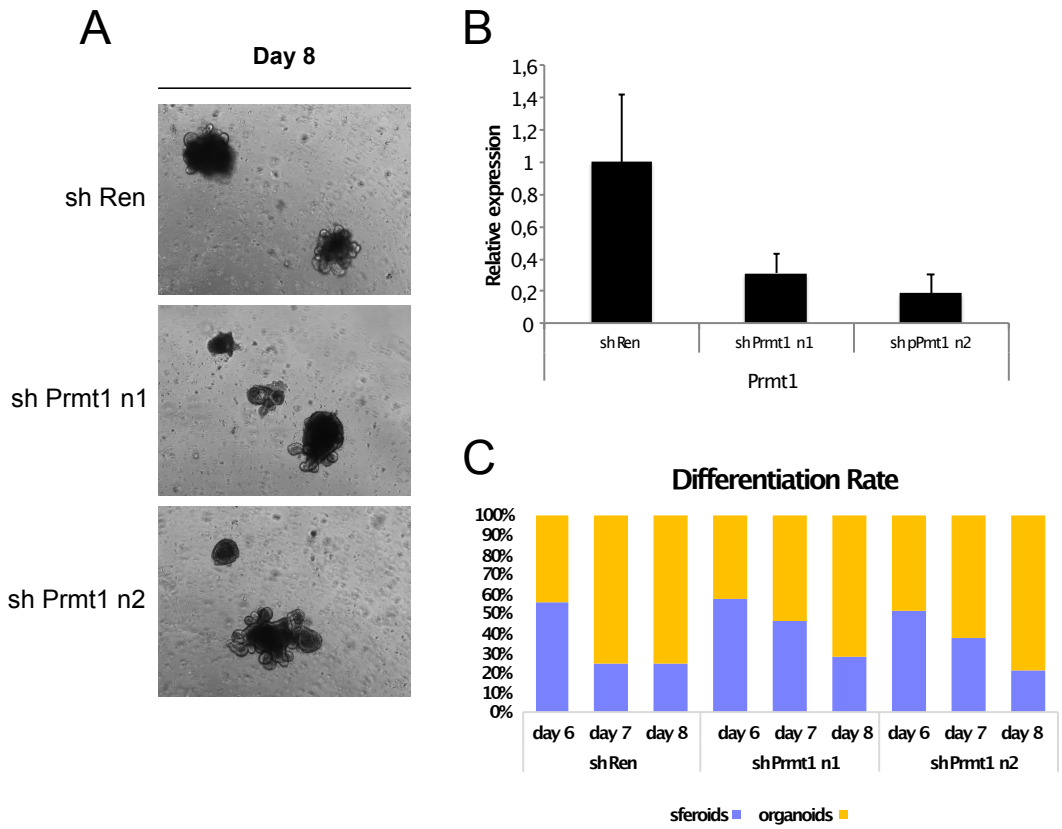

Fig. 7 Prmt1 loss does not impair WT organoids growth. A) Pictures representing WT organoids infected with shRen and shRNA targeting Prmt1. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 8 days growth curve analysis. B) RT-qPCR analysis showing the efficacy of the knockdown mediated by the shRNAs. C) Histogram depicting the organoids differentiation rate in organoids transduced with control shRNA and two independent shRNA targeting Prmt1.

2.8 Pharmacological inhibition of Prmt1 mimics the knockdown in β-catenin stabilized organoids

PRMT1 is the predominant mammalian arginine type I methyltransferase[173,174]. It catalyzes the formation of monomethylarginine and asymmetric dimethylarginine and it is responsible for about 85% of total protein arginine methylation activity[175].

To address if the catalytic activity of Prmt1 is involved in the phenotype observed I treated β-catenin stabilized organoids with MS023, a PRMTs type I inhibitor[176].

First, I checked if the treatment worked on intestinal organoids. To do so, I used as a control of the efficiency of the inhibition exerted by MS023 the global levels of H4R3me2a, known

to be one of the several substrates of Prmt1 enzymatic activity. As depicted by the Western Blot in Fig. 8, this target was not methylated anymore also with low doses of the drug.

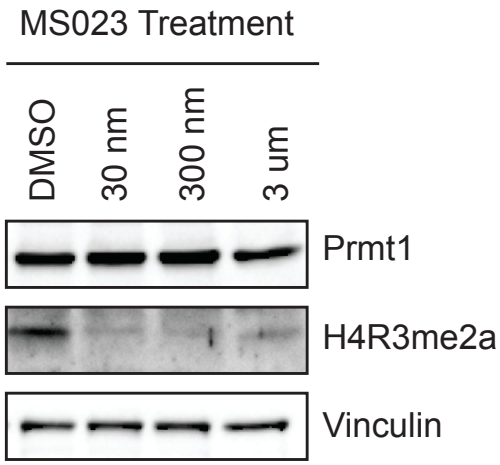

Fig. 8 MS023 treatment is efficient on β-catenin stabilized organoids. Western Blot of lysates of organoids treated with increased concentrations of MS023. H4R3me2a levels are decreased upon 72 hours of treatment.

Then, I evaluated the effect of the treatment on intestinal organoids growth. First, I seeded an equal number of single living cells in all the conditions allowing the formation of organoids. After 2 days, I treated the organoids for 72 hours with one dose of MS023 compound in three different increasing concentrations (Fig. 9 B), observing growth defects similar to those previously observed by knocking down Prmt1 (Fig 9 C, D). Prmt1 inhibition by MS023 induced impairment of the organoids growth in a dose dependent manner. These results suggest that the catalytic activity of Prmt1 is crucial for organoids growth, raising the possibility that the methylation deposited by this protein could be necessary for β-catenin stabilized organoids homeostasis.

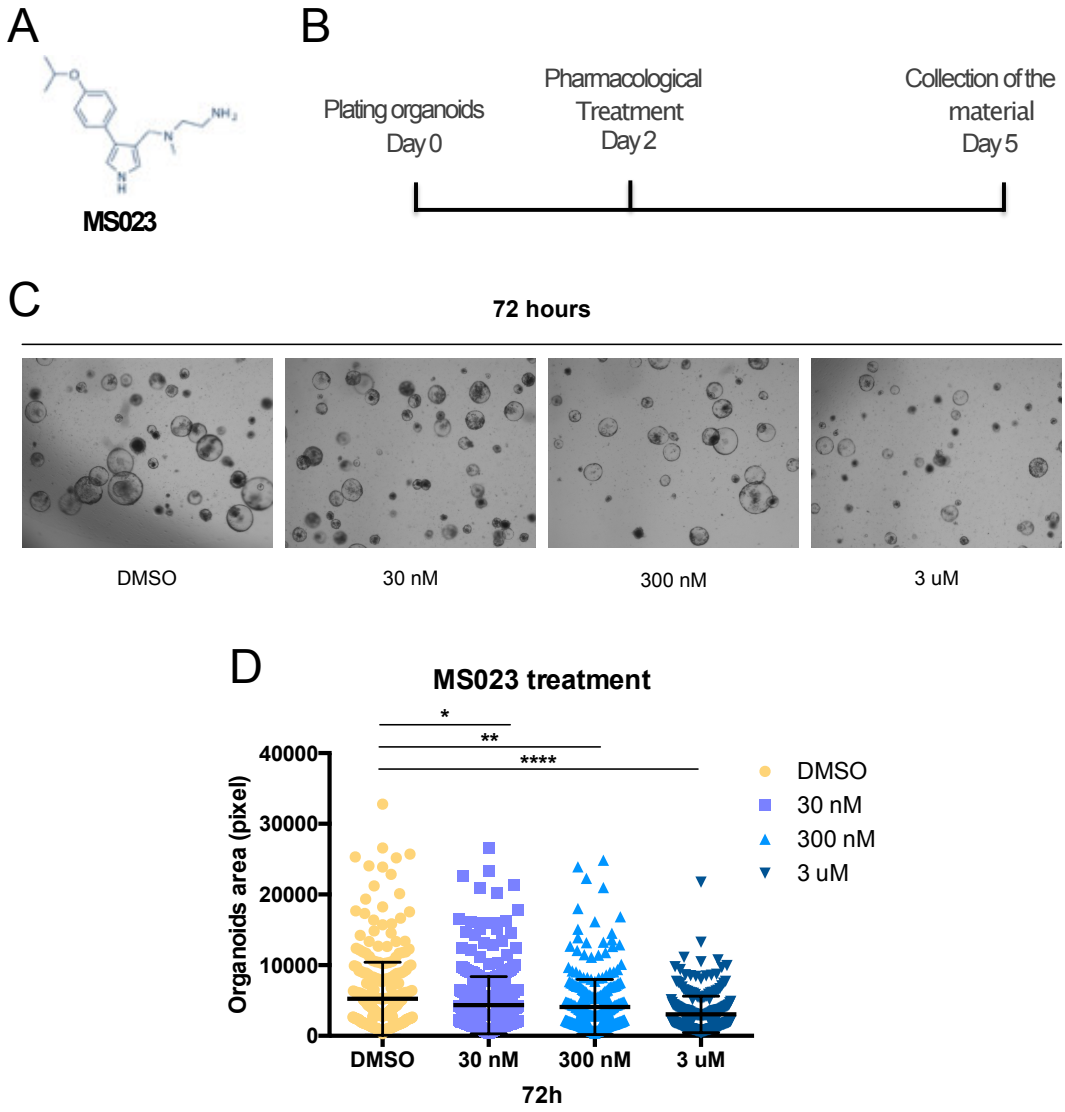

Fig. 9 MS023 treatment impairs growth rate in β-catenin stabilized organoids. A) Schematic representation of MS023 molecule. B) Workflow of MS023 treatment. C) Pictures of organoids treated with increased concentration of MS023 at 72 hours. D) Dot-plot depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

2.9 MS023 treatment decrease the growth rate in APC KO organoids

I successively compared the effect of Prmt1 silencing with the pharmacologic treatment with MS023 in APC KO organoids, using the same approach described for β-catenin stabilized organoids. Briefly, I treated APC KO organoids for 72 hours with increasing concentrations of MS023 (Fig.10 B). Although APC KO organoids seemed from a statistical point of view to be less sensitive to MS023 treatment in comparison with the knockdown of Prmt1, it is

still possible to notice an effect on growth both in 300 nM and at 3 μM conditions (Fig.10 A,C).

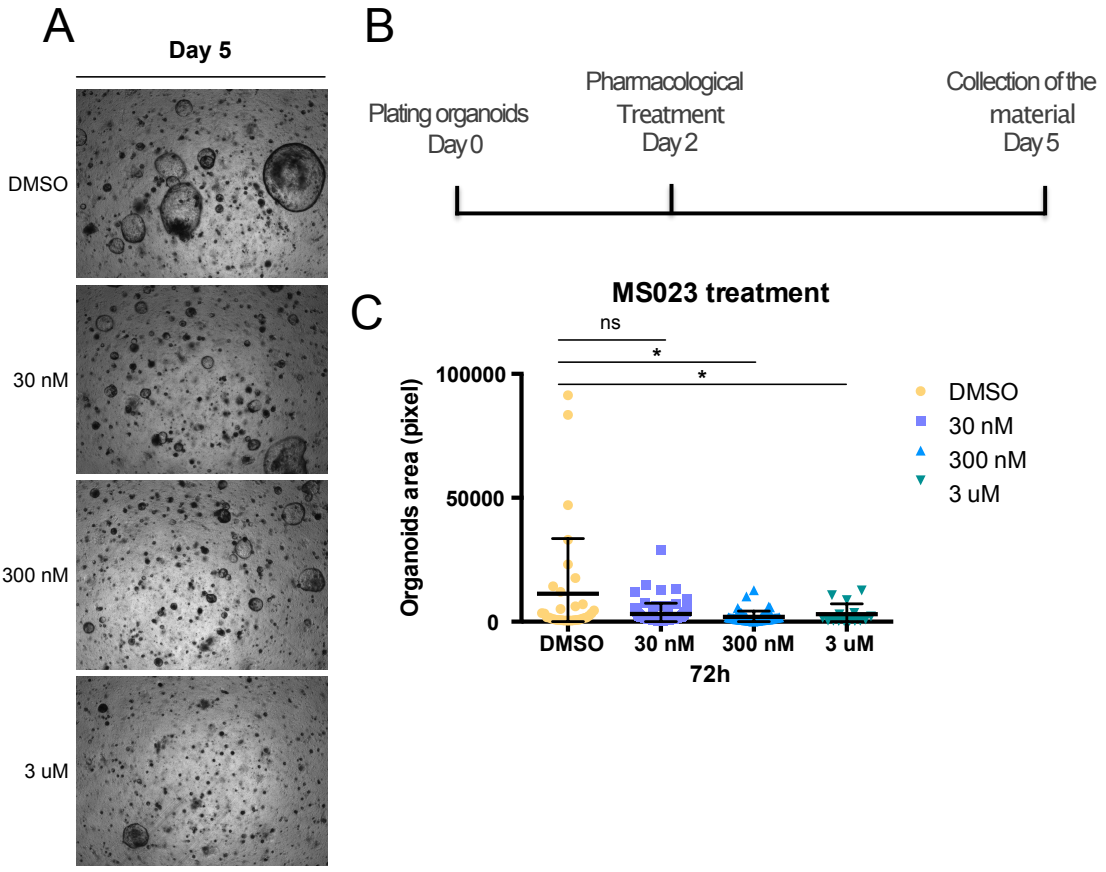

Fig. 10 Pharmacologic inhibition of Prmt1 catalytic activity causes a decrease in growth in APC KO organoids. A) Pictures of organoids treated with increased concentration of MS023 at 72 hours. B) Workflow of MS023 treatment. C) Dot-plot depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

2.10 Pharmacological inhibition of Prmt1 does not impair WT organoids growth as observed with its silencing

To assess whether the catalytic inhibition of Prmt1 recapitulates the results obtained with shRNA mediated interference experiment also in WT organoids, I needed to modify the treatment protocol. In fact, WT organoids grow definitely slower than the organoids bearing Wnt signaling activating mutations. Thus, I had to extend the treatment by adding a second MS023 dose. So, I seeded an equal number of single living cells per plate and I let organoids to form. After 48 hours, I treated organoids for 72 hours with a first dose of this compound in three different increasing concentrations. Since treated organoids were still not fully

differentiated at this time point (Fig. 11 B), I applied a second dose of MS023 for another 72 hours (Fig. 11 A). As expected, treated WT organoids did not show any sign of block in growth and in differentiation. It is indeed possible to note that such organoids treated also with high concentration of the drug display similar size when compared to the vehicle. Furthermore, they are able to fully differentiate, recapitulating both the crypt-like domain and the villus like domain (Fig 11 C). Overall, these results suggest that treatment with PRMTs inhibitors does not cause any side effect in normal organoids, opening a scenario in which these compounds could be promising in treatment of colorectal cancer.

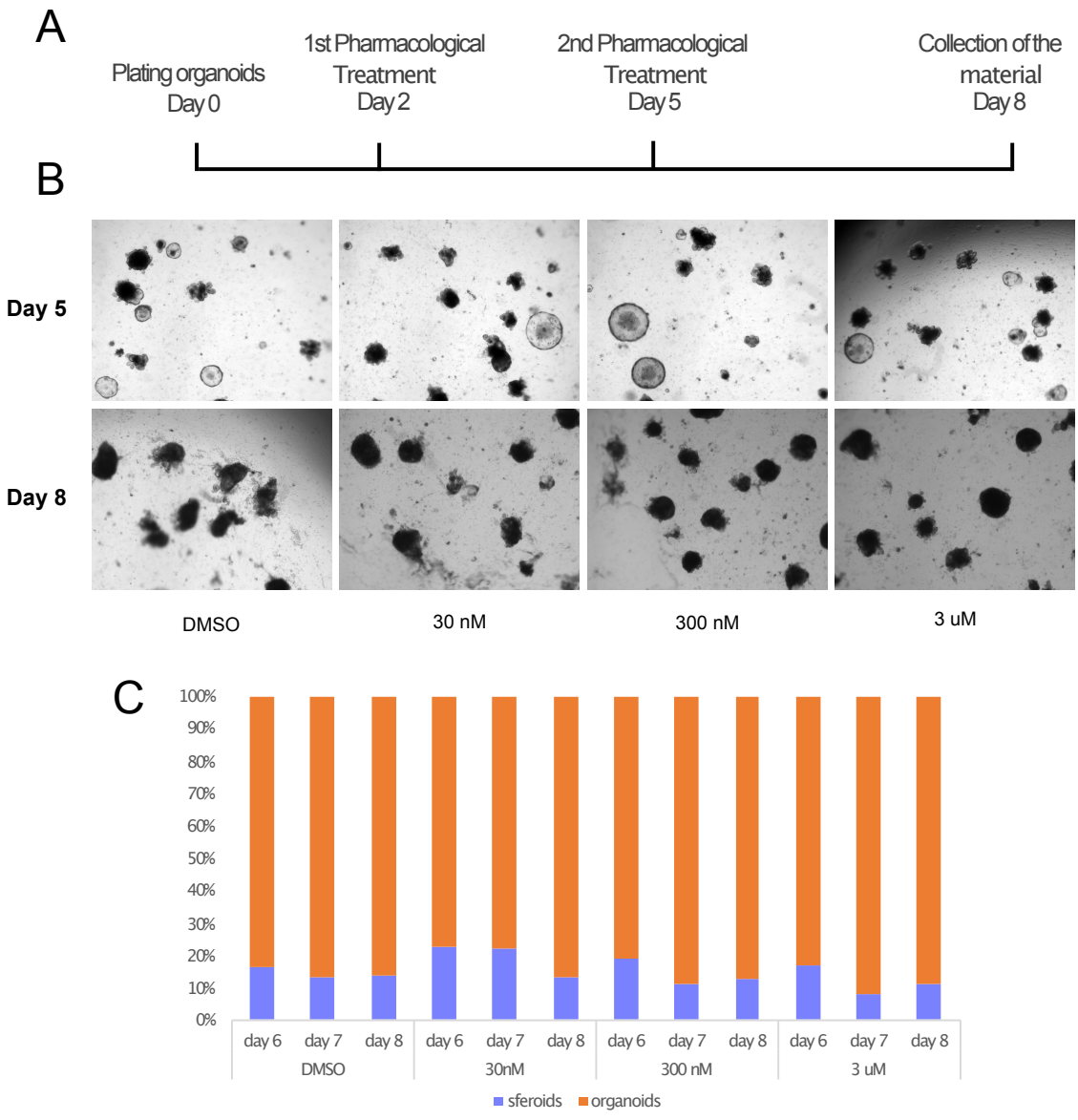

Fig. 11 Inhibition of Prmt1 activity does not compromise growth in WT organoids. A) Workflow of MS023 treatment. B) Pictures of organoids treated with increased concentration of MS023 after 72 hours of

first dose(day 59 and after 72 hours of the second dose (day 8). C) Histogram depicting the differentiation rate in organoids treated with MS023 at day 6,7,8..

2.11 Human metastatic CRC organoids show a dramatic decrease in growth upon PRMT1 silencing

Late stages of colorectal cancer are characterized by the presence of further oncogenic mutations that gives additional growth advantages to the tumor. Since there is still an open therapeutic window for CRC tumor at advanced stages, I decided to investigate whether also human patients-derived metastatic organoids were sensitive to Prmt1 loss or other mutations that occur at advanced stages of this tumor could instead confer resistance.

To accomplish this task, I performed the knockdown of PRMT1 in three different patients-derived-organoids characterized by different set of gene mutations, such as TP53, SMAD4 or RAS (see tables in Fig.12). Interestingly, all three patient-derived organoids showed immediately a global arrest in cells growth upon PRMT1 silencing (Fig. 12). These data provide further evidences that PRMT1 could be an attractive druggable target also for therapeutic purposes for CRC at late stages.

