UNIVERSITAT DE
BARCELONA

Biomarkers and Combinatorial Drug Targets
for a Personalized Therapy in Colorectal Cancer

Cristina Moreta Moraleda

Biomarkers and Combinatorial Drug Targets for a Personalized Therapy in Colorectal Cancer

Cristina Moreta Moraleda

PhD THESIS
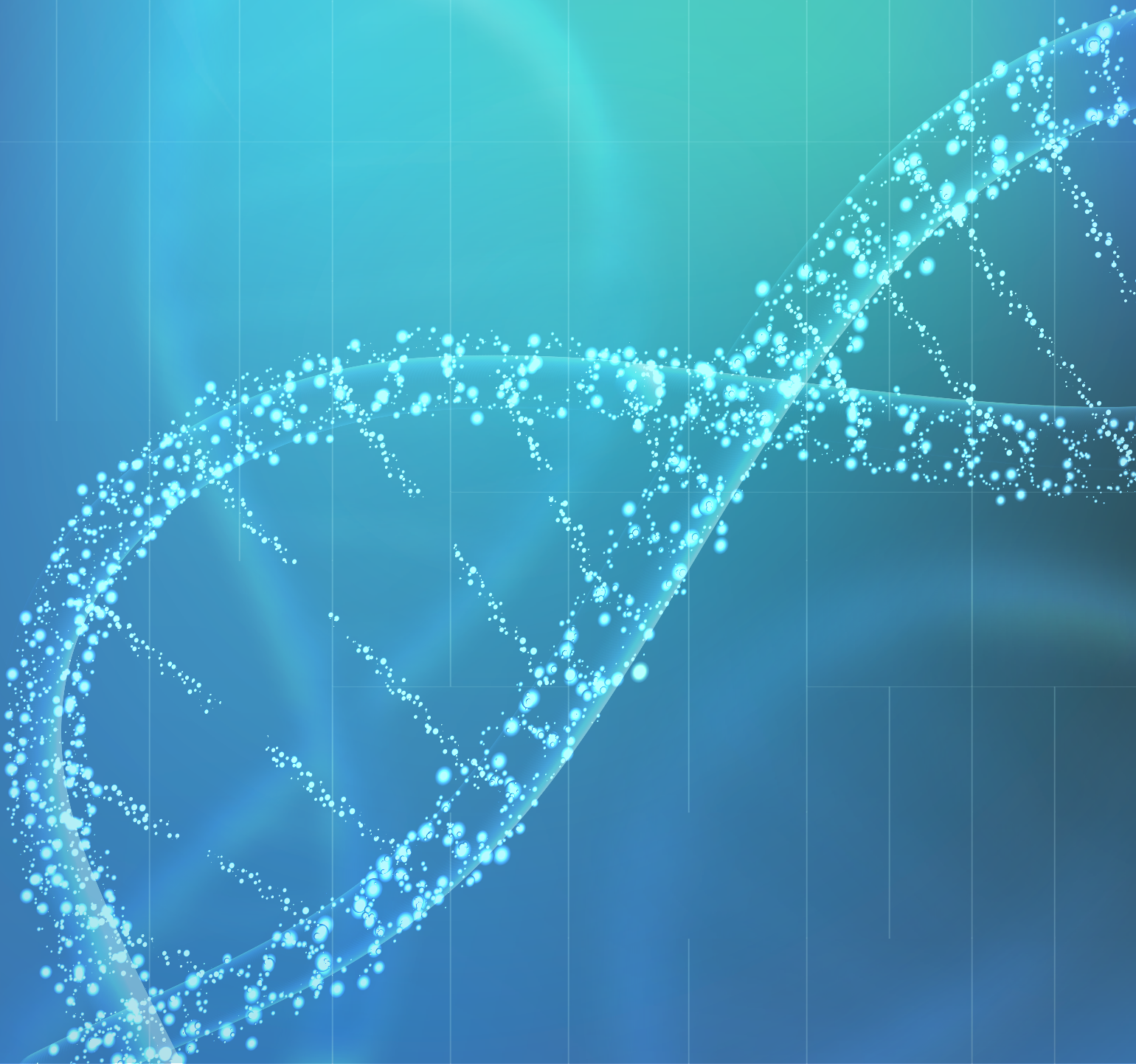
Universitat de Barcelona

Facultat de Medicina

Programa de Doctorat en Biomedicina

Biomarkers and Combinatorial Drug Targets for a Personalized Therapy in Colorectal Cancer

Memòria presentada per

Cristina Moreta Moraleda

per optar al grau de doctor per la Universitat de Barcelona

Tesi realitzada a la Unitat d'Immunologia, Departament de Patologia i Terapèutica Experimental, Facultat de Medicina (Campus Bellvitge)

Directora

Director

Sonia-Vanina Forcales Fernández

Manuel Perucho Martinez

Tutora

Doctoranda

Concepció Soler Prat

Cristina Moreta Moraleda

Barcelona, 30 de Setembre de 2021

UNIVERSITAT DE BARCELONA

Als pilars de la meva vida:

Família

Amistat

Amor

ACKNOWLEDGEMENTS
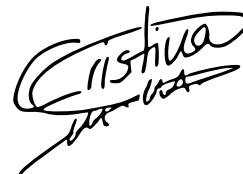
Quan vaig decidir que volia realitzar un doctorat havia escoltat opinions de tot tipus. I, certament, ha estat un camí complicat. Durant aquests 4 anys he passat moments de molta frustració, dubte i, de vegades, improvisació. També està ple de records d'una alegria immensa quan un resultat sortia, per petit que fos. Però, sobretot, són 4 anys plens de persones que m'han fet créixer i recolzat fins a l'infinit. Ha sigut una gran experiència perquè he tingut la sort de rodejar-me de gent brillant, treballadora i que han estat presents sempre que les he necessitat.

En primer lloc, a qui haig d'agrair enormement haver arribat fins aquí és a tu, *Sonia*. Vas creure en mi des del principi i has lluitat tot el que ha estat a les teves mans perquè pogués realitzar aquesta tesi. Tots sabem quant costa trobar finançament a la ciència... Tot i que les dues sabíem que seria molt difícil obtenir una beca, tu vas moure cel i terra per aconseguir que pogués desenvolupar aquest projecte. Només per això, t'estaré sempre agraïda. Però tu no només has estat la meva directora de tesi o la meva "jefi", no, has sigut molt més: has sigut un referent, una mentora. Sempre has estat present per ensenyar-me tot el que sabies, i repetir-m'ho totes les vegades que calgués; la teva porta sempre ha estat oberta perquè vingués a molestar-te quan calgués, inclús quan tenies muntanyes de feina, inclús quan potser no necessitava parlar de ciència... La veritat, havia escoltat opinions de tot respecte fer una tesi, però del que estic segura és que sense tu no hagués evolucionat ni la meitat del que ho he fet tan a nivell professional com, sobretot, personal. La meva gran sort és saber que d'aquesta tesi m'he endut una persona com tu; gràcies de debò, *Sonia*.

A qui li dec la meitat d'aquesta tesi és a tu, *Cris*. Tu vas aparèixer com un tornado d'alegria i motivació a la meva vida. Hem treballat colze a colze fins que ja no podíem més i la neurona ens anava a explotar... Aquest projecte el vam tirar endavant les dues, on no arribava una hi era l'altre. No et pots imaginar com enyoro parlar de ciència i planificar experiments amb tu; tot i que ara ja no estem enganxades l'una a l'altre en el dia a dia, sé que sempre hi ets. Des del principi vam tenir la sort d'entendre'ns i compenetrar-nos de meravella al laboratori, i aquella va ser la gran sort per tirar endavant tota la feinada que hi ha darrere d'aquest projecte. Hem compartit molts desastres al laboratori, però també hem celebrat molt. Però, sobretot, el que vaig aprendre amb tu és a relativitzar, a valorar les coses que són realment importants, i vaig créixer com a persona enormement al teu costat. Tu i jo vam crear un vincle fort allà, a

la sala de cultius, la nostra segona casa; però és un vincle que va arribar per quedar-se. Ja saps que t'enyoro moltíssim, però que estic aquí, sempre, per tot el que necessitis.

También me gustaría dar las gracias a *Manuel*, que me acogiste muy amablemente en tu laboratorio cuando yo no sabía ni cómo pipetear. Valoro enormemente que hayas querido ser director en esta tesis, y todos los consejos que me has dado. Guardo con gran respeto y admiración todas tus consideraciones.

A qui estic molt agraïda també és a tota la gent involucrada en el "Response Project". *Eva*, moltes gràcies per tenir el despatx obert sempre que t'he necessitat, i mil gràcies per tots els consells rebuts durant aquests anys, has sigut una segona "jefi" per mi. También quería agradecerte *Marcus* todo el esfuerzo que has hecho para que este proyecto tirase adelante, por todos los consejos y por haber pensado en mi siempre que has podido ayudarme; muchas gracias de verdad. Gracias también a *Jeannine* por ir siempre un paso por delante en el proyecto para guiarnos y ayudarnos en el análisis de todos los datos que generamos. I moltes gràcies a la *Raquel*, per donar-nos tot el suport que vam necessitar al principi del projecte i estar sempre disposada a ajudar-nos en el que calgués.

Muchísimas gracias a mis niñas de los "Desayunos con Amor", de verdad que no sabía la suerte que tenía cuándo os cruzasteis en mi vida. *Sara*, ay *Sarita*, ¿y qué te voy a decir yo a ti? Que te adoro, con esa mirada tan transparente y esa sonrisa siempre en los labios. Eres todo dulzura y por eso se hace muy fácil querer estar a tu lado. *Tania*, mi *Tania*, no sabes cuánto me he llegado a reír contigo; pocas veces se encuentran a personas tan auténticas, sinceras y con tan buen corazón como tú, de verdad. *Carmeta*, tu sempre seràs la meva pollo! Va ser un plaer poder-te ensenyar i m'encanta veure com estàs creixent, tant dins com fora del lab. Cada día, cuando por la mañana me llegan esos mensajes de que vais a desayunar, me lleno de todos los buenos recuerdos que tenemos en esos sofás; siempre os estaré inmensamente agradecida por acompañarme en esta tesis como lo habéis hecho.

Voldria agrair a tota la gent que va passar per l'IGTP i que, d'una manera o altra, em van ajudar o, simplement, van fer que els meus records allà fossin els millors xerrant sobre qualssevol cosa mentre escalfàvem el tupper. Penso en gent com la *Mar*, en *David Llige*, l'*Eli*, la *Sara Bystrup*, la *Maria*, la *Bea* i en *Sergio*, la *Julia*, la *Carla* en *Vicenc* la *Mace*, la *Nuria*, la *Mireia* l'*Helena*, la *Marguerite-Marie*, en *David Corujo*,

en *David Izquierdo*, en *Quim*... I també a en *Luca* i la *Milena*, que van passar pel lab uns mesos però van deixar gravada Sardenya dins nostre. A tots vosaltres, gràcies.

Per últim, m'agradaria agrair enormement a tota la unitat d'immunologia com m'han rebut, sempre amb les mans obertes. Moltes gràcies *Concepció, Juanjo, Lejuan, Ingrid, Dani, Javi, Maria* i *Miquel*, de debò. Des del primer dia m'heu integrat com una més i m'heu posat sempre les coses molt fàcils, de debò, mil gràcies. I would especially want to thank you, *Lejuan*, for being always able to help me and explain me how everything was working in the lab, even though you could be busy. You always have a smile for me and it has been a pleasure sharing the bench with you. També et vull agrair especialment a tu, *Ingrid*, que em vas acollir amb els braços oberts en un moment de canvi professional i personal molt difícil per mi; gràcies per sempre trobar un moment per explicar-me com funcionava tot i gràcies per les nostres converses a les 8 del matí. Sense tu, aquest canvi hagués estat molt més complicat, ha sigut una sort trobar-te.

A nivell personal, vull agrair a tota la meva família i amics per haver-me acompanyat en aquest camí, per creure en mi. Gràcies a tota la família de la comarca, en especial l'*Arnau*, la *Meri* i la petita *Mariona*, perquè, tot i haver d'estar lluny més del que ens hagués agradat últimament (maleït Covid), sempre els he sentit molt a prop. Gràcies a la meva "família política", a l'*Antonio*, la *Montserrat*, l'*Ascension*, en *Lluis*, en *Josep Maria*, la *Merce* i la petita *Arlet*; gràcies per haver entrat a la meva vida per quedar-vos. Gracias *Yaya* por siempre preocuparte qué tal me iba en el laboratorio, si el proyecto avanzaba y, sobretodo, preocuparte por mi cómo solo tú sabes hacerlo, desde el cariño más infinito. Gracias, *Mama* y *Papa*, por apoyarme siempre en todo y convertirme lo que me ha llevado a estar aquí hoy, escribiendo esta tesis; sin vosotros, yo no estaría aquí y esto no habría sido posible. Y gracias a vosotros, *Andrea* y *Sergi*, por escucharme hablar de cosas que seguramente ni entendíais (a veces una necesita hablar de ciencia sí o sí) pero, aún y así, escucharme. I, sobretot, gràcies a tu, *Lluis*, per entendre sempre tot el que aquesta tesi ha implicat i significat per mi; simplement, gràcies per seu el meu company de viatge.

TABLE OF CONTENTS

Abbreviations

ABBREVIATIONS

5-FU	5-Fluorouracil
ABC	ATP-binding cassette
ACF	ATP-utilizing chromatin assembly and remodeling factor
ACS	American Cancer Society
AML	Acute myeloid leukemia
ASRi	Age-standardized incidence rate
ASRm	Age-standardized mortality rate
BAF	BRG1/BRM-associated factor
BCL-xL	Bcl-2-like protein 1 isoform
BER	Base excision repair
bp	Base pairs
CCiTUB	Centre Científics i Tencològics de la Universitat de Barcelona
ccRCC	Clear cell renal cell carcinoma
CHD	Chromodomain helicase DNA-binding complex
CHRAC	Chromatin assembly complex
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CMS	Consensus molecular subtype
CPM	Count per million
CPT-11	Irinotecan
CR	Complete response
CRC	Colorectal cancer
CT	Chemotherapy
CTR1	Copper transporter 1
DACH	1,2-diaminocyclohexane
DNMT	DNA methyltransferase
DNMTi	DNA methyltransferase inhibitors
DP	Double positive
DPD	Dihydropyrimidine dehydrogenase
DSBs	Double-strand breaks
dTMP	Deoxythymidine monophosphate
dUMP	Deoxyuridine monophosphate
EcoR	Ecotropic receptor
EMT	Epithelial-to-mesenchymal transition
ERCC1	Excision repair cross-complementing group 1
FDA	US Food and Drug Administration
FdUMP	Fluorodeoxyuridine monophosphate
FdUTP	Fluorodeoxyuridine triphosphate

FOLFIRI	Leucovorin + 5-Fluorouracil + Irinotecan
FOLFOX	Leucovorin + 5-Fluorouracil + Oxaliplatin
FUIRI	5-Fluorouracil + Irinotecan
FUOX	5-Fluorouracil + Oxaliplatin
FUTP	Fluorouridine triphosphate
G418	Geneticin
GBAF	GLTSCR1/1L-associated BAF
GCO	Global Cancer Observatory
GO	Gene Ontology
gRNA	guide RNA
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HDACi	Histone deacetylase inhibitors
HDM	Histone demethylase
HMT	Histone methyltransferase
HOTAIR	HOX transcript antisense RNA
HR	Homologous recombination
IARC	International Association for Research on Cancer
iBETs	Inhibitors of the bromodomain and extra-terminal motif proteins
ISWI	Imitation switch complex
KO	Knock-out
L-OHP	Oxaliplatin
LOF	Loss-of-function
lncRNA	Long non-coding RNA
MBD	Methyl-CpG-binding domain
MDM2	Mouse double minute 2 homolog
MDS	Myelodysplastic syndrome
MMR	Mismatch repair
MRP	Multidrug resistance-associated protein
MSI	Microsatellite instable
MSS	Microsatellite stable
ncBAF	Non-canonical BAF
ncRNA	Non-coding RNA
NF-κB	Nuclear factor kappa B
NER	Nucleotide excision repair
NGS	Next Generation Sequencing
NHEJ	Non-homologous end joining
NSCLC	Non-small cell lung cancer
NURF	Nucleosome remodeling factor
OS	Overall survival

OXA	Oxaliplatin
PBAF	Polybromo-associated BAF
PcG	Polycomb group proteins
PD	Progressive disease
PFS	Progression free survival
piRNA	PIWI-interacting RNA
Plat-E	Platinum-E
PMP	Paramagnetic particle
PR	Partial response
PRC2	Polycomb repressive complex 2
RSF	Remodelling and spacing factor
RT	Room temperature
SCNAs	Somatic copy number alterations
SN-38	Active form of Irinotecan
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
SWI/SNF	Switch-sucrose non-fermentable complex
TCGA	The Cancer Genome Atlas
TET	Ten-eleven translocation
TNF-α	Tumor necrosis factor-alpha
TS	Thymidylate synthase
TSS	Transcriptional start site
TP	Thymidine phosphorylase
UGT1A	UDP-glucoronosyl-transferase 1A enzymes
UTR	Untranslated regions
XPF	Xeroderma pygmentosum group F
WB	Western Blot

List of Figures

LIST OF FIGURES

List of Tables

LIST OF TABLES

Introduction

1. Colorectal cancer

1.1. Incidence and mortality

According to the Global Cancer Observatory (GCO) of the International Association for Research on Cancer (IARC), in 2020 colorectal cancer (CRC) represented the second-most and the third-most common cancer in women and men, respectively. Combined, in both sexes, CRC is the third-most common cancer, accounting for 10% of all cases (Figure 1A) (1). More than half of these cases occur in most-developed countries. The age-standardized incidence rate (ASRi) of CRC in the world is higher in men (23.4 per 100.000 individuals) than in woman (16.2 per 100.000 individuals). Incidence varies geographically, where Australia and New Zealand have the highest incidence while Western Africa and South-Central Asia have the lowest rates (Figure 1B) (1).
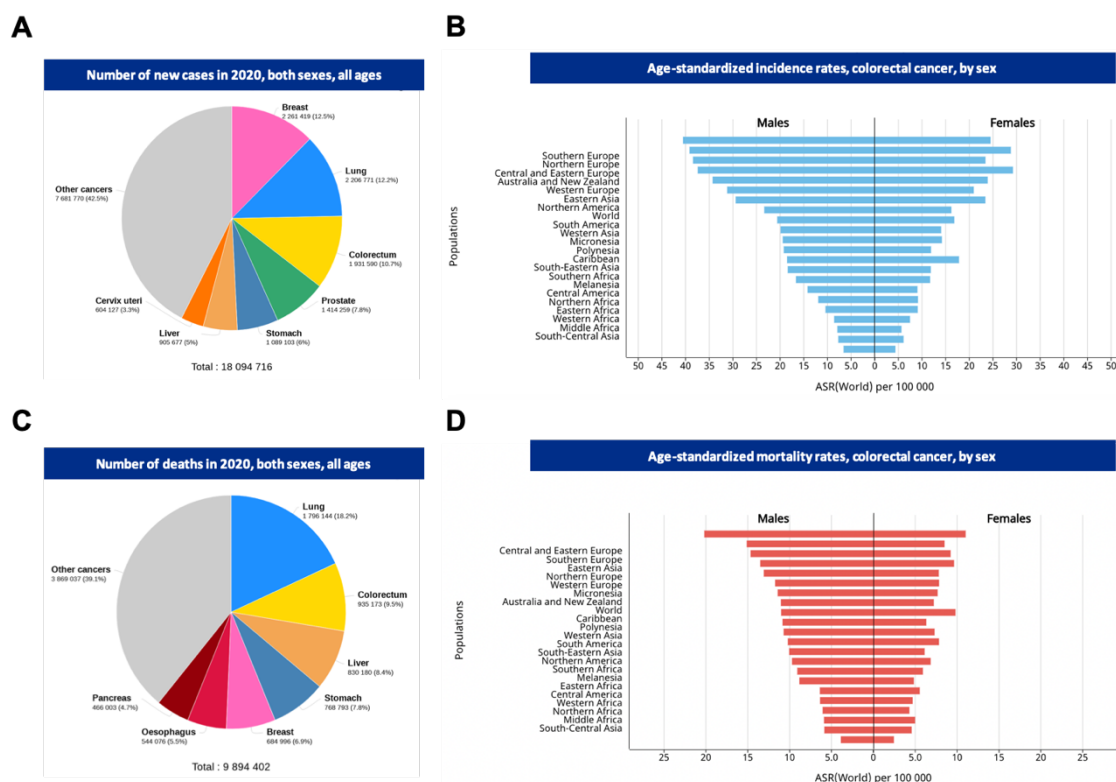
Figure 1. Incidence and mortality statistics of colorectal cancer in 2020 (Source: GLOBOCAN 2020). (A) Number of new cancer cases in both sexes worldwide. **(B)** Age-standardized incidence rates of colorectal cancer in the world divided by sex. **(C)** Number of cancer deaths in both sexes worldwide. **(D)** Age-standardized mortality rates of colorectal cancer in the world divided by sex.

In 2020, 935.173 people died as a result of CRC worldwide, making this disease the second-most common of cancer-related deaths, only after lung cancer (Figure 1C). The age-standardized mortality rate (ASRm) of CRC in different countries reflects incidence rates' values. For instance, ASRm continues to be higher in men (20.2 per 100.000 individuals) than in women (11 per 100.000 individuals). Furthermore, ASRm rates are higher in more-developed countries in comparison with less-developed countries, such

as Western Africa or South-Central Asia (Figure 1D) (1). Nevertheless, mortality of CRC depends of the stage of the tumor at the moment of diagnosis, which is influenced by the availability of population screening programs and the level of care of each country (2).

1.2. Prognosis

Based on data from the American Cancer Society (ACS) (3), the 5-year survival rate approximates 90% when the colon cancer is detected at a "localized" stage (e.g., stages I, IIA and IIB), which means there is no sign that the tumor has spread outside the colon or the rectum. By contrast, when the tumor is diagnosed at an "advanced" stage (e.g. stage IV), where the tumor has spread to other organs of the body, this 5-year survival rate decreases to 14%. This data confirms the importance of the tumor stage at diagnosis for the consequent survival of the patient: when tumors are diagnosed at an early stage the probability to overcome them is much higher than when the diagnosis is delayed (4).