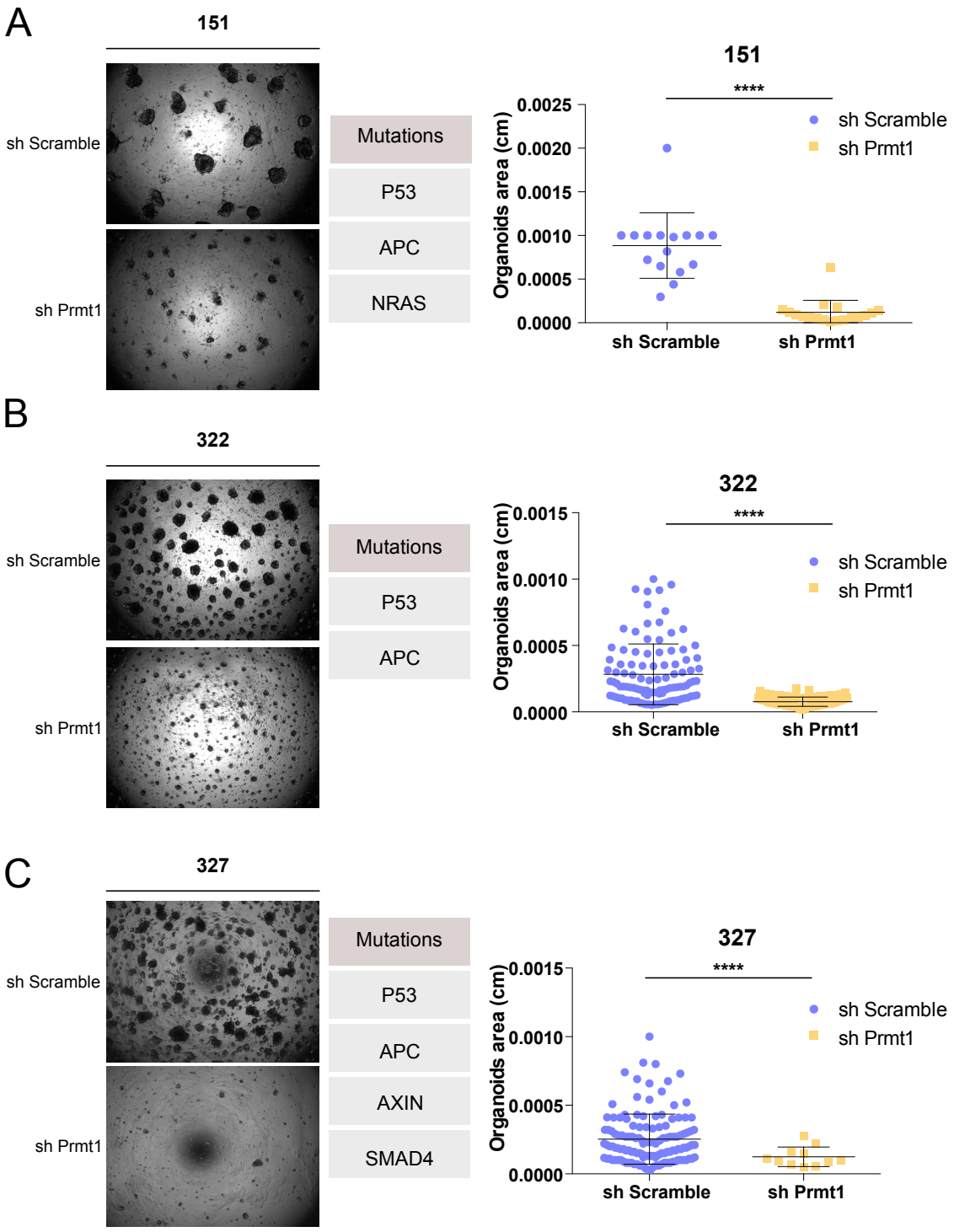

Fig. 12 Prmt1 loss results in global decrease in human organoids. A,B,C) Left: Pictures representing organoids infected with sh Scramble and shRNA targeting Prmt1. Center: table-representing mutations harbored by 151,322, 327 patient-derived organoid. Right: Dot-plot graphs depicting the distribution of organoids area at day 4. To assess statistically significative differences, Mann-Whitney test was applied. For all the three patients cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a growth curve analysis.

3. Second approach: Reverse genetic screen

3.1 Set up of the reverse genetic screen

3.1.1 The epigenetic library

As second approach to uncover new potential therapeutic targets in CRC treatment, I performed a functional shRNA screen on β-catenin stabilized mouse organoids taking advantage of a barcoded epigenetic library composed by 240 genes, 10 shRNA per each gene in order to minimize the off-target effects.

The library is divided in two different modules, mouseEpi1 and mouseEpI2 (mEPI1 and mEPI2). I performed two biological replicates for mEPI1 and three for mEPI2. This two different modules contain several writers, readers and erasers of epigenetic marks, transcription factors, DNA methyltranferases, proteins involved in chromatin remodeling or DNA repair processes, histone chaperones, and as internal controls 3 essential genes (Pcna, Psma1, Rpl30) and the luciferase (Fig. 13 B). The backbone retains the possibility to visualize transduced cells via GFP expression and allows the antibiotic selection of infected cells with puromycin resistance. Furthermore, each shRNA of the library contains both an unambiguous sequenceable barcode that allows the identification of the shRNA during High-throughput (HT) sequencing (Fig.13 A).

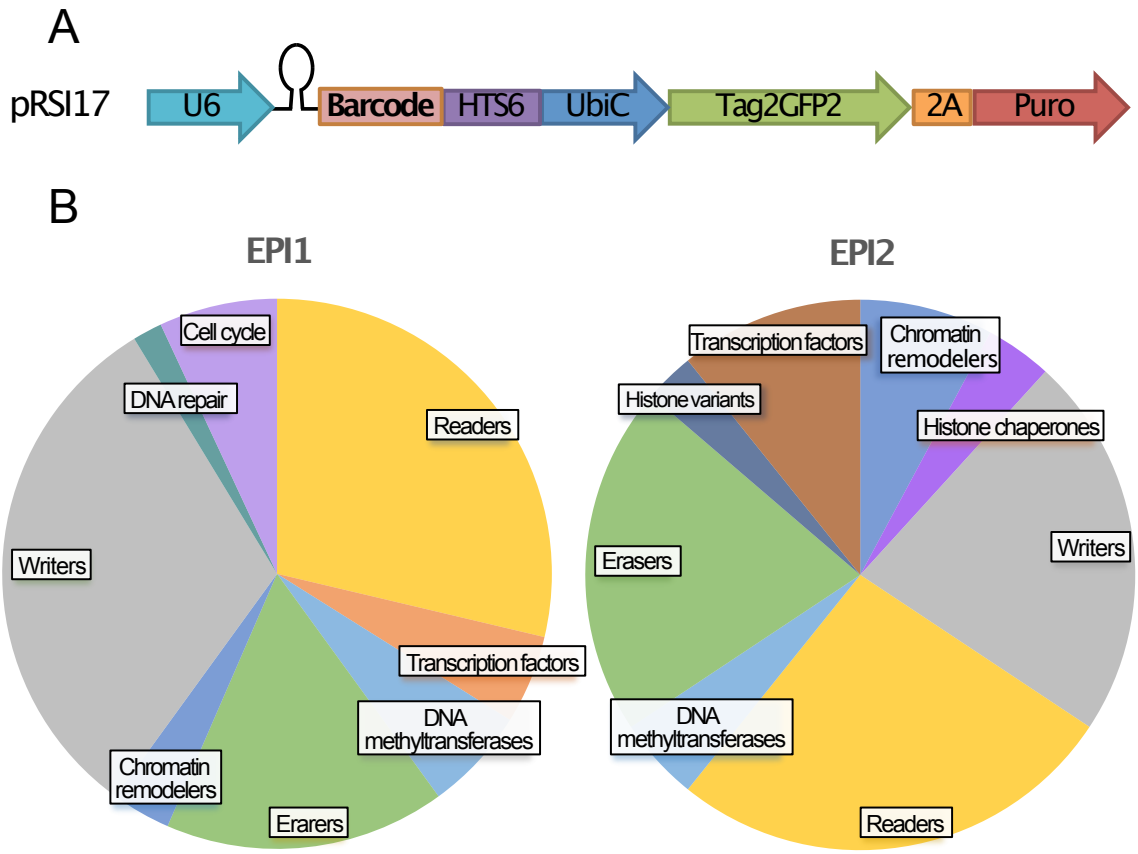

Fig. 13. The epigenetic library. A) Structure of the backbone of the epigenetic library. B) Left: mEPI1 composition Right: mEPI2 composition.

3.1.2 The screen protocol

β-catenin stabilized organoids were infected with the shRNA library at low multiplicity of infection (MOI=0.1) to assure that each cell carried one single viral integrant. The correct amount of viral particle delivered to organoids was assessed by FACS analysis via visualization of GFP expression by a technician of the laboratory (Fig.14 A).

I then defined the screen procedure: I first infected organoids as single cells overnight following the procedure described in paragraph 2.2, and, after 16 hours, I collected the reference. This is due to the fact that this window of time allows shRNAs packed in lentivirus to integrate in the genome of the host cells, but not to exert their inhibitory effect. The lack of functionality was evaluated taking advantage of the GFP cassette, indeed organoids at 16 hours after infection were GFP negative. The remaining seeded organoids were then collected after seven days when the selective pressure occurred. Then, genomic DNAs of

the reference and the experimental sample were extracted taking advantage of Phenol Clorophrom and subjected to PCR amplification of the Barcodes associated to the shRNA. The products were analyzed by Next Generation Sequencing (NGS) for barcodes (BCs) quantification, focusing the attention on depleted barcodes or drop-out hits, that represents genes essential for cell growth (Fig. 14B).

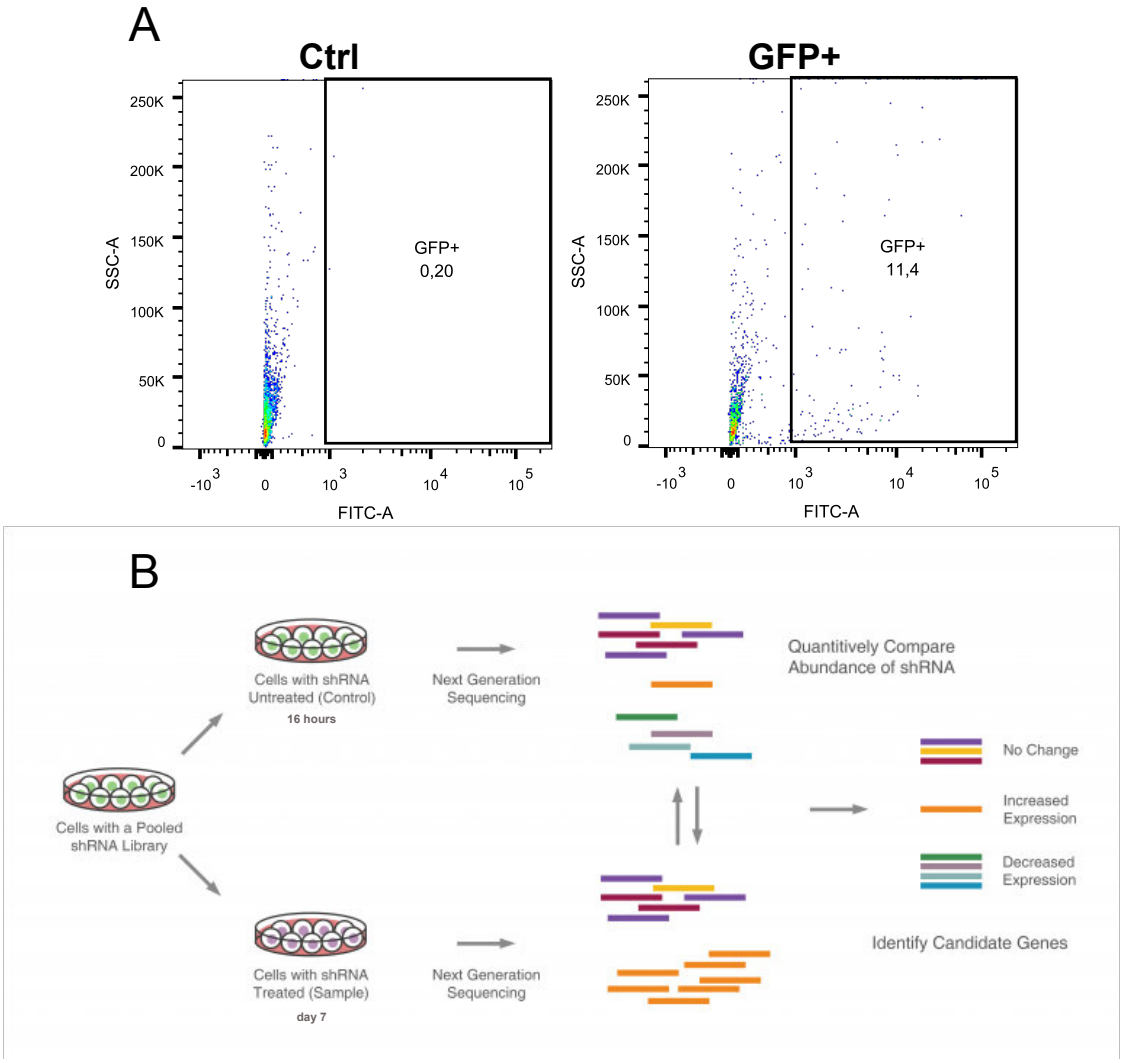

Fig. 14 Functional screen settings. A) Lentiviral MOI of the library established through FACS analysis. B) Schema of functional screen on organoids (adapted from coderegenesis website).

3.2 The screen analysis

Since 3D organoids represent a sophisticate biological model and the functional screen was performed in few biological replicates, a bioinformatician of the lab took advantage of several different statistical analyses to obtain hits as robust as possible.

In particular candidates to be validated were obtained by merging five different methods: Robust Z-score, Roast, Camera, Mageck and SSMD.

The Robust Z-score identified as hits genes with at least 7 barcodes under the median of the total distribution of the barcodes. In Roast, Camera and Mageck methods, genes were ranked according to the significance only the genes that scored with a significant p-value ($p<0.1$) were considered[152-154,177]. SSMD analysis instead provides a meaningful and interpretable criterion to classify the size of barcode depletion, and genes localized in the range between very strong and fairly moderate ($3 > |ssmd| \geq 1$) were contemplated[150]. I then started to focus on genes that scored in at least three out of the five methods (Fig.15 A).

Then, since I was particularly interested in understanding if there were multiple subunits of functional epigenetic complexes within the depleted barcodes identified I took advantage of String and Biogrid biological databases, annotating the biological interactions between hits previously identified and the reservoir of hits displayed by the less stringent method, the Robust Z-score. When I performed this analysis, I was able to identify several subunits of different functional complexes (Fig.15 B).

First, I identified Cohesin complex. Cohesin is a chromosome-associated multi-subunit protein complex highly conserved in eukaryotes, which mediates cohesion between replicated sisters chromatids and is therefore essential for chromosome segregation in dividing cells, for efficient repair of damaged DNA and for regulation of gene expression in both proliferating and post-mitotic cells[178]. Notably it scores as depleted in several screens among different tissues indicating its essentiality in several cell lines, thus providing an important evidence of the reliability of the intestinal organoids screen procedure (https://depmap.org/portal/gene/SMC3?tab=dependency).

Then, I was able to detect several Sin3a interactors: Brms1, Sin3b and Kdm5a. These proteins are part of the Sin3-Hdac chromatin remodeling complexes, involved in the regulation of transcription of many genes achieved by dynamic changes in histone modification, not only at promoters, but also at transcribed regions[179]. Next, I found two

distinct interactors of Wdr5 part of two distinct functional complexes. The first one is Kat8, which notably is another hit scoring in three out of the five statistical methods. Kat8 is a lysine acetyltransferase (KAT), which is able to acetylate lysine 16 of histone H4 (H4K16ac), a modification that decompacts chromatin structure. Kat8 and its non-specific lethal (NSL) complex members have been shown to localize to gene promoters and enhancers in the nucleus, as well as to microtubules and mitochondria to regulate key cellular processes[180]. The other subunit is Kmt2d, the major mammalian histone H3K4me1 methyltransferase that, in association with Wdr5 and other proteins, forms a stable complex, which is a major regulator of enhancer activity[181]. Thus, I decided to include one of the subunits of the Sin3 complex, Brms1, and then Kat8 and Kmt2d in the validation.

A

Gene	ROAST	Camera	Zscore	SSMD	Mageck	Calls
Dnmt1		y	y	y		3
Dnmt3l		y	y	y		3
Kat2b		y	y	y	y	4
Kat7	y		y		y	3
Kat8			y	y	y	3
Phf21a	y			y	y	3
Phf5a		y	y	y		3
Setdb2	y		y	y		3
Sin3a		y	y	y	y	4
Sirt4		y	y	y		3
Smc3		y	y	y	y	4
Sp140		y	y	y	y	4
Wdr5		y	y	y		3

B

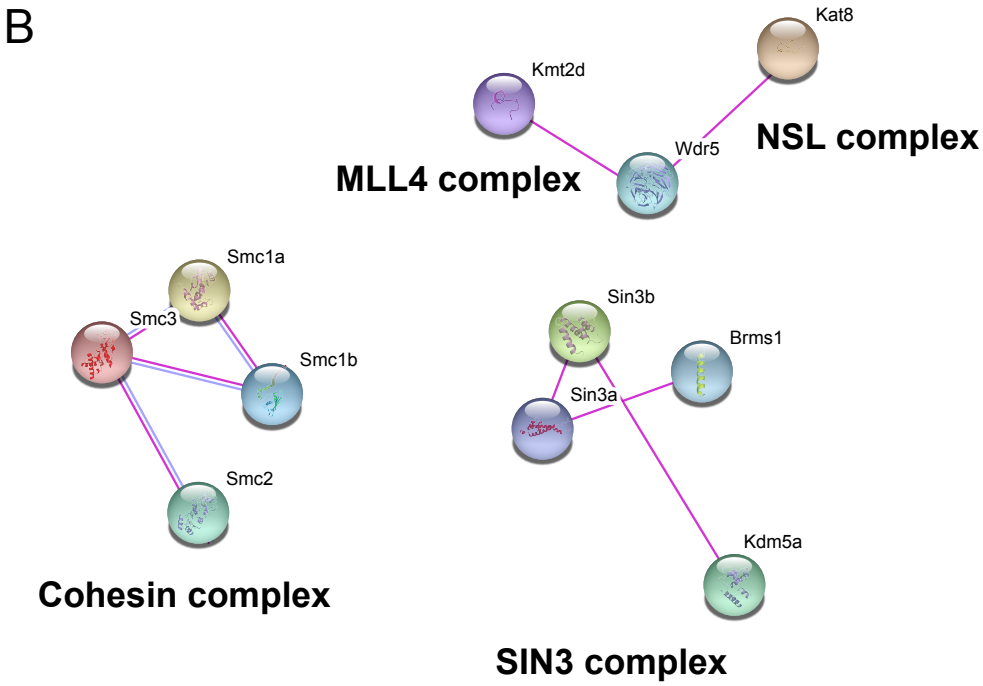

Fig. 15 Statistical analyses identified several robust hits part of functional epigenetic complexes. A) Table representing all the candidates that scored in at least three out of the five statistical methods. B) Schematic representation of interaction between selected genes identified in al least three different statistical methods and the reservoir of hits displayed by Robust Z-score only.

3.3 Validation of shRNA sequences of selected hits in E14 cells

To further implement the validation process, I took advantage of shRNA sequences that were not contained in the two epigenetic modules where possible, and I cloned them in a different backbone, the pLKO.1. I was able to obtain two independent sequences per gene exerting an efficient knockdown with the exception of Kat8. Since it was not possible to obtain any

good silencing of this hit, I decided to discard it from the validation process and to proceed with other candidates.

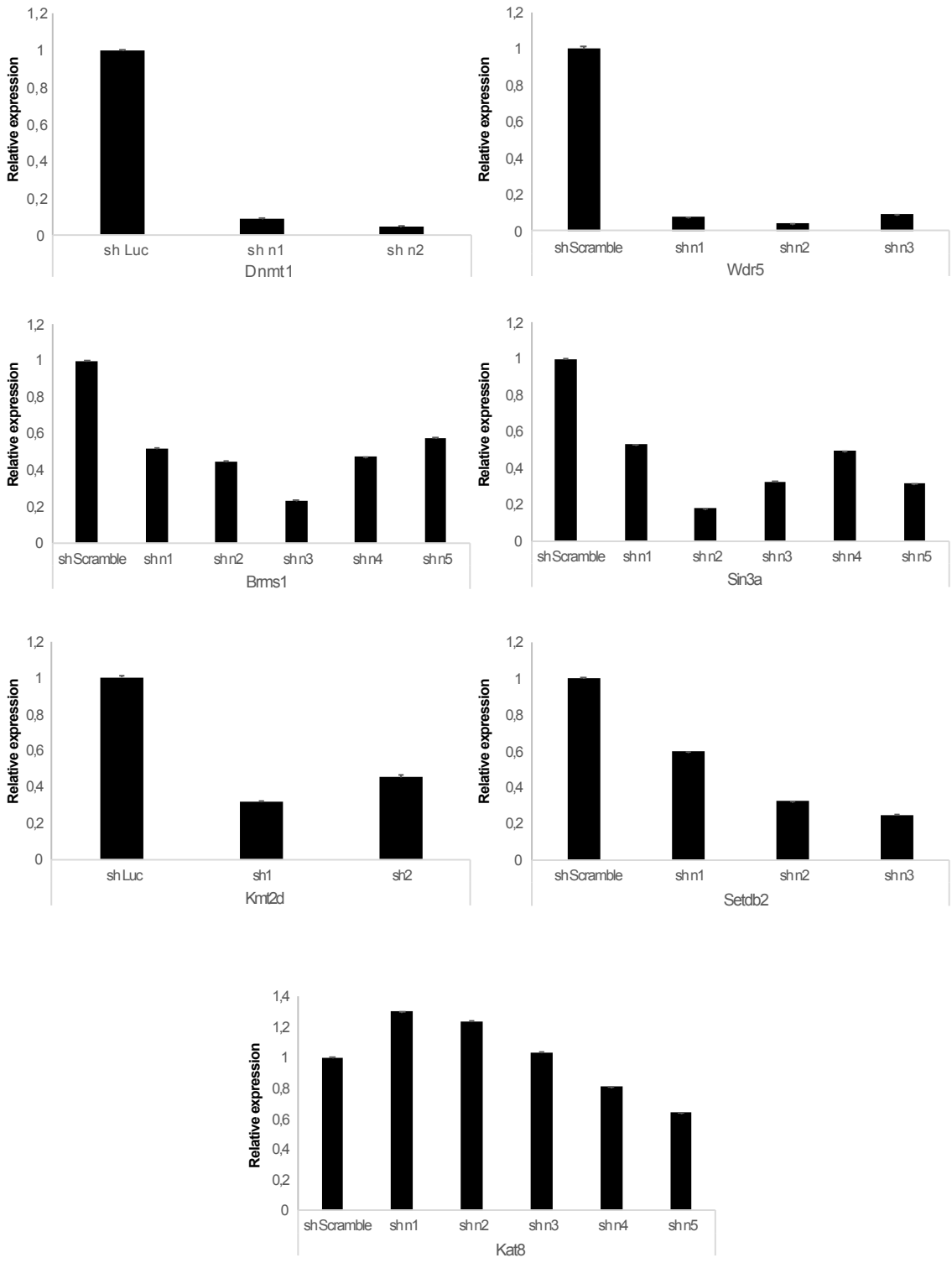

Fig. 16 Knockdown efficiency test of shRNAs on E14 cell line. RT-qPCR analysis in E14 showing the knockdown efficiency mediated by the shRNAs. Error bars correspond to mean± SD.The two sequences for each gene that mediated the best knockdown were then used in the validation.

3.4 β-catenin stabilized organoids show impairment in growth upon Dnmt1 and Setdb2 silencing

Although I was mainly interested in the validation of epigenetic complexes, in order to enhance the robustness of the validation process, I included two selected hits that did not show any interaction, Dnmt1 and Setbd2.

Dnmt1 is the principal DNA methyltransferase in mammalian cells, a highly dynamic enzyme with multiple regulatory features that can control DNA methylation in cells, regulating the process of maintenance of DNA methylation in cells[182].

Setdb2 instead is a histone methyltransferase able to contribute to chromatin architecture by methylation of histone H3 at lysine 9 (H3K9me), an important player in the formation of heterochromatin, chromatin condensation, and transcriptional repression[183].

Drop-out hits silencing should cause a detrimental effect on organoids growth. In order to address if the selected candidates knockdown could exert any effect on organoids growth, I took advantage of the procedure described in paragraph 2.4, first infecting organoids and then replating them as same number of single cells both in the control and in the knockdown of interest to study the organoids growth with a dedicated assay. All the validation experiments were performed in biological triplicate.

As depicted in the pictures (Fig.17 A, B) and in the dot-plot analysis (Fig 17 E,F), both the knockdown of the two candidates showed effects consistent with the expected phenotype and the silencing of the specific gene expression occurred as confirmed by RT-qPCR (Fig.17 C,D). So, I considered these two genes as an example of successful validation and I further proceed with the validation of the epigenetic complexes.

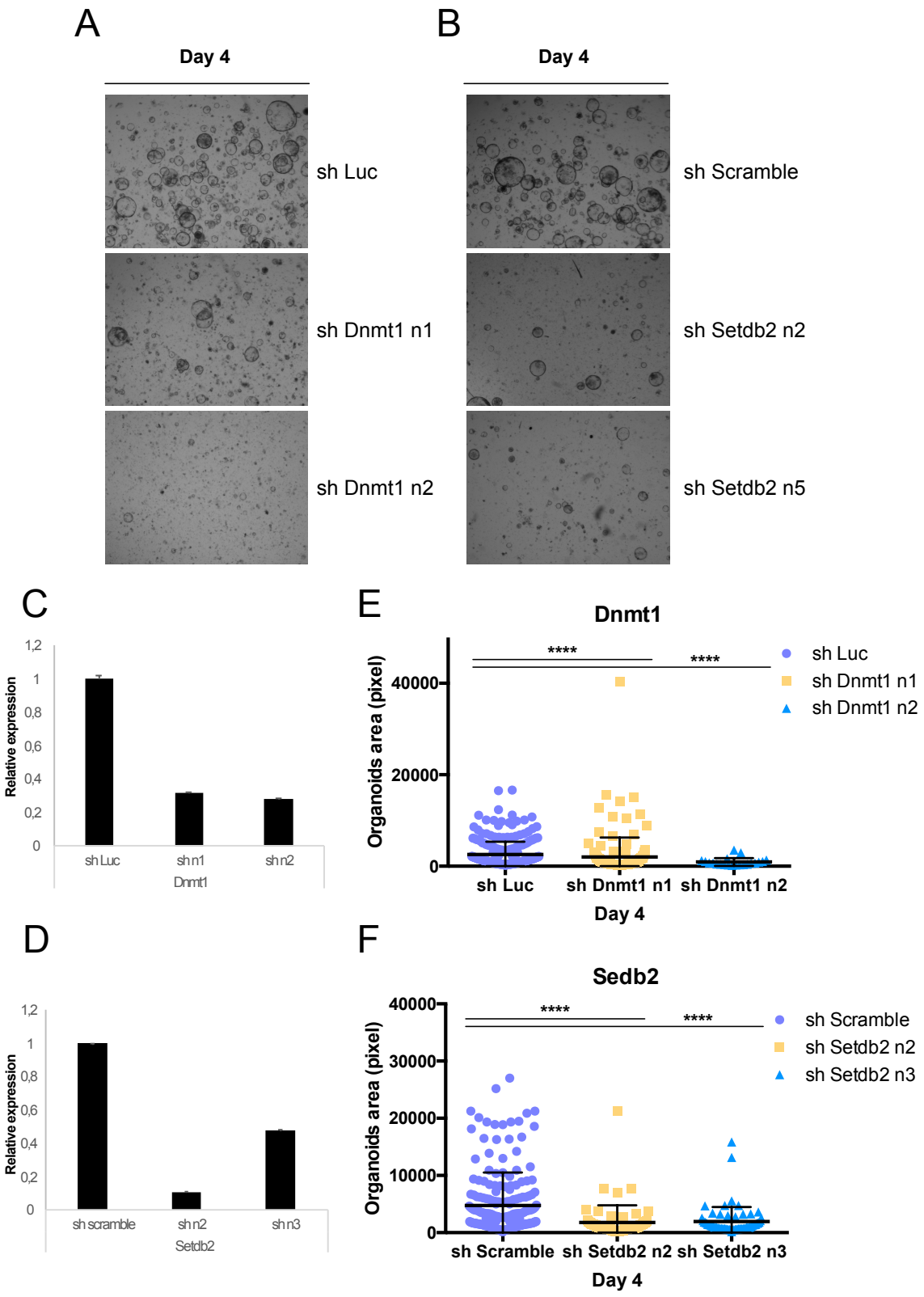

Fig. 17 Dnmt1 and Setdb2 knockdown cause a decrease in growth in β-catenin stabilized organoids. A,B**)** Pictures representing organoids infected with shRNA targeting Dnmt1 and Setdb2 and their control shRNA, shLuc and shScramble, respectively. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 4 days growth curve analysis. C,D) RT-qPCR analysis showing the efficacy of the knockdown mediated by the shRNAs. Error bars correspond to mean± SD E,F) Dot-plot depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

3.5 Sin3a and Brms1 knockdown drastically decreases proliferation in β-catenin stabilized organoids

First, I validated subunits the two subunit of Sin3 epigenetic complex: Sin3a and Brms1. Concerning the first one I was able to obtain an efficient knockdown of two independent shRNA (Fig. 18 C) that caused a dramatic decrease of organoid size respect to the control (Fig. 18 A), opening the hypotheses a role of this protein in intestinal organoids growth.

I was not able instead to obtain two independent knockdown of Brms1 (Fig.18 D). As depicted in Fig. 16, when I tested the sequences on E14 cells, shRNA n3 was the only one able to provide a gene silencing below 50%. Furthermore, since organoids integrate few viral copies due to their low efficiency of infection, I even obtained less efficient knockdown in comparison to those observed in E14. But since the working shRNA targeting Brms1 induced an appreciable decrease in organoids growth, I considered Brms1 as validated (Fig.18 B, F). These results highlight a possible role of this complex in organoids growth.

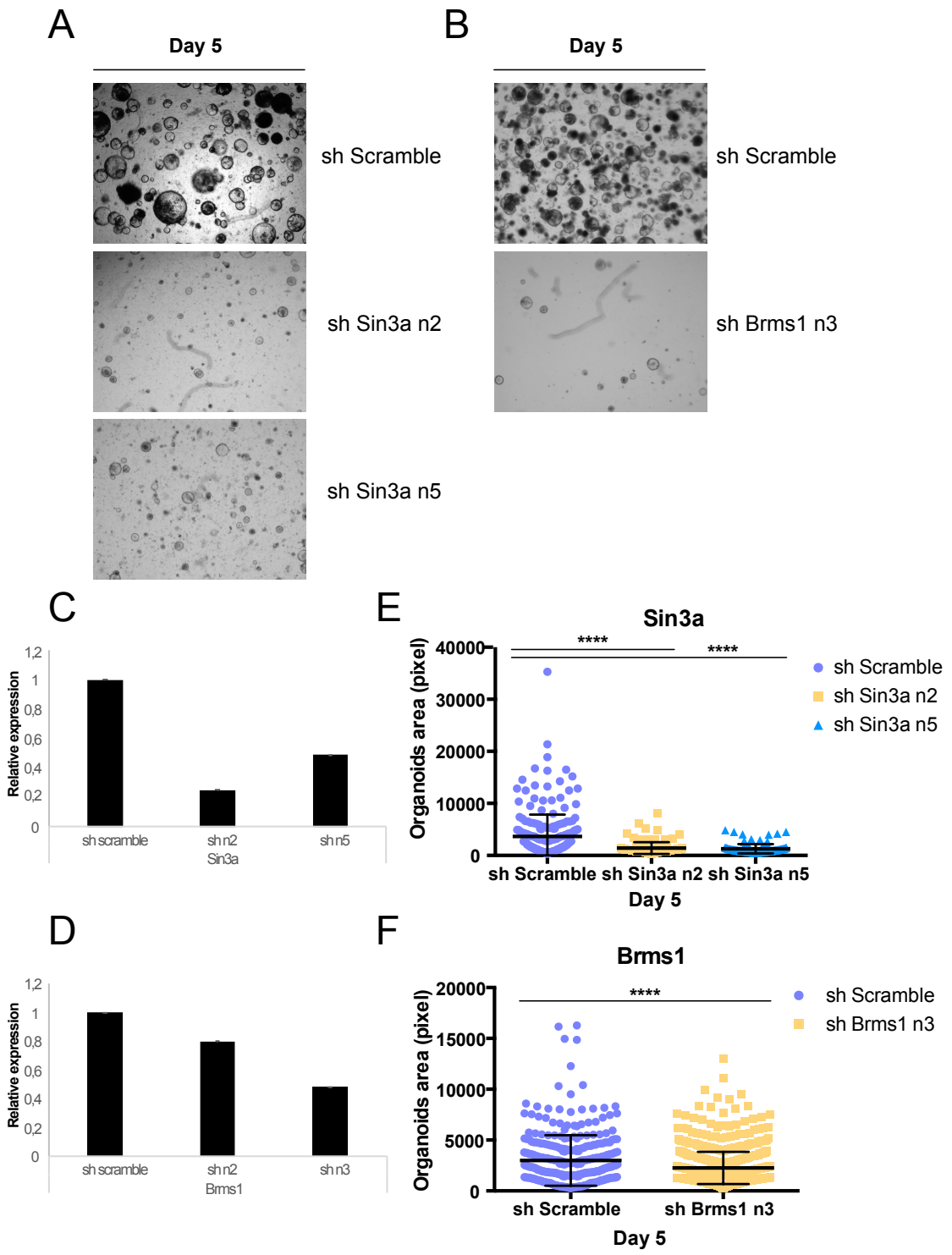

Fig. 18 The knockdown of both the two subunits of Sin3 complex show a decrease in growth in β-catenin stabilized organoids. A,B) Pictures representing organoids infected with shRNA targeting Sin3a and Brms1 and their control shRNA. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 5 days growth curve analysis. C,D) RT-qPCR analysis showing the efficacy of the knockdown mediated by the shRNAs. E,F) Dot-plot depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

3.6 Reduced expression of Mll4 complex subunits compromises organoids growth

Next, I tested the two subunits belonging to Mll4 complex: Wdr5 and Kmt2d.

Upon the knockdown with their dedicated shRNAs, both Kmt2d and Wdr5 loss displayed a really strong phenotype, in which only few organoids were able to form and grow. Since the number of organoids was drastically lower respect to the control, as depicted in the Fig.19B, it was not possible to perform significant statistical analysis by comparing the control conditions with the shRNA (Fig.19 E). RT-qPCR analysis confirmed the efficacy of the knockdown event (Fig.19 C). Overall, I considered both Kmt2d and Wdr5 hit as validated since the effect of their silencing caused drastic reduction in organoids growth raising the possibility of the involvement of Mll4 epigenetic complex in organoids homeostasis.

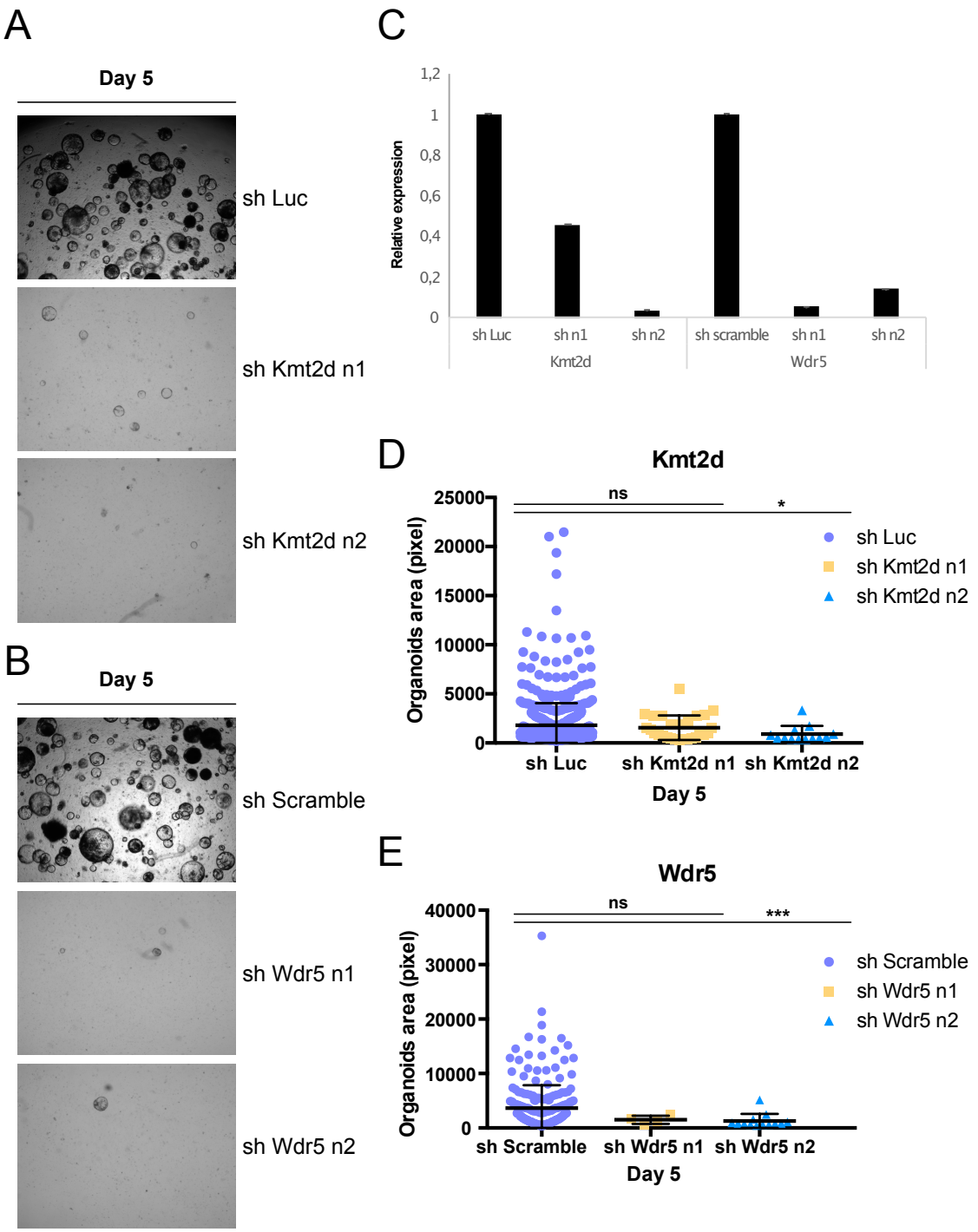

Fig. 19 Silencing the two subunits of Mll4 complex cause growth arrest in β-catenin stabilized organoids. A,B**)** Pictures representing organoids infected with shRNA targeting Kmt2d and Wdr5 and their control shRNA, shLuc and shScramble, respectively. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a 5 days growth curve analysis. C) RT-qPCR analysis showing the efficacy of the knockdown mediated by the shRNAs. D,E) Dot-plot depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

3.7 The silencing of MLL4 and SIN3 complex subunits cause a global detrimental effect in proliferation in human HCT-116 CRC cell line

In order to better characterize the role of Sin3 and Mll4 complexes in colon cancer, and to perform mechanistic studies that are not possible in organoid cultures due to huge issues in terms of feasibility in experiments requiring high number of cells, I took advantage of a colorectal cancer cell line, HCT-116, that harbors driver mutation on β-catenin locus that cause a constitutive activation of WNT pathway as well as the murine organoids previously utilized. HCT-116 is a cell line of a primary colorectal cancer that is easier to handle and so it can provide a more quantitative assessment of the phenotype observed, through classical growth curve and flow cytometric analyses, and furthermore it is a feasible platform to perform mechanistic experiments.

First, I investigated whether these cells could show the same phenotype observed in organoids through growth curve analysis.

Briefly, I transduced HCT-116 with human shRNA targeting the gene of interest and I collected them after puromycin selection. Next I counted cells in order to seed the same number of cells/well for each condition to perform a growth curve assay.

The analysis confirmed that the block in cell growth observed in β-catenin stabilized organoids was present also in HCT-116 cells. WDR5, SIN3A and KMT2D in particular showed a massive arrest in cell growth (Fig.20 A, B, D), while BRMS1 knockdown displayed a milder, but still relevant effect (Fig.20 C). Overall the effect pursued the silencing of the four candidates provides indication of a potential role of the two epigenetic complexes in colorectal cancer cell growth.