1.3. Risk factors

It has been widely-reported that several factors may contribute to the etiology and development of CRC. These factors could be divided into three main groups: genetic, epigenetic and environmental.

The majority of CRC are sporadic (non-hereditary), thus present a negative family history; however, there is a small subgroup in the patient population that present a hereditary CRC syndrome, accounting for 5-10% of all patients. In this regard, the two most common hereditary syndromes are Lynch syndrome and the familial adenomatous polyposis coli. Lynch syndrome is caused by a mutation in one of the DNA mismatch-repair genes (*MLH1*, *MSH2*, *MSH6*, *PMS2* and *EPCAM*). During replication, as mismatch-repair mechanisms are impaired in the cell, DNA mutations start to accumulate, especially in microsatellite DNA repeats, which consist in arrays of tandemly repeated (i.e. adjacent) DNA motifs that range in length from one to six or up to ten nucleotides. This predisposition to mutation (genetic hypermutability or mutator phenotype) has been designated as microsatellite instability (MSI), and plays an important role in the development of CRC, also in sporadic cases (2,5–8). In the case of the familial adenomatous polyposis coli, most patients develop a large number of adenomas at a very young age due to mutations in *APC* gene, which is highly implicated in WNT pathway (2,7,8).

There is also evidence that epigenetic silencing might also contribute to the early formation of sporadic adenomas, that would later result into carcinomas following the

model suggested by Vogelstein (9,10). In the late 1980s, he described a multistep model for the development of malignant colorectal tumors from adenomas through the sequential accumulation of mutations in oncogenes (i.e. mutations on *RAS* or *MYC* genes) and tumor suppressor genes such as *TP53*. More recently, *MGMT* promoter has been reported to be aberrantly methylated in the first stages of colon adenomas, suggesting its potential as a marker of early-stage tumors (11–13). Moreover, a subset of CRC tumors severely hypermethylated have been subclassified by a CpG island methylator phenotype (CIMP) status (14), suggesting that defects in the maintenance of global DNA methylation patterns may contribute to a specific subgroup of CRC tumors (11).

Regarding the importance of the environment in CRC, aspects such as smoking, excessive alcohol intake, high consumption of red and processed meat or obesity have been proven to negatively impact the disease. Furthermore, the risk of suffering CRC is higher in patients with type 2 diabetes mellitus or inflammatory bowel disease. By contrast, daily physical activity and the consumption of fresh fruits and vegetables, calcium, fiber and vitamin D are factors that contribute to reduce the probability of developing CRC (2,7).

1.4. Pathophysiology

CRC is developed when normal colonic epithelial cells acquire several hallmarks of cancer (15,16) through the accumulation of gene mutations and epigenetic alterations that activate oncogenes and inactivate tumor suppressor genes.

CRC is a process that takes over 10-15 years to occur. In the most frequent model of CRC development (Figure 2, top part), firstly dysplastic adenomas appear, which are the most common form of premalignant precursor lesions. Mutations in *APC* gene are an early event in this formation process, occurring in almost 70% of all non-hereditary cases of colorectal adenomas. These adenomas can progress to advanced adenomas that can finally result into colorectal carcinomas. This adenoma-carcinoma sequence is further promoted by the accumulation of mutations in oncogenes such as *KRAS* or tumor suppressor genes such as *TP53*. This characteristic sequence of gene mutations is often accompanied by chromosomal instability (CIN), which accelerates the accumulation of mutations and epigenetic alterations in oncogenes and tumor suppressor genes and contributes to select by clonal expansion those cells with the most malignant behavior (2,7,17).

Nevertheless, approximately 15% of sporadic CRC develop through different molecular pathways from a subset of polyps, called serrated polyps (Figure 2, bottom part). This type of polyps, when arise in the right colon, commonly present MSI and CIMP phenotype; by contrast, when appear in the left colon, these polyps are microsatellite stable (MSS) yet frequently present mutations in the oncogene *KRAS* and some of them present an attenuated form of CIMP (2,7,17).
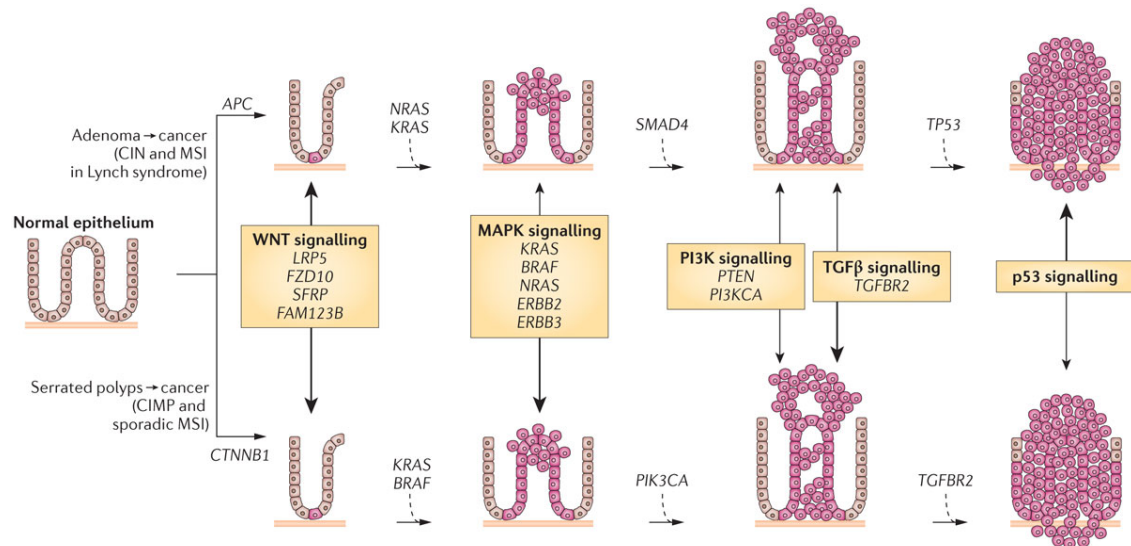
Figure 2. The polyp to colorectal cancer sequences. The traditional pathway (top) involves the development of adenomas that can progress to adenocarcinomas from normal colonic epithelium. In the bottom part, how serrated polyps progress to serrated colorectal cancer, which affects 15% of sporadic CRC tumors. Image from: Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, Van De Velde CJH & Watanabe T. Colorectal cancer. *Nat. Rev. Dis. Prim.* **1**: 1–25 (2015) (2).

In general terms, the most common somatic mutations in CRC occur in the following genes: *APC*, *CTNNB1*, *KRAS*, *BRAF*, *SMAD4*, *TGFBR2*, *TP53*, *PIK3CA*, *ARID1A*, *SOX9*, *FAM123B* and *ERBB2*. All of them impact distinct pathways such as the WNT–β-catenin, MAPK or the PI3K and TGF-β signaling pathways. In addition to gene mutations, epigenetic alterations strongly contribute to the cancer progression. DNA methylation alterations can result in two main issues: when CpG-rich regions (CpG islands) in gene promoters become hypermethylated they can favor gene silencing of tumor suppressor genes; when other genomic regions become hypomethylated they can contribute to the expression of oncogenes (18). Moreover, global hypomethylation has been observed for all cancers (19,20), mostly affecting repetitive elements (21), which can facilitate chromosomal instability by mechanisms not completely understood (2,22–25).

1.5. Diagnosis

Diagnosis of CRC patients can result from either an evaluation of a patient that already presents symptoms or as a result of a screening program. The most common symptoms

a patient can report include blood in the stools, abdominal pain or change in bowel habits, among others. In the case of screening programs (colonoscopies and fecal occult blood tests) it should be taken into account that CRC is more suitable for this type of assessment than other tumors as has a long preclinical stage (over 10 years) and high incidence in the population (2).

Nevertheless, final diagnosis of CRC should be histologically from biopsy samples taken during colonoscopy, which is the gold standard for early diagnostic method in CRC. It has a high accuracy and can perfectly assess the location of the tumor or tumors in the colon. Indeed, colonoscopy provides also a therapeutic effect, as gives the opportunity to remove adenomas, thus preventing their further development into carcinomas. Therefore, colonoscopy has been proven to reduce cancer incidence and mortality (2,7). For instance, several studies with follow-up data of more than 20 years have demonstrated the reduction of colorectal cancer-related mortality when colonoscopy screening is applied to a cohort of patients (26,27). However, another important aspect in CRC diagnosis is to determine the presence or not of distant metastases. The most frequent metastases locations are liver and lungs, despite this last is much less common than liver metastases; for this reason, liver imaging by computed tomography scan is recommended in all patients with CRC (7).

1.6. Management

1.6.1. Surgery

Surgery is the main curative treatment for patients with non-metastasized CRC. In colon cancer, the tumor and the corresponding lymph vessels are removed. The extent of the surgery strongly depends on the tumor localization and the supplying blood vessels. It has been proved that either open surgery or laparoscopy resection are safe and present the same long-term results (2,7).

1.6.2. Adjuvant therapy

Stage II colon cancer patients are associated with statistically significant better disease-free survival and overall survival ratios than stage III patients. It has been proven that stage II patients have a reduced benefit in survival when treated with adjuvant chemotherapy (CT); for this reason, adjuvant therapies are only recommended in patients with high relapse risk. This is the case of poorly differentiated tumors or when there is vascular, lymphatic or perineural tumor invasion, among others. By contrast, adjuvant therapies are standard for stage III tumors, when the tumor might spread to lymph nodes or even nearby organs or tissues. In these cases, an intravenous

combination of 5-fluorouracil, leucovorin and oxaliplatin (FOLFOX protocol) or 5-fluorouracil, leucovorin and irinotecan (FOLFIRI protocol), in combination or not with targeted therapies, are the most common approaches currently used in the clinics (2,7).

1.6.3. Metastatic disease

Survival of patients with metastatic disease (stage IV) improved in the last decades, achieving in clinical trials a median overall survival of 30 months. This improvement was due to the implementation of chemotherapeutic treatments and the introduction of targeted therapies. In the case of chemotherapeutic combinations, the first-line treatment for metastatic disease has been established as an intravenous combination of FOLFOX protocol or FOLFIRI protocol (2,7).

Regarding targeted therapies, they can be classified into four main groups:
a. Monoclonal antibodies against EGFR (cetuximab and panitumumab): more than 75% of CRC tumors overexpress EGFR, and this overexpression is associated with a reduced survival and increased risk of metastases. Monoclonal antibodies that block this receptor have proven their efficacy in patients that did not respond to CT combinations. However, for the correct efficiency of these antibodies, is crucial that tumors do not present activating mutations in *KRAS* and/or *NRAS* (2,7), which are downstream effectors of the EGFR pathway.
b. Monoclonal antibodies against VEGF-A (bevacizumab): one of main hallmarks of cancer is to establish a new vascular network for the tumor; thus, VEGF-A is a key glycoprotein secreted during tumor angiogenesis. Bevacizumab has demonstrated in several studies higher efficacy in combination with FOLFOX or FOLFIRI than CT combinations alone; however, the mechanisms of action are still poorly understood (2,7,28,29).
c. Fusion proteins that target multiple pro-angiogenic growth factors (aflibercept) or small-molecule-based kinase inhibitors (regorafenib) (2,7).
d. Immunotherapy with checkpoint inhibitors: clinical trials testing antibodies against PD1 (pembrolizumab or nivolumab) and CTLA-4 (ipilimumab) have shown beneficial effects in MSI patients (those CRCs with higher mutational burden) by inhibiting the blockade imposed to tumor-infiltrated T cells (30–34).

1.7. The consensus molecular subtypes of CRC

Despite many efforts invested on classifying the different subtypes of colorectal cancers by establishing subcategories such as CIN, MSI or CIMP tumors (2,7,17), there were not well-defined groups that may explain the differences between patients in terms of

response to the treatment or overall survival ratios, among others. Thus, other classifications were investigated based on global gene expression profiles. In 2015, Guinney et al. (35) formed an international consortium and published an extensive study based on large-scale data from 18 sources using six different classification methods that revealed four consensus molecular subtypes (CMSs) with distinguishing features, which are summarized in Figure 3.

CMS1 MSI immune	CMS2 Canonical	CMS3 Metabolic	CMS4 Mesenchymal
14%	37%	13%	23%
MSI, CIMP high, hypermutation	SCNA high	Mixed MSI status, SCNA low, CIMP low	SCNA high
BRAF mutations	*KRAS* mutations		
Immune infiltration and activation	WNT and MYC activation	Metabolic deregulation	Stromal infiltration, TGF-β activation, angiogenesis
Worse survival after relapse			Worse relapse-free and overall survival

Figure 3. Classification of colorectal cancer tumors in four CMSs. Table with the main characteristics of the four CMSs of CRC, reflecting significant biological differences on gene expression throughout the different types. Image from: Guinney J, Dienstmann R, Wang X, De Reyniès A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**(11): 1350–1356 (2015) (35).

1.7.1. CMS1 (MSI immune subtype)

Samples of CMS1 group represent 14% of CRC patients. They present a high number of mutations and a low prevalence of somatic copy number alterations (SCNAs), containing the majority of MSI tumors, which also frequently present mutations on *BRAF* gene. This type of tumors have also an overexpression of proteins involved in DNA damage repair as well as a defective DNA mismatch repair mechanism. Methylation profiles of CMS1 tumors show an extensive hypermethylation pattern, corresponding with high CIMP status. Moreover, this subtype has the characteristic of an increased expression of genes associated with a diffuse immune infiltrated, mainly T_H1 and cytotoxic T cells, together with strong activation of immune evasion pathways, which seems to be associated with MSI tumors (36).

In terms of epidemiology, CMS1 tumors are most commonly diagnosed in females with right-sided lesions and present higher histopathological grade. This population presents a poor survival rate after relapse, which has been proven to correlate with MSI and *BRAF*-mutated tumors (37,38).

1.7.2. CMS2 (canonical subtype)

CMS2 subtype includes 37% of all CRC cases. In contrast to CMS1 subtype, CMS2 tumors present high ratio of SCNAs, consistent with CIN. Indeed, they present more copy number gains in oncogenes while copy number losses occur in tumor suppressor genes. Moreover, these tumors exhibit epithelial differentiation and upregulation of WNT and MYC downstream targets, two pathways classically involved in CRC pathogenesis. Furthermore, CMS2 tumors mainly appear in the left side of the colon and show better survival rates after relapse than CMS1 or CMS4 subtypes.

1.7.3. CMS3 (metabolic subtype)

Molecular subtype CMS3 is the smallest category, corresponding to 13% of CRC patients. Samples of CMS3 present a characteristic global genomic and epigenomic profile that clearly differs from the other CMS groups: *KRAS* gene is often mutated, have low SCNAs, low prevalence of CIMP status as well, and 30% of samples show hypermutated patterns, reminding of MSI status. In addition, enrichment for metabolism signatures was found in CMS3 CRCs, which is in agreement with the occurrence of *KRAS*-activating mutations, that have been described to induce strong metabolic adaptation (39–41). Furthermore, this is supported by a low expression of the let-7 miR family, which is always accompanied by high *KRAS* expression levels (42,43).

1.7.4. CMS4 (mesenchymal subtype)

CMS4 subtype represents 23% of colorectal tumors. As CMS2 tumors, CMS4 also presents an elevated number of SCNAs, correlating with high CIN status. However, this subtype clearly shows an upregulation of genes implicated in epithelial-to-mesenchymal transition (EMT), in the complement-mediated inflammatory system, and in the activation of TGF-β signaling, angiogenesis or matrix remodeling pathways. Moreover, CMS4 tumors have gene expression profiles related to stromal infiltration and invasion, overexpression of extracellular matrix proteins and higher genetic content of non-cancer cells. Indeed, miR-200 family, which has been associated with EMT regulation (44,45), is downregulated in CMS4 subtype. Furthermore, CMS4 patients are usually diagnosed at more advanced stages (stages III and IV) and result in worse overall survival and worse relapse-free survival ratios than other CMSs.

Intriguingly, a 13% of samples showed mixed features, highlighting a transition phenotype or the intra-tumoral heterogeneity.

2. Epigenetics and colorectal cancer

DNA and histones form a macromolecular complex, named chromatin, which provides a scaffold for the packaging of our genome. The functional unit of chromatin is the nucleosome, which is formed by 147 base pairs of DNA wrapped around a histone octamer unit, composed by each of two histones H2A, H2B, H3 and H4 (46,47).

In general terms, chromatin can be subdivided into two different states: euchromatin and heterochromatin. Euchromatin corresponds to an open and transcriptionally permissive or active conformation, containing most of the active genes, while heterochromatin presents a highly condensed and transcriptionally inert conformation, where inactive genes are located (46,48). Indeed, heterochromatin has been proven to protect DNA from being accessed by complexes that promote gene transcription. Furthermore, heterochromatin has been subclassified into facultative heterochromatin, which mostly contains genes that should be kept silent during all or some developmental stages, and constitutive chromatin that contains gene-poor permanently silent regions such as pericentromeres, centromeres or telomeres (48).

On the whole, there exists a complex network that regulates gene expression and chromatin status, and epigenetics has arisen as a key regulator mechanism in these processes (Figure 4). The term of "epigenetics" was firstly described by the developmental biologist Conrad H. Waddington in 1942 (49) and was lately stablished as "the study of heritable changes in gene expression mediated by mechanisms other than alterations in the DNA sequence" (46,50–52). However, nowadays epigenetics is understood as the combination of mechanisms that regulate DNA biology, without necessarily being inherited. This epigenetic landscape includes modifications of the DNA itself, cross-talk of post-translational histone modifications and an interplay of chromatin topology-dependent factors (47). All these modifications alter chromatin structure by changing interactions between nucleosomes. In general terms, this process is comprised by *initiators*, such as some long non-coding RNAs (53), *writers*, which establish the different epigenetic marks, *readers*, that recognize and interpret those epigenetic marks, and regulatory elements, such as *remodelers*, which can reposition nucleosomes, or *insulators*, that form boundaries between domains (46,54).

2.1. DNA methylation

DNA methylation is the most widely studied epigenetic modification. The best characterized mark is the enzymatic addition of a methyl group to the 5'-position of
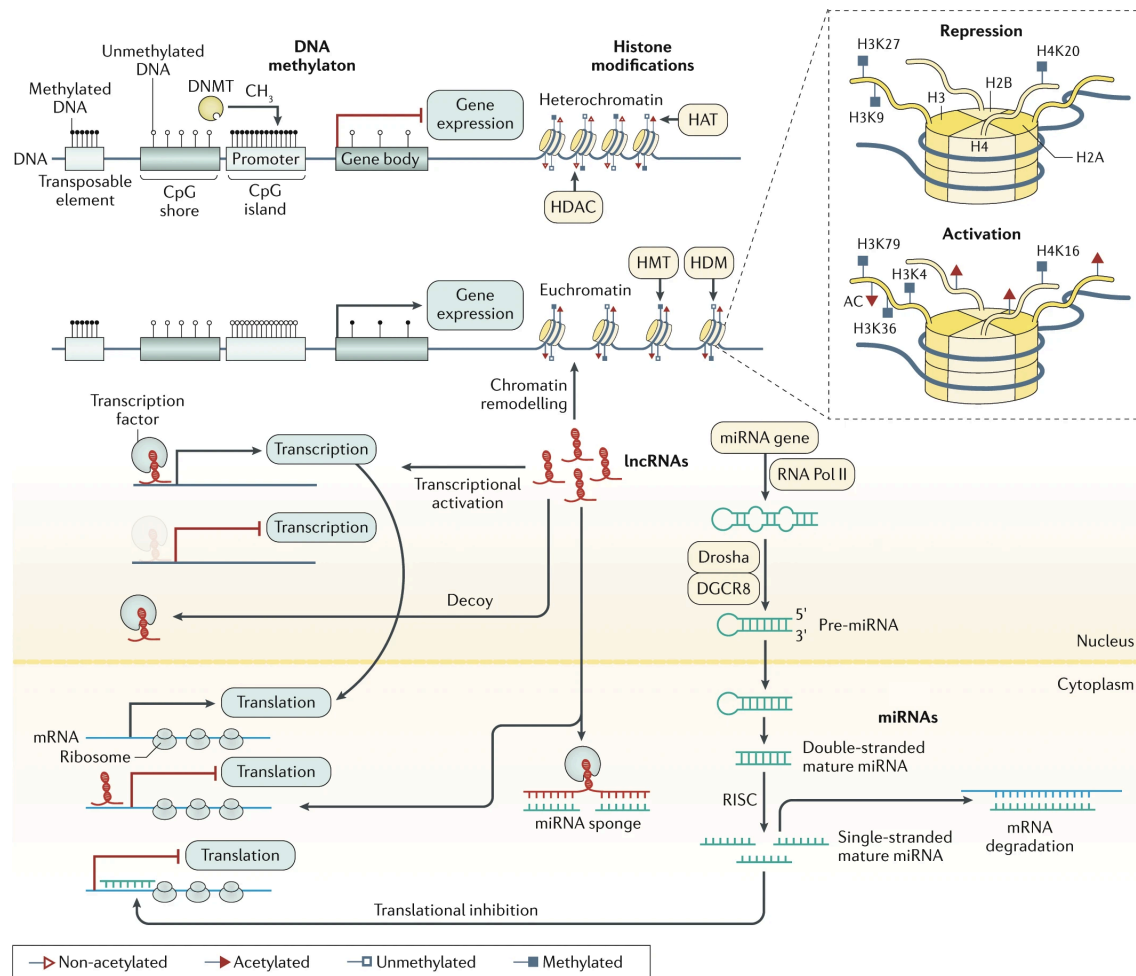
Figure 4. Epigenetic landscape of regulation in colorectal cancer. The main epigenetic regulatory mechanisms involved in CRC are presented in this figure, including DNA methylation, histone modifications, lncRNAs and miRNAs. Image modified from: Jung G, Hernández-Illán E, Moreira L, Balaguer F & Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat. Rev. Gastroenterol. Hepatol.* **17**(2): 111–130 (2020) (52).

cytosine by DNA methyltransferases (DNMTs) to produce 5-methyl-cytosine, a common base in DNA. In humans, DNA methylation occurs predominately in cytosines that precede guanines, forming a CpG dinucleotide. CpG sites are usually distributed in the genome heavily methylated, including those within the gene bodies, to act as repressive regulatory elements of transcription. However, there exists CpG-rich regions, also called CpG islands, more likely associated to the 5' region of vertebrate genes. They generally are 200-2.000 base pairs long, have a CG content over 50%, and are commonly demethylated in normal cells (17,50–52), therefore, they are considered predictors of active or potentially active promoters.