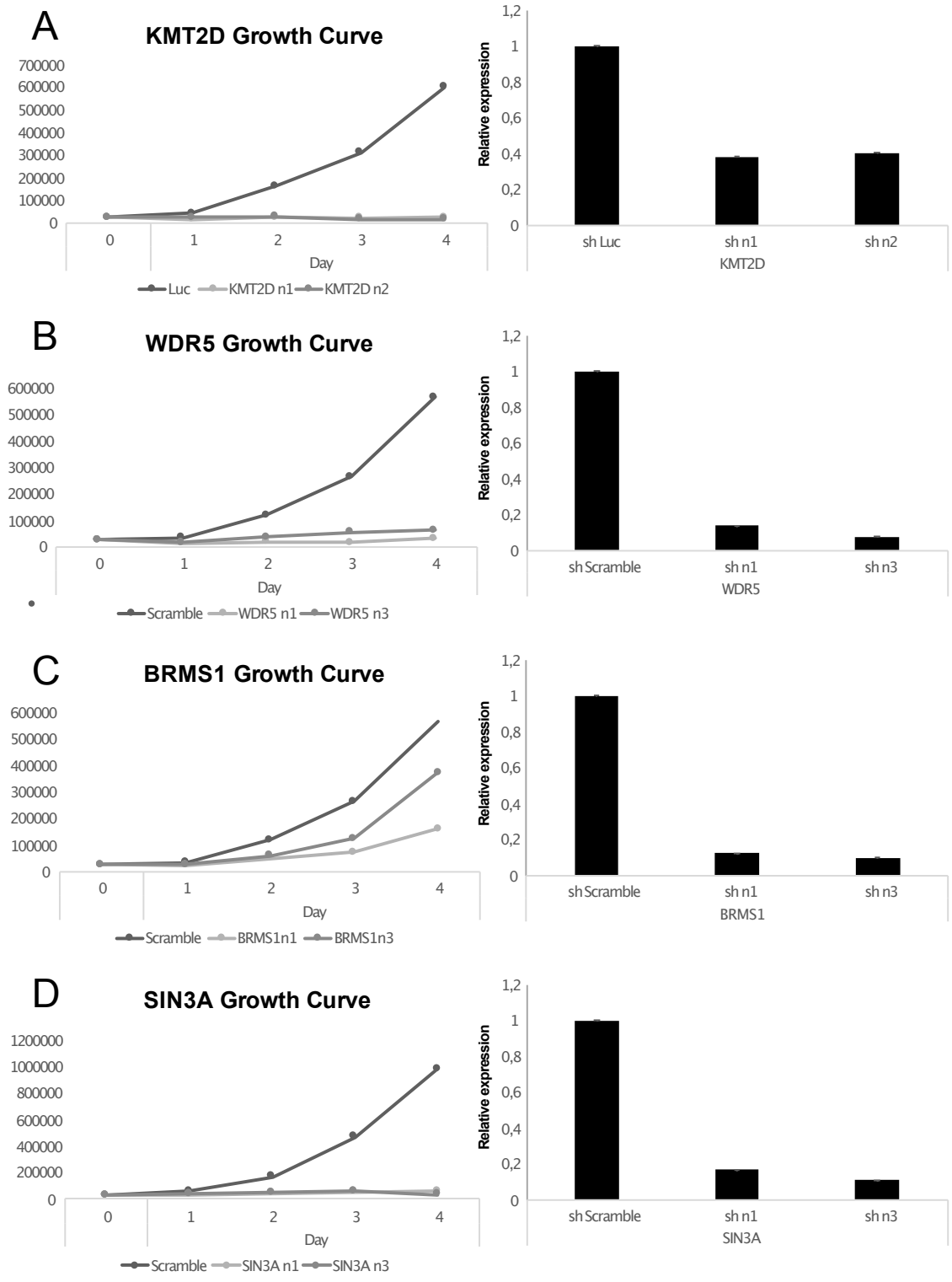

Fig. 20 BRMS1, KMT2D, SIN3A and WDR5 growth curves in human HCT-116 CRC cell line: **A,B,C,D**) Left: Growth curves representing the number of alive cells per day, from day 0 to day 4; Right: RT-qPCR analysis of knockdown efficiency of shRNA constructs targeting each gene.

3.8 SIN3A loss induces a slight destabilization of SIN3A-HDAC1 complex

To dissect at a molecular level the effect of SIN3-HDAC silencing, I decided to focus my attention to SIN3A subunit. I investigated through Western Blot analysis if upon SIN3A knockdown the stability of the complex was affected, testing the level of one of its catalytic subunit, HDAC1. Surprisingly, I observed that although HDAC1 is part of multiple functional complexes[184], SIN3A silencing slightly induces its destabilization, highlighting the possibility of a role of SIN3A in mediating the stabilization of the complex.

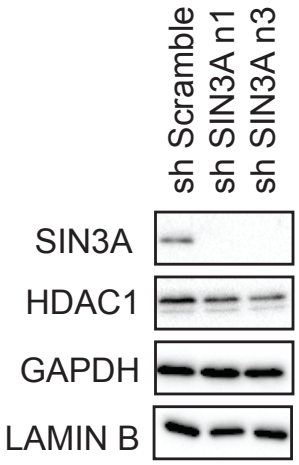

Fig.21 Western Blot analysis upon SIN3A silencing. Western blot of HCT-116 lysates infected with two independent shRNA targeting SIN3A.

3.9 SIN3A silencing in human HCT- 116 CRC cell line does not influence the levels of histones modifications

To test the possibility of fluctuations in histone modifications landscape upon SIN3A silencing I performed Western Blot analysis focusing my attention on histone modifications of H3 tail that exert either activating or repressing effects on transcription regulation.

SIN3A-HDAC1 complex has two catalytic cores: the first consists in HDAC1/2, involved in deacetylation of multiple residues on histone tails[172] while the second, JARID1A, is involved in the specific demethylation of H3K4me3 and –me2[185]. Thus, I analyzed levels of

methylation on H3K4 in bulk without detecting any significant change in such modifications.

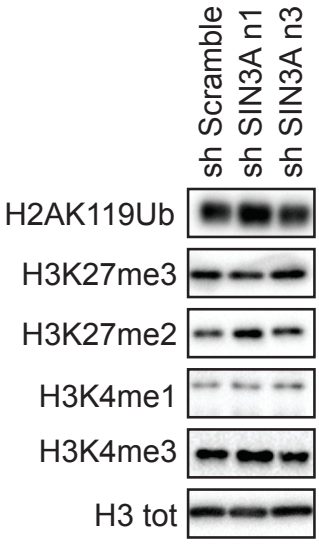

Fig.22 SIN3A silencing did not cause any change in histone modifications levels. Western blot of HCT-116 lysates infected with two independent shRNA targeting SIN3A.

3.10 SIN3A is associated with actively transcribed regions in HCT-116

I next wished to determine the genome-wide chromatin binding profile of SIN3A in HCT-116 cells. To do this, I took advantage of previously published data available in ENCODE of genome-wide enrichment profiles of SIN3A, H3K4me1, H3K4me3 and H3K27me3 (Fig.23). This analysis revealed that SIN3A binds on the majority of H3K4me3-positive promoters in HCT-116 cells. Interestingly, the majority of SIN3A sites are located at promoter regions, but a significant part is also bound to enhancers.

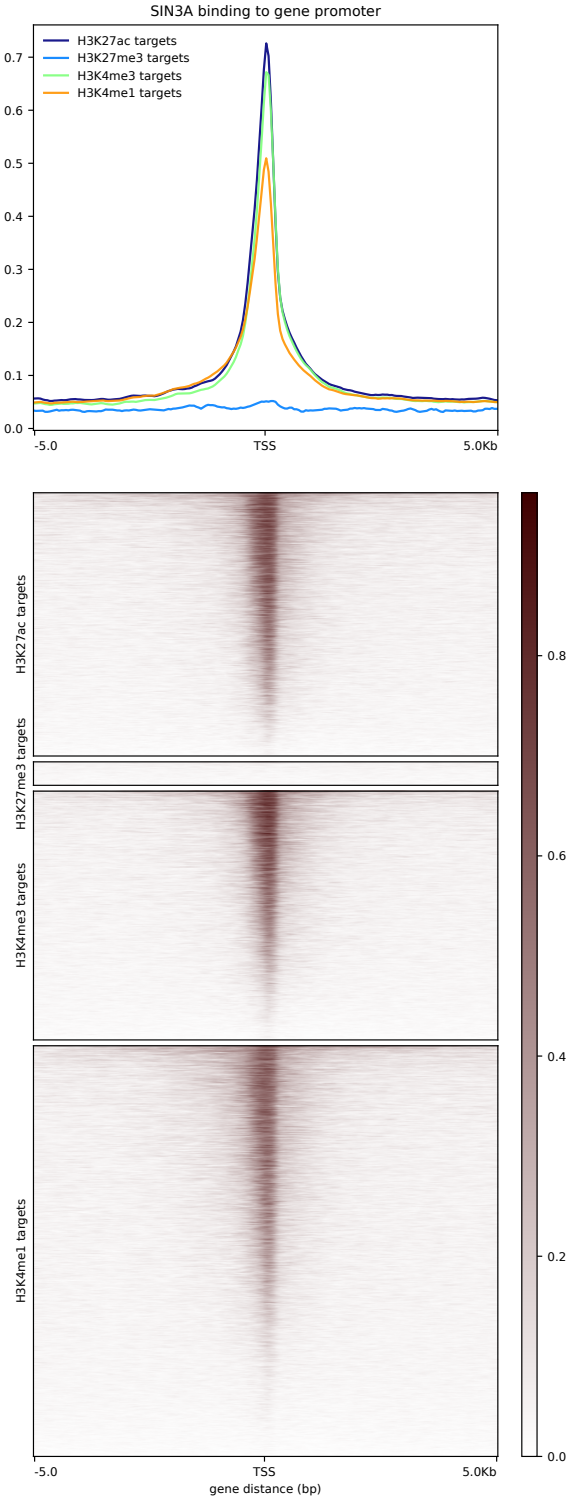

Fig.23 SIN3A localize at active promoters and enhancers. Heatmap analysis representing previously published ChIP-Seq data sets for SIN3A and the histone modifications H3K4me1, H3K4me3 and H3K27me3. The sequence reads 5,000 bp up- and downstream of the transcriptional start site are shown, and the relative intensities are indicated in brown.

3.11 Bulk RNA-seq in HCT-116 cells revealed one biological process downregulated upon SIN3A silencing

To further investigate the phenotype arising from the knockdown of SIN3A in HCT-116 cells, and to discovery if it was linked to global transcriptional changes, I performed a total RNA sequencing. The analysis performed by a bioinformatician in the lab, depicted in Fig.24, revealed that only few genes were differentially expressed.

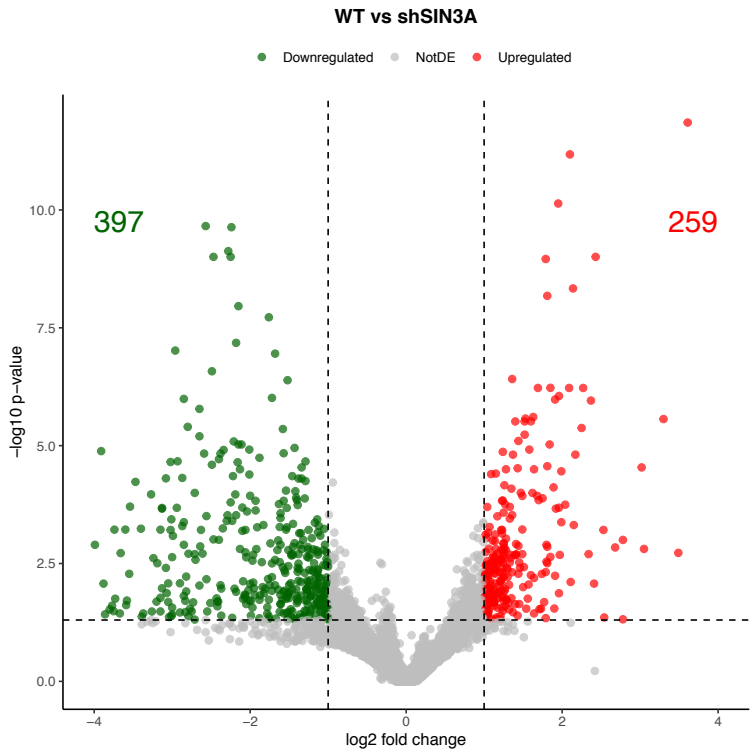

Fig.24 SIN3A knockdown does not result in huge transcriptional changes. Volcano-plot representing upregulated and downregulated genes of HCT-116 cells upon SN3A silencing.

I then decided to deeply analyze the specific regulated genes. The analysis revealed that among upregulated genes there were no enriched pathways. The downregulated genes instead revealed a biological process significantly downregulated, represented by proteins involved in antigen presentation.

3.12 KMT2D loss arrests cells growth but it does not affect the stability of WDR5 subunit

To dissect at a molecular level the effect of MLL4 complex loss, I focused on KMT2D subunit. First, I performed Western Blot on HCT-116 cell line analyzing the levels of its scaffold protein, WDR5, to check whether KMT2D loss affect the stability of other subunits of MLL4 complex. I could not detect any significant change in its levels (Fig.26 A). This could be explained by the fact that this protein still interacts to others to form several epigenetic complexes involved in different functions[186].

Moreover, in order to investigate if the phenotype observed in these cells could be related to cell cycle block, I investigated the levels of P21, observing a significant increase of this protein upon knockdown of KMT2D cells.

Given this fact, to understand if p21 upregulation could be also involved in the huge decrease in size detected in β-catenin stabilized organoids, I performed RT-qPCR for p21 expression on cDNA extracted of two independent replicates of organoids infected with two shRNA targeting Kmt2d and the control shRNA. I observed that also β-catenin stabilized organoids shows an increase in the expression of p21 (Fig.26 B). This suggest a scenario in which upon KMT2D loss, cells could arrest in a mechanism in which p21 could be involved.

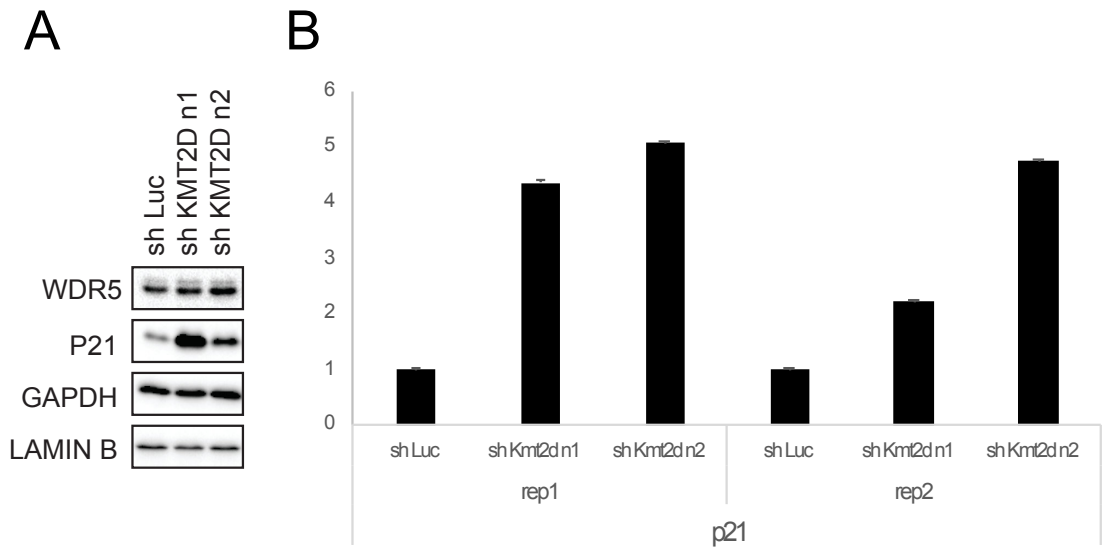

Fig.25 KMT2D silencing cause an upregulation of P21 in HCT-116 and in β-catenin stabilized organoids. A) Western blot of HCT-116 lysates infected with two independent shRNA targeting KMT2D. B) RT-qPCR analysis of β-catenin stabilized organoids infected with two independent shRNA targeting Kmt2d, showing upregulation of p21 compared to control.

3.13 Loss of enhancer-associated Compass complex leads to an increase of Polycomb activity

Since KMT2D is the major mono-methyltransferase that mediates H3K4 mono-methylation at the level of active enhancers[181] I investigated H3K4me1 and H3K4me3 modifications levels together with other histone modifications whose effect is to affect transcriptional regulation in a negative fashion.

While there were no significant changes in H3K4me1, probably explained by the partial knockdown of KMT2D (Fig. 27A), surprisingly I observed a huge increase in H2AK119Ub and H3K27me3 repressive modifications (Fig.27 B). I analyzed if the levels of BAP1 deubiquitinase and RING1B ubiquitin ligase, which respectively control the eradication and the deposition of this specific modification where affected. This was done to understand if the upregulation of the H2AK119Ub and H3K27me2/3 was related either to enhanced activity of PRC1 and PRC2 Polycomb Repressive Complexes, the two epigenetic complexes that mediate the deposition of such modifications, or by a deregulation of the expression of the epigenetic players associated with the deposition of such marks. I did not detect any

change in quantity of these specific proteins, concluding that the increase of the modification is linked to a hyperactivation of PRC1/2. Importantly, these results suggested that KMT2D silencing cross-talks with a strong increase of the H3K27me2/3 and H2AK119Ub levels, without affecting the overall levels of PRC1 and PR-DUB components.

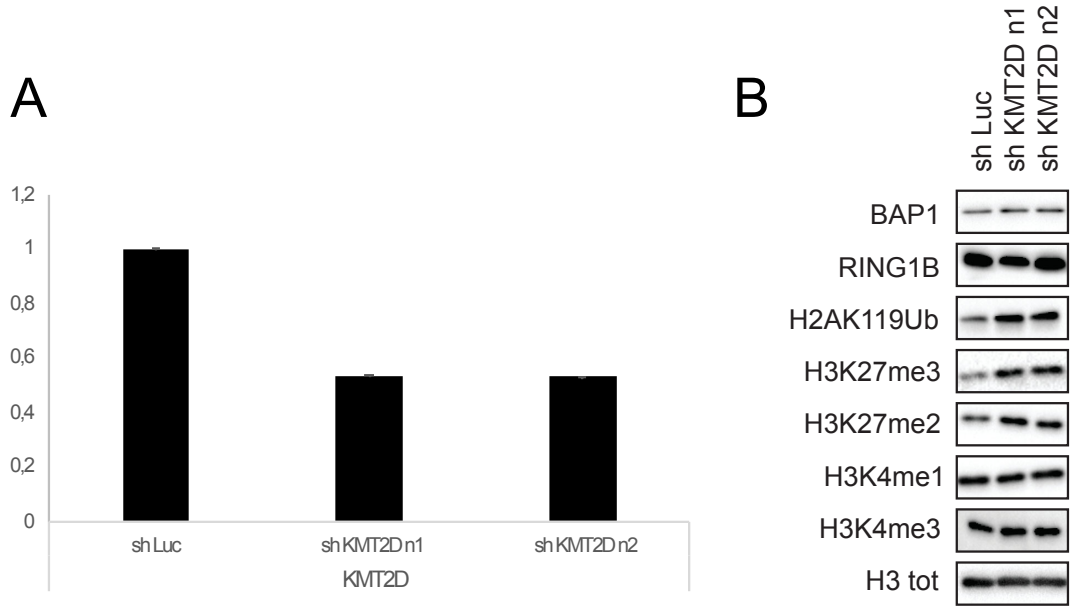

Fig.26 KMT2D silencing cause an increase in activity in Polycomb activity. B) Western blot of HCT-116 lysates infected with two independent shRNA targeting KMT2D.

3.14 Bulk RNA-seq in HCT-116 cells upon KMT2D silencing revealed the upregulation of P53 signaling pathway

To further investigate the block in cell growth arising from the knockdown of KMT2D in HCT-116 cells, and to discovery if it was linked to a regulation of global transcription, I performed a total RNA sequencing on such cells. The outcome of this experiment, analyzed by a bioinformatician in the lab, depicted in Fig.28, revealed little changes in the transcription of these cells. To understand if the few regulated genes belong to specific pathways I took advantage of Enrichr online tool. Among the upregulated genes, I was able to find proteins related to TP53 signaling pathway. This raises the possibility in which HCT-116 cells could undergo a massive block in growth mediated by TP53, that could activate its

target p21, that I previously observed as upregulated both in murine intestinal organoids and HCT-116 cells (Fig.26 A,B).

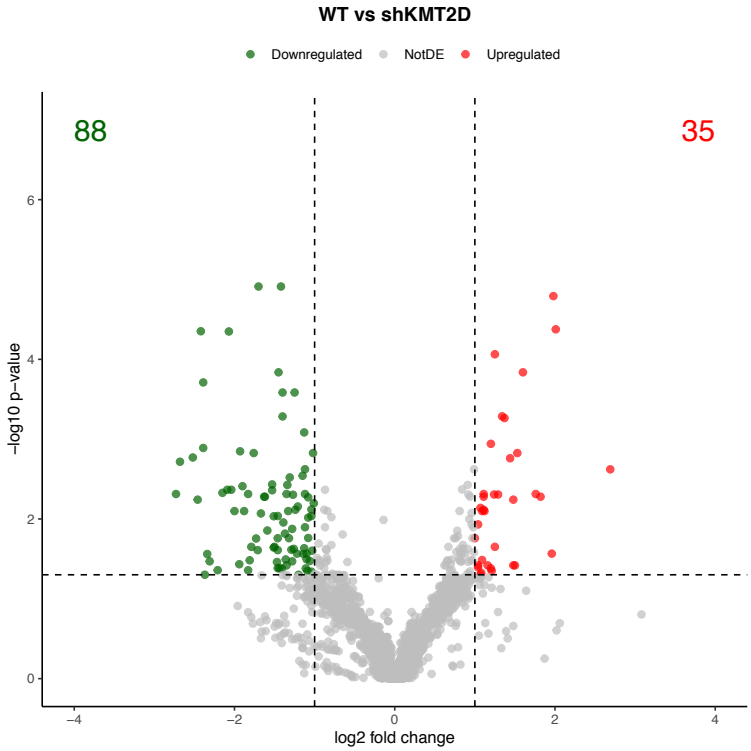

Fig.27 KMT2D knockdown causes very small transcriptional changes. Volcano-plot representing upregulated and downregulated genes of HCT-116 cells upon KMT2D silencing.

3.15 SIN3A and MLL4 complexes are essential for proliferation of patient derived metastatic organoids

The implementation an increasing number of therapeutic options for the treatment of metastatic colorectal cancer has substantially improved the overall survival, but until now, activating RAS mutations is the only validated biomarker that precludes patients from anti-EGFR treatment. Despite the advances so far, life expectancy unfortunately continues to be limited in the majority of patients with metastatic colorectal cancer. Late stages of tumorigenesis are characterized by the presence of additional oncogenic mutations and new clinical and molecular biomarkers to predict efficacy and tolerability are urgently needed. To address the therapeutic potential of the selected genes on metastatic colorectal cancer I investigated the effects induced by the silencing of the candidates in human patients-derived metastatic organoids. Given that murine organoids only harbor one driver mutation, I

wondered if the reduced proliferation observed in β-catenin stabilized organoids upon Sin3a, Brms1 and Kmt2d knockdown could be observed also in human metastatic intestinal organoids. To accomplish this purpose, I took advantage of two patients-derived organoids lines characterized by different gene mutations reported in Fig.12, respectively 322 and 327 patient-derived lines.

Organoids infected with SIN3A knockdown showed immediately severe phenotypic effects causing a global arrest in cells growth (Fig.29 A) in both patient lines. KMT2D and BRMS1 loss caused negative effects on cells growth comparable with the silencing in murine organoids and in HCT-116 cell line (Fig. 29 B,C). The results provided preliminary evidences of the therapeutic potential of these candidates in CRC treatment.

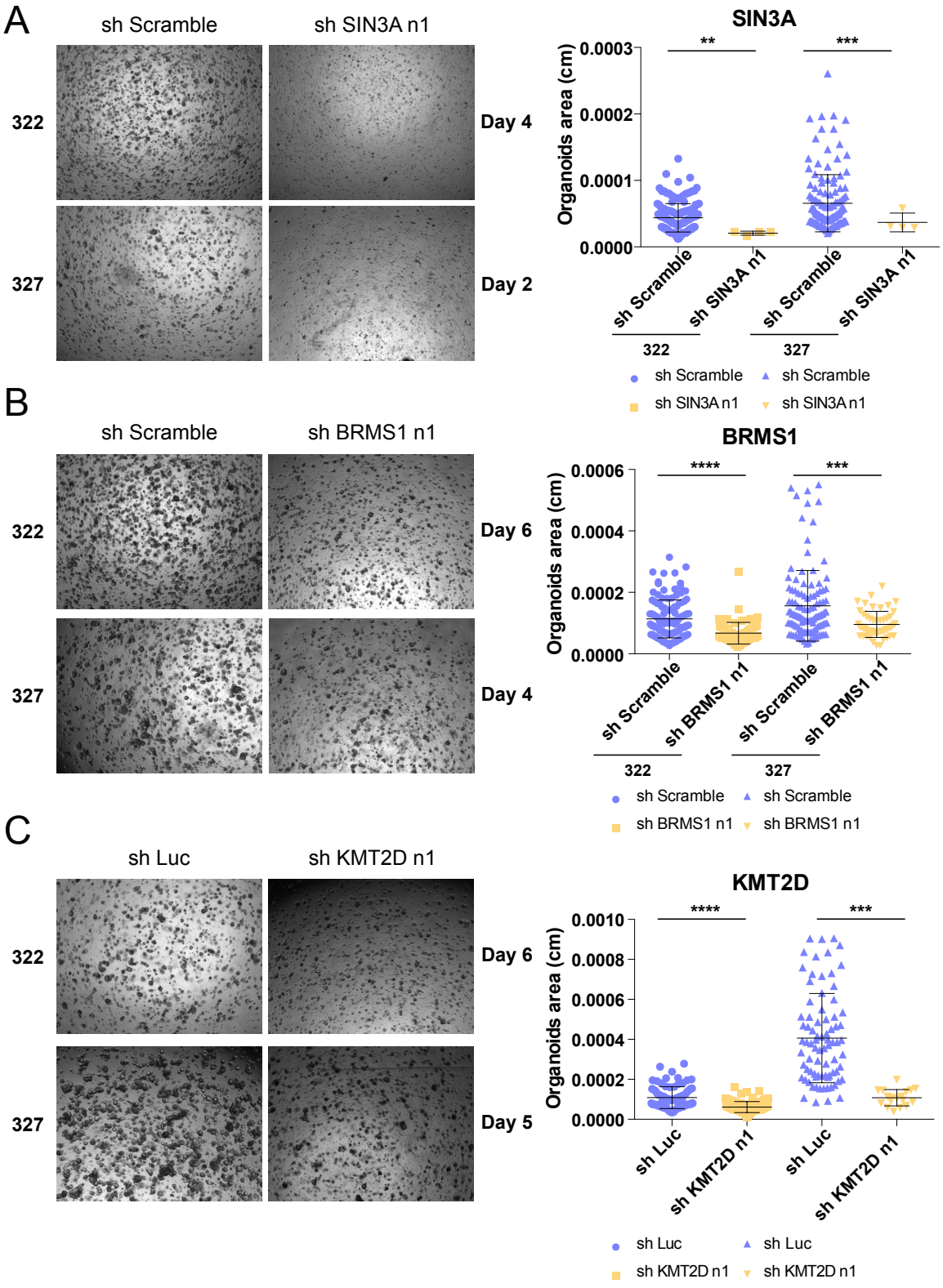

Fig.28 Patients-derived organoids infection with shRNAs targeting BRMS1: Left: Pictures of organoids at day 6 of the growth curve assay after organoids infection with control shRNA and two independent shRNA targeting BRMS1. Briefly, cells were infected and selected with puromycin for 72 hours, then organoids were collected, disrupted into single cells and replated to perform a growth curve analysis. Right: Dot-plots depicting the representation of the distribution of all the organoids areas. To assess statistically significative differences, Mann-Whitney test was applied.

3.16 Pharmacological inhibition of histone deacetylase activity causes detrimental effect on growth in human metastatic organoids mimicking SIN3A and BRMS1 knockdown

Another task of the project was to investigate the therapeutic potential of the candidates in CRC treatment trough the comparison of SIN3A, BRMS1 with the treatment with an epigenetic inhibitor, the Valproic Acid.

To do so, I treated metastatic organoids for 72 hours with this drug, an FDA-approved anticonvulsant which acts as a direct pan-histone deacetylase (HDAC) inhibitor[187]. This is due to the fact that no compounds targeting specifically SIN3A are available in the market, so to pharmacologically inhibit the complex, I targeted one of its catalytic cores, HDAC1/2.

First, I assessed the efficacy of the treatment through Western Blot analysis of samples collected at 24 hours, utilizing H3K27 acetylation as indicator of global inhibition of deacetylase activity. Notably, I observed that even at a low drug concentration this modification resulted increased in both patients (Fig 30 B,C).

Treatment carried out at 72 hours revealed that increasing concentrations of this epigenetic drug globally affects both 322 and 327 organoids growth (Fig.30 A,F).

Next, I wondered whether the effects in growth observed could be related to apoptosis. To understand this a technician in the lab performed FACS analysis of treated organoids at 72 hours stained with DAPI. This revealed that the percentage of death cells was low (Fig.30 D,E), suggesting that the observed phenotype could be caused by a growth arrest rather than cells death.

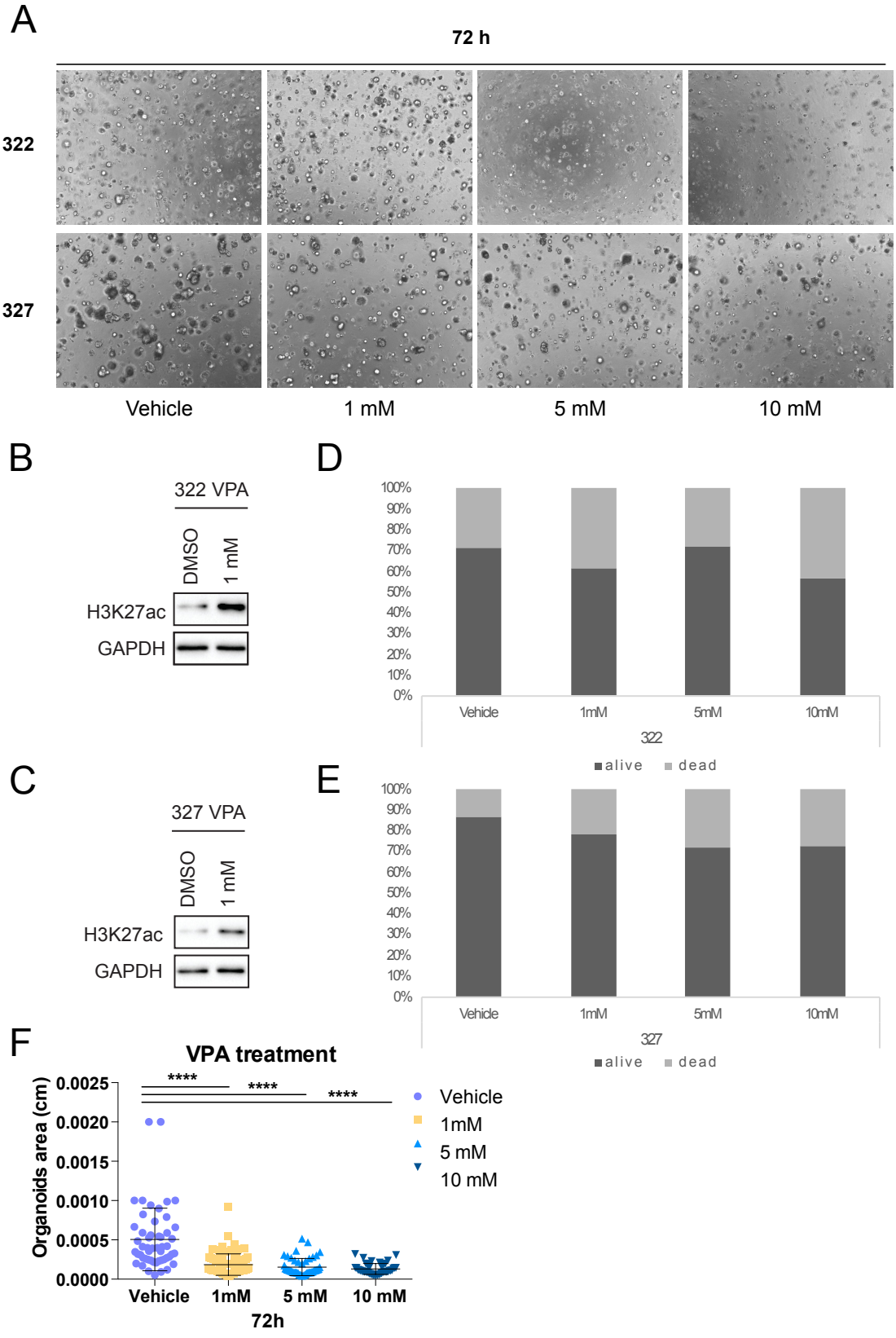

Fig.29 Patients derived organoids treated with Valproic Acid show a global decrease in growth. A) Pictures of treated organoids at 72 hours B,C) Western Blot assay verifying the efficacy of treatment with Valproic Acid. D,E) FACS analysis of DAPI-stained treated organoids 72 hours. F) Dot-plot depicting the representation of the distribution of all the organoids areas at 72 hours post-treatment. To assess statistically significative differences, Mann-Whitney test was applied.

3.17 Treatment with JQ-1 results in global decrease in metastatic organoids growth displaying similar effects of KMT2D knockdown

JQ-1 represents a potent, highly specific and competitive inhibitor of the Bromodomain and Extra-Terminal motif (BET family)[188], which are involved in the activation of enhancer activity through reading H3K27ac.

KMT2D is also involved in regulation of enhancer activity through the deposition of a specific histone modification, H3K4me1. However, no drugs are currently available to target KMT2D directly. To inhibit the same enhancer pathway through other means I made use of JQ-1 on metastatic organoids in increasing concentrations to investigate its effect on organoids growth. To assess the efficacy of the treatment, I collected organoids at 48 hours and performed RT-qPCR analysis assessing the levels of IFRD2, a MYC target, since it is reported that JQ-1 effect on cell growth has been linked to MYC downregulation[189]. In both the patients, IFRD2 levels were significantly downregulated with the treatment (Fig. 27 B,C). Notably, JQ-1 treatment resulted in a global impairment on organoids growth after 72 hours (Fig.27 A).

FACS analysis (Fig.27 D,E) revealed that the number of dead cells was definitely lower compared to the number of alive cells, supporting a scenario in which the phenotype observed is caused by cell cycle arrest.

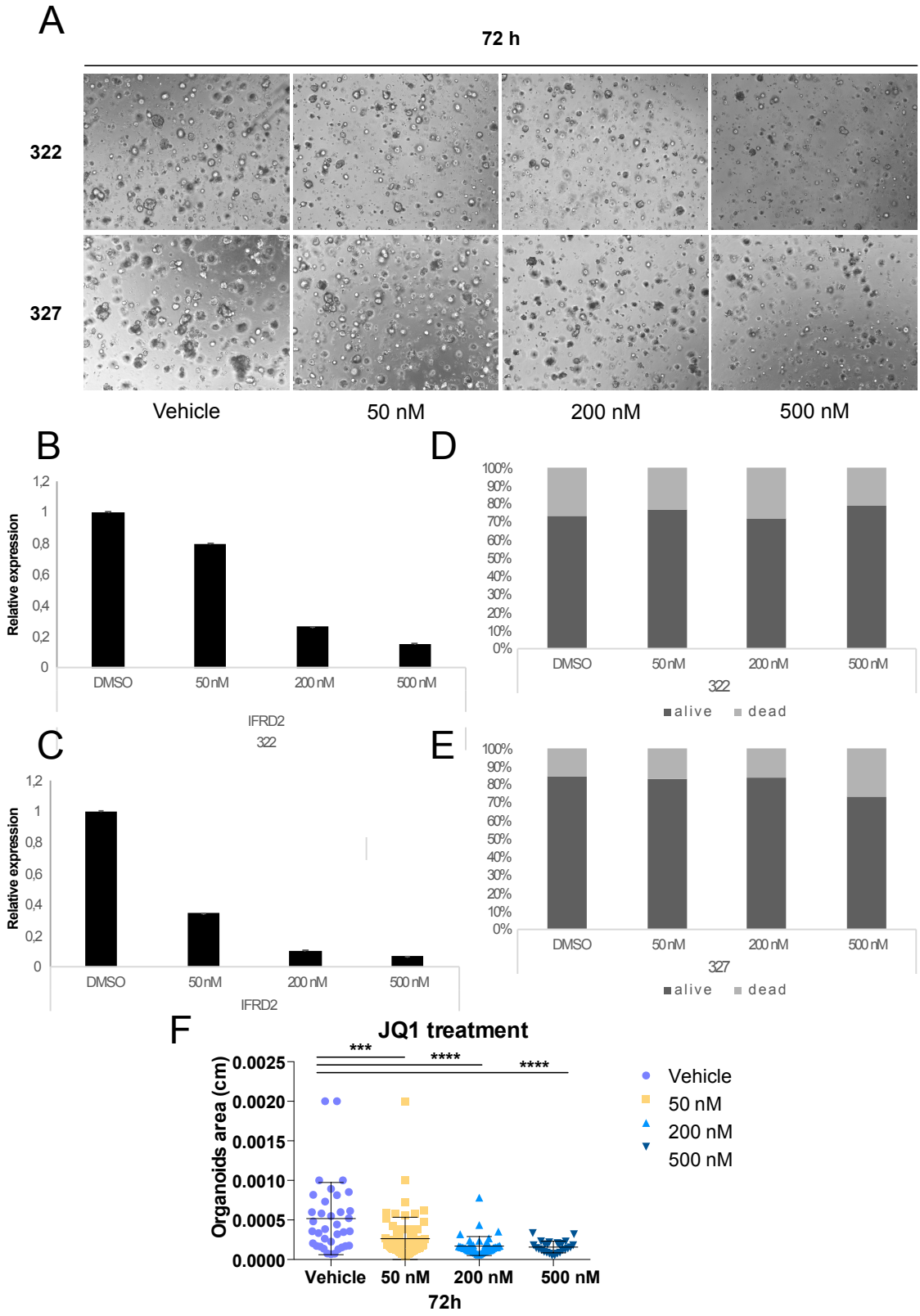

Fig.30 Treatment with JQ1 of patient derived organoids revealed a huge block in cell growth A) Pictures of treated organoids at 72 hours B,C) Western Blot assay verifying the efficacy of treatment with Valproic Acid. D,E)FACS analysis of DAPI-stained organoids 72 hours post-treatment. F) Dot-plot depicting the representation of the distribution of all the organoids areas at 72 hours post-treatment. To assess statistically significative differences, Mann-Whitney test was applied.

DISCUSSION

Despite impressive improvements in colorectal cancer survival over the last decades, the 5 years survival rate for metastatic disease remains approximately 10%. The necessity to dissect the molecular mechanisms that induce CRC formation and relapse as well as to find new drugs to treat more effectively this type of cancer are major open questions. Inhibition of upstream signaling pathways often leads to unspecific effects due to the high grade of cross talk and, therefore, alternative strategies are required to generate more direct and specific treatments. Since chromatin regulators are specific downstream players of these signals by modulating gene expression, they are attractive and little explored targets for cancer growth also for pharmacological intervention. Chromatin modifiers are an heterogeneous class of proteins with diverse druggable activities, useful to identify new regulators of ISCs self-renewal and proliferation as well as novel regulators that integrate in the nucleus the signals of WNT pathway, the contributor to CRC development.

Here I show that I have identified novel potential epigenetic targets by means of RNAi pooled and arrayed screens performed *in vitro* in 3D intestinal organoids.

With the arrayed approach I validated some of the epigenetic players regulated in the mutant models *in vitro* and I have demonstrated that Prmt1 is essential for colorectal cancer growth. In particular, it has already been demonstrated that Prmt1 is implicated in tumor progression[190,191], but up to now no mechanistic data are available in colorectal cancer. In my system, Prmt1 silencing significantly reduces cell proliferation in APC KO and β-catenin stabilized 3D organoids, suggesting that its inhibition can be important to block cancer progression.

Despite the high selectivity of targeted therapy, unpredictable side effects and toxicity in normal cells can emerge[192]. Therefore a crucial point is to test the effects of Prmt1 silencing on normal cells. My data indicates that Prmt1 depletion does not influence the proliferative potential of the wild-type 3D organoids, suggesting that Prmt1 is selectively responsible of cell survival and proliferation in cells with WNT pathway stably activated.