In pathological scenarios, such as cancer, it has been observed dramatic changes in DNA methylation patterns in normal vs. tumor tissue. Tumor cells present a global hypomethylated pattern in relation with its normal tissue, resulting in transcriptional

activation of repetitive elements, differential inclusion of introns involved in alternative splicing or expression of genes that should be silent, such as oncogenes (46,50). The opposite phenomena also occurs in cancer; malignant cells often show hypermethylation of CpG islands, which has been demonstrated to transcriptionally repress key tumor suppressor and DNA repair genes, thus contributing to cancer development (46,52). For instance, hypermethylation of tumor suppressor genes has been characterized including *APC*, *MLH1* or *CDKN2A* (17,50–52).

This is specially the case of CRC, where genome-wide hypomethylation was one of the first aberrant methylation events detected in all stages of the disease. While it has been well-described an hypomethylated status at promoter regions of oncogenes, such as *MYC* or *HRAS* (55,56), and in super-enhancers of the gene encoding for β-catenin, hypomethylation of repetitive elements, which accounts for most of the global hypomethylation observed in cancer, is less understood. Repetitive elements include an heterogeneous class of sequences such as LINE-1 or macrosatellites, among others, which should be normally silenced in adult somatic tissues. LINE-1 hypomethylation is inversely correlated with MSI and CIMP status (50, 55), however, the consequences of LINE-1 hypomethylation in the oncogenic process are not clear; it can be hypothesized that if activated by hypomethylation, might act as retrotransposons that could be inserted in genomic unstable regions to promote genomic instability, and indeed retrotransposon proteins have been found in epithelial carcinomas (57). Macrosatellite hypomethylation also associates with genomic instability in several cancer types (58–60), however mechanisms explaining this connection are for the most part missing. Our group described a novel lncRNA from demethylated NBL2 macrosatellites in a subset of CRC patients; however, whether it contributes to the disease remains to be addressed.

Furthermore, as explained in previous sections, alterations in DNA methylation profiles are associated with different malignant features and survival ratios in CRC. In this regard, genomic instability through MSI has been widely detected in a subset of CRC tumors. Most of MSI tumors usually arise from the inactivation of a DNA mismatch repair gene, such as *MLH1*, by aberrant promoter hypermethylation. Absence of these genes, results in the accumulation of DNA replication errors, especially in repetitive microsatellite sequences, and has been associated with poor survival rates after relapse (35,51). Additionally, it has been also characterized another group of tumors that are highly hypermethylated, presenting a CIMP status, together with frequent mutations in *KRAS* and *BRAF* genes, demonstrating the importance of DNA methylation patterns in the prognosis and development of CRC (14,17,51).

There has been lots of efforts on integrating all these alterations' data together with the different features observed across colorectal tumors though combining genetic, epigenetic and molecular alterations. Figure 5 shows an scheme trying to integrate all the characteristics known to date (61); however, in the future years, this type of meta-analysis will be complemented with new data since our capacity on improving high-throughput techniques is increasing day by day.
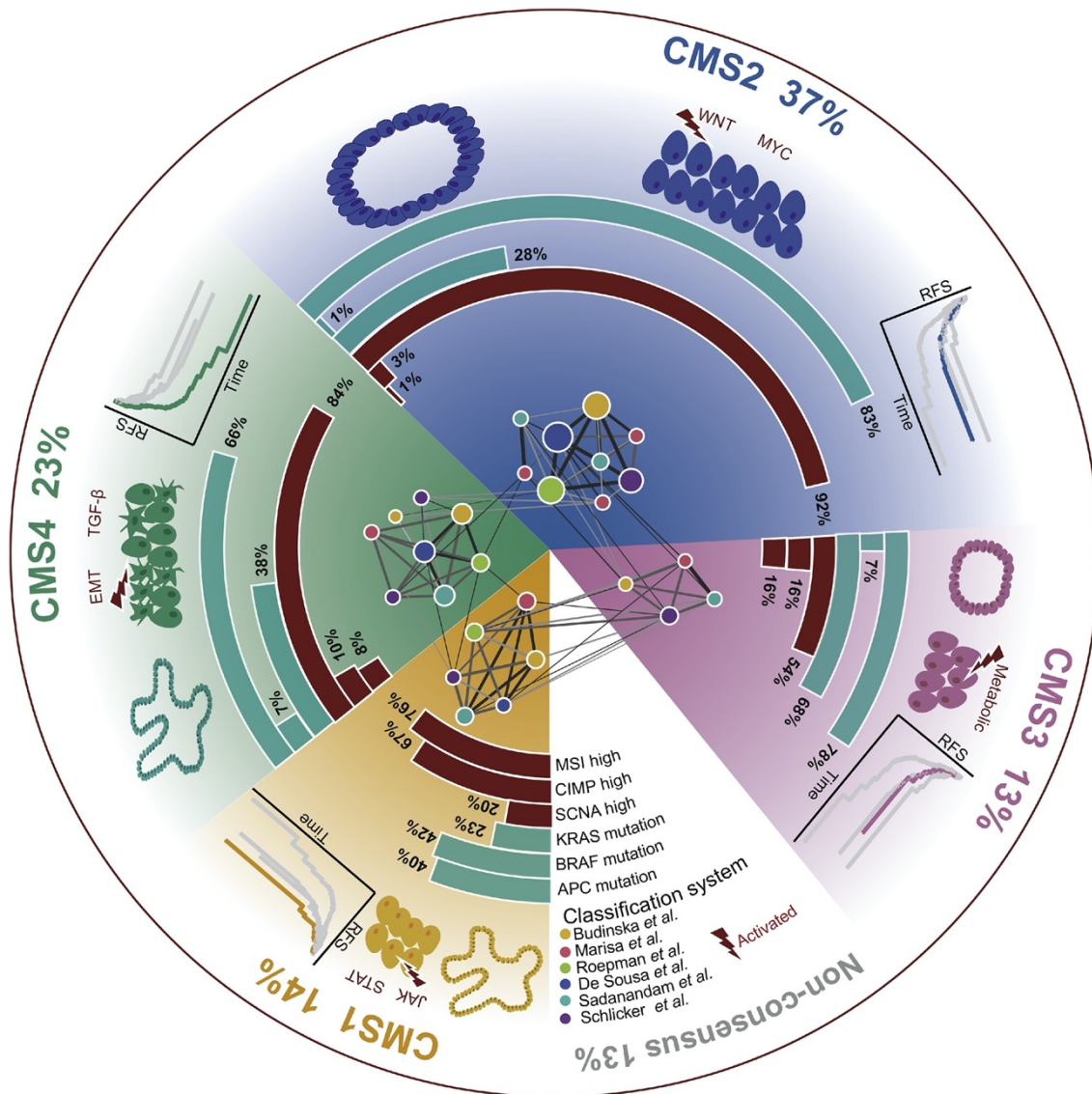
Figure 5. Meta-analysis integrating all the data from the different CMSs in CRC. Integration of genetic, epigenetic and transcriptomic data known to date in the different molecular subtypes of CRC. Image from: Wang W, Kandimalla R, Huang H, Zhu L, Li Y, Gao F, Goel A & Wang X. Molecular subtyping of colorectal cancer: Recent progress, new challenges and emerging opportunities. *Semin. Cancer Biol.* **55**: 37–52 (2019) (61).

Nevertheless, the mechanisms leading to altered DNA methylation in cancer are still not well understood. As mentioned before, DNA methylation is catalyzed through DNMTs. Four active DNMTs have been discovered: DNMT1, DNMT3A, DNMT3B and DNMT3L (17,46,54,62). *De novo* methylation at CpG sites during embryogenesis is performed by DNMT3A and DNMT3B enzymes in cooperation with DNMT3L, which has been suggested to increase and/or modulate their activity at target sites (63,64). DNMT1 acts as a maintenance methyltransferase, therefore it recognizes hemimethylated DNA generated during replication and methylates the newly synthesized CpG dinucleotides whose partners on the parental strand are already methylated. DNTM3A and DNTM3B have been described as well to participate in the maintenance of DNA methylation in addition to DNMT1 (17,46,54,62). However, global DNA hypomethylation observed in virtually all cancer types does not associate clearly with mutations or alterations in the expression of DNMTs, as these mutations are not frequent.

On the other hand, DNA demethylating enzymes for the moment have been only observed in plants (65); passive DNA demethylation, though, might occur during DNA replication in the absence of DNMT1 maintenance activity or proper recruitment to genomic loci (46,62). However, in recent years, the ten-eleven translocation (TET) family of proteins has emerged as an alternative to partially explain DNA demethylation. This group of DNA hydroxylases are in charge of 5-methylcytosines oxidation, generating 5-hydroxymethylcytosines that further derivate to 5-formylmethylcytosines and 5-carboxylmethylcytosines. Despite all the implications of this process are still not well understood, they seem to play an important role on regulating transcription: these final products can be actively removed during DNA repair or passively disappear during replication, thus contributing to DNA demethylation. In addition, they might change the recruitment of some chromatin factors: for instance, they favor or preclude the binding of several methyl binding proteins (MBD proteins) (66), and involve both insulator and enhancer regions, therefore the TET family may be playing a dual role in transcriptional regulation (46,62).

2.2. Histone modifications

Another level of epigenetic regulation that our genome presents comprises histone proteins and all their modifications, that change depending on the needs of the cell. Modifications on histone proteins directly impact DNA conformation by altering the three-dimensional structure of nucleosomes; therefore, they might create a more compacted state of chromatin to transcriptionally inactivate these regions whereas they can also

open the chromatin conformation in other zones, thus promoting the transcription of genes from those areas (51,67).

This type of modifications occur in the tails of all histone proteins and their variants, which protrude from the nucleosome core, and have been mainly identified in lysine, arginine and serine residues, involving different types of chemical groups, the most common being methyl, acetyl, phosphate, ubiquitin and citrullin. However, other modifications comprising biotinylation, sumoylation, ADP ribosylation, propionylation, butyrylation, or glycosylation, to mention a few, are less understood with their functional meaning still being under intense investigation. In contrast, histone methylation has been deeply studied, identifying several degrees of methylation: monomethylation, dimethylation and trimethylation. In general terms, histone acetylation has been associated with transcriptional activation while histone methylation effects may depend on the type of amino acid and its position in the histone tail (50). For instance, active transcriptional states of chromatin are characterized by di- and trimethylation of histone H3 at lysine 4 (H3K4me2 and H3K4me3) and at lysine 36 (H3K36me2 and H3K36me3), and acetylation of histones H3 and H4 at lysines 4, 9 and 14; by contrast, trimethylation of histone H3 at lysine 9 and 27 (H3K9me3 and H3K27me3) are generally associated with inactive transcription as are enriched in heterochromatin (51). However, a particular combination of these modifications at a given loci ultimately shapes the transcriptional status of a gene. Despite all, preferential combinations have been revealed as well as their functional meaning. For instance, H3K4me3 and H3K27me3 co-occur in promoters of silent loci, usually in genes that may become either activated (by losing H3K27me3) or repressed (by losing H3K4me3) during embryonic development or cell differentiation. Thus, understanding the histone code is far more complex than previously anticipated.

Furthermore, it should be taken into consideration that, besides the catalytic function these modifications represent, the enzymes catalyzing them display "reader" domains to recognize and bind specific areas (containing a particular combination of histone marks), susceptible of being targeted. Moreover, these enzymes are controlled by upstream signaling cascades. All these features may open new therapeutic windows; on one hand, targeting the binding pocket could interfere on their catalytic or reader activity, whereas targeting the residues outside the binding pocket could modify the histone sequence specificity or impair their activation. Overall, it implies that histone proteins could dock at different modified residues or at the same amino acid but promoting different modifications and, as such, distinct epigenetic states (46).

2.2.1. Histone acetylation

The acetylation of histones occurs at lysine residues and neutralizes the positive charge of histone tails, thus weakens the interaction between the DNA and histones and promotes a less-compacted state of the chromatin. This process is catalyzed by histone acetyltransferases (HATs) and usually results in gene transcription activation through exposing gene promoter sites to the transcriptional machinery. On the other hand, this reaction can be reversed by histone deacetylases (HDACs) (46,52,67). Additionally, this mark is mainly recognized by a family of proteins that contain bromodomains, present in chromatin remodelers, transcriptional coactivators, HATs or histone methyltransferases (HMTs) (46,67).

Importantly, this modification seems to play a crucial role in cancer development and progression, since hyperacetylation of histones is associated with the aberrant activation of oncogenes, whereas hypoacetylation can silence tumor suppressor genes as well, therefore contributing to tumor growth and expansion (52). For instance, in CRC has been demonstrated a higher level of HDAC2 in adenocarcinomas in comparison to adenomas, suggesting a role on colorectal cancer progression (68,69); in addition, there is also evidence pointing that colorectal cancer progression is accompanied by a reduction on H4K16 acetylation (68,70).

2.2.2. Histone methylation

Histone tails can be methylated at arginine, lysine and histidine residues. The enzymes in charge of catalyzing histone methylation and demethylation are HMTs and histone demethylases (HDMs), respectively. Lysine residues may be mono-, di- or trimethylated whereas arginine residues can be symmetrically or asymmetrically methylated. The different states of histone methylation promote different biological outcomes: histone methylation not only changes the compaction of DNA but also creates sites that can be recognized by different protein complexes, such as the transcription initiation factor TFIID subunit 3 (TAF3), which activates WNT–β-catenin target genes through PHD domain when recognizes H3K4me3 mark (71–73). This may result on transcriptional consequences that strongly influence normal cell differentiation. Indeed, alterations on these transcription processes may lead to activation of oncogenes and potent silencing of tumor suppressor genes, therefore it may also have a potential role in carcinogenesis and tumor progression (46,52,67).

There are several marks that have been deeply studied and are well-characterized. Methylation of H3K4, H3K36 and H3K79 have been associated mostly with active genes in euchromatin, whereas methylation of H3K9, H3K27 and H4K20 are more present in

heterochromatic regions with silenced genes. As mentioned above, the methylation state affects the role of each mark as well. For instance, when H3K4 is monomethylated (H3K4me1) is frequently associated with active enhancers, while H3K4me2/3 tend to be more present at transcriptional start sites (TSSs). Likewise, H3K9 seems to have a dual role: H3K9me3 is associated with gene repression whereas H3K9me1 is present at active genes (46).

Regarding histone-associated proteins, one of the most studied complexes has been Polycomb repressive complex 2 (PRC2), which belongs to the Polycomb group proteins (PcG). The main core units of this complex are EZH2, EED, SUZ12 and JARID2; EZH2 is a H3K27 methyltransferase, thus, PRC2 complex is mainly involved in transcription repression through depositing H3K27me3 mark, impacting different cell functions such as cell proliferation and differentiation (48,54,74). Furthermore, due to its role on regulating transcription, PRC2 has been reported to be implied in cancer: EZH2 is commonly overexpressed in cancer, promoting aberrant histone methylation and, therefore, silencing gene expression through modulating DNA compaction (17,75–77). Interestingly, high expression of two PRC2 components, EZH2 and SUZ12, have been correlated with better prognosis in CRC patients (52,78). Thus, the contribution of PRC2 complex to cancer is context-dependent, as often occurs with many alterations found in oncogenic onset and progression.

2.3. Non-coding RNAs

During the past decades, it was firmly accepted that only the 2% of coding genome was functional; the rest was considered as "junk DNA" or completely worthless. In the last years, due to the implementation of high-throughput genomic platforms, the ENCODE Project Consortium (79) revealed that 80% of the genome was transcribed, although was not subsequently translated into protein, but nevertheless could have critical roles in regulating a wide range of cell biology processes, from embryonic development to malignant proliferation, among many others (80). These mediators are currently known as non-coding RNAs (ncRNAs) and can be subdivided into two different categories based on their size (46,51,52,80):

i. Small ncRNAs: their sequence is less than 200 nucleotides, and comprise miRNAs, siRNAs, piRNAs and snoRNAs.

ii. Long ncRNAs (lncRNAs): are longer than 200 nucleotides.

2.3.1. miRNAs

miRNAs are single-stranded RNA fragments of 18-25 nucleotides in length that regulate gene expression by binding to complementary sequences in the 3'-UTR regions of target mRNAs to promote their degradation or to inhibit the mRNA translation into protein. It has been reported that miRNAs can target specific mRNAs or regulate gene expression profiles by targeting several genes simultaneously (50–52).

miRNAs expression might be deregulated by genetic alterations, such as mutations or deletions; moreover, epigenetics strongly influence miRNAs expression as well through DNA methylation. For this reason, deregulated miRNAs have demonstrated to be potentially oncogenic by inhibiting tumor suppressor genes, while some of them can acquire a tumor suppressive function by inhibiting oncogene's expression (52).

Interestingly, several miRNAs have been shown to be up- or downregulated in pathological scenarios, contributing to diseases such as cancer (50,51). In CRC, several families of miRNAs have demonstrated to be involved in the development and progression of the tumor. For instance, the miR-17-92a cluster, miR-135b or miR-145 regulate the WNT–β-catenin pathway, which is involved in CRC initiation (51,81–83); likewise, p53, which is known to be a key tumor suppressor protein, is also regulated by miR-34a/b/c, miR-133a or miR-143 (51,84–87).

2.3.2. lncRNAs

lncRNAs are an heterogeneous group of ncRNAs involved in many biological processes. They seem to have a critical function at chromatin, as they might act as molecular chaperones or scaffolds for chromatin regulators, thus modifying chromatin access and/or regulating nuclear architecture (46,52). Additionally, they can act as positive or negative regulators of transcription by interacting with promoters or enhancers, and regulating mRNA stability through direct interaction with target mRNAs or acting as a binding site for multiple miRNAs (52).

lncRNAs have demonstrated to be involved in several biological processes such as cell proliferation, differentiation or apoptosis. For this reason, deregulation of these ncRNAs has been associated in many cancer-related pathways (52). One of the best-studied lncRNA is HOTAIR, which acts as a scaffold for the PRC2 and the LSD1-containing CoREST/REST complex, and is aberrantly expressed in CRC. It has been reported that overexpression of HOTAIR correlates with the presence of liver metastases and worst prognosis in CRC patients (46,88,89). Likewise, variations in HOTAIR levels directly alter

PRC2 occupancy, which has been associated with changes on the invasive potential of CRC malignant cells (46,90).

2.4. Chromatin remodelers

As mentioned above, chromatin might be regulated through different mechanisms: DNA modifications, histone modifications as well as several protein complexes that change chromatin architecture, also known as "remodelers". These mechanisms can function individually or can act together to modulate genome-wide topology and gene expression, thus regulating many processes such as cell division and differentiation (91). Moreover, it has been widely studied that, when chromatin regulation processes fail, it has an important impact in development and disease. In this regard, data from The Cancer Genome Atlas Research Network or the International Cancer Epigenome Consortium has pointed out that mutations in chromatin remodelers are present in 50% of cancers; indeed, in some tumors, like the pediatric glioblastoma multiforme, these mutations represent the sole genetic abnormalities found, thus confirming the role of these genes in oncogenic-causative functions (91–93).

Chromatin remodelers are multi-subunit complexes that use the energy derived from ATP hydrolysis to reposition, eject, slide or alter the composition of nucleosomes, thus modifying DNA structure to facilitate the access of DNA-binding proteins and the transcriptional machinery to promote gene expression (47,91,94,95).

Nevertheless, the role that remodelers play in the cell is much more complex. Some remodelers are enzymes that ensure the correct density and spacing of nucleosomes, thus may also contribute to gene repression; other groups of remodelers cooperate with transcription factors and histone-modification enzymes to move or eject histones, improving the binding of transcription factors to DNA. Furthermore, a set of remodelers is in charge of creating chromosomal regions where canonical histones can be replaced by histone variants (95,96).

Based on phylogenetic and functional assays, chromatin remodelers have been classified in four different groups: switch/sucrose non-fermentable (SWI/SNF) complex, imitation switch (ISWI) complex, chromodomain helicase DNA-binding (CHD) complex, and INO80 complex (91,95). One aspect that these complexes have in common is the catalytic activity, which is based on a SWI/SNF2-like core ATPase/helicase; then, each complex contains accessory subunits that include DNA and histone-binding motifs, which provides and extensive complex diversity on each family (91). These four groups of remodelers will be further explained in the next sections.

2.4.1. SWI/SNF complex

The SWI/SNF complex, which was firstly characterized in *S. cerevisiae*, is conserved throughout eukaryotes and has been the most studied remodeler complex. It contains an ATPase subunit from the SNF2 family that alters chromatin organization through sliding nucleosomes or evicting histones from chromatin to promote chromatin accessibility during transcription and DNA repair mechanisms (97,98).

In mammals, SWI/SNF complex can be subdivided into three major complexes (Figure 6): a) cBAF (canonical BRG1/BRM-associated factor), b) PBAF (polybromo-associated BAF), and c) ncBAF or GBAF (non-canonical BAF or GLTSCR1/1L-associated BAF) (91,98,99).
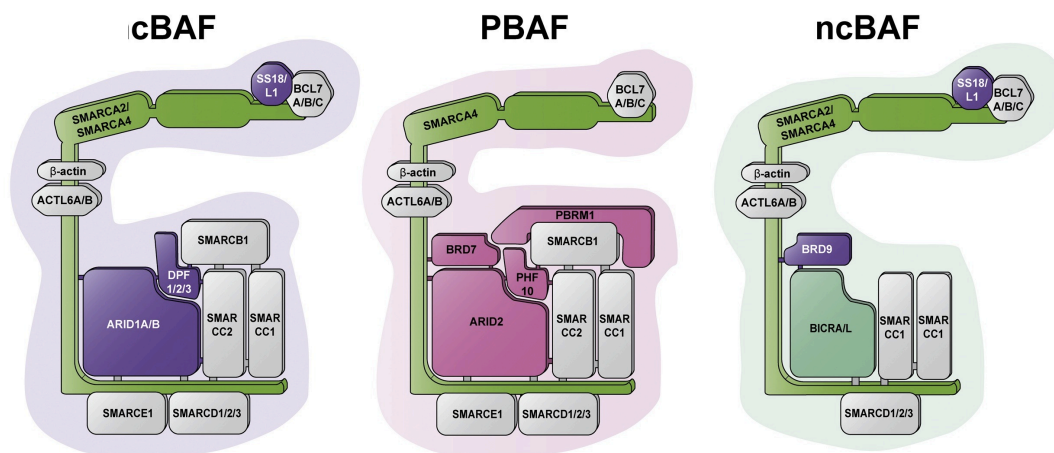
Figure 6. Structure of the different SWI/SNF complexes cBAF, PBAF and ncBAF. Structure with the main subunits of each SWI/SNF complex. Image modified from: Harrod A, Lane KA & Downs JA. The role of the SWI/SNF chromatin remodelling complex in the response to DNA double strand breaks. *DNA Repair (Amst).* **93**: 102919 (2020) (98).