I then dissected its therapeutic potentiality in patients expanding my observations in a subset of human metastatic colorectal cancer organoids that were available in the laboratory. The genetic depletion of Prmt1 was still able to arrest the growth also in presence of other mutations, revealing that Prmt1 could be a promising target also at late stage of carcinogenesis, potentially filling the gap in the therapeutic strategies present up to now in such advantage stage cancer.

Treatment with a PRMTs type I inhibitor, MS023, revealed that the catalytic activity of Prmt1 exerts a central role in the growth inhibition observed in murine organoids.

Although the drug used in these experiments is not specific for Prmt1, it is still the best option to inhibit this arginine methyltransferase so far. Other inhibitors available in the market, such as TC-E 5003 and C21, show a higher specificity towards Prmt1 activity, but really low inhibition efficiency for the methyltransferase activity comparing to MS023[193,194]. Indeed, TC-E 5003 and C21 display an IC_{50} of 1,5 μM and 1,8 μM, respectively, whereas MS023 inhibits Prmt1 activity at nanomolar (IC_{50} of 30 nM) concentrations. However, nowadays MS023 does not represent a therapeutic option due to its lack of specificity.

Without a clear knowledge related to the implications of the inhibition of the catalytic activity of Prmt1, it is still too preliminary to define the reason why Prmt1 inhibition blocks organoids growth. It is particularly difficult to speculate which could be the key Prmt1 targets involved in the block of growth observed. In fact, such enzyme is responsible for at least 85% of all asymmetric arginine dimethylation reactions within the cell and other type I arginine methyltransferases do not show major compensatory activity in terms of re-accumulation of asymmetric dimethylation in absence of Prmt1[175].

To overcome this, future experiments that will combine transcriptomic analyses of Prmt1 depletion in WT, APC KO and β-catenin stabilized organoids will be performed to unravel the precise effect of the loss of this protein in a well-defined genetic background and to understand an eventual cross-talk between WNT deregulated pathway and this specific arginine methyltransferase. The interplay between PRMT1 and β-catenin is indeed still not

clear. It is reported that in 293T cells PRMT1 acts interacting and methylating the 378th arginine residue of Axin, a negative regulator of Wnt signaling and a key scaffold protein for the β-catenin destruction complex. Knockdown of endogenous PRMT1 by short hairpin RNA reduced the level of Axin, thus enhancing the level of cytoplasmic β-catenin as well as β-catenin-dependent transcription activity[195]. On the other side, in esophageal squamous cell carcinoma (ESCC) the elevated expression of PRMT1 facilitated the expression of stem cell-like properties and tumorigenicity activating Wnt/β-catenin and Notch signaling pathways, whereas knockdown of PRMT1 significantly diminished the self-renewal properties of ESCC[196]. This two works raise the hypothesis of a tissue specificity influencing the methylation pattern of PRMT1, but further experiments are needed to address the interplay between PRMT1 and WNT pathway in the intestinal compartment.

Moreover, the evaluation of the mechanisms through trascriptomic analyses on a plethora of human metastatic organoids characterized by different mutational background both in presence of Prmt1 silencing and upon treatment with MS023 will provide some insights on the molecular mechanism on a clinical relevant context.

Finally it is possible to envision dissecting thoroughly the mechanism of this inhibition taking advantage also of proteomic approaches performed on colorectal cancer cell lines, such as immune affinity enrichment for asymmetric dimethylated peptides, to understand through analysis of methylated peptides which could be the key target differentially methylated upon Prmt1 silencing.

The second approach, a pooled epigenetic screen on 3D intestinal organoids, has also unravelled key epigenetic players in colorectal cancer tumorigenesis. Since this has never been attempted before, I established such experimental procedure with intestinal organoids, generating both important technological advances that could be applied to other tissues and tumors as well as novel knowledge related to the molecular circuits that regulate colorectal cancer development. I took advantage of a published protocol in which MCF10DCIS.com

cells were screened with this specific epigenetic library[197]. Although I used a 3D mouse model that presents a lot of technical challenges, I was able to represent the complexity of the system and contextually reduce the possibility of detecting false positive depleted genes, as shRNAs not belonging to the epigenetic library were included in the validation analyses where possible. Moreover, I showed that a relevant portion of identified genes is represented by several subunits of coherent functional complexes.

The first epigenetic complex that came out from the screen is Sin3-HDAC complex. This complex is highly conserved and function in distinct cellular processes including cell-cycle regulation[198], maintenance of stem cell pluripotency[199], self-renewal and cellular differentiation. Mammalian Sin3 complex is composed by six core subunits: HDAC1/2, RbAp46, RbAp48, SIN3A/SIN3B, SAP18 and SAP30[200]. More recent studies identified additional subunits among with BRMS1[201]. This complex is known to associate also to RBP2/KDM5A in order to cooperate in repressing target genes trough, deacetylation, demethylation and hypothetically, repositioning of nucleosomes[202].

I validated two different subunits of this complex: Brms1 and Sin3a. Interestingly, other proteins belonging to Sin3-HDAC functional complex came from the functional screening: Sin3b and Jarid1a, a Jumonji H3K4 demethylase constituting one of the two different catalytic cores of such epigenetic complexes. Due to the promising data emerged from this project these two candidates will be for sure included in the validation process and investigated in future experiments.

Breast metastasis suppressor gene 1 (BRMS1) was discovered as a potential tumor suppressor gene, significantly reducing the metastatic capacity of breast cancer cells[203]. Further studies identified BRMS1 involved in metastases suppression not only in breast cancer, but also in non-small cell lung cancer[204], ovarian cancer[205] and glioma[206]. However, the role of such protein in colorectal cancer has never been investigated.

I surprisingly confirmed in the validation of the selected candidates identified in the screen that Brms1 silencing compromised cell growth both in β-catenin stabilized murine organoids, in HCT-116 cell line and also in human metastatic organoids in all the independent biological replicates performed. This was totally unexpected considering what is reported in the current literature, and since its exact function in Sin3-HDAC complex has not been described yet, further experiments are needed to thoroughly elucidate the role of BRMS1 in CRC pathogenesis.

The other subunit of Sin3-HDAC that was validated is SIN3A. This protein harbors 6 different domains crucial for its interaction with a large number of transcriptional factors, among with are present four paired amphipathic helix (PAH) domains, one Highly conserved Region (HCR) and an Histone Deacetylase interaction domain (HID)[42]. PAH1 and PAH2 domains serve as interaction domains with various transcription factors, while HID, PAH3 and PAH4 shows scaffolding functions interacting with other subunits of the co-repressor complex[207].

The silencing of this protein showed huge detrimental effects on growth in murine organoids, HCT-116 cells, and also in human metastatic organoids. Further studies could possibly address the therapeutic potential of this complex for the development of CRC treatments.

The role of Sin3a in cancer is still ambiguous as independent studies reported this protein as both oncogene and tumor suppressor. Several studies reported Sin3a as essential for survival of transformed cells[208], and interference with its function leaded to epigenetic reprogramming and differentiation in breast cancer cells[209]. Another work revealed Sin3a as a negative regulator for tumor progression in *Drosophila*[210]. Its role in colorectal cancer it is not yet characterized.

SIN3A knockdown in HCT-116 cells slightly induced HDAC1 destabilization, although this enzyme takes part in multiple functional complexes including NuRD, CoREST and NCoR/SMRT complexes[184], thus revealing a potential function of SIN3A scaffold in stabilizing the catalytic core. Additional analyses on several histone modifications levels

revealed that no significant changes were detected upon SIN3A inhibition, raising the possibility that other epigenetic complexes may compensate the levels of modifications regulated by this specific complex.

To gain more insights on the phenotype observed, global transcriptional changes were observed trough RNA sequencing on HCT-116 cells. The results surprisingly revealed that only few genes were differentially expressed: no enriched pathways were identified, while among down-regulated genes were identified components involved in antigen presentation. In particular almost all the components of major histocompatibility complex (MHC) class II were down regulated.

MHC II is traditionally involved in presenting peptides from antigens located in the extracellular space, binding to CD4 T-cells, thus leading to their activation. It is known that the class II transactivator (CIITA), the master transcriptional regulator of MHC class II promoters is regulated in a negative fashion by HDAC1/HDAC2. Inhibition of HDACs enhances IFN-γ-induced MHC class II expression, Notably, Sin3a and Sin3b are important to reinforce the inhibitory action of histone deacetylases[211,212].

These results could be against the current literature, but it is important to stress that SIN3A alone was knocked down, whereas SIN3B is still active. It is also important to mention that I have no information about the status of HDAC2 upon SIN3A knockdown, so it could still exert its negative regulation on MHC class II promoters.

To demonstrate the therapeutic potential of SIN3-HDAC complex in future clinic applications, I took advantage of an epigenetic drug, Valproic Acid (VPA) that has been approved for treatment of hematologic malignancies[213] and is currently in clinical trial for colorectal cancer. Since no direct inhibitors for SIN3A and BRMS1 are available on the market I took advantage of this well-known inhibitor of class I and II histone deacetylases[187], to target one of the two catalytic cores of SIN3-HDAC complex. Although Valproic Acid does not constitute a specific inhibitor for HDAC1/2, the two enzymes constituting the

deacetylase core of this complex, it is one of the most specific HDAC inhibitors available on the market so far. Current HDAC inhibitors indeed act specifically against a class of HDACs, or against all types of HDACs (pan-inhibitors)[149].

Overall, the treatment clearly shows that the pharmacological inhibition of histone deacetylase activity mimics SIN3A knockdown, causing detrimental effects on growth in human CRC metastatic organoids, thus raising a hypothesis in which this complex could represent a promising target in the treatment of late stages of CRC. Further experiments analyzing global transcriptional changes of human metastatic organoids in presence of both SIN3A knockdown and VPA treatment could highlight the common regulated pathways and thus provide mechanistic hints about the mechanism causing a block in growth in such tumor upon regulation of SIN3A-HDAC complex.

The second epigenetic complex that underwent the validation process was Mll4 complex. MLL4 in contained in one of the two COMPASS-like complexes and is associated with development-specific genes, principally contributing to monomethylation of H3K4[181,214]. KMT2C and KM2D (MLL3 and MLL4 respectively) are the principal monomethyltransferases of H3K4, and they generally exert their function in a non-redundant manner, although there may be a significant overlap on enhancers in various cells and tissues. Active enhancers will come in contact with related promoters through a mediator complex and cohesin proteins, activating RNApolII activity and thus facilitating transcription. Enhancer regions are enriched with modifications on H3K4 and H3K27. Active enhancers contain both H3K4me1 and H3K27ac marks, whereas inactive enhancers are enriched with H3K4me1 and H3K27me3[215].

I validated two different proteins of this H3K4 monomethytransferase-containing complex: Wdr5 and Kmt2d.

WD-repeat protein-5 (WDR5) is a highly conserved WD40 repeat-containing protein essential for the regulation of multiple biological processes. Is a member of several

epigenetic complexes, among with the NSL, a histone acetyltransferase complex, and the COMPASS complexes. COMPASS complexes are composed by six non-redundant mammalian SET1/MLL HMTs, each with a distinct regulatory role[216-218]. Each of these comprises a common core set of proteins known as WRAD: WDR5, RBBP5, ASH2L and DPY30[219]. WDR5 overexpression has been found overexpressed in several tumors, like leukemias[220], breast cancer[221], or bladder cancer[222]. However, although WDR5 could appear a promising therapeutic target, its inhibition shows an impact also on normal cells due to its role in many biological processes, such as regulating the activation of HOX genes in human cells[223] or preserving stemness and self-renewal in mESCs[224].

WDR5 silencing in murine organoids and HCT-116 cells induce dramatic effects on cells viability, but since it is established that it has an essential role both in cancer and normal cells I prevalently focused my attention on the other subunit of the Mll4 complex, KMT2D. The role of KMT2D in cancer is controversial: some evidences suggest that MLL3 and MLL4/COMPASS family could act as tumor suppressors[225] and their mutations promote tumorigenesis in a tissue-specific manner. Trr, the *Drosophila* homolog of MLL3/MLL4, was reported to negatively regulate cell proliferation[226]. Another *in vivo* study identified MLL3/MLL4 as a coactivator for p53 to induce expression of p53 target[227]. More recent works provided a different angle of MLL4 involvement in cancer: it was shown that KMT2D reduction in medulloblastoma and colon cancer cells resulted in a global reduction of proliferation[228,229]. Taken together, all these studies clearly provide an indication that role of MLL4 in cancer is context dependent.

KMT2D silencing in murine organoids harboring β-catenin stabilization and in colon cancer cell line drastically reduces cell growth, in a mechanism sustained by the upregulation of P21. This is further sustained by global transcriptional data that although there was not a significant increase in P21, clearly indicates an upregulation of the p53-signaling pathway. Silencing KMT2D in metastatic CRC organoids infection allowed verifying if the block in proliferation was maintained in human patient-derived metastatic organoids. These

organoids importantly carry additional oncogenic mutations, among with p53 loss of function, which could confer resistance to the phenotype observed in the other models utilized. Strikingly, KMT2D depletion completely arrests cell growth and cell proliferation also in metastatic patient-derived organoids.

Both the human metastatic organoids lines utilized in KMT2D silencing experiments were mutated on TP53 locus and, although the status of P21 has not been investigated yet, it is possible to envision two different scenarios.

The first possibility is that p21 could be still activated in a mechanism p53-independent. In the majority of the cases p53-mediated growth inhibition is dependent on induction of p21, an inhibitor of cyclin-dependent kinases that is required for cell cycle progression. Usually, in cells in which p53 is null, expression of p21 diminishes strikingly. However, an absolute correlation between p21 expression and the status of the p53 gene in cancers is lacking and is consistent with other studies that have suggested that p21 may also be regulated by p53-independent pathways under different stimuli. Examples include Chk2-dependent senescence and p21 transcriptional induction in breast carcinoma and immortalized keratinocytes, which are p53-defective[230]. Or alternatively, ascofuranone administration in human cancer cells deficient in p53 that results in p21 upregulation through p53-independent suppression of c-Myc expression, leading to cytostatic G1 arrest[231].

A second explanation can be provided by a recent work in which it has been demonstrated that breast cancer cells harboring an TP53 gain of function mutant that resulted from a missense mutation in its DNA-binding domain, the same mutation present in both patient-derived organoids utilized in the project, interacts with a set of chromatin regulators among with MLL1 and MLL4. Knockdown of MLL1 or MLL4 in gain-of-function p53 breast cancer cells severely reduced cell growth, phenocopying knockdown of gain-of-function p53[232].

To gain some molecular insights about the mechanism of growth inhibition, I investigated the stability of the complex, concluding that WDR5 stability is not affected by KMT2D depletion. This could probably due to the fact that WDR5 plays a central scaffolding role in the other five SET1/MLL histone methyltransferase complexes[186]. Then, the effect of KMT2D knockdown on chromatin marks was investigated, thus generating several observations.

First, although HCT-116 cell line harbors an inactivating mutation in *KMT2C* gene[181], I was not able to detect any difference in H3K4me1 signal, but this is probably due to the fact that the levels of knockdown obtained were not sufficient to observe global alterations in this modification upon KMT2D loss. Quantitative ChIP experiments will be useful to understand if there is a modulation for H3K4me1 signal in defined genomic regions, such as enhancers. Other PTMs investigated were chromatin repressive marks, like ubiquitination of H2A on lysine 119 (H2AK119Ub) and di-/tri-methylation of H3 on lysine 27 (H3K27me2/me3). Surprisingly, I observed consistent increase in the levels of such modifications without any change in the levels of Really Interesting New Gene 1B (RING1B), the ubiquitin E3-ligase of PRC1 that mediates the deposition of H2AK119Ub, and BAP1, the subunit of Polycomb Repressive-Deubiquitinase (PR-DUB) complex that erases this mark[233].

One of the possible explanations for such increase is that WDR5 is also a known interactor of one of the non-canonical PRC1 (PRC1.6)[234], raising the hypothesis that WDR5, free from KMT2D binding, could be recruited more to PRC1.6 specific complex. This could explain the increasing levels of H2AK119Ub observed in HCT-116 without any effect on the protein levels of BAP1 deubiquitinase and RING1B ubiquitin ligase. Since non-canonical PRC1 complexes have been found to recruit at chromatin level PRC2 through H2AK119Ub deposition[235], the concomitant increase in H3K27me2/3 can be explained as an expected consequence.

A different possible scenario relies on the stability of Mll4 complex. Among the KMT2D interactors in the complex, only WDR5 subunit stability was examined. Another subunit of

Mll4 complex is represented by UTX, an H3K27me3 demethylase. Both MLL3 and MLL4 complexes carry H3K27 demethylase activity[236]. Absence of the KMT2D protein results in the collapse of the KMT2D complex and the destabilization of UTX in cells[237], which could be causative of an increase of H3K27me3 levels. ChIP-seq analysis of the pattern of deposition of global H3K27me3 in combination with binding profiles with UTX could address and explain such increase in repressive mark.

To finally demonstrate the potentiality of Mll4 complex as clinical target in the treatment of CRC, I compared the effects of the silencing the treatment with an epigenetic drug available in the market, JQ1.

KMT2D in known to be involved in the deposition of a mark that mediates the regulation of enhancers, but no drug targeting specifically this protein are available so far. JQ1, a highly specific and competitive inhibitor of BRD4, one member of the Bromo-domain and extra-terminal motif (BET) family, is able to regulate the enhancer activity[188]. The use of JQ1 has demonstrated efficacy in targeting tumor cells, among with prostate cancer,[238]pancreatic adenocarcinomas[239], ovarian cancer[240]. Data demonstrated that the treatment with JQ1 is similar to KMT2D knockdown, causing a global impairment in metastatic organoids growth. However, in order to define them properly as therapeutic targets, it will be necessary to evaluate the effects of candidates silencing also in a wild-type context.

In conclusion, I demonstrated the efficacy of 3D murine and human organoid models as a powerful platform to elucidate molecular events governing cancer initiation and development, trough bot pooled and arrayed screen, thus generating a new technological advances that could be extended to other tissues and tumors.

My data provided new molecular insights on the role of specific epigenetic complexes acting in a WNT pathway constitutive activated background. The results demonstrated that not only SIN3A, KMT2D and BRMS1 silencing drastically stops cell proliferation in β-catenin stabilized murine organoids and human CRC cell line, but are also causative of negative

effects in patient-derived metastatic organoids growth, thus gaining a therapeutic potential in a tumor context in which efficient therapies are still lacking.

BIBLIOGRAPHY

1 Jemal, A. *et al.* Global cancer statistics. *CA: a cancer journal for clinicians* **61**, 69-90, doi:10.3322/caac.20107 (2011).

2 Haggar, F. A. & Boushey, R. P. Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clinics in colon and rectal surgery* **22**, 191-197, doi:10.1055/s-0029-1242458 (2009).

3 Arnold, C. N., Goel, A., Blum, H. E. & Boland, C. R. Molecular pathogenesis of colorectal cancer: implications for molecular diagnosis. *Cancer* **104**, 2035-2047, doi:10.1002/cncr.21462 (2005).

4 Kwong, L. N. & Dove, W. F. APC and its modifiers in colon cancer. *Advances in experimental medicine and biology* **656**, 85-106, doi:10.1007/978-1-4419-1145-2_8 (2009).

5 Vaughn, C. P., Zobell, S. D., Furtado, L. V., Baker, C. L. & Samowitz, W. S. Frequency of KRAS, BRAF, and NRAS mutations in colorectal cancer. *Genes, chromosomes & cancer* **50**, 307-312, doi:10.1002/gcc.20854 (2011).

6 Fleming, N. I. *et al.* SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer research* **73**, 725-735, doi:10.1158/0008-5472.CAN-12-2706 (2013).

7 Sansom, O. J. *et al.* Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes & development* **18**, 1385-1390, doi:10.1101/gad.287404 (2004).

8 Huang, D. *et al.* Mutations of key driver genes in colorectal cancer progression and metastasis. *Cancer metastasis reviews* **37**, 173-187, doi:10.1007/s10555-017-9726-5 (2018).

9 Dow, L. E. *et al.* Apc Restoration Promotes Cellular Differentiation and Reestablishes Crypt Homeostasis in Colorectal Cancer. *Cell* **161**, 1539-1552, doi:10.1016/j.cell.2015.05.033 (2015).

10 Marshman, E., Booth, C. & Potten, C. S. The intestinal epithelial stem cell. *BioEssays : news and reviews in molecular, cellular and developmental biology* **24**, 91-98, doi:10.1002/bies.10028 (2002).

11 Barker, N., van Oudenaarden, A. & Clevers, H. Identifying the stem cell of the intestinal crypt: strategies and pitfalls. *Cell stem cell* **11**, 452-460, doi:10.1016/j.stem.2012.09.009 (2012).

12 Merlos-Suarez, A. *et al.* The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell stem cell* **8**, 511-524, doi:10.1016/j.stem.2011.02.020 (2011).

13 Barker, N. *et al.* Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608-611, doi:10.1038/nature07602 (2009).

14 Komiya, Y. & Habas, R. Wnt signal transduction pathways. *Organogenesis* **4**, 68-75, doi:10.4161/org.4.2.5851 (2008).

15 Gordon, M. D. & Nusse, R. Wnt signaling: multiple pathways, multiple receptors, and multiple transcription factors. *The Journal of biological chemistry* **281**, 22429-22433, doi:10.1074/jbc.R600015200 (2006).

16 He, X., Semenov, M., Tamai, K. & Zeng, X. LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way. *Development* **131**, 1663-1677, doi:10.1242/dev.01117 (2004).

17 MacDonald, B. T., Tamai, K. & He, X. Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Developmental cell* **17**, 9-26, doi:10.1016/j.devcel.2009.06.016 (2009).

18 Mao, J. *et al.* Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Molecular cell* **7**, 801-809, doi:10.1016/s1097-2765(01)00224-6 (2001).

19 Zeng, X. *et al.* Initiation of Wnt signaling: control of Wnt coreceptor Lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions. *Development* **135**, 367-375, doi:10.1242/dev.013540 (2008).
20 Wharton, K. A., Jr. Runnin' with the Dvl: proteins that associate with Dsh/Dvl and their significance to Wnt signal transduction. *Developmental biology* **253**, 1-17, doi:10.1006/dbio.2002.0869 (2003).
21 Wallingford, J. B. & Habas, R. The developmental biology of Dishevelled: an enigmatic protein governing cell fate and cell polarity. *Development* **132**, 4421-4436, doi:10.1242/dev.02068 (2005).
22 Hatsell, S., Rowlands, T., Hiremath, M. & Cowin, P. Beta-catenin and Tcfs in mammary development and cancer. *Journal of mammary gland biology and neoplasia* **8**, 145-158 (2003).
23 Clevers, H. Wnt/beta-catenin signaling in development and disease. *Cell* **127**, 469-480, doi:10.1016/j.cell.2006.10.018 (2006).
24 Kinzler, K. W. *et al.* Identification of FAP locus genes from chromosome 5q21. *Science* **253**, 661-665, doi:10.1126/science.1651562 (1991).
25 Knudsen, A. L., Bisgaard, M. L. & Bulow, S. Attenuated familial adenomatous polyposis (AFAP). A review of the literature. *Familial cancer* **2**, 43-55 (2003).
26 Kinzler, K. W. & Vogelstein, B. Lessons from hereditary colorectal cancer. *Cell* **87**, 159-170, doi:10.1016/s0092-8674(00)81333-1 (1996).
27 Korinek, V. *et al.* Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**, 1784-1787, doi:10.1126/science.275.5307.1784 (1997).
28 Liu, W. *et al.* Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. *Nature genetics* **26**, 146-147, doi:10.1038/79859 (2000).
29 Morin, P. J. *et al.* Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* **275**, 1787-1790, doi:10.1126/science.275.5307.1787 (1997).
30 Bass, A. J. *et al.* Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nature genetics* **43**, 964-968, doi:10.1038/ng.936 (2011).
31 Sato, T. *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262-265, doi:10.1038/nature07935 (2009).
32 Pinto, D., Gregorieff, A., Begthel, H. & Clevers, H. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & development* **17**, 1709-1713, doi:10.1101/gad.267103 (2003).
33 Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274-284, doi:10.1016/j.cell.2013.07.004 (2013).
34 Haramis, A. P. *et al.* De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. *Science* **303**, 1684-1686, doi:10.1126/science.1093587 (2004).
35 Wong, V. W. *et al.* Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. *Nature cell biology* **14**, 401-408, doi:10.1038/ncb2464 (2012).
36 Hofmann, C. *et al.* Cell-cell contacts prevent anoikis in primary human colonic epithelial cells. *Gastroenterology* **132**, 587-600, doi:10.1053/j.gastro.2006.11.017 (2007).
37 Sato, T. *et al.* Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **141**, 1762-1772, doi:10.1053/j.gastro.2011.07.050 (2011).

38 van de Wetering, M. *et al.* Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933-945, doi:10.1016/j.cell.2015.03.053 (2015).
39 Crespo, M. *et al.* Colonic organoids derived from human induced pluripotent stem cells for modeling colorectal cancer and drug testing. *Nature medicine* **23**, 878-884, doi:10.1038/nm.4355 (2017).
40 Fujii, M. *et al.* A Colorectal Tumor Organoid Library Demonstrates Progressive Loss of Niche Factor Requirements during Tumorigenesis. *Cell stem cell* **18**, 827-838, doi:10.1016/j.stem.2016.04.003 (2016).
41 Drost, J. *et al.* Sequential cancer mutations in cultured human intestinal stem cells. *Nature* **521**, 43-47, doi:10.1038/nature14415 (2015).
42 Kadamb, R., Mittal, S., Bansal, N., Batra, H. & Saluja, D. Sin3: insight into its transcription regulatory functions. *European journal of cell biology* **92**, 237-246, doi:10.1016/j.ejcb.2013.09.001 (2013).
43 Roper, J. *et al.* In vivo genome editing and organoid transplantation models of colorectal cancer and metastasis. *Nature biotechnology* **35**, 569-576, doi:10.1038/nbt.3836 (2017).
44 Bernards, R. Finding effective cancer therapies through loss of function genetic screens. *Current opinion in genetics & development* **24**, 23-29, doi:10.1016/j.gde.2013.11.007 (2014).
45 Sims, D. *et al.* High-throughput RNA interference screening using pooled shRNA libraries and next generation sequencing. *Genome biology* **12**, R104, doi:10.1186/gb-2011-12-10-r104 (2011).
46 Cellecta. (2014).
47 Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868-871, doi:10.1126/science.184.4139.868 (1974).
48 Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *Journal of molecular biology* **319**, 1097-1113, doi:10.1016/S0022-2836(02)00386-8 (2002).
49 Luger, K. Structure and dynamic behavior of nucleosomes. *Current opinion in genetics & development* **13**, 127-135, doi:10.1016/s0959-437x(03)00026-1 (2003).
50 Ausio, J. The shades of gray of the chromatin fiber: recent literature provides new insights into the structure of chromatin. *BioEssays : news and reviews in molecular, cellular and developmental biology* **37**, 46-51, doi:10.1002/bies.201400144 (2015).
51 Wallrath, L. L., Lu, Q., Granok, H. & Elgin, S. C. Architectural variations of inducible eukaryotic promoters: preset and remodeling chromatin structures. *BioEssays : news and reviews in molecular, cellular and developmental biology* **16**, 165-170, doi:10.1002/bies.950160306 (1994).
52 Brown, S. W. & Nur, U. Heterochromatic Chromosomes in the Coccids. *Science* **145**, 130-136 (1964).
53 Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260, doi:10.1038/38444 (1997).
54 Zheng, C. & Hayes, J. J. Structures and interactions of the core histone tail domains. *Biopolymers* **68**, 539-546, doi:10.1002/bip.10303 (2003).
55 Kim, J. A., Hsu, J. Y., Smith, M. M. & Allis, C. D. Mutagenesis of pairwise combinations of histone amino-terminal tails reveals functional redundancy in budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5779-5784, doi:10.1073/pnas.1203453109 (2012).

56 Schuster, T., Han, M. & Grunstein, M. Yeast histone H2A and H2B amino termini have interchangeable functions. *Cell* **45**, 445-451 (1986).
57 Ling, X., Harkness, T. A., Schultz, M. C., Fisher-Adams, G. & Grunstein, M. Yeast histone H3 and H4 amino termini are important for nucleosome assembly in vivo and in vitro: redundant and position-independent functions in assembly but not in gene regulation. *Genes & development* **10**, 686-699 (1996).
58 Cosma, M. P. Ordered recruitment: gene-specific mechanism of transcription activation. *Mol Cell* **10**, 227-236 (2002).
59 Levine, M. & Tjian, R. Transcription regulation and animal diversity. *Nature* **424**, 147-151, doi:10.1038/nature01763 (2003).
60 Sudarsanam, P. & Winston, F. The Swi/Snf family nucleosome-remodeling complexes and transcriptional control. *Trends Genet* **16**, 345-351 (2000).
61 Tyler, J. K. & Kadonaga, J. T. The "dark side" of chromatin remodeling: repressive effects on transcription. *Cell* **99**, 443-446 (1999).
62 Singleton, M. R. & Wigley, D. B. Modularity and specialization in superfamily 1 and 2 helicases. *J Bacteriol* **184**, 1819-1826 (2002).
63 Hargreaves, D. C. & Crabtree, G. R. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res* **21**, 396-420, doi:10.1038/cr.2011.32 (2011).
64 Versteege, I. *et al.* Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**, 203-206, doi:10.1038/28212 (1998).
65 Grand, F. *et al.* Frequent deletion of hSNF5/INI1, a component of the SWI/SNF complex, in chronic myeloid leukemia. *Cancer research* **59**, 3870-3874 (1999).
66 Allfrey, V. G., Faulkner, R. & Mirsky, A. E. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **51**, 786-794 (1964).
67 Cosgrove, M. S., Boeke, J. D. & Wolberger, C. Regulated nucleosome mobility and the histone code. *Nature structural & molecular biology* **11**, 1037-1043, doi:10.1038/nsmb851 (2004).
68 Voigt, P. *et al.* Asymmetrically modified nucleosomes. *Cell* **151**, 181-193, doi:10.1016/j.cell.2012.09.002 (2012).
69 Watt, F. & Molloy, P. L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & development* **2**, 1136-1143, doi:10.1101/gad.2.9.1136 (1988).
70 Rothbart, S. B. & Strahl, B. D. Interpreting the language of histone and DNA modifications. *Biochimica et biophysica acta* **1839**, 627-643, doi:10.1016/j.bbagrm.2014.03.001 (2014).
71 Blackledge, N. P. & Klose, R. CpG island chromatin: a platform for gene regulation. *Epigenetics* **6**, 147-152, doi:10.4161/epi.6.2.13640 (2011).
72 Klose, R. J. & Bird, A. P. Genomic DNA methylation: the mark and its mediators. *Trends in biochemical sciences* **31**, 89-97, doi:10.1016/j.tibs.2005.12.008 (2006).
73 Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935, doi:10.1126/science.1170116 (2009).
74 Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300-1303, doi:10.1126/science.1210597 (2011).

75 Shen, L. *et al.* Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692-706, doi:10.1016/j.cell.2013.04.002 (2013).
76 Phillips, D. M. The presence of acetyl groups of histones. *The Biochemical journal* **87**, 258-263 (1963).
77 Zentner, G. E. & Henikoff, S. Regulation of nucleosome dynamics by histone modifications. *Nature structural & molecular biology* **20**, 259-266, doi:10.1038/nsmb.2470 (2013).
78 Unnikrishnan, A., Gafken, P. R. & Tsukiyama, T. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nature structural & molecular biology* **17**, 430-437, doi:10.1038/nsmb.1780 (2010).
79 Allis, C. D. *et al.* New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-636, doi:10.1016/j.cell.2007.10.039 (2007).
80 Zeng, L. & Zhou, M. M. Bromodomain: an acetyl-lysine binding domain. *FEBS letters* **513**, 124-128 (2002).
81 Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nature genetics* **40**, 897-903, doi:10.1038/ng.154 (2008).
82 Tie, F. *et al.* CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* **136**, 3131-3141, doi:10.1242/dev.037127 (2009).
83 Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21931-21936, doi:10.1073/pnas.1016071107 (2010).
84 Bedford, M. T. & Clarke, S. G. Protein arginine methylation in mammals: who, what, and why. *Molecular cell* **33**, 1-13, doi:10.1016/j.molcel.2008.12.013 (2009).
85 Schwartz, Y. B. & Pirrotta, V. Polycomb silencing mechanisms and the management of genomic programmes. *Nature reviews. Genetics* **8**, 9-22, doi:10.1038/nrg1981 (2007).
86 Schmitges, F. W. *et al.* Histone methylation by PRC2 is inhibited by active chromatin marks. *Molecular cell* **42**, 330-341, doi:10.1016/j.molcel.2011.03.025 (2011).
87 Schneider, R. *et al.* Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nature cell biology* **6**, 73-77, doi:10.1038/ncb1076 (2004).
88 Santos-Rosa, H. *et al.* Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Molecular cell* **12**, 1325-1332 (2003).
89 Pray-Grant, M. G., Daniel, J. A., Schieltz, D., Yates, J. R., 3rd & Grant, P. A. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433**, 434-438, doi:10.1038/nature03242 (2005).
90 Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77-88, doi:10.1016/j.cell.2007.05.042 (2007).
91 Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326, doi:10.1016/j.cell.2006.02.041 (2006).
92 Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics* **39**, 311-318, doi:10.1038/ng1966 (2007).
93 Kizer, K. O. *et al.* A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular*

and cellular biology **25**, 3305-3316, doi:10.1128/MCB.25.8.3305-3316.2005 (2005).
94 Li, J., Moazed, D. & Gygi, S. P. Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *The Journal of biological chemistry* **277**, 49383-49388, doi:10.1074/jbc.M209294200 (2002).
95 Sun, X. J. *et al.* Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *The Journal of biological chemistry* **280**, 35261-35271, doi:10.1074/jbc.M504012200 (2005).
96 Yuan, W. *et al.* Heterogeneous nuclear ribonucleoprotein L Is a subunit of human KMT3a/Set2 complex required for H3 Lys-36 trimethylation activity in vivo. *The Journal of biological chemistry* **284**, 15701-15707, doi:10.1074/jbc.M808431200 (2009).
97 Carrozza, M. J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581-592, doi:10.1016/j.cell.2005.10.023 (2005).
98 Joshi, A. A. & Struhl, K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Molecular cell* **20**, 971-978, doi:10.1016/j.molcel.2005.11.021 (2005).
99 Keogh, M. C. *et al.* Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**, 593-605, doi:10.1016/j.cell.2005.10.025 (2005).
100 Pradeepa, M. M., Sutherland, H. G., Ule, J., Grimes, G. R. & Bickmore, W. A. Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS genetics* **8**, e1002717, doi:10.1371/journal.pgen.1002717 (2012).
101 Kolasinska-Zwierz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nature genetics* **41**, 376-381, doi:10.1038/ng.322 (2009).
102 Peters, A. H. *et al.* Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Molecular cell* **12**, 1577-1589 (2003).
103 Robertson, K. D. DNA methylation and human disease. *Nature reviews. Genetics* **6**, 597-610, doi:10.1038/nrg1655 (2005).
104 Kareta, M. S., Botello, Z. M., Ennis, J. J., Chou, C. & Chedin, F. Reconstitution and mechanism of the stimulation of de novo methylation by human DNMT3L. *The Journal of biological chemistry* **281**, 25893-25902, doi:10.1074/jbc.M603140200 (2006).
105 Fatemi, M., Hermann, A., Pradhan, S. & Jeltsch, A. The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *Journal of molecular biology* **309**, 1189-1199, doi:10.1006/jmbi.2001.4709 (2001).
106 Bestor, T. H. & Verdine, G. L. DNA methyltransferases. *Current opinion in cell biology* **6**, 380-389, doi:10.1016/0955-0674(94)90030-2 (1994).
107 Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257, doi:10.1016/s0092-8674(00)81656-6 (1999).
108 Bestor, T. H. The DNA methyltransferases of mammals. *Human molecular genetics* **9**, 2395-2402, doi:10.1093/hmg/9.16.2395 (2000).
109 Liang, G. *et al.* Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Molecular and cellular biology* **22**, 480-491, doi:10.1128/mcb.22.2.480-491.2002 (2002).

110 Fatemi, M., Hermann, A., Gowher, H. & Jeltsch, A. Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *European journal of biochemistry* **269**, 4981-4984, doi:10.1046/j.1432-1033.2002.03198.x (2002).
111 Jenuwein, T., Laible, G., Dorn, R. & Reuter, G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cellular and molecular life sciences : CMLS* **54**, 80-93 (1998).
112 Nguyen, A. T. & Zhang, Y. The diverse functions of Dot1 and H3K79 methylation. *Genes & development* **25**, 1345-1358, doi:10.1101/gad.2057811 (2011).
113 Collins, R. E. *et al.* In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. *The Journal of biological chemistry* **280**, 5563-5570, doi:10.1074/jbc.M410483200 (2005).
114 Zhang, X., Zhou, L. & Cheng, X. Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *The EMBO journal* **19**, 3509-3519, doi:10.1093/emboj/19.14.3509 (2000).
115 McBride, A. E. & Silver, P. A. State of the arg: protein methylation at arginine comes of age. *Cell* **106**, 5-8, doi:10.1016/s0092-8674(01)00423-8 (2001).
116 Cook, J. R. *et al.* FBXO11/PRMT9, a new protein arginine methyltransferase, symmetrically dimethylates arginine residues. *Biochemical and biophysical research communications* **342**, 472-481, doi:10.1016/j.bbrc.2006.01.167 (2006).
117 Krause, C. D. *et al.* Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential. *Pharmacology & therapeutics* **113**, 50-87, doi:10.1016/j.pharmthera.2006.06.007 (2007).
118 Marmorstein, R. & Roth, S. Y. Histone acetyltransferases: function, structure, and catalysis. *Current opinion in genetics & development* **11**, 155-161, doi:10.1016/s0959-437x(00)00173-8 (2001).
119 Roth, S. Y., Denu, J. M. & Allis, C. D. Histone acetyltransferases. *Annual review of biochemistry* **70**, 81-120, doi:10.1146/annurev.biochem.70.1.81 (2001).
120 Voss, A. K. & Thomas, T. MYST family histone acetyltransferases take center stage in stem cells and development. *BioEssays : news and reviews in molecular, cellular and developmental biology* **31**, 1050-1061, doi:10.1002/bies.200900051 (2009).
121 Chen, H. *et al.* Nuclear receptor coactivator ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CBP/p300. *Cell* **90**, 569-580, doi:10.1016/s0092-8674(00)80516-4 (1997).
122 Spencer, T. E. *et al.* Steroid receptor coactivator-1 is a histone acetyltransferase. *Nature* **389**, 194-198, doi:10.1038/38304 (1997).
123 Blanco-Garcia, N., Asensio-Juan, E., de la Cruz, X. & Martinez-Balbas, M. A. Autoacetylation regulates P/CAF nuclear localization. *The Journal of biological chemistry* **284**, 1343-1352, doi:10.1074/jbc.M806075200 (2009).
124 Dawson, M. A., Kouzarides, T. & Huntly, B. J. Targeting epigenetic readers in cancer. *The New England journal of medicine* **367**, 647-657, doi:10.1056/NEJMra1112635 (2012).
125 Yun, M., Wu, J., Workman, J. L. & Li, B. Readers of histone modifications. *Cell research* **21**, 564-578, doi:10.1038/cr.2011.42 (2011).
126 Jacobs, S. A. & Khorasanizadeh, S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080-2083, doi:10.1126/science.1069473 (2002).
127 Taniguchi, Y. The Bromodomain and Extra-Terminal Domain (BET) Family: Functional Anatomy of BET Paralogous Proteins. *International journal of molecular sciences* **17**, doi:10.3390/ijms17111849 (2016).