Table 1 depicts the different subunits of the SWI/SNF complexes indicating their role, gene and protein names, which not always coincide. All three complexes contain core subunits that include Ini1, BAF170 and BAF155, and one of the ATPases (BRG1 or BRM); however, they also contain several "accessory" subunits that provide a distinct identity to each of them (97,99). For instance, cBAF complex, which is a 12-member complex, contains one of two mutually exclusive ARID proteins, ARID1A or ARID1B; combination of BRM and ARID1A has been proven to be required for efficient non-homologous end joining (NHEJ) repair pathway, which is one of the main DNA repair mechanisms, by promoting the association of KU70 and KU80 with double-strand breaks (DSBs) (100–102). On the other hand, PBAF complex binds a distinct ARID protein, ARID2, and also binds exclusively to PBRM1 and BRD7; PBAF has been involved in homologous recombination (HR) through promoting sister chromatin cohesion in

combination with Cohesin complex at DNA DSBs (100,103,104). More recently, it was described the ncBAF complex, characterized for uniquely bind to the accessory subunits GLTSCR1 and GLTSCR1L (100,105,106).

TYPE OF SUBUNIT	GENE NAME	PROTEIN NAME	TYPE OF COMPLEX		
			cBAF	PBAF	ncBAF
Core Subunits	SMARCB1	Ini1	+	+	−
	SMARCC1	BAF155	+	+	+
	SMARCC2	BAF170	+	+	+
ATPases	SMARCA4	BRG1	+	+	+
	SMARCA2	BRM	+	+	+
Accessory Subunits	ARID1A (BAF250A)		+	−	−
	ARID1B (BAF250B)		+	−	−
	ARID2 (BAF200)		−	+	−
	PBRM1		−	+	−
	BRD7		−	+	−
	BAF57 (SMARCE)		+	+	−
	BAF60A (SMARCD1)		+	+	+
	BAF60B (SMARCD2)		+	+	+
	BAF60C (SMARCD3)		+	+	+
	GLTSCR1 (BICRA)		−	−	+
	GLTSCR1L (BICRAL)		−	−	+

Table 1. Classification of different SWI/SNF subunits according to their function and to the SWI/SNF complex type that they belong to (+ or –). Some have synonym terms for the protein, while others show different gene and protein names whose acceptance has changed over the years.

SWI/SNF complex and its subunits are well-known to be involved in the regulation of several cellular functions. For instance, mutations or loss of SWI/SNF subunits have been reported to lead to sensitivity to DNA DSBs inducing agents in several cell types. When a DSB occurs in the DNA, SWI/SNF complexes are rapidly recruited and have been shown to be involved in two main DNA repair pathways: through NHEJ and by HR. In this regard, subunits such as ARID1A and BRM of cBAF complex, or BRG1 from PBAF complex, have demonstrated to be required for an efficient NHEJ as they seemed to re-organize the chromatin flanking the DSB to improve binding of repair factors to DNA. Moreover, PBAF complex seems to be required in response to DNA DSBs to efficiently repress transcription as PBAF functions downstream of ATM, a well-known protein to be involved in DNA DSBs repair. On the other hand, there is clear evidence that HR is not efficient in cells deficient for SWI/SNF subunits; however, the exact mechanism by which DNA repair is not optimal is still poorly understood (98,99).

Due to the many roles that remodelers have, should not be surprising that mutations in members of SWI/SNF complex are involved in pathogenesis, specially the development of cancer. The first strong link with this disease was the identification of the biallelic inactivation of *SMARCB1* (Ini1 subunit) as the main cause of ~98% of malignant rhabdoid

tumors. Afterwards, several exome-wide sequencing studies have revealed that more than 20% of all human cancers contain mutations in SWI/SNF-encoding genes. Some of these mutations have been proven to be key drivers in the oncogenesis of different types of tumors (91,99) mostly by impairing tumor suppressor functions (107). For instance, inactivating mutations on *PBRM1*, an exclusive member of the PBAF complex, are present in more than 40% of cases of clear cell renal cell carcinoma (ccRCC), and seem to be associated with a better response in front of immune checkpoint therapy (96,99,100). The mechanisms of action are still under investigation, but for BRG1 mutations there is a clear impairment of TOP2A recruitment to the chromatids during mitosis, resulting in anaphase bridge formations that lead to chromosomal aberrations, a hallmark of cancer development (108). However, upregulation of the expression of selected SWI/SNF subunits in some cancers indicates that the role of this complex in tumorigenesis is more intricate than previously anticipated, and probably genomic location, timing, co-occurrent alterations and concomitant pathways may impact the final output of SWI/SNF alterations.

2.4.2. ISWI complex

Chromatin remodeling complexes containing the ISWI ATPase were firstly discovered in *Drosophila melanogaster* yet it was lately found in many organisms, including yeast and humans. In *Drosophila*, it represents the catalytic core of three types of complexes: NURF (nucleosome remodeling factor), CHRAC (chromatin accessibility complex), and ACF (ATP-utilizing chromatin assembly and remodeling factor) (47,109,110). By contrast, human ISWI remodelers might have two ATPase subunits, SNF2H (SMARCA5) and SNF2L (SMARCA1); furthermore, SNF2H is the ATPase catalytic core subunit equivalent to ACF, CHRAC and RSF (remodeling and spacing factor), whereas SNF2L forms the human NURF complex (47).

Regarding the function of ISWI complexes, they are mainly involved in the assemble and regulation of nucleosomes' space, thus limiting chromatin accessibility and, consequently, gene expression; however, NURF complex may also facilitate the access to chromatin to promote transcription (95,109). In addition, some studies have shown an implication of ACF complex in regulation of chromatin folding into loop domains, therefore contributing to nucleosome positioning to organize large chromatin domains (109). Moreover, there is evidence supporting that mammalian ISWI complex should be required for DNA replication in highly condensed heterochromatin regions, specially pericentromeric areas (109,111). Another potential function recently described is related to the phosphorylation of histone H2A.X (γ-H2A.X): two ISWI complexes, WICH and

CERF, that contain SMARCA5 and SMARCA1 ATPases respectively, have demonstrate to affect H2A.X phosphorylation, which is a key event in the detection and response to DNA damage, specially DSBs (112–114).

2.4.3. CHD complex

Chromodomain helicase DNA-binding (CHD) proteins were initially identified as mammalian DNA-binding factors with a SWI-like helicase domain (112); lately, it was discovered that were composed of two tandem chromodomains in the N-terminal part and the ATPase domain (115). Indeed, the ATPase domain of the CHD complex is highly similar with the one of ISWI complexes; it only differs that bears on its amino terminus these two chromodomains in tandem. For this reason, in general terms, CHD remodelers present three main functions: a) nucleosome assembly and spacing, b) increase gene access through exposing promoter areas, and c) nucleosome editing, mainly by incorporating histone H3.3 (95).

The different enzymes that form CHD complexes can be subdivided in 3 groups: class I (CHD1 and CHD2), class II (CHD3, CHD4 and CHD5), which is also known as the NuRD complex, and class III (contains CHD6-CHD9). In this manner, CHD1 has been shown to promote the stabilization of H2A.X and increase the efficiency of DSBs repair through HR; moreover, CHD2 protein interacts with PARP1 to facilitate histone H3.3 deposition in NHEJ DNA-repair regions. Additionally, subfamily III of CHD remodelers have demonstrated to interact with several transcription factors and to post-translationally modify histone H3 as well (112).

Regarding the most-studied CHD subfamily of remodelers, the NuRD complex, it should be remarked that has been reported its repressor role on transcription when binds DNA due to its associated histone deacetylases; furthermore, NuRD complex interacts with GATAD2A/B proteins to activate a downstream pathway where GATAD2 members interact with MDB2/3 to recruit HDAC1 and/or HDAC2 (95,112). For instance, it is well-described that HDAC1 and HDAC2, when recruited through NuRD complex, can remove acetyl groups on histone tails of proteins highly relevant in many cell processes, such as the tumor suppressor p53. Nevertheless, the role of NuRD complex in diseases such as cancer remains unclear due to presenting opposite effects both promoting or inhibiting tumor growth and metastasis depending on the tissue; this contradiction might be explained by the NuRD complex capacity to associate or modulate the activity of not only tumor suppressor genes, but also oncogenes such as Bcl-6 (116–118).

2.4.4. INO80 complex

The INO80 subfamily of complexes was originally identified in a screen that pointed out genes required to be activated and expressed in response to inositol starvation in *S. cerevisae*; here, it was also identified the INOsitol-requiring *INO80* gene as the one encoding for the ATPase subunit of this complex (95,119,120).

Later on, it was identified that all members of INO80 complex contain actin-related proteins (ARPs) (121). In this regard, it was also characterized a distinguishing feature of INO80 subfamily: it has a split ATPase domain on its catalytic subunit. This unique aspect enables INO80 catalytic subunit to associate at the same time with RVB1 and RVB2 DNA helicases, which recruit ARP5 to the complex when DNA damage occurs and needs to be repaired, or to regulate chromatin transcription (94).

Since its discovery, INO80 complex has been identified to be implicated in several processes related to chromatin and DNA, such as transcription regulation, DNA replication, DNA damage repair (especially when DSBs occur) or nucleosome sliding (47,120).

Despite some of these functions may seem similar to other remodelers, INO80 complex presents unique functions as well. For instance, INO80 subfamily is mainly in charge of all nucleosome editing processes, which involves the incorporation or removal of histone variants to create specialized chromatin regions in a replication-independent manner. In this way, it should be highlighted the role of INO80 complex to incorporate the H2A variant H2A.Z. Several studies showed how INO80C is required to remove H2A-H2B dimers and replace them by H2A.Z-H2B; moreover, INO80C is also implicated on nucleosome sliding to catalyze the eviction and replacement of these H2A.Z-H2B dimers, thus demonstrating a role not only on facilitating chromatin access but also on editing chromatin (95).

Furthermore, INO80 complex has been also described to influence DNA damage checkpoint pathways and to regulate mitotic stability. Checkpoint pathways function in cooperation with DNA repair pathways by altering cell cycle kinetics to facilitate the repair of damaged DNA; this process is tightly related to the appearance of γ-H2A.X around DNA damage areas to recruit other checkpoint proteins. Since it has been widely observed that INO80 complex immediately binds γ-H2A.X after a DSB, should not be surprising altered checkpoints responses in mutants of INO80 complex (119,122). Furthermore, mechanistic studies in INO80 complex mutants also described alterations of chromatin structure around centromere regions, where there is an increase of histone

variant H2A.Z, which correlates with defects in chromosome segregation and polyploidy (119,123).

Due to the fact that INO80 complex presents several, yet highly important, regulatory functions, its malfunction due to mutations or deregulations seems to result in pathogenic situations, such as cancer. For instance, upregulation of INO80 has been correlated with tumor proliferation in lung cancer cell lines; in the same line, high expression levels of INO80 have been identified in melanoma patients (94,124). Nevertheless, since INO80 complex was discovered more "recently", in the early 2000s, there is still few evidence on the exact mechanisms where INO80 complex is involved during tumorigenesis.

2.5. Therapeutics

Since the discovery of epigenetics, understanding its chromatin regulation mechanisms has contributed to what is called "epigenetic drug therapy". As mentioned above, epigenetics are not fixed modifications on the DNA sequence, yet they are potentially reversible; thus, promoter hypermethylation observed in tumor suppressor genes might be a promising target through the inhibition of DNMTs with azacytidine, resulting in cell cycle arrest and increase of apoptosis in cancer cells (125–127).

There are two different types of drugs that target the epigenome currently on clinical trials (128):

- Broad reprogrammers: also called genomic medicines, these type of drugs promote large-scale changes in gene expression; this group includes DNMT inhibitors (DNMTi), HDAC inhibitors (HDACi) and inhibitors of the bromodomain and extra-terminal motif proteins (iBETs).
- Targeted therapies: drugs developed to an specific patient subset, shifting towards the concept of "precision medicine". For instance, EZH2 inhibitors are used to treat lymphomas with EZH2 activating mutations (129); additionally, some IDH inhibitors have been effective in acute myeloid leukemia (AML) clinical trials, where *IDH1* and *IDH2* are frequently mutated (130).

Nevertheless, it should be remarked that clinical studies evaluating epigenetic therapies in CRC have been tested in patients with advanced stages of disease that have failed in other treatments. In general terms, advanced tumors seem to be more heterogeneous and present a higher accumulation of both genetic and epigenetic alterations. For this reason, to increase its efficacy, it might be considered the use of epigenetic therapies as

adjuvant treatment at earlier stages of the tumor since epigenetic alterations manifest early during tumor development and they present less genomic alterations as well (52).

2.5.1. DNA methyltransferase inhibitors (DNMTi)

DNMTi can be divided into two groups: nucleoside analogs, which covalently trap DNMTs onto DNA, and non-nucleoside analogues, that directly bind the catalytic regions of DNMTs. Currently, there are two DNMTi approved by the US Food and Drug Administration (FDA) for the treatment of myelodysplastic syndrome (MDS) and AML: azacytidine (also known as 5-azacytidine) and decitabine (also known as 5-aza-2'-deoxycytidine), which both of them are nucleoside analogues (125,126,131).

In the case of CRC, DNMTi azacytidine has demonstrated a promising synergistic effect in combination with chemotherapeutic agents such as 5-FU, irinotecan and oxaliplatin in CRC cell lines (52,132). In a phase I/II clinical trial, refractory CIMP-high metastatic CRC patients were treated with azacytidine in combination with CAPOX (capecitabine and oxaliplatin), obtaining a good toleration with high rates of stable disease, although no objective responses were reported (52,133).

Nevertheless, one of the main issues with these drugs is their specificity and toxicity. Since these are agents that cause genome-wide decrease in DNA methylation levels, genes can be randomly activated, including potential oncogenes. Indeed, it has also been observed that DNA re-methylates after drug removal, thus sustaining therapeutic drug levels to obtain a clinical benefit is challenging. Moreover, these compounds seem to be cytotoxic for normal cells, interfering with DNA synthesis, and have shown low efficacy in solid tumors due to a less efficient drug penetration (126,128,131).

For this reason, non-nucleoside analogues have gained interest since they are natural molecules, thus they may be less toxic as do not require incorporation into the DNA (126,131). For instance, sulforaphane, which is an isothiocyanate obtained by eating cruciferous vegetables, seems to present anticancer properties through epigenetic regulation. It has been demonstrated that sulforaphane reduces global methylation by inhibiting DNMT1 and DNMT3A, resulting in demethylated CpG islands at CTCF binding regions. Moreover, sulforaphane has also been shown to inhibit HDACs, thus promoting cell cycle arrest, autophagy and apoptosis in colon cancer cells (126,134,135).

2.5.2. Histone deacetylase inhibitors (HDACi)

The benefits of HDACi have been widely addressed in clinical trials. FDA approved vorinostat, also known as SAHA (suberoylanilidehydroxamic acid) and depsipeptide (romidepsin) for treating cutaneous T cell lymphomas, belinostat for the treatment of peripheral T cell lymphoma, and panobinostat was recently approved for the treatment of drug-resistance multiple myeloma in combination with the proteasome inhibitor bortezomib (125,128,131).

HDACi are drugs that aimed to alter the balance between acetylation and deacetylation of histone lysines; they have been tested to treat cancer in several clinical trials since deacetylation of histones is precisely a cause of abnormal gene repression in cancer (125). Despite the exact mechanism by which some genes are more susceptible than others to be affected by these drugs, HDACi seem to reactivate the transcription of tumor suppressor genes such as *P21* through increasing histone acetylation; moreover, HDACi also modulate the activity and stability of key genes in the development of tumors, such as p53 or the nuclear factor kappa B (NF-κB), as HDACs have demonstrated to deacetylate non-histone proteins as well (125,131). Furthermore, it should be also considered that HDACi present a high dose-dependent effect and, thus, may be cytotoxic: at high doses, HDACi have demonstrated to induce DNA damage, especially DSBs, promoting cell cycle arrest and apoptosis of normal cells (125,136).

Regarding to CRC, similarly to results with DNMTi, vorinostat and belinostat presented synergistic effects in combination with FOLFOX or FOLFIRI in preclinical studies that evaluated cell survival through MTT and immunohistochemistry techniques, and phase I/II trials (52,137). In addition, vorinostat demonstrated to be safe in a phase I/II trial in combination with 5-FU and folinic acid in patients with refractory metastatic CRC, although the efficacy was very limited (52,138). Another example of a non-nucleoside analogue would be the case of curcumin, an anti-inflammatory agent that has been widely proved to inhibit cyclooxygenase and epigenetic enzymes such as both HDACs and acetyltransferases (126,139–144); it seems to also regulate genome-wide DNA methylation patterns in colon cancer cells (126,145). Nevertheless, non-nucleoside analogues have not yet entered in clinical trials, thus further research is needed.

2.5.3. Inhibitors of the bromodomain and extra-terminal motif proteins (iBETs)

iBETs bind irreversibly to the bromodomains of BET proteins, which in turn bind to acetylated histone lysine residues to regulate functions such as histone acetylation, chromatin remodeling or transcriptional regulation (128,131). BET family of proteins have

shown an oncogenic role since they may translocate with other genes to alter gene regulation and expression and, thus, to promote the activation of potential oncogenes (131). One of the BET proteins that has been proven to be involved in tumor development is BRD4, which translocate with NUTM1 to form an oncoprotein that inhibits epithelial differentiation and promotes proliferation by upregulating the expression of *MYC* through an enhancer activity (128,131). In this regard, several drugs have been synthesized to target BRD4. For instance, JQ1, originally created to treat autoimmune diseases, displaces the BRD4-NUT oncoprotein from chromatin through competitive binding, thus inhibiting the proliferation in NUT midline carcinoma cells both *in vitro* and *in vivo* (131,146).

2.5.4. Targeted therapy

Targeted therapy has the purpose to find a specific drug for a subset of patients with defined characteristics. One example would be targeting specific genetic activating mutations identified in a type of cancer: in this regard, it has been reported activating mutations in the H3K27 histone N-methyltransferase EZH2 in lymphomas; EZH2 inhibitors have demonstrated to induce a selective killing of cell lines carrying these specific mutations (128,129,131,147).

Another strategy of targeted therapy is the so called "synthetic lethality", where is aimed to inhibit a specific gene in the presence of an inactivating mutation in another gene, thus inhibition of both genes together leads cells to apoptosis. In other words, the loss of either one alone of the genes has a little effect on cell viability yet the simultaneous loss in both genes results in cell death (131). For instance, it has been shown *in vitro* that drugs that inhibit H2K79 N-methyltransferase DOT1L increase the apoptosis of leukemia cells that present *MLL* (also known as KMT2A) gene translocations (128,131,148). Furthermore, SWI/SNF subunits PBRM1 and BRG1 have also been reported to be synthetically lethal with EZH2 inhibition in several cancer cell lines (131,149).

2.5.5. Combined therapy

One of the epigenetic combinations that has been explored is combining DNMTi with HDACi to simultaneously inhibit DNA methylation and histone deacetylation. The rationale behind this idea is that highly methylated regions of the genome are associated with transcriptionally repressed chromatin, which is usually accompanied by deacetylated histone lysines. For this reason, several studies have shown that re-expression of these genes is increased by initial treatment with low doses of DNMTi followed by administration of HDACi (125,128,150). Moreover, multiple clinical trials are

testing these combinations, mainly in hematopoietic diseases such as MDS or AML, despite results are still unclear (125,151–153).

Another combination widely assessed in *in vitro* studies is the combination of DNMTi and/or HDACi with other cancer therapies, such as cytotoxic drugs already given in the clinics to patients. In this way, it should be remarked the idea that many acquired resistance mechanisms to cytotoxic drugs might be related to epigenetic regulation, therefore this resistance could be reversible with drugs that inhibit DNA methylation and/or histone deacetylation (125,128). For instance, there are clinical trials suggesting that the combination of vorinostat, an HDACi, with carboplatin and paclitaxel improved response rates, progression-free and overall survival in patients with metastatic non-small cell lung cancer (NSCLC) (125,154). Additionally, there are also clinical trials showing promising results on sensitizing ovarian cancer to a combination of standard cytotoxic drugs with DNMTi (128,155–157).

2.6. Biomarkers

The term "biomarker" has a dynamic definition that is constantly under revision due to our evolving understanding of cancer. Currently, biomarkers could be described as substances that: a) are easily measured to identify a patient's cancer, b) identify patient's prognosis, and/or c) predict patient's response to a treatment (51).

2.6.1. DNA methylation

In CRC, one of the most studied non-invasive DNA methylation biomarker for diagnosis is the methylation of *SEPT9* gene in plasma, which encodes a GTP-binding protein involved in actin dynamics, cytoskeletal remodeling, vesicle trafficking and exocytosis. The study of this biomarker in large cohorts of CRC patients gave a sensitivity ranging from 48% to 90% and a specificity from 73% to 97%. For this reason, this biomarker was approved in 2016 by the FDA to be commercialized as the first molecular blood-based assay for CRC screening under the name of Epi proColon test (Epigenomics). A second generation of this test, the Epi proColon 2.0, achieved to increase the sensitivity up to 70-90% and specificity increased above 90%. Nevertheless, all studies also shown methylation of *SEPT9* had a limited sensitivity to detect advanced adenomas, which means CRC patients at stages III/IV (51,52,126,158).

Another non-invasive DNA biomarker widely studied to detect CRC is the methylation status of vimentin (*VIM*), a gene highly methylated in adenomas and CRC tissues. Methylation of vimentin can be detected in blood despite is in stool samples where

sensitivity and specificity reach 81% and 95%, respectively, independently of the CRC stage (51,52,159).

Nevertheless, there is an epigenetic biomarker that is currently used in CRC clinical practices: the analysis of somatic *MLH1* promoter methylation. Despite *MLH1* is mainly known to be involved in mismatch repair (MMR) pathways, it has been identified to cause Lynch syndrome as well, which is a specific CRC hereditary syndrome. Moreover, it has been observed that the most frequent cause of *MLH1* inactivation is due to a somatic inactivation of both alleles through promoter methylation. For this reason, evaluating the hypermethylation of *MLH1* has been implemented in the clinics to differentiate between Lynch syndrome patients and sporadic CRC with MMR deficiency (52).

Regarding prognostic roles of DNA methylation biomarkers, despite there is still few evidence, some interesting results have been reported. For instance, hypomethylation of LINE-1 elements have been associated with poor survival outcomes in CRC patients (52,160). Moreover, hypomethylated LINE-1 elements detected in plasma of CRC patients were also correlated with disease progression, especially in patients with larger tumors, advanced lymph node stages and with distant metastasis (52,161,162).

2.6.2. Histone modifications

In the case of histone modifications as biomarkers, there is little evidence mainly due to technical limitations in quantitative analysis and because of their lack of specificity in different cancer types.

Nevertheless, several histone marks have arisen as potential diagnostic biomarkers in CRC. For instance, methylation of lysine 9 on H3 (H3K9) is higher in adenomas and CRC respect to normal colonic mucosa. In the same way, acetylation of H3K27 and H4K12 has been also reported to be increased in CRC (51,52,69,163,164). On the other hand, reduced levels of H3K9me3 and H3K20me3 in circulating nucleosomes were observed in CRC patients when compared with control individuals (52,165).

2.6.3. Non-coding RNAs

In the last years, ncRNAs, miRNAs and lncRNAs have arisen as future biomarkers for diagnosis, prognosis and response prediction in CRC because of a higher stability across different types of samples, such as tissue, blood or stool, and due to their easy identification and quantification through routine laboratory techniques. Nevertheless, in contrast to methylation biomarkers, commercialization and implementation of ncRNAs biomarkers into clinical practice still needs large-scale validation studies (52).

For instance, a study carried out in CRC tissue samples and ten different CRC cell lines identified the top ten differentially deregulated miRNAs in CRC. miR-21 and miR-143, the most abundantly expressed, might be playing key roles in CRC development since they have oncogenic and tumor suppressor properties, respectively (52,166). Moreover, miR-21 detection has demonstrated high sensitivity and specificity, 64% and 85% respectively, in both blood and stool samples in terms of diagnostic purposes, although seems to be a good biomarker for prognosis and survival (52,158,167–169). Indeed, miR-21 seems to be deregulated at early stages of the adenoma-carcinoma sequence, opening the possibility to identify patients that still have not developed a malignant tumor (51,170).

3. Chemoresistance in colorectal cancer

3.1. Chemoresistance and 5-Fluorouracil

5-Fluorouracil (5-FU) is a drug developed more than 50 years ago (171) and widely used at the beginning to treat several types of cancer, such as breast or head and neck tumors, despite best results were always obtained in CRC, especially in combination with other chemotherapeutic agents such as oxaliplatin or irinotecan (172). Nevertheless, resistance to 5-FU and other fluoropyrimidines is a major issue to succeed in advanced CRC patients' therapy.

5-FU is an antimetabolite drug, an analogue of uracil with a fluorine atom at the C-5 position in place of hydrogen. Figure 7 represents the pathways involved in 5-FU metabolism; when enters in the cells, is rapidly converted to three different metabolites: fluorodeoxyuridine monophosphate (FdUMP), fluorodeoxyuridine triphosphate (FdUTP) and fluorouridine triphosphate (FUTP). These active metabolites block the main molecular target of 5-FU, which is thymidylate synthase (TS), encoded by *TYMS* gene (172–175). TS catalyzes the methylation of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP) by using 5,10-methylenetetrahydrofolate as a cofactor. Thus, TS a key enzyme in thymidylate *de novo* synthesis, an essential precursor for DNA replication and repair (172,173).

Figure 7. Metabolism of 5-FU. Pathways involved in the metabolic activation of 5-FU to promote DNA damage. Image from: Longley DB, Harkin DP & Johnston PG. 5-Fluorouracil: Mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**(5): 330–338 (2003) (172).

A critical point for fluoropyrimidines activity is to form an inhibitory ternary complex between 5-FdUMP, *TYMS* and the folate cofactor. This complex inhibits *TYMS* activity to decrease thymidylate levels and, consequently, block DNA synthesis in cancer cells. Low availability of folate cofactor in tumors has been demonstrated to increase intrinsic resistance to 5-FU. For this reason, 5-FU is always administrated in combination with folinic acid (leucovorin) in CRC patients (173,176).

Nevertheless, it has been observed that an increased TS expression could be the main molecular mechanism of 5-FU resistance. Several studies have tried to elucidate the mechanisms by which TS expression is altered, but the exactly pathway remains unclear. In some cases, an increase in TS activity may be explained by an increased copy number of the gene or as a result of translational upregulation. Also, altered

structural forms of TS with lower affinity for FdUMP have been associated with 5-FU resistance (173,177).

Another key enzyme widely studied in 5-FU resistance is dihydropyrimidine dehydrogenase (DPD), which is the main enzyme in charge of 5-FU catabolism. On one hand, it has been reported that absence of DPD expression in CRC patients generates a severe 5-FU toxicity due to decreased drug catabolism. On the other hand, it has been stablished a correlation between high DPD mRNA expression levels and resistance to 5-FU in CRC tumors (172,173).

Furthermore, thymidine phosphorylase (TP), also known as the angiogenic platelet-derived endothelial cell growth factor, is another enzyme involved in the metabolism, and consequently the resistance, of 5-FU. Among several functions, it degrades 5-fluorodeoxyuridine into a less potent form of 5-FU, but it can also catalyze the reverse reaction. Thus, TP levels can modulate the sensitivity of cancer cells to fluoropyrimidines. For instance, higher TP levels have been correlated with tumor growth, invasion and metastasis in clinical studies, all of them related to worse prognosis features (172,173).

3.2. Chemoresistance and Oxaliplatin

Oxaliplatin (OXA) is a third-generation platinum drug developed to overcome resistance against cisplatin and carboplatin, first- and second-generation platinum drugs, respectively. In OXA, amine groups of cisplatin are substituted by 1,2-diaminocyclohexane (DACH) ligand, which is translated in higher water solubility, less toxicity and lack of cross-resistance with cisplatin, together with higher cytotoxicity. Moreover, differently from cisplatin and carboplatin, oxaliplatin is widely used in the clinics in combination with 5-FU to treat adjuvant and metastatic CRC with proven clinical efficacy and a good safety profile (173,178–180).

When oxaliplatin enters the cell is solubilized, and, consequently, able to interact with nucleophilic molecules such as DNA, RNA and proteins, despite DNA is the preferred target (181,182). The main mechanism of action is to form DNA adducts, through both intra- and inter-strand crosslinks, between two adjacent guanine residues or, in less frequency, between a guanine and an adenine; thus, oxaliplatin impairs DNA synthesis, replication and transcription to promote apoptosis of cancer cells (173,175,176,182).

Cell sensibility to this chemotherapeutic agent might be influenced by cell efficiency on repairing these DNA lesions. Several repair molecular mechanisms have been proposed to be involved in oxaliplatin resistance (173,176,182):

- Direct reversal repair mechanism.
- Decrease in cellular uptake and/or increase in cellular efflux.
- Base excision repair (BER) system.
- Nucleotide excision repair (NER) system.
- DNA MMR mechanism.
- Double-strand breaks repair mechanism.

For instance, some works presented that overexpression of CTR1 (copper transporter 1), which participates in the uptake of oxaliplatin, sensitized lung cancer cells to oxaliplatin, although its specific role on resistance remains unclear (178,182,183). Moreover, intracellular ATPases ATP7A and ATP7B have proven a role in resistance to platinum drugs as demonstrated the capacity to sequester oxaliplatin into subcellular compartments, thus limiting its cytotoxicity (182,184). Indeed, there is also evidence that mRNA and protein expression of ATP7B is correlated with response to oxaliplatin in metastatic CRC tumors. Patients with low ATP7B expression of both protein and mRNA levels have the maximum benefit from FOLFOX treatment in comparison with patients that presented high expression of ATP7B (185).

Regarding the BER system, it has been well-characterized a missense variation on *XRCC1* gene, 28152A>G (R399Q, rs25487), a critical subunit of the pathway. This polymorphism has been correlated with a decreased repair activity of the system, thus impairing the outcome of patients treated with FOLFOX CT (175,176,186).

Nevertheless, oxaliplatin-DNA adducts are mainly repaired through NER pathway, which involves several steps and complex enzymes. The most studied NER mediator has been *ERCC1*, together with its catalytic partner *XPF*, that are in charge of an excision step of the damaged DNA fragment followed by the synthesis of a new DNA strand (175,187). In some *in vitro* studies, down-regulation of *ERCC1* has been strongly correlated with sensitivity to oxaliplatin (187–189). Additionally, there is evidence that *ERCC1* could act as a predictive marker of poor response to oxaliplatin when is highly expressed in patients, despite further validations in more clinical trials are still required (175,186,190).

Another oxaliplatin effect is to induce both intrinsic and extrinsic apoptosis pathways, despite the exactly mechanism of action is still unclear. It seems that the most important

component of this pathway would be the tumor-suppressor protein p53, which activates cell-cycle control checkpoints after DNA damage to promote cell death. For instance, gain-of-function mutations or loss of p53, that occur in a wide range of human tumors, have been associated with intrinsic resistance to oxaliplatin in cancer cells (187).

Regarding epigenetics and resistance of oxaliplatin, there is evidence that, for instance, inactivation of *SRBC* through its promoter hypermethylation is correlated with acquired oxaliplatin resistance and poor outcome in both *in vitro* and *in vivo* studies. Since SRBC interacts with BRCA1, a key participant in the repair of DNA DSBs caused by platinum drugs, it might be considered an activating role of SRBC towards BRCA1, leading to an opposite effect of a BRCA1 loss and, thus, promoting the acquisition of oxaliplatin resistance (187,191).

3.3. Chemoresistance and Irinotecan

Irinotecan (7-ethyl-10-[4-(1-piperidino)-1-piperidino]-carbonyloxycamptothecin), also known as CPT-11, is a camptothecin analog developed for the treatment of different types of cancer in the clinics. The target of irinotecan is topoisomerase Iβ (*TOP1*), an enzyme involved in the relief of the torsional stress developed during DNA replication by inducing single strand breaks. Indeed, irinotecan binds to this topoisomerase I-DNA complex to prevent re-ligation of these breaks, resulting in irreversible DSBs and, thus, leading to cell death (192–195). In other words, topoisomerases are a type of nuclear enzymes involved in the maintenance of DNA topology during transcription and replication, thus, reducing DNA twisting and supercoilings. They form covalent links to DNA to allow the passage of single- or double-stranded DNA; specifically, topoisomerase I binds to single-stranded DNA breaks (193,195). Despite the topoisomerase I-irinotecan-DNA union is not lethal by itself, when impacts with the replication forks promotes the formation of DSBs, leading to an irreversible arrest of these replication forks and facilitating cell death (195).

Nevertheless, it has been recently observed that the mechanism of action of irinotecan may also depend on the cell type: in cells with a quiescent proliferating rate but rich in mitochondria, such as hepatocytes, the drug induces mitochondrial dysfunction and oxidative stress; on the other hand, is in highly proliferative cells, such as cancer cells, where the main target of irinotecan is topoisomerase I (193,196).

When irinotecan enters the cell, it has to be hydrolyzed to become its active metabolite, SN-38, which is 1000-fold more potent than irinotecan, through the action of

carboxylesterases CES1 and CES2 (192,195,197). SN-38 is deactivated through a glucoronidation by UGT1A enzyme family (mainly UGT1A1 but also UGT1A7 and UGT1A9), converting it in the liver to the inactive form SN-38G (SN-38 glucoronide) (176,195,197). It is well-known that the missense variant *UGT1A1*28*, characterized by an extra TA repeat in the promoter region of the gene, is responsible of a less efficient bilirubin glucuronidation and, thus, contributes to severe irinotecan-associated toxicity, such as severe neutropenia and diarrhea. This variant was the first one to be related to Gilbert's syndrome (176,192,195,197), and there is enough supporting evidence that indicates *UGT1A1* polymorphisms may be of clinical interest to predict, and consequently prevent, irinotecan toxicity (192).

As irinotecan needs to be converted into SN-38, it requires several enzymes, such as the above mentioned carboxylesterases. For this reason, it has been well established an *in vitro* correlation between carboxylesterases levels and sensitivity to irinotecan in human non-small cell lung cancer cell lines (195,198). Furthermore, several studies pointed out that the main enzyme involved in irinotecan hydrolysis is CES2. In this way, CRC HT-29 cell line that overexpress CES2 increased irinotecan hydrolysis ratio and, thus, was more sensitive to the treatment than the same cell line with increased CES1 expression (192,199).

Another mechanism involved in irinotecan sensitivity includes ABC transmembrane transporters. For instance, MRP member actively transports CT agents out of the cell and has been proven to participate in the active efflux of both irinotecan and SN-38 (195).

Furthermore, failure on repairing irinotecan-induced DNA damage might contribute to irinotecan resistance in cells in combination with RNA transcription processes. In this regard, it has been reported that the collision between RNA polymerase complex and topoisomerase I cleavable complex results in transcription arrest and the appearance of single-strand breaks, thus, promoting topoisomerase I degradation through the proteasome; when topoisomerase I is degraded, repair of the single-strand breaks might occur. For this reason, it has been proposed that tumor sensitivity to proteasome may serve as an indicator of irinotecan sensitivity or resistance. However, studies on CRC cell lines have shown that the reduction of topoisomerase I caused by irinotecan is defective, thus, impairing topoisomerase I degradation though proteasome; whether this defect on topoisomerase I reduction contributes to irinotecan resistance in CRC is still unclear (195,200).

Additionally, NF-κB seems to play a central role in the treatment of irinotecan. NF-κB is an ubiquitous transcription factor that controls transcription through a wide range of genes involved in inflammation and immunity. When cells receive irinotecan to promote DNA damage, NF-κB gets activated and acts as an antiapoptotic factor. In this way, NF-κB inhibits the apoptotic cascade induced by tumor necrosis factor-alpha (TNF-α), oncogenes and/or chemotherapeutic agents, especially irinotecan. For this reason, NF-κB inhibition may contribute to irinotecan-induced apoptosis and, thus, contributing to the therapy with this chemotherapeutic agent (195).

Nevertheless, recent studies have suggested that topoisomerase I is not the unique target of irinotecan. It has been observed that SN-38 can interact with MDM2, a protein involved in cell death mediated through p53 protein, and is also related to the anti-apoptotic BCL-xL protein. In this way, in hepatocellular carcinoma cells, SN-38 seems to induce p53 expression and phosphorylation, which activates a down-stream pathway that includes proteins such as the apoptosis regulator BAX, caspase-3 and caspase-9, together with a reduction on BCL-xL levels, thus triggering cell apoptosis in a p53-dependent mechanism (193,201).

Antecedents

As widely mentioned in the introduction, chromatin remodelers and epigenetic factors seem to play key roles on regulating processes that drive the development of tumors, although the mechanisms of action are still poorly understood in many cases.

Nevertheless, our group and some of our collaborators had found strong evidence that chromatin factor genes might be deregulated in CRC. Our laboratory performed an exome sequencing of 8 CRC samples and many chromatin remodelers, especially SWI/SNF members, presented mutations; furthermore, the same gene could be mutated at different sites within the same s SWI/SNF protein levels by Wes (theoretically a nuclear stable p histone H1 expression, where completely missing in the more m modifications, such as citrullinat
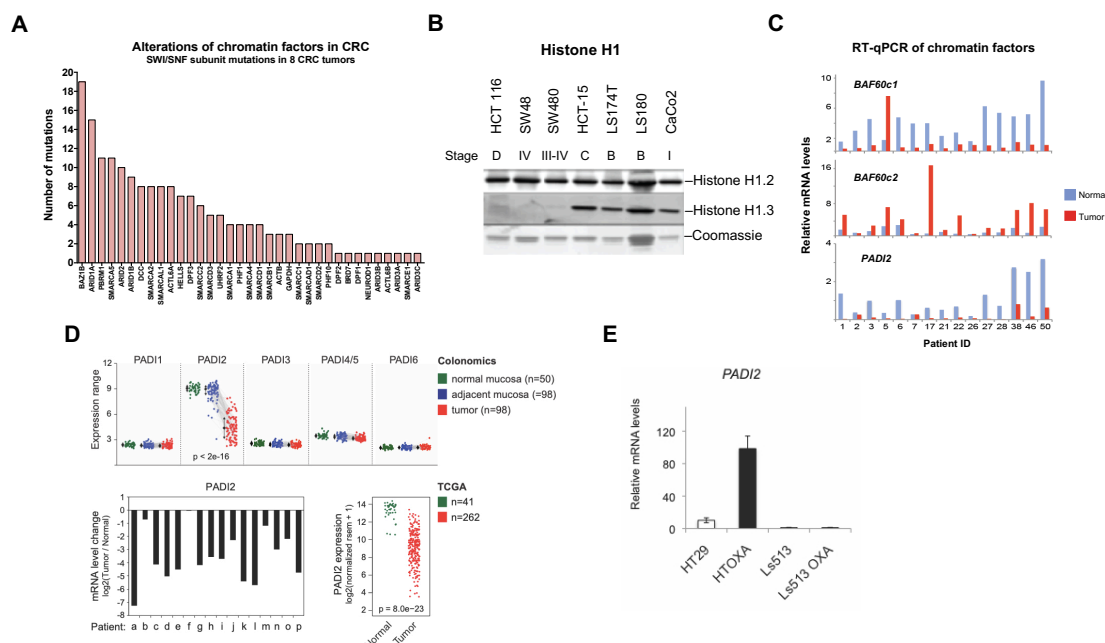
Figure 8. Preliminary results of chromatin remodelers and epigenetic factors altered in CRC. (A) Exome sequencing identified several mutations in chromatin remodelers and epigenetic factors in 8 CRC tumor samples, especially in SWI/SNF members. **(B)** Western Blot of histone H1 subtypes 1.2 and 1.3 throughout different CRC cell lines. Image extracted from: Terme J-M, Millán-Ariño L, Mayor R, Luque N, Izquierdo-Bouldstridge A, Bustillos A, Sampaio C, Canes J, Font I, Sima N, Sancho M, Torrente L, Forcales S, Roque A, Suau P & Jordan A. Dynamics and dispensability of variant-specific histone H1 Lys-26/Ser-27 and Thr-165 post-translational modifications. *FEBS Lett.* **588**(14): 2353–2362 (2014) (202). **(C)** RT-qPCR of BAF60c1, BAF60c2 and PADI2 chromatin factors comparing their expression in normal (in blue) vs. tumor (in red) patient samples (gift of Dr. Peinado's group). **(D)** Expression of PADIs' chromatin factors comparing normal (in green), adjacent (in blue) and tumor tissue (in red) according to Colonomics (top); changes of PADI2 expression comparing tumor vs. normal sample in a cohort of CRC patients (bottom left); PADI2 expression in normal vs. tumoral tissue from a TCGA CRC cohort (bottom right). Image from: Cantariño N, Musulén E, Valero V, Peinado MA, Perucho M, Moreno V, Forcales S-V, Douet J & Buschbeck M. Downregulation of the Deiminase PADI2 is an Early Event in Colorectal Carcinogenesis and Indicates Poor Prognosis. *Mol. Cancer Res.* **14**(9): 841–848 (2016) (204). **(E)** RT-qPCR of PADI2 levels comparing HT-29 and Ls513 CRC cell lines to the same cell lines resistant to OXA (HTOXA and LS513 OXA).

structure that could facilitate replication during uncontrolled proliferation (203). Thus, it became more evident that H1 implication in the oncogenic process could be more relevant than previously thought. Citrullination is promoted by PADI enzymes, and two groups in our institute were also studying the implication of PADIs in cell fate (Dr. Buschbeck's group) and in colorectal cancer (Dr. Martinez-Balibrea's group). Their data indicated that PADI2 RNA was significantly downregulated in CRC tumors (Figures 8C and 8D) (204). Intriguingly, when some CRC cell lines, such as HT-29, were cultured *in vitro* to become resistant to oxaliplatin, PADI2 was upregulated (Figure 8E). Additionally, Dr. Forcales previous work during her postdoctoral research found that histone H1.3 and H1.5 bind different SWI/SNF subunits during proliferation and differentiation, specifically two variants of the BAF60c subunit. Those variants also presented altered expressions in tumor samples compared to normal matched tissue (Figure 8C).

Altogether, taking into account published data, data from our collaborators and our own, indicated an altered chromatin pathway involving subunits of SWI/SNF complex, histone H1, and PADIs, in CRC tumors. Whether these alterations contribute to the onset, progression and/or malignancy of the disease remained unclear; as well as if they could represent diagnostic or prognostic markers. Moreover, while these alterations could be "passengers", but nevertheless they could give novel vulnerabilities to cytotoxic drugs whose mode of action impacts DNA biology. Therefore, these data inspired the main hypothesis of the present thesis.

Hypothesis and Objectives

Figure 9. Scheme of how chromatin factors may facilitate or impair the DNA access of chemotherapeutic drugs in CRC. Our and others' lab data indicate a regulatory pathway between SWI/SNF complex, histone H1 and PADIs. On one hand, our data (unpublished) demonstrates that SWI/SNF complex interacts with several histone H1 variants; on the other hand, PADIs citrullinate histone H1 to promote a less condensed status of chromatin. Therefore, chromatin factors such as PADIs might be influencing DNA structure by promoting or preventing interactions of histones with chromatin remodelers (e.g. SWI/SNF complex), thus affecting different aspects of DNA biology such as gene expression, DNA replication and DNA damage response. Moreover, alterations in chromatin structure may facilitate or impair the cytotoxic effects of chemotherapeutic drugs.

Therefore, the global aim of this thesis was to identify chromatin factors that could represent novel therapeutic targets for the treatment of advanced CRC by synergizing with chemotherapy, to prevent resistance to treatment, which is often the cause of deficient therapies.

To achieve this goal, the specific aims of this thesis were:

1. To perform a pool approach loss-of-function screen using an improved retroviral library of more than 7.300 shRNAs against 912 epigenetic and chromatin factors in the presence of two chemotherapeutic approaches given in the clinics for advanced CRC (FOLFOX and FOLFIRI).

2. To individually validate chromatin factors that could sensitize or impair chemotherapy action.

3. Study the mechanisms of action of selected genes in terms of chemotherapy synergistic or antagonistic effects by *in vitro* functional assays.

4. To explore the biomarker significance, in terms of predictive value for response to treatment, for top hit genes identified in the screen.

Results

1. Setting up of a loss-of-function screening

Genetic screenings are powerful tools to discover novel players in different model settings. We decided to use a pool-approach loss-of-function (LOF) screen using a shRNA library to uncover novel chromatin factors whose deregulation could synergize with chemotherapy (CT). Nevertheless, screenings are noisy by nature and present a high background. To minimize this issue and assure that the genes arisen from this screening may be potential good candidates to overcome chemoresistance, a battery of experiments were performed to set up this type of methodology.

1.1. Safe approach

The shRNA library targets several genes for 912 human chromatin factors (Annex I), some of them with potential tumor suppressor roles. To prevent harmful effects for the researcher, we packaged the retroviral vectors containing the shRNAs into an ecotropic packaging cell line (Plat-E), which produces viral particles that can only infect murine cell lines. Therefore, we had to introduce a murine ecotropic receptor (EcoR) in the human colorectal cancer cell line HT-29 by lentiviral infection. In this way, human HT-29 cells were converted to "infectable" by a retroviral vector specific for murine cells.

Figure 10A shows the expression of EcoR receptor by RT-qPCR in HT-29 EcoR infected cells with adequate controls. HT-29 EcoR cells and those selected with puromycin showed the band corresponding to EcoR receptor, coincident with the band that appears in the viral supernatant. Negative controls (no RNA or no RT) indicated that the band obtained was specifically from cDNA synthesis and not from genomic or environmental contamination. Moreover, another human CRC cell line, such as HCT116, did not show amplification, as expected. To test whether this receptor was completely functional, HT-29 +/- EcoR cells were infected with the ecotropic produced viral supernatants containing a retroviral vector with a GFP (pMSCV-LEPG) or a mCherry fluorophore (pMSCV-LENC). Figures 10B and 10C present the cytometry analysis and a table with the resulting values of pMSCV-LEPG and pMSCV-LENC infections. In this regard, HT-29 EcoR cells were infected at 42.5% and 12.7% for GFP and mCherry vectors, respectively, in comparison with HT-29 cell line (not infected). A positive control was also used, the murine myoblast cell line C2C12, which was infected at higher efficiencies for both pMSCV-LEPG (89.7%) and pMSCV-LENC (28.9%) vectors.

In conclusion, the EcoR receptor was functional and we could use the HT-29 EcoR cell line as our model system for the screening in a safe manner.

A

B

C

	HT-29	HT-29 EcoR	C2C12
pMSCV-LEPG	0%	42.5%	89.7%
pMSCV-LENC	0%	12.7%	28.9%
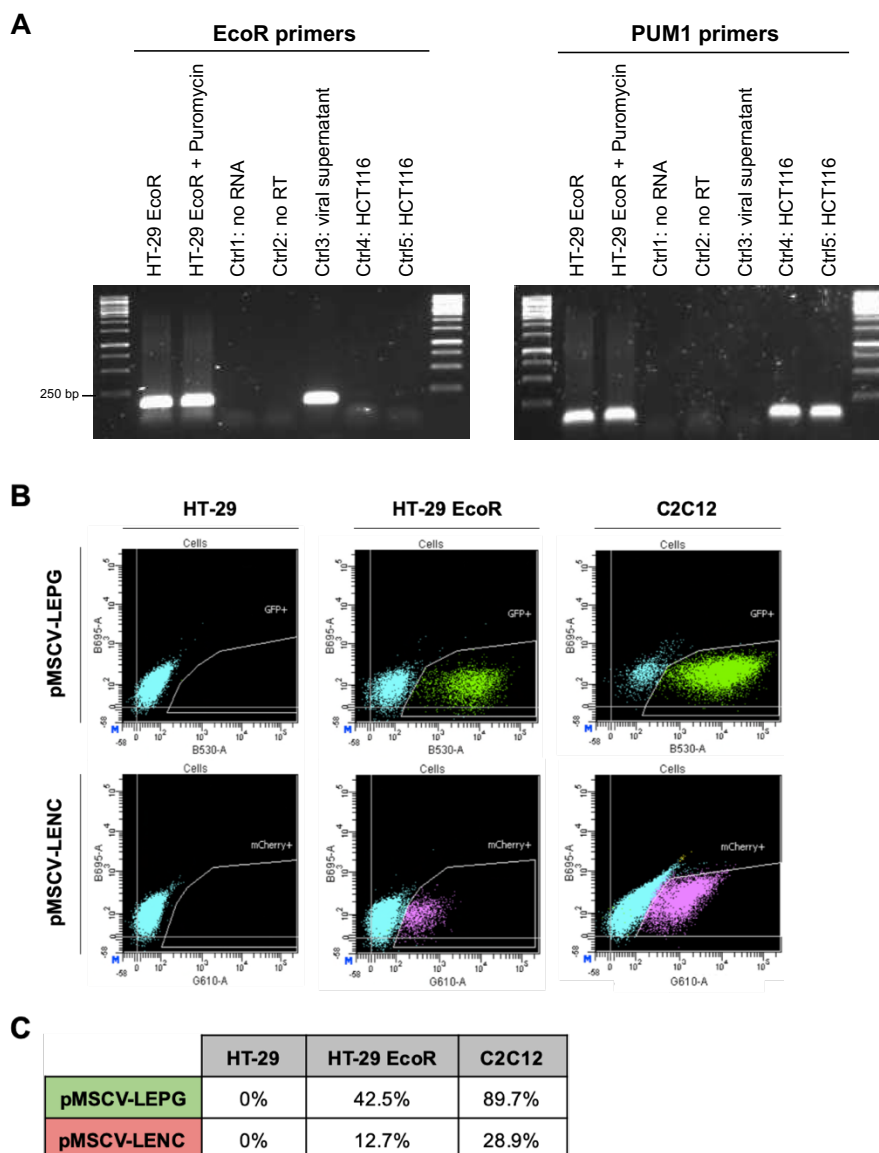
Figure 10. Introduction and validation of the ectotropic receptor EcoR in HT-29 human colorectal cancer cell line. (A) 1% Agarose gel showing EcoR amplification by RT-qPCR. EcoR receptor was correctly introduced and expressed in HT-29 EcoR cells. As a positive control, EcoR was amplified from viral supernatant (Ctrl3); as negative controls, HCT116 cells not infected with EcoR lentivirus (Ctrl4 and Ctrl5), blank (Ctrl1), and no RT (Ctrl2). **(B)** Cytometry analysis of cells infected with pMSCV-LEPG and pMSCV-LENC vectors. **(C)** Table detailing the infection percentages of section B. HT-29 EcoR cells had 42.5% and 12.7% of GFP and mCherry positive cells, respectively, whereas HT-29 were not infected; positive control C2C12 cell line was infected at 89.7% and 28.9% for pMSCV-LEPG and pMSCV-LENC, respectively.

1.2. Determination of viral supernatant dilution to obtain one shRNA per cell

The shRNA library was constructed by our collaborators in a modified backbone, which is ten times more efficient in downregulating its targets than a regular one (205). This potency allows for a pool-approach screening, since even at high supernatant dilution rates (to avoid co-infections) the effects of the shRNAs should be observed. To determine the dilution at which an infection of 1 shRNA/cell is achieved, we co-infected HT29-EcoR cells with two retroviral vectors containing GFP or mCherry fluorophores

(pMSCV-LEPG and pMSCV-LENC, respectively) at different dilutions. Also, vectors were infected at the same proportion (1:1) and at 1:3 proportion, as mCherry vector was observed to infect with less efficiency (compare in 1:1 proportion, where GFP gives around 16% and mCherry 1% in replicate 1 and 9.5% in GFP and 2% in replicate 2).

Table 2 presents a summary of the flow cytometry analysis, indicating the percentages of GFP positive, mCherry positive, and double positive (DP) infected cells. The dilution at which DP cells disappear was slightly different in two independent replicates, ranging between 1/3 and 1/6 of viral supernatant dilutions. Therefore, to ideally obtain no DP cells, we decided to be stringent and the viral supernatant had to be diluted until no higher than 1% of cells were infected.

VIRAL SUPERNATANT		REPLICATE 1			REPLICATE 2		
		GFP (%)	mCherry (%)	DP (%)	GFP (%)	mCherry (%)	DP (%)
Mock		0.2	0	0	0.1	0	0
pMSCV-LEPG		29.75	0	0	17.35	0	0
pMSCV-LENC		0	3.4	0	0	4.2	0
1:1 proportion	2X	16.65	1.1	0.5	9.5	2	0.15
	1X	16	1.05	0.4	8.5	1.75	0.15
1:3 proportion	1X	10.95	2	0.45	5.55	2.6	0.1
	2/3 dilution	9.85	1.85	0.25	3.75	1.8	0.1
	1/3 dilution	6.15	1.05	0.1	2.1	0.95	0
	1/6 dilution	3.7	0.65	0	1.2	0.5	0
	1/15 dilution	1.55	0.35	0	0.55	0.2	0

Table 2. Co-infection assay of pMSCV-LEPG and pMSCV-LENC vectors. Percentage of fluorescent cells by cytometry analysis for GFP, mCherry and DP cells in two independent co-infection assays in HT-29 EcoR cells. As controls, non-infected HT-29 EcoR cells (Mock) and HT-29 EcoR cells only infected with pMSCV-LEPG or pMSCV-LENC vectors. In 1:1 proportion, 2X and 1X supernatant quantities were used, obtaining low levels of mCherry infected cells but still obtaining DP cells. To obtain comparable infections, supernatants from pMSCV-LEPG and pMSCV-LENC were used in a 1:3 ratio. Supernatant dilutions higher than 1/3 resulted in 0% DP cells.

1.3. Validation of RNAi pathway with control shRNAs

Due to screenings' high background, it was necessary to include robust internal controls that minimize false positives. As we would monitor the presence and/or absence of clonal populations of cells, adequate controls such as shRNAs that will kill all target cells (positive controls) and shRNAs that do not bear any effect (negative controls) were necessary. These shRNAs set up the threshold window where different degrees of effects can be obtained by many different shRNAs. Moreover, this approach makes sure that the siRNA pathway is not impaired in this cell line; otherwise, this screen could not be performed.

Figure 11 presents a kill curve assay analysed by flow cytometry, where all three control vectors worked as expected. The strong positive control (in red), which has a shRNA against *RPA3* gene (a RNA polymerase II subunit), killed almost all cells; another positive control containing a shRNA against *MYC* gene (in blue), killed steadily dividing cells. As a negative control, we used a shRNA that contains the luciferase gene from Renilla (in green), not present in human cells, thus, having no effect on HT-29 EcoR cells. To have an additional control vector, we also monitored the effects of HT-29 EcoR cells infected with pMSCV-LENC alone. Therefore, controls behaved as expected.
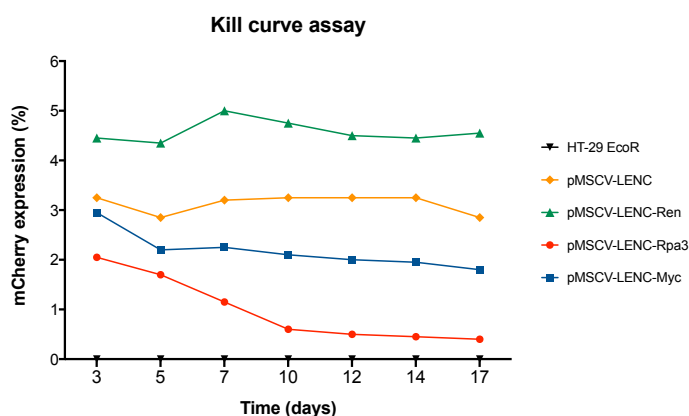
Figure 11. Kill curve assay. mCherry fluorescent levels (%) were monitored by flow cytometry during 17 days after infecting HT-29 EcoR cells with different control vectors that will provide us the threshold window of harmful effects that we should expect in the LOF screening. In red, pMSCV-LENC-Rpa3 is the strongest positive control, killing almost all cells after 10 days in culture; in blue, another positive control, pMSCV-LENC-Myc, killed all dividing cells. In green, the negative control, pMSCV-LENC-Ren, is represented; also, pMSCV-LENC was used as an infection control (in orange).

1.4. Determination of chemotherapy concentrations to mimic clinical regimes

One of the main aims of this project is to find potential new therapeutic targets to overcome patients' chemoresistance to FOLFOX and FOLFIRI, the two major chemotherapeutic treatments in CRC. These treatments consist in combinations of 5-FU and Leucovorin with OXA (FOLFOX) or SN-38 (FOLFIRI). For this purpose, was necessary to establish a culture protocol that mimics the regimes given in the clinics, while setting the adequate concentrations allowing to identify synergistic effects of the chemotherapeutic drugs and the shRNAs. For *in vitro* cultures, FUOX mimics FOLFOX treatment without Leucovorin, which is given to patients for better absorption purposes and, thus, it is not necessary in cell cultures; in this same way, FUIRI mimics FOLFIRI treatment without Leucovorin. The first step to establish the desired concentrations of FUOX and FUIRI for the screenings was to identify by XTT individual IC_{50}s of the drugs composing these treatments: 5-FU, OXA and SN-38. Figure 12 shows dose-response curves of 5-FU (in blue), OXA (in red) and SN-38 (in yellow) at different drug concentration points; we determined IC_{50} values for these drugs at 10 µM for 5-FU, 2 µM for OXA, and at 5.5 nM for SN-38.

Dose-response curves
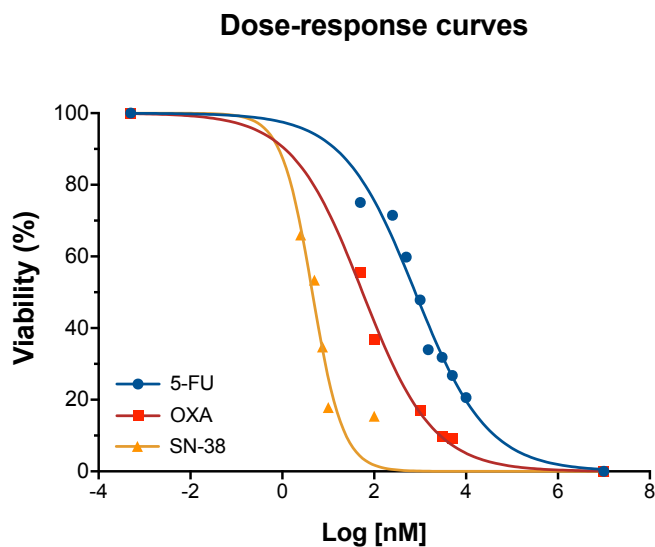
Figure 12. Dose-response curves of 5-FU, OXA and SN-38 at different drug concentrations. IC_{50}s were determined at 10 µM for 5-FU (in blue), 2 µM for OXA (in red), and 5.5 nM for SN-38 (in yellow).

However, using the individual IC_{50}s concentrations in combinatorial treatments (FUOX and FUIRI) resulted in higher mortality rates than expected. This data indicated that we had to test directly several combinations of drug concentrations to determine IC_{50}s for FUOX and FUIRI treatments.

Nevertheless, after several discussions with oncologists and researchers, we decided to set up our screening using the concentrations of combined treatments that kill approximately 20% of the cells (IC_{20}) during a relatively long period (we arbitrarily set it at three weeks), which comprised four cycles of treatment. Briefly, if harsh treatments are used, such as IC_{80}s, they will result in survival of cells enriched for shRNAs that tolerate high CT levels; therefore, favoring the identification of "resistant genes" (when downregulated or absent they confer resistance, resulting in cell survival). However, in these treatments, relevant sensitizer genes would be difficult to be identified since many shRNAs would drop-out from the culture, without being necessarily good sensitizers; they might be still present in mild treatments. In mild treatments such as IC_{20}s, drop-out shRNAs would identify the most susceptible genes that when absent potentiate CT effects resulting in cell death, what we would call "sensitizer genes". In other words, IC_{80}s treatments would favor the identification of resistant profiles, whereas IC_{20}s would favor the identification of sensitizer profiles.

Instead of using XTT, we decided to monitor cell survival by flow cytometry using dyes that stain live cells (DiOC) and dead cells (DAPI). In this manner, transitory phenotypes for apoptotic cells could also be quantified allowing for a more detailed analysis of the combinatorial drug effects (and future synergies with shRNAs). Moreover, interactions between CTs and XTT reagents could also interfere with the final result (206–208). Figure 13 shows the percentage of live and dead cells for several drug concentrations

of FUOX (left) and FUIRI (right). FUOX IC_{20} was achieved after four cycles of a combination of 0.1 µM of 5-FU + 0.02 µM of OXA, and FUIRI IC_{20} was achieved after four cycles of a combination of 0.1 µM of 5-FU + 0.055 nM of SN-38. Additionally, we performed a screening with harsh CT conditions (IC_{80} achieved at 1.67 µM of 5-FU + 0.33 µM of OXA for FUOX, and at 1.67 µM of 5-FU + 0.92 nM of SN-38 for FUIRI).

Determination of IC_{20} for FUOX and FUIRI after 4 cycles of treatment
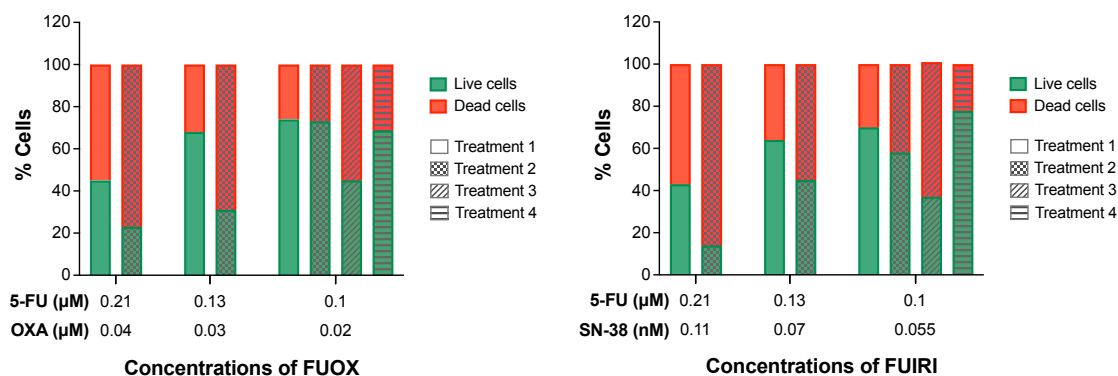
Figure 13. Determination of IC_{20} for FUOX and FUIRI after 4 cycles of treatment. Live (in green) and dead (in red) cells at different concentrations of FUOX (left) and FUIRI (right) after 4 cycles of treatment (marked with different patterns) to establish IC_{20} of combined therapies. In the case of the higher drug concentrations tested, experiments were stopped after treatment 2 due to a high mortality % of cells. In the case of FUOX, IC_{20} was established at 0.1 µM 5-FU + 0.02 µM OXA, whereas IC_{20} for FUIRI was achieved at 0.1 µM 5-FU + 0.055 nM SN-38.

2. Identification of chromatin regulators that affect cell survival in response to chemotherapy

2.1. Loss-of-function screenings

One of the hurdles of a pool-approach method is that we had to scale up the quantity of cells to obtain a full representation of all shRNAs, while trying to obtain individually integrated molecules per cell.

Figure 14 summarizes how the experiment was performed; firstly, HT-29 EcoR cells were infected with the improved retroviral library that contains 7.300 shRNAs against 912 chromatin factors. To obtain a 1.000X representativeness of each shRNA when DNA is sequenced, we had to recover at least $7.3 \cdot 10^6$ live cells (7.300 shRNAs x 1.000) after 4 cycles of chemotherapy. Taking into account that we had previously determined that 1% infected cells will ensure achieving an individual shRNA per cell (Table 2), we consequently had to seed 100 times more cells, giving a seeding number of 730 million HT-29 EcoR cells at the beginning of the experiment (section 3 of Materials and Methods). After 21 days of treatment with FUOX and FUIRI, genomic DNA was extracted and sequenced in parallel with Illumina HiSeq to compare the abundance of shRNAs in control and treated cells.
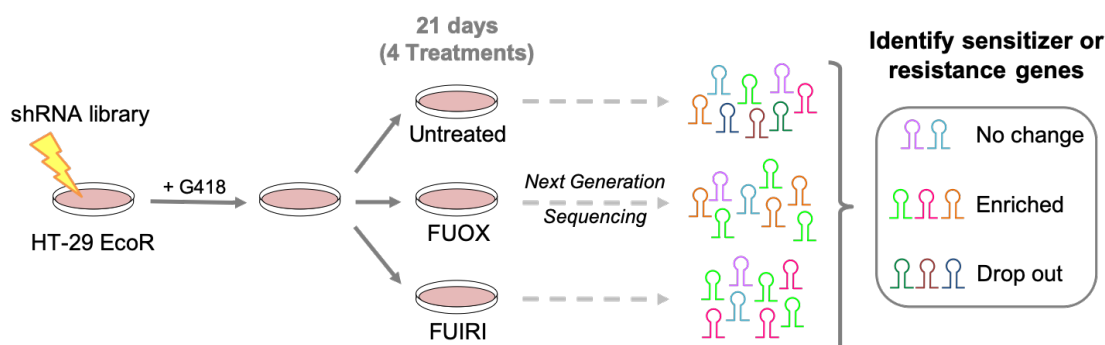
Figure 14. Scheme of how loss-of-function screenings were performed. HT-29 EcoR cell line was infected with a shRNA library that contained 7.300 shRNAs against 912 chromatin factor genes. Two types of screening were performed to identify possible drug targets and biomarkers: with low chemotherapeutic regimens (IC_{20}), which puts more focus identifying shRNAs that disappear in treated conditions respect to controls (drop-outs), and high chemotherapy (IC_{80}) to identify enriched shRNAs that favor resistance in treated conditions respect to controls.

2.2. Analysis of sequenced data

Figure 15 presents volcano plots of differentially expressed shRNAs with top candidate hits for FUOX (Figure 15A) and FUIRI (Figure 15B) treatments at IC_{20}s. It should be taken into account that there are:

- shRNAs that disappear (drop-outs) in the treated conditions respect to the controls (left side of the volcano plot), which are shRNAs that target "sensitivity genes" (required for survival): depletion of these genes sensitizes the cells to the

chemotherapy and, thus, contribute to cell death. These genes are those suitable to be considered as novel combinatorial drug targets.

- shRNAs that are enriched in the treated conditions respect to the controls (right side of the volcano plot), which target "resistance genes": downregulation of these genes confers resistance to the cells in the presence of chemotherapy, thus, their decrease contributes to cell survival. These genes could be considered potential biomarkers that when absent or low expressed in primary tumors could predict worst response (resistance to treatment).

A

B
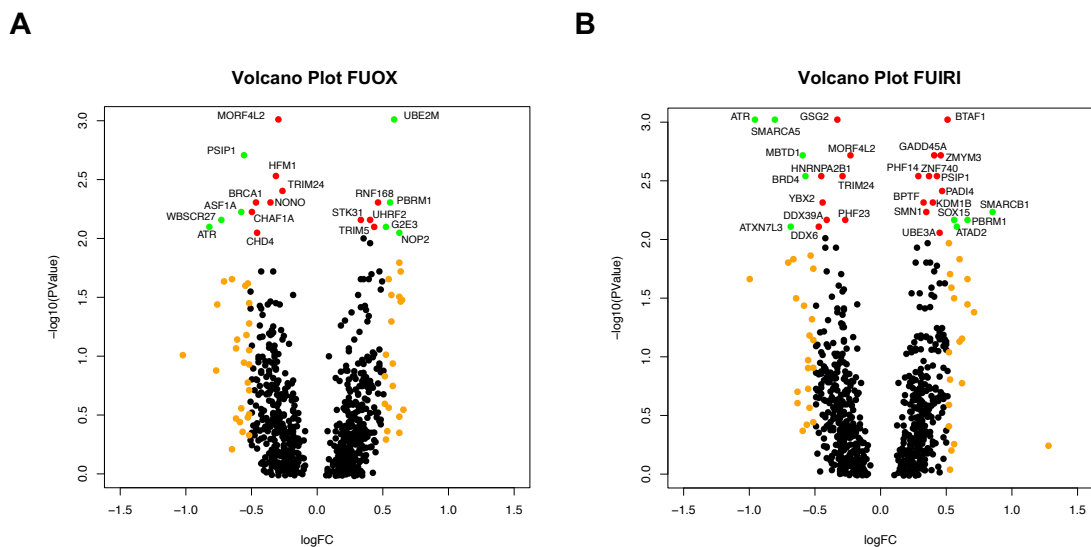
Figure 15. Volcano plots of FUOX (A) and FUIRI (B) top candidate hits from the screening. In red, genes with *p*-value <0.1; in orange, genes with log2FC >0.5; and in green, genes with *p*-value <0,1 and log2FC >0.5.

665 genes were identified to influence cell survival upon FUOX treatment: 338 genes showed their shRNAs depleted and 327 genes whose shRNAs were enriched; for FUIRI treatment shRNAs were depleted for 353 genes and enriched for 315 genes (Figure 16A). Tables 12-15 (Annex II) summarize top 50 enriched and dropped-out shRNAs for FUOX and FUIRI treatments. From these lists, we filtered the data by genes that had at least 6 of the shRNAs behaving in the same direction; in other words, genes that contained at least 6 out of 8 shRNAs disappearing or over-represented in treated cells in comparison to control condition. In this manner, we obtained several top candidate genes distributed as Figure 16B and Tables 16 and 17 (Annex II) show. When performing Gene Ontology (GO) analysis with these gene lists, sensitizer genes (Figure 16C) appeared to be mainly involved in regulation of organelle assembly and response to DNA damage in FUOX (in blue) whereas genes related to regulation of metabolic processes were the most represented in FUIRI (in orange). In the case of resistant genes (Figure 16D), FUOX genes (in blue) were involved in several pathways such as nucleosome

disassembling or ATP-dependent chromatin remodeling, while FUIRI genes (in orange) were not only implicated in regulation of metabolic processes but also in ATP-dependent chromatin remodeling and regulation of DSBs repair.
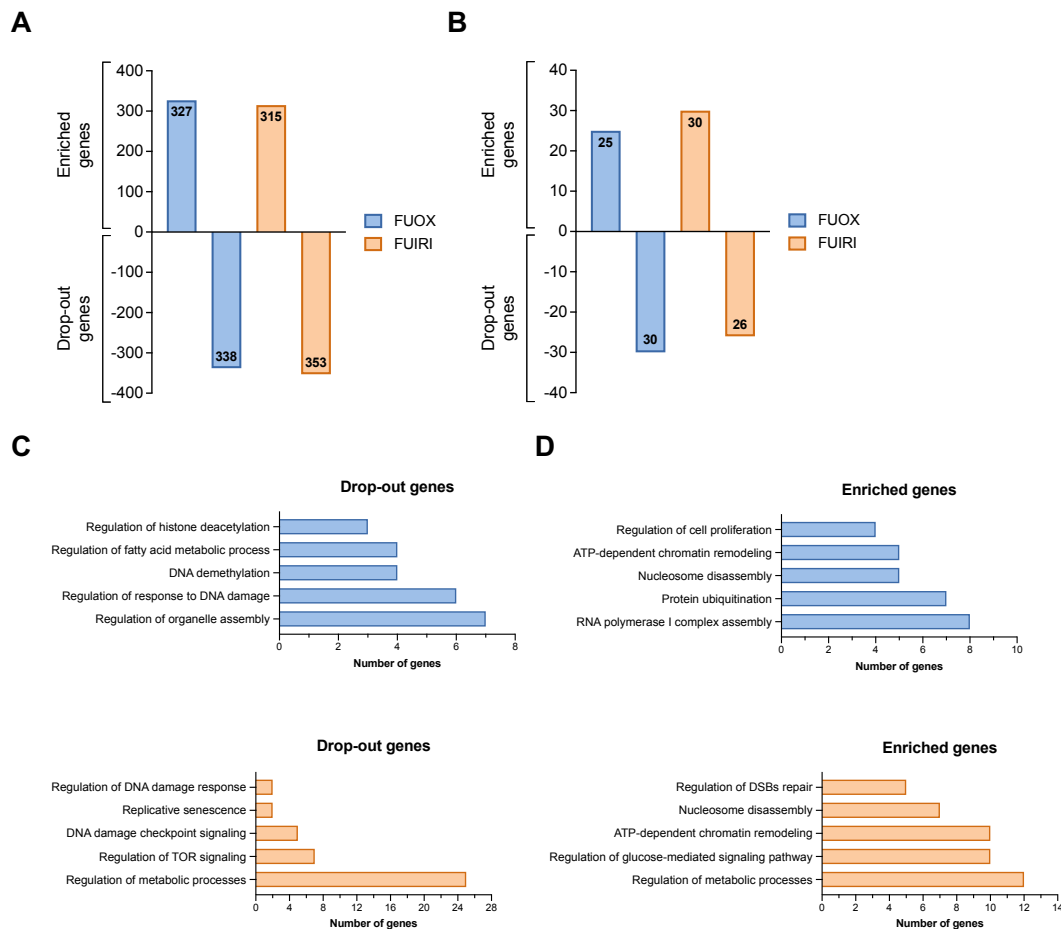
Figure 16. Analysis of enriched and drop-out genes arisen from the screening. (A) Bar graph representing the number of all enriched and depleted genes obtained from the screening for FUOX (in blue) and FUIRI (in orange) treatments. **(B)** Bar graph representing the number of enriched and depleted genes for FUOX (in blue) and FUIRI (in orange) treatments after filtering the data to obtain genes that had at least 6 out of 8 shRNAs behaving in the same direction. **(C and D)** Gene Ontology analysis of the biological processes where depleted (C) and enriched (D) genes (FUOX genes in blue and FUIRI genes in orange) were involved. GO analysis was performed with PANTHER Overrepresentation Test performed using a Fisher's Exact test without any correction; reference lists, which contained all enriched or depleted genes obtained from the screening analysis (section A data), are in Annex II.

At the end, we reduced the list of interest to 21 candidate genes for "sensitizing" to FUOX treatment, 18 for FUIRI, and 8 candidate genes that were common for both treatments (Figure 17A, left); the names of the sensitizers are shown in Figure 17B. Regarding "resistant" genes, we selected 10 possible targets for FUOX, 11 for FUIRI, and only 1 common for both treatments (Figure 17A, right). These lists of resistant genes were also compared with data from the screening experiment where we used high doses of FUOX and FUIRI (IC_{80}); most of the cells died; however, we could recover enough material to perform one sample for sequencing. Therefore, in absence of biological and technical replicates, we cannot show the volcano plots. However, the lists of genes are shown in

Tables 18 and 19 (Annex II). Comparing both lists of resistant genes, the final list of selected candidate genes is represented on Figure 17C.
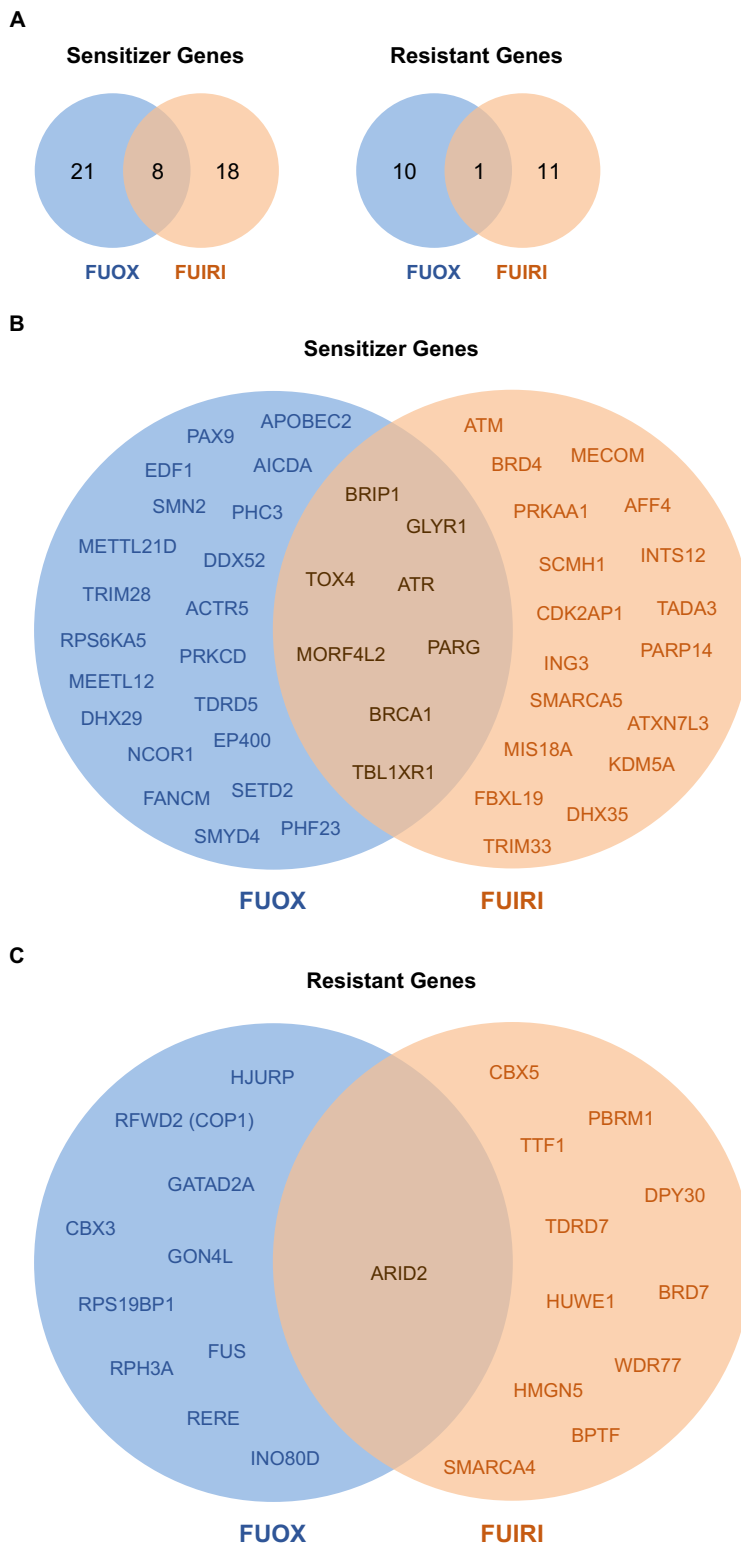
Figure 17. Pie charts specifying the top candidate sensitizer or resistant genes obtained in the LOF screenings for FUOX and FUIRI treatments. (A) Pie chart presenting the distribution of the number of sensitizer and resistant genes for FUOX and FUIRI treatments obtained in the LOF screenings. **(B and C)** Pie charts showing the name of all the top candidate sensitizers (B) or resistant (C) genes classified by FUOX, FUIRI or common.

3. Individual validation of each candidate gene

In total, we selected 69 top candidate sensitizer and resistance genes from the screening; however, not all of them could be deeply assessed. For this reason, based on the known characteristics of each gene and published bibliography, we narrowed the list to 22 genes that showed interest based on diverse criteria such as:

1. Unknown functions and no available drugs,
2. Unknown mechanisms yet part of known complexes involved in DNA repair,
3. Roles in resistance in other cancer types.

In this manner, we could cover several scenarios in which some unprecedented but high-risk targets could be further evaluated as well as more conventional ones but still new as sensitizers for CRC. Once we decided our possible candidate genes from the screening, they needed to be individually validated to verify their potential value as new biomarkers or possible drug targets, or if they represented false positives that arose from the noise of the screenings.

3.1. Validation of 70% of the candidate genes (15 out of 22)

Individual validation of the selected 22 candidate genes was first approached by RT-qPCR and cell viability assays to test the efficiency of the individual knock-downs at the RNA level, and whether their influence on cell survival when treated individually with FUOX or FUIRI was still the same as predicted by the screening. Taking into account these two types of approaches, four different scenarios were considered:

1) <u>RT-qPCR and cell viability assay worked as expected</u>: when successful results were obtained in both assays (RNA downregulation and high mortality of sensitizer or low mortality for resistant genes), we considered those genes as validated; consequently, we assumed their potential role as biomarkers or drug targets, despite a further study would be required. 15 genes fell in this category (Table 3).

2) <u>RT-qPCR demonstrated downregulation of targeted gene yet cell viability assay did not present expected mortality tendencies</u>: genes included in this group were considered false positives arisen from the screenings' background; thus, they were discarded. 5 were found in this category (Table 3).

3) <u>RT-qPCR did not show downregulation of targeted gene but cells presented expected mortality tendencies</u>: this group probably contained off-target effects. These genes could not be directly discarded, as their downregulation was not

achieved; with this group of genes we decided to analyze the effect of a second shRNA. 2 genes appeared in this category although, after analyzing a 2nd shRNA, only 1 gene remained in this section (Table 3).

4) <u>Neither RT-qPCR nor cell viability assays functioned as predicted</u>: genes whereby these two assays did not work could not be directly discarded although they were not good candidates; again, we tried the second shRNA to decide whether those genes were false positives or not. 3 genes were inside this group although, after analyzing a 2nd shRNA, only 1 gene was maintained (Table 3).

Figure 18 presents all the validation results for sensitizer genes. In blue are represented FUOX genes whereas FUIRI genes are in orange. On the left (Figure 18A), RT-qPCR with two different normalizer genes (*PUM1* and *MRPL9*) was used to measure the knock-down efficiency at RNA level. Successful individual knock-downs were considered when there was approximately less than 50% of expression respect to the shControl. Figures 18B and 18C present graphs of how live cells respond to chemotherapy at different doses measured by flow cytometry (Figure 18B) and normalized graphs (Figure 18C) of these values to better observe the scale of the response. The selected chemotherapeutic doses ranged from twice the combination of the individual IC$_{50}$s, several dilutions of the IC$_{50}$s, and the combined IC$_{20}$.
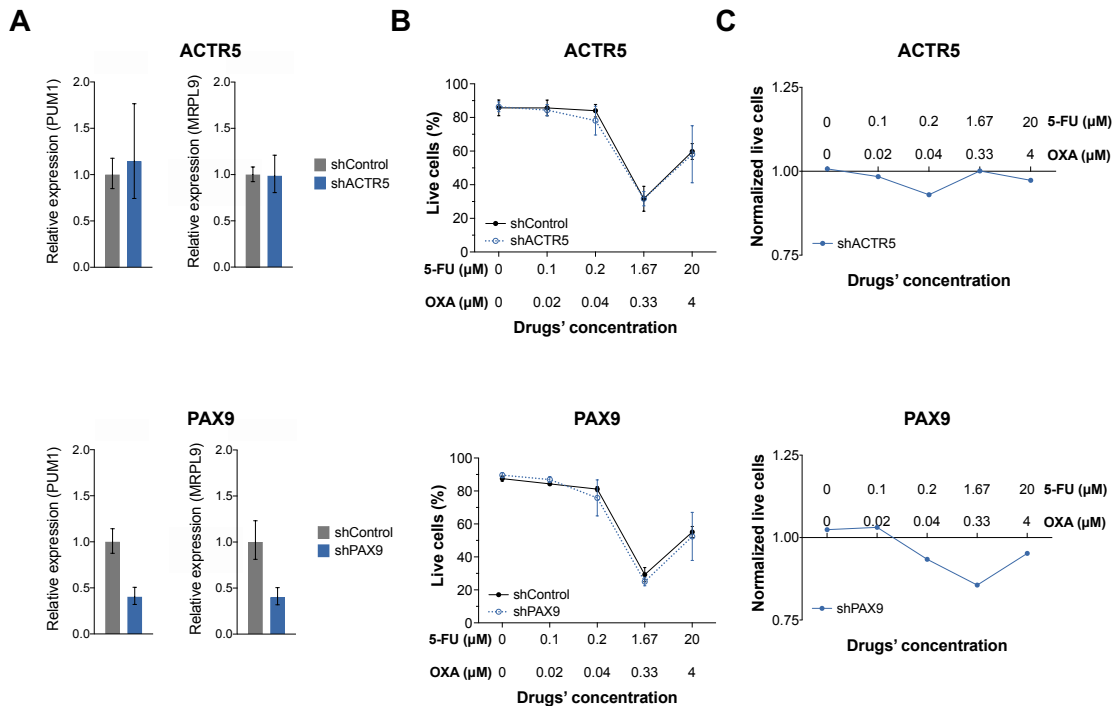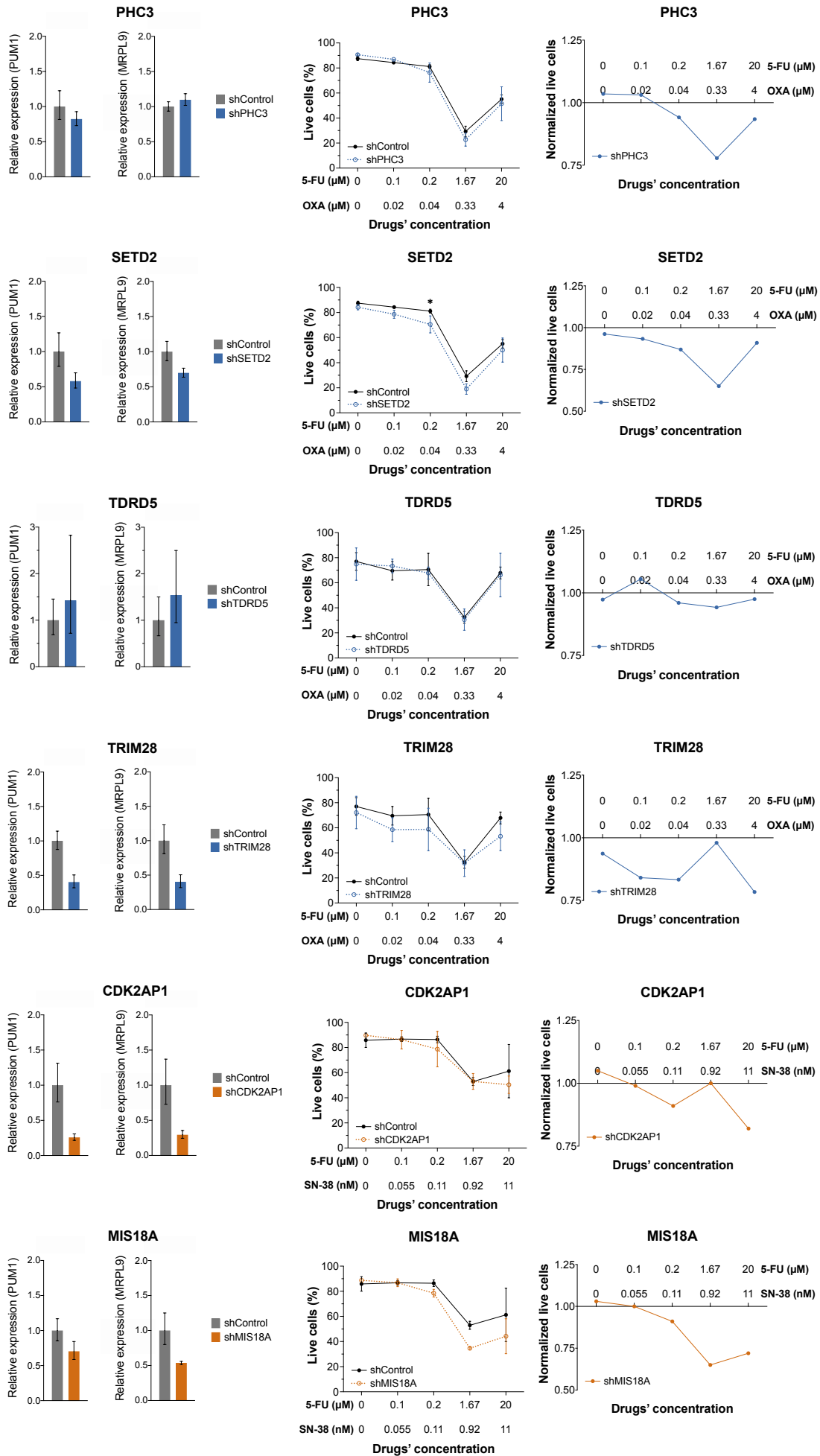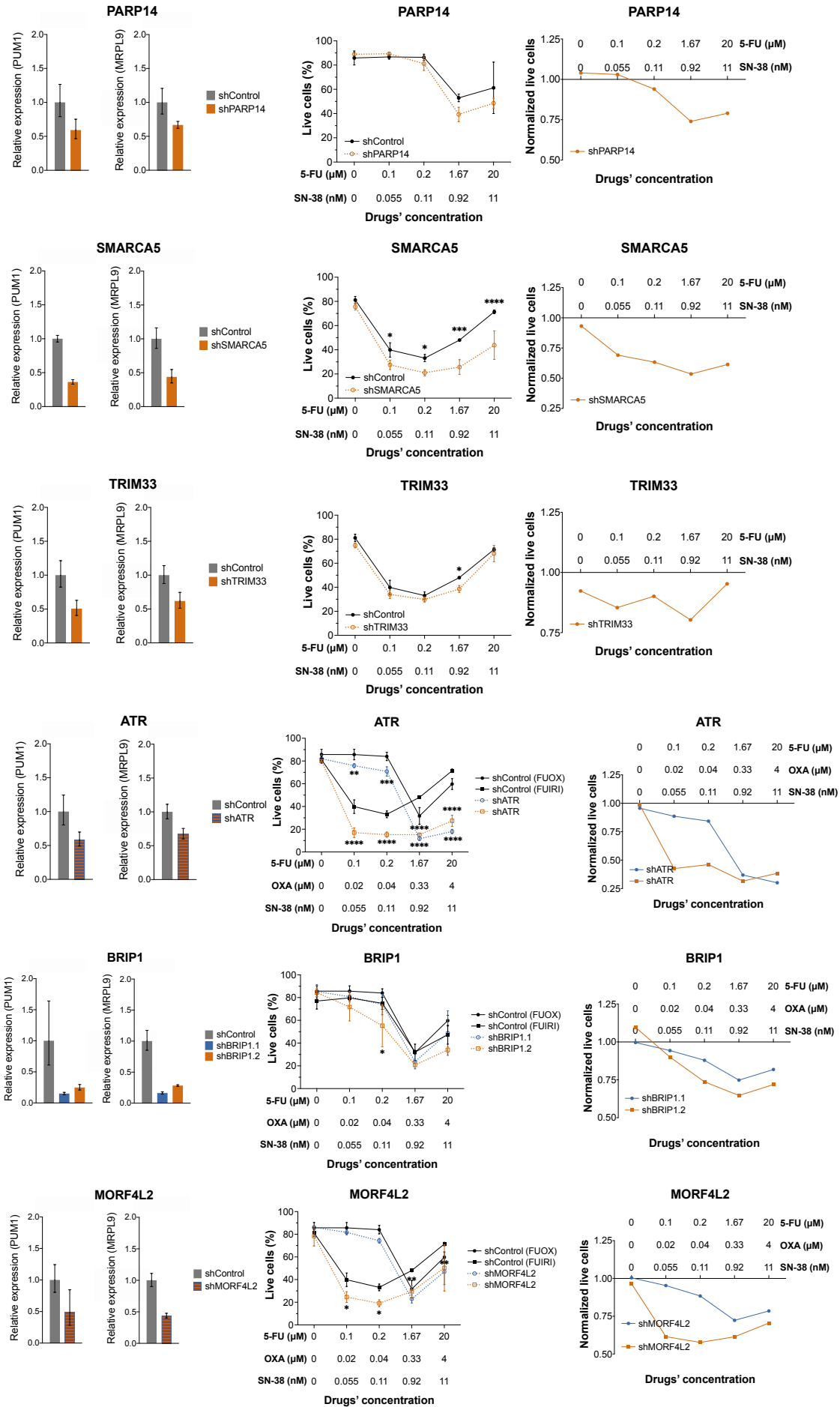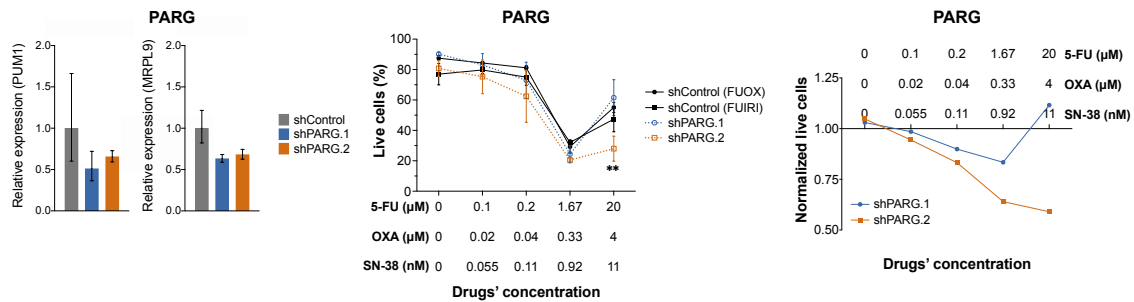
Figure 18. Individual validation of top candidate sensitizer genes. (A) Relative expression measured by RT-qPCR in HT-29 EcoR cells infected with a shControl vs. shRNA of our candidate genes. Two normalizer genes were used: *PUM1* and *MRPL9*. In blue, shRNAs most efficient for FUOX treatment; in orange, shRNAs most efficient for FUIRI treatment. In both colors, shRNAs that were most efficient for both treatments. **(B)** Percentage of live cells in HT-29 EcoR cells infected with a shControl vs. shRNA of our candidate gens after FUOX or FUIRI treatment at different doses measured by flow cytometry in three biological replicates. Two-way ANOVA statistic test with multiple comparisons was applied to all of the genes; differences on cell viability within the same treatment condition that were statistically significant are represented as * (p-value < 0.05), ** (p-value < 0.005), *** (p-value < 0.0005) or **** (p-value < 0.0001). **(C)** Normalized graphs of section's B data after dividing live cells' mean of shGene between live cells' mean of shControl cells. For (B) and (C), in blue are cells treated with FUOX and in orange cells treated with FUIRI.

Firstly, in FUOX-related genes, neither *ACTR5* or *TDRD5* presented any downregulation at the RNA level, accompanied with an irrelevant mortality tendency on cell viability assays. *PAX9* presented a downregulation of around 60% at RNA level, and cell viability experiments showed a slightly increase on cell mortality (around 15%) in comparison to shControl cells. Similarly, *TRIM28* was downregulated correctly by the shRNA; however, cell viability experiments presented an unclear tendency of cell mortality. Surprisingly, viability assays of *PHC3* showed the expected mortality tendency, around 25% more cells dying than shControl cells, despite there was no downregulation at RNA level. Finally, despite expression levels showed only a downregulation of around 30% for *SETD2*, it resulted in 30% more cell death when compared to shControl cells in cell viability assays, especially at the highest doses of FUOX.

In the case of genes assessed for FUIRI treatment, *CDK2AP1* showed a downregulation of RNA expression of over 70% although had an unclear mortality tendency. *MIS18A* and *PARP14* graphs presented a decrease on RNA expression of 50%, resulting in the expected mortality tendencies in cell viability assays, where at highest doses of FUIRI there was an increase of cell death of 30% and 25%, respectively, in comparison to shControl cells. *SMARCA5* showed a reduction on expression of 60%, which was translated in 40-50% more cell death than control cells in viability assays at all FUIRI concentrations. Lastly, *TRIM33* downregulation of 50% resulted in an increase of 25% in cell death when treated at 1.67 µM 5-FU + 0.92 nM SN-38.

Regarding the genes that appeared in the screening for both treatments, *ATR* downregulation reduced its expression around 40%, which produced a strong increase on cell mortality (around 70%) when cells were treated with FUOX or FUIRI in cell viability

assays; it should be remarked that mortality induced by FUIRI treatment was maintained over 50% throughout all the different concentrations. Similarly, downregulation of 50% of *MORF4L2* expression resulted in 35% less viability at the highest doses of chemotherapy; interestingly, cell viability was decreased and maintained around 35% more than control cells throughout all FUIRI treatments. In the case of *BRIP1*, which had a different top targeting shRNA when cells were treated with FUOX or FUIRI, both of them presented a reduction on RNA levels of 75%; moreover, cell viability assays demonstrated the expected mortality tendencies since shBRIP1 cells presented around 50% more cell mortality in comparison to shControl cells for both treatment conditions, especially at highest chemotherapy concentrations. Finally, around 25% downregulation of *PARG* was achieved with two different shRNAs (best targeting shRNA was different for FUOX and FUIRI treatments); regarding cell viability assays, when FUIRI was administered there was almost 50% less cell viability in comparison to control cells whereas FUOX showed a decrease in cell viability of only 20%.

Taken all together, we considered as validated *PAX9* and *SETD2* genes for FUOX treatment and *MIS18A*, *PARP14*, *SMARCA5* and *TRIM33* for FUIRI chemotherapy; in the case of common genes for both treatments (*ATR*, *BRIP1*, *PARG* and *MORF4L2*), all of them were successfully validated. For these genes we obtained at least 50% of downregulation in gene expression and, in viability assays, knock-down cells had lower viability than control cells; in other words, cells infected with shRNAs against these potential candidate genes died more than shControl cells in the presence of FUOX or FUIRI (Figure 18C). Cells infected with shRNAs against genes such as *ATR*, *BRIP1* or *SMARCA5* behaved similar to control cells at dose 0 of chemotherapy; under chemotherapy pressure there was a huge decrease on cell viability, especially at the highest doses.

On the other hand, *TRIM28* and *CDK2AP1* were discarded since cell viability experiments presented unclear tendencies, although downregulation at RNA level was confirmed. *PHC3*, despite there was no downregulation at RNA level, showed an effect on cells' viability when treated with FUOX, which indicates an off-target effect; however, it cannot be directly discarded since the gene was not downregulated, thus, a second shRNA was tested (Figure 20). In the case of *ACTR5* and *TDRD5* genes, they were also not discarded since their potential value could not be assessed due to a poor downregulation effect on the RT-qPCRs; instead, a second shRNA was also tested.

In the same way, Figure 19 presents all the validation results for resistant genes. Figure 19A shows RT-qPCR results; Figures 19B and 19C present the results of the viability cells measured by flow cytometry.
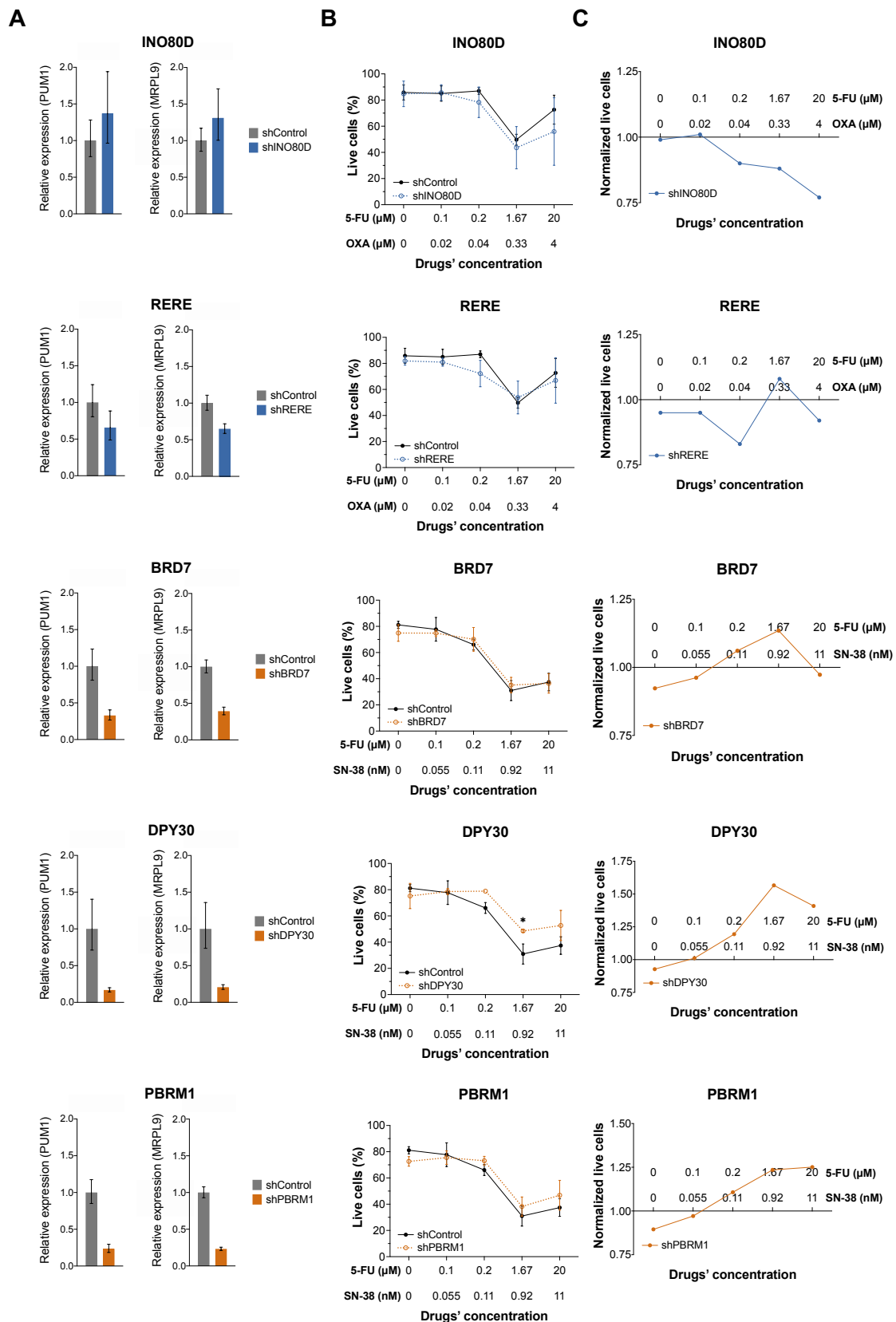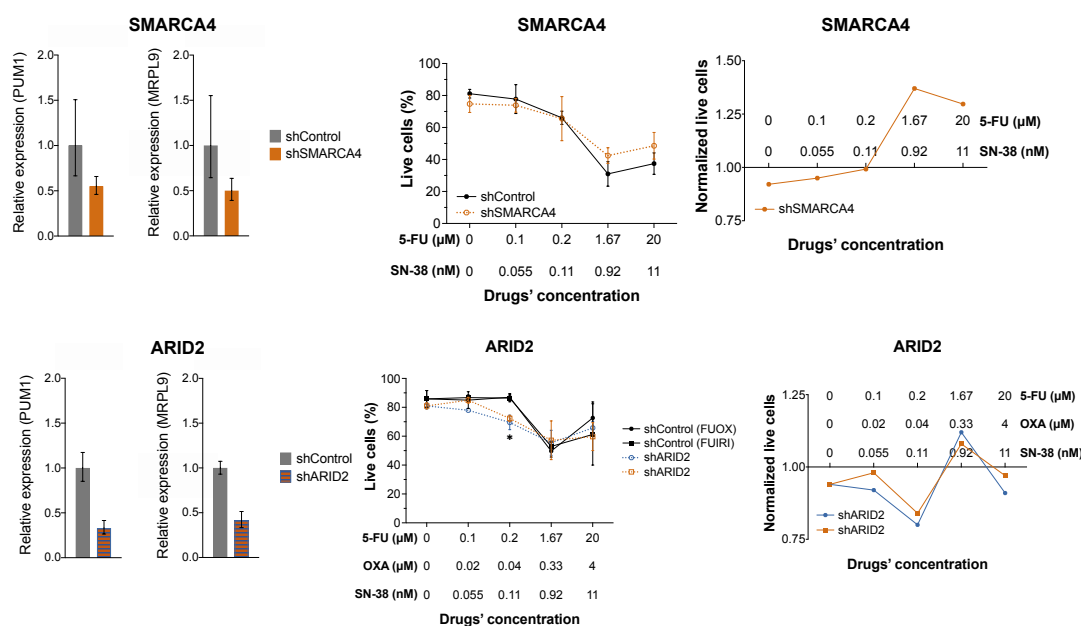
Figure 19. Individual validation of top resistant candidate genes. (A) Relative expression measured by RT-qPCR in HT-29 EcoR cells infected with a shControl vs. shRNA of our candidate genes. Two normalizer genes were used: *PUM1* and *MRPL9*. In blue, shRNAs most efficient for FUOX treatment; in orange, shRNAs most efficient for FUIRI treatment. In both colors, shRNAs that were most efficient for both treatments. **(B)** Percentage of live cells in HT-29 EcoR cells infected with a shControl vs. shRNA of our candidate gens after FUOX or FUIRI treatment at different doses measured by flow cytometry in three biological replicates. Two-way ANOVA statistic test with multiple comparisons was applied and differences on cell viability within the same treatment condition that were statistically significant are represented as * (p-value < 0.05). **(C)** Normalized graphs of section's B data after dividing live cells' mean of shGene between live cells' mean of shControl cells. For (B) and (C) in blue are cells treated with FUOX and in orange cells treated with FUIRI.

Regarding resistant genes assessed for FUOX treatment, it can be observed that *INO80D* was not downregulated by the shRNA although cell viability assays presented more mortality in some FUOX doses in comparison to control cells, which is the opposite expected tendency. In the case of *RERE*, despite having a reduction on RNA expression of 30%, cell viability assays did not present an increase on cell survival at any FUOX concentration.

For resistant genes that appeared with FUIRI treatment, *BRD7* was well-downregulated with 60% reduction at RNA level, which resulted in a slightly increase of 15% in cell viavility in comparison to shControl cells at one of the highest FUIRI concentrations. Concerning *DPY30*, a strong downregulation was achieved, reducing the expression over 80%, which indeed was accompanied by 50% more cell survival of shDPY30 than shControl cells at highest doses of FUIRI. Similarly, *PBRM1* was downregulated around 80% at RNA level, although cell survival increased 25% in comparison to control cells in cell viability assays. Moreover, a decrease on expression of 50% of *SMARCA4* resulted in 35% higher cell viability in comparison to shControl cells at the highest FUIRI doses. *ARID2*, which was a resistant candidate gene common for FUOX and FUIRI treatments, was correctly downregulated by shRNA, obtaining a reduction on expression of 60%; however, cell viability assays did not show a clear increase on survival tendency,

although an slight increase of 10% in the combination of 1.67 μM 5-FU + 0.33 μM OXA for FUOX and 1.67 μM 5-FU + 0.92 nM SN-38 for FUIRI was observed.

In this case, we considered as validated those genes where the knock-down made cells more resistant to chemotherapy, thus, cells grew more than shControl condition, together with a downregulation of at least 50% at RNA level. *BRD7*, *DPY30*, *PBRM1* and *SMARCA4* genes, which all of them were tested for FUIRI chemotherapy, were correctly validated; as we previously observed, in genes such as *DPY30* or *SMARCA4* there was a relevant increase in cell survival of more than 50% respect to control cells. On the other hand, a good downregulation of *ARID2* did not show a potent effect on increasing cell viability, thus it was discarded as a good candidate. As *RERE* had not a potent downregulation at RNA level yet seems to present a correct viability tendency, it was decided to test a second shRNA. In the case of *INO80D*, as none of the validation assays worked as expected, a second shRNA was also tested.

Figure 20 presents the results of HT-29 EcoR cells infected with a second shRNA against genes that were not validated with the 1st shRNA, such as *ACTR5*, *PHC3*, *TDRD5*, *INO80D*, *RERE* and *MIS18A* genes; in the case of *MIS18A*, although it was correctly validated with the first shRNA, we decided to try a second shRNA to improve the downregulation at RNA level as we were especially interested on this gene for further analysis. As mentioned above, Figure 20A shows RT-qPCRs and Figures 20B and 20C present the results of the viability cells measured by flow cytometry.
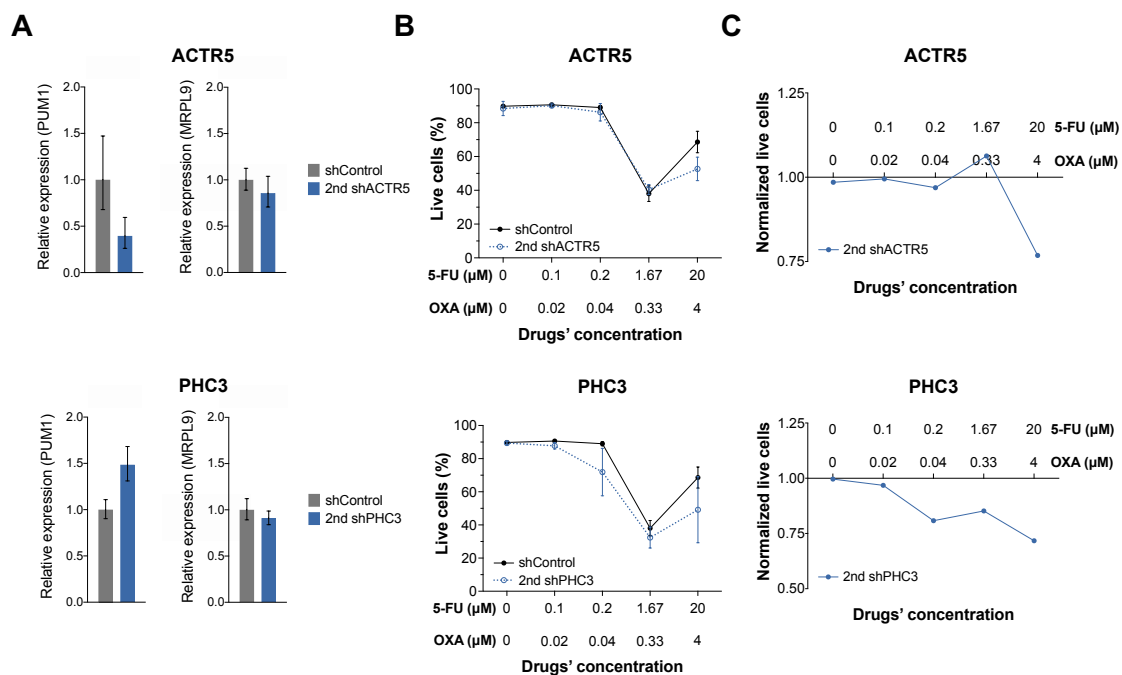