128 Sanchez, R. & Zhou, M. M. The PHD finger: a versatile epigenome reader. *Trends in biochemical sciences* **36**, 364-372, doi:10.1016/j.tibs.2011.03.005 (2011).
129 Zhang, H. *et al.* TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell research* **20**, 1390-1393, doi:10.1038/cr.2010.156 (2010).
130 Xu, Y. *et al.* Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Molecular cell* **42**, 451-464, doi:10.1016/j.molcel.2011.04.005 (2011).
131 Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development* **30**, 733-750, doi:10.1101/gad.276568.115 (2016).
132 Cuthbert, G. L. *et al.* Histone deimination antagonizes arginine methylation. *Cell* **118**, 545-553, doi:10.1016/j.cell.2004.08.020 (2004).
133 Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941-953, doi:10.1016/j.cell.2004.12.012 (2004).
134 Lee, M. G., Wynder, C., Cooch, N. & Shiekhattar, R. An essential role for CoREST in nucleosomal histone 3 lysine 4 demethylation. *Nature* **437**, 432-435, doi:10.1038/nature04021 (2005).
135 Wissmann, M. *et al.* Cooperative demethylation by JMJD2C and LSD1 promotes androgen receptor-dependent gene expression. *Nature cell biology* **9**, 347-353, doi:10.1038/ncb1546 (2007).
136 Karytinos, A. *et al.* A novel mammalian flavin-dependent histone demethylase. *The Journal of biological chemistry* **284**, 17775-17782, doi:10.1074/jbc.M109.003087 (2009).
137 Cloos, P. A., Christensen, J., Agger, K. & Helin, K. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes & development* **22**, 1115-1140, doi:10.1101/gad.1652908 (2008).
138 He, J., Kallin, E. M., Tsukada, Y. & Zhang, Y. The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15(Ink4b). *Nature structural & molecular biology* **15**, 1169-1175, doi:10.1038/nsmb.1499 (2008).
139 Yang, W. M., Yao, Y. L., Sun, J. M., Davie, J. R. & Seto, E. Isolation and characterization of cDNAs corresponding to an additional member of the human histone deacetylase gene family. *The Journal of biological chemistry* **272**, 28001-28007, doi:10.1074/jbc.272.44.28001 (1997).
140 Hu, E. *et al.* Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. *The Journal of biological chemistry* **275**, 15254-15264, doi:10.1074/jbc.M908988199 (2000).
141 Glozak, M. A. & Seto, E. Acetylation/deacetylation modulates the stability of DNA replication licensing factor Cdt1. *The Journal of biological chemistry* **284**, 11446-11453, doi:10.1074/jbc.M809394200 (2009).
142 Du, J. *et al.* Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* **334**, 806-809, doi:10.1126/science.1207861 (2011).
143 Santi, D. V., Norment, A. & Garrett, C. E. Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proceedings of the National Academy of Sciences of the United States of America* **81**, 6993-6997, doi:10.1073/pnas.81.22.6993 (1984).
144 Cheng, J. C. *et al.* Inhibition of DNA methylation and reactivation of silenced genes by zebularine. *Journal of the National Cancer Institute* **95**, 399-409, doi:10.1093/jnci/95.5.399 (2003).
145 Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904-917, doi:10.1016/j.cell.2011.08.017 (2011).

146 Mertz, J. A. *et al.* Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 16669-16674, doi:10.1073/pnas.1108190108 (2011).
147 Beguelin, W. *et al.* EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer cell* **23**, 677-692, doi:10.1016/j.ccr.2013.04.011 (2013).
148 Yap, D. B. *et al.* Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117**, 2451-2459, doi:10.1182/blood-2010-11-321208 (2011).
149 Eckschlager, T., Plch, J., Stiborova, M. & Hrabeta, J. Histone Deacetylase Inhibitors as Anticancer Drugs. *International journal of molecular sciences* **18**, doi:10.3390/ijms18071414 (2017).
150 Zhang, X. D. Illustration of SSMD, z score, SSMD*, z* score, and t statistic for hit selection in RNAi high-throughput screens. *Journal of biomolecular screening* **16**, 775-785, doi:10.1177/1087057111405851 (2011).
151 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).
152 Wu, D. *et al.* ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176-2182, doi:10.1093/bioinformatics/btq401 (2010).
153 Wu, D. & Smyth, G. K. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic acids research* **40**, e133, doi:10.1093/nar/gks461 (2012).
154 Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome biology* **15**, 554, doi:10.1186/s13059-014-0554-4 (2014).
155 Oumard, A., Qiao, J., Jostock, T., Li, J. & Bode, J. Recommended Method for Chromosome Exploitation: RMCE-based Cassette-exchange Systems in Animal Cell Biotechnology. *Cytotechnology* **50**, 93-108, doi:10.1007/s10616-006-6550-0 (2006).
156 Desta, Z., Ward, B. A., Soukhova, N. V. & Flockhart, D. A. Comprehensive evaluation of tamoxifen sequential biotransformation by the human cytochrome P450 system in vitro: prominent roles for CYP3A and CYP2D6. *The Journal of pharmacology and experimental therapeutics* **310**, 1062-1075, doi:10.1124/jpet.104.065607 (2004).
157 Harada, N. *et al.* Intestinal polyposis in mice with a dominant stable mutation of the beta-catenin gene. *The EMBO journal* **18**, 5931-5942, doi:10.1093/emboj/18.21.5931 (1999).
158 Onuma, K. *et al.* Genetic reconstitution of tumorigenesis in primary intestinal cells. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11127-11132, doi:10.1073/pnas.1221926110 (2013).
159 Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature protocols* **9**, 171-181, doi:10.1038/nprot.2014.006 (2014).
160 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
161 Faust, G. G. & Hall, I. M. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics* **30**, 2503-2505, doi:10.1093/bioinformatics/btu314 (2014).
162 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

163 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
164 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic acids research* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).
165 Consortium, E. P. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**, 636-640, doi:10.1126/science.1105136 (2004).
166 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).
167 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
168 Zhu, L. J. *et al.* ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**, 237, doi:10.1186/1471-2105-11-237 (2010).
169 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165, doi:10.1093/nar/gkw257 (2016).
170 Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell reports* **5**, 1704-1713, doi:10.1016/j.celrep.2013.11.020 (2013).
171 Serra, D. *et al.* Self-organization and symmetry breaking in intestinal organoid development. *Nature* **569**, 66-72, doi:10.1038/s41586-019-1146-y (2019).
172 Sato, T. & Clevers, H. Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. *Science* **340**, 1190-1194, doi:10.1126/science.1234852 (2013).
173 Tang, J., Gary, J. D., Clarke, S. & Herschman, H. R. PRMT 3, a type I protein arginine N-methyltransferase that differs from PRMT1 in its oligomerization, subcellular localization, substrate specificity, and regulation. *The Journal of biological chemistry* **273**, 16935-16945, doi:10.1074/jbc.273.27.16935 (1998).
174 Frankel, A. *et al.* The novel human protein arginine N-methyltransferase PRMT6 is a nuclear enzyme displaying unique substrate specificity. *The Journal of biological chemistry* **277**, 3537-3543, doi:10.1074/jbc.M108786200 (2002).
175 Pawlak, M. R., Scherer, C. A., Chen, J., Roshon, M. J. & Ruley, H. E. Arginine N-methyltransferase 1 is required for early postimplantation mouse development, but cells deficient in the enzyme are viable. *Molecular and cellular biology* **20**, 4859-4869, doi:10.1128/mcb.20.13.4859-4869.2000 (2000).
176 Eram, M. S. *et al.* A Potent, Selective, and Cell-Active Inhibitor of Human Type I Protein Arginine Methyltransferases. *ACS chemical biology* **11**, 772-781, doi:10.1021/acschembio.5b00839 (2016).
177 Dai, Z. *et al.* edgeR: a versatile tool for the analysis of shRNA-seq and CRISPR-Cas9 genetic screens. *F1000Research* **3**, 95, doi:10.12688/f1000research.3928.2 (2014).
178 Peters, J. M., Tedeschi, A. & Schmitz, J. The cohesin complex and its roles in chromosome biology. *Genes & development* **22**, 3089-3114, doi:10.1101/gad.1724308 (2008).
179 Hayakawa, T. & Nakayama, J. Physiological roles of class I HDAC complex and histone demethylase. *Journal of biomedicine & biotechnology* **2011**, 129383, doi:10.1155/2011/129383 (2011).
180 Sheikh, B. N., Guhathakurta, S. & Akhtar, A. The non-specific lethal (NSL) complex at the crossroads of transcriptional control and cellular homeostasis. *EMBO reports* **20**, e47630, doi:10.15252/embr.201847630 (2019).

181 Hu, D. *et al.* The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Molecular and cellular biology* **33**, 4745-4754, doi:10.1128/MCB.01181-13 (2013).
182 Svedruzic, Z. M. Dnmt1 structure and function. *Progress in molecular biology and translational science* **101**, 221-254, doi:10.1016/B978-0-12-387685-0.00006-8 (2011).
183 Falandry, C. *et al.* CLLD8/KMT1F is a lysine methyltransferase that is important for chromosome segregation. *The Journal of biological chemistry* **285**, 20234-20241, doi:10.1074/jbc.M109.052399 (2010).
184 Laugesen, A. & Helin, K. Chromatin repressive complexes in stem cells, development, and cancer. *Cell stem cell* **14**, 735-751, doi:10.1016/j.stem.2014.05.006 (2014).
185 Christensen, J. *et al.* RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell* **128**, 1063-1076, doi:10.1016/j.cell.2007.02.003 (2007).
186 Guarnaccia, A. D. & Tansey, W. P. Moonlighting with WDR5: A Cellular Multitasker. *Journal of clinical medicine* **7**, doi:10.3390/jcm7020021 (2018).
187 Gottlicher, M. *et al.* Valproic acid defines a novel class of HDAC inhibitors inducing differentiation of transformed cells. *The EMBO journal* **20**, 6969-6978, doi:10.1093/emboj/20.24.6969 (2001).
188 Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067-1073, doi:10.1038/nature09504 (2010).
189 Donato, E. *et al.* Compensatory RNA polymerase 2 loading determines the efficacy and transcriptional selectivity of JQ1 in Myc-driven tumors. *Leukemia* **31**, 479-490, doi:10.1038/leu.2016.182 (2017).
190 Chuang, C. Y. *et al.* PRMT1 expression is elevated in head and neck cancer and inhibition of protein arginine methylation by adenosine dialdehyde or PRMT1 knockdown downregulates proliferation and migration of oral cancer cells. *Oncology reports* **38**, 1115-1123, doi:10.3892/or.2017.5737 (2017).
191 Altan, B. *et al.* Nuclear PRMT1 expression is associated with poor prognosis and chemosensitivity in gastric cancer patients. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* **19**, 789-797, doi:10.1007/s10120-015-0551-7 (2016).
192 Widakowich, C., de Castro, G., Jr., de Azambuja, E., Dinh, P. & Awada, A. Review: side effects of approved molecular targeted therapies in solid cancers. *The oncologist* **12**, 1443-1455, doi:10.1634/theoncologist.12-12-1443 (2007).
193 Shen, N. Y., Ng, S. Y., Toepp, S. L. & Ljubicic, V. Protein arginine methyltransferase expression and activity during myogenesis. *Bioscience reports* **38**, doi:10.1042/BSR20171533 (2018).
194 Obianyo, O. & Thompson, P. R. Kinetic mechanism of protein arginine methyltransferase 6 (PRMT6). *The Journal of biological chemistry* **287**, 6062-6071, doi:10.1074/jbc.M111.333609 (2012).
195 Cha, B. *et al.* Methylation by protein arginine methyltransferase 1 increases stability of Axin, a negative regulator of Wnt signaling. *Oncogene* **30**, 2379-2389, doi:10.1038/onc.2010.610 (2011).
196 Zhao, Y. *et al.* PRMT1 regulates the tumour-initiating properties of esophageal squamous cell carcinoma through histone H4 arginine methylation coupled with transcriptional activation. *Cell death & disease* **10**, 359, doi:10.1038/s41419-019-1595-0 (2019).
197 D'Alesio, C. *et al.* RNAi screens identify CHD4 as an essential gene in breast cancer growth. *Oncotarget* **7**, 80901-80915, doi:10.18632/oncotarget.12646 (2016).

198 Rayman, J. B. *et al.* E2F mediates cell cycle-dependent transcriptional repression in vivo by recruitment of an HDAC1/mSin3B corepressor complex. *Genes & development* **16**, 933-947, doi:10.1101/gad.969202 (2002).
199 Liang, J. *et al.* Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nature cell biology* **10**, 731-739, doi:10.1038/ncb1736 (2008).
200 Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex. *Cell* **89**, 357-364, doi:10.1016/s0092-8674(00)80216-0 (1997).
201 Meehan, W. J. *et al.* Breast cancer metastasis suppressor 1 (BRMS1) forms complexes with retinoblastoma-binding protein 1 (RBP1) and the mSin3 histone deacetylase complex and represses transcription. *The Journal of biological chemistry* **279**, 1562-1569, doi:10.1074/jbc.M307969200 (2004).
202 van Oevelen, C. *et al.* A role for mammalian Sin3 in permanent gene silencing. *Molecular cell* **32**, 359-370, doi:10.1016/j.molcel.2008.10.015 (2008).
203 Seraj, M. J., Samant, R. S., Verderame, M. F. & Welch, D. R. Functional evidence for a novel human breast carcinoma metastasis suppressor, BRMS1, encoded at chromosome 11q13. *Cancer research* **60**, 2764-2769 (2000).
204 Yang, J. *et al.* Breast cancer metastasis suppressor 1 inhibits SDF-1alpha-induced migration of non-small cell lung cancer by decreasing CXCR4 expression. *Cancer letters* **269**, 46-56, doi:10.1016/j.canlet.2008.04.016 (2008).
205 Zhang, S., Lin, Q. D. & Di, W. Suppression of human ovarian carcinoma metastasis by the metastasis-suppressor gene, BRMS1. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society* **16**, 522-531, doi:10.1111/j.1525-1438.2006.00547.x (2006).
206 Mei, P. *et al.* BRMS1 suppresses glioma progression by regulating invasion, migration and adhesion of glioma cells. *PloS one* **9**, e98544, doi:10.1371/journal.pone.0098544 (2014).
207 Grzenda, A., Lomberk, G., Zhang, J. S. & Urrutia, R. Sin3: master scaffold and transcriptional corepressor. *Biochimica et biophysica acta* **1789**, 443-450, doi:10.1016/j.bbagrm.2009.05.007 (2009).
208 Ellison-Zelski, S. J. & Alarid, E. T. Maximum growth and survival of estrogen receptor-alpha positive breast cancer cells requires the Sin3A transcriptional repressor. *Molecular cancer* **9**, 263, doi:10.1186/1476-4598-9-263 (2010).
209 Farias, E. F. *et al.* Interference with Sin3 function induces epigenetic reprogramming and differentiation in breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 11811-11816, doi:10.1073/pnas.1006737107 (2010).
210 Das, T. K., Sangodkar, J., Negre, N., Narla, G. & Cagan, R. L. Sin3a acts through a multi-gene module to regulate invasion in Drosophila and human tumors. *Oncogene* **32**, 3184-3197, doi:10.1038/onc.2012.326 (2013).
211 Zika, E., Greer, S. F., Zhu, X. S. & Ting, J. P. Histone deacetylase 1/mSin3A disrupts gamma interferon-induced CIITA function and major histocompatibility complex class II enhanceosome formation. *Molecular and cellular biology* **23**, 3091-3102, doi:10.1128/mcb.23.9.3091-3102.2003 (2003).
212 Chou, S. D., Khan, A. N., Magner, W. J. & Tomasi, T. B. Histone acetylation regulates the cell type specific CIITA promoters, MHC class II expression and antigen presentation in tumor cells. *International immunology* **17**, 1483-1494, doi:10.1093/intimm/dxh326 (2005).

213 Brodie, S. A. & Brandes, J. C. Could valproic acid be an effective anticancer agent? The evidence so far. *Expert review of anticancer therapy* **14**, 1097-1100, doi:10.1586/14737140.2014.940329 (2014).
214 Herz, H. M. *et al.* Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes & development* **26**, 2604-2620, doi:10.1101/gad.201327.112 (2012).
215 Calo, E. & Wysocka, J. Modification of enhancer chromatin: what, how, and why? *Molecular cell* **49**, 825-837, doi:10.1016/j.molcel.2013.01.038 (2013).
216 Yokoyama, A. *et al.* Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Molecular and cellular biology* **24**, 5639-5649, doi:10.1128/MCB.24.13.5639-5649.2004 (2004).
217 Glaser, S. *et al.* Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* **133**, 1423-1432, doi:10.1242/dev.02302 (2006).
218 Ford, D. J. & Dingwall, A. K. The cancer COMPASS: navigating the functions of MLL complexes in cancer. *Cancer genetics* **208**, 178-191, doi:10.1016/j.cancergen.2015.01.005 (2015).
219 Ernst, P. & Vakoc, C. R. WRAD: enabler of the SET1-family of H3K4 methyltransferases. *Briefings in functional genomics* **11**, 217-226, doi:10.1093/bfgp/els017 (2012).
220 Ge, Z. *et al.* WDR5 high expression and its effect on tumorigenesis in leukemia. *Oncotarget* **7**, 37740-37754, doi:10.18632/oncotarget.9312 (2016).
221 Dai, X. *et al.* WDR5 Expression Is Prognostic of Breast Cancer Outcome. *PloS one* **10**, e0124964, doi:10.1371/journal.pone.0124964 (2015).
222 Chen, X. *et al.* Upregulated WDR5 promotes proliferation, self-renewal and chemoresistance in bladder cancer via mediating H3K4 trimethylation. *Scientific reports* **5**, 8293, doi:10.1038/srep08293 (2015).
223 Wysocka, J. *et al.* WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell* **121**, 859-872, doi:10.1016/j.cell.2005.03.036 (2005).
224 Ang, Y. S. *et al.* Wdr5 mediates self-renewal and reprogramming via the embryonic stem cell core transcriptional network. *Cell* **145**, 183-197, doi:10.1016/j.cell.2011.03.003 (2011).
225 Dawkins, J. B. *et al.* Reduced Expression of Histone Methyltransferases KMT2C and KMT2D Correlates with Improved Outcome in Pancreatic Ductal Adenocarcinoma. *Cancer research* **76**, 4861-4871, doi:10.1158/0008-5472.CAN-16-0481 (2016).
226 Kanda, H., Nguyen, A., Chen, L., Okano, H. & Hariharan, I. K. The Drosophila ortholog of MLL3 and MLL4, trithorax related, functions as a negative regulator of tissue growth. *Molecular and cellular biology* **33**, 1702-1710, doi:10.1128/MCB.01585-12 (2013).
227 Lee, J. *et al.* A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8513-8518, doi:10.1073/pnas.0902873106 (2009).
228 Guo, C. *et al.* KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget* **4**, 2144-2153, doi:10.18632/oncotarget.1555 (2013).
229 Kim, J. H. *et al.* UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. *Cancer research* **74**, 1705-1717, doi:10.1158/0008-5472.CAN-13-1896 (2014).

230 Aliouat-Denis, C. M. *et al.* p53-independent regulation of p21Waf1/Cip1 expression and senescence by Chk2. *Molecular cancer research : MCR* **3**, 627-634, doi:10.1158/1541-7786.MCR-05-0121 (2005).
231 Jeong, J. H. *et al.* p53-independent induction of G1 arrest and p21WAF1/CIP1 expression by ascofuranone, an isoprenoid antibiotic, through downregulation of c-Myc. *Molecular cancer therapeutics* **9**, 2102-2113, doi:10.1158/1535-7163.MCT-09-1159 (2010).
232 Zhu, J. *et al.* Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. *Nature* **525**, 206-211, doi:10.1038/nature15251 (2015).
233 Sahtoe, D. D., van Dijk, W. J., Ekkebus, R., Ovaa, H. & Sixma, T. K. BAP1/ASXL1 recruitment and activation for H2A deubiquitination. *Nature communications* **7**, 10292, doi:10.1038/ncomms10292 (2016).
234 Scelfo, A. *et al.* Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities. *Molecular cell* **74**, 1037-1052 e1037, doi:10.1016/j.molcel.2019.04.002 (2019).
235 Kahn, T. G. *et al.* Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. *Nucleic acids research* **44**, 10132-10149, doi:10.1093/nar/gkw701 (2016).
236 Swigut, T. & Wysocka, J. H3K27 demethylases, at long last. *Cell* **131**, 29-32, doi:10.1016/j.cell.2007.09.026 (2007).
237 Lee, J. E. *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife* **2**, e01503, doi:10.7554/eLife.01503 (2013).
238 Tan, Y. *et al.* Inhibition of BRD4 suppresses tumor growth in prostate cancer via the enhancement of FOXO1 expression. *International journal of oncology* **53**, 2503-2517, doi:10.3892/ijo.2018.4577 (2018).
239 Miller, A. L. *et al.* The BET inhibitor JQ1 attenuates double-strand break repair and sensitizes models of pancreatic ductal adenocarcinoma to PARP inhibitors. *EBioMedicine* **44**, 419-430, doi:10.1016/j.ebiom.2019.05.035 (2019).
240 Qiu, H. *et al.* JQ1 suppresses tumor growth through downregulating LDHA in ovarian cancer. *Oncotarget* **6**, 6915-6930, doi:10.18632/oncotarget.3126 (2015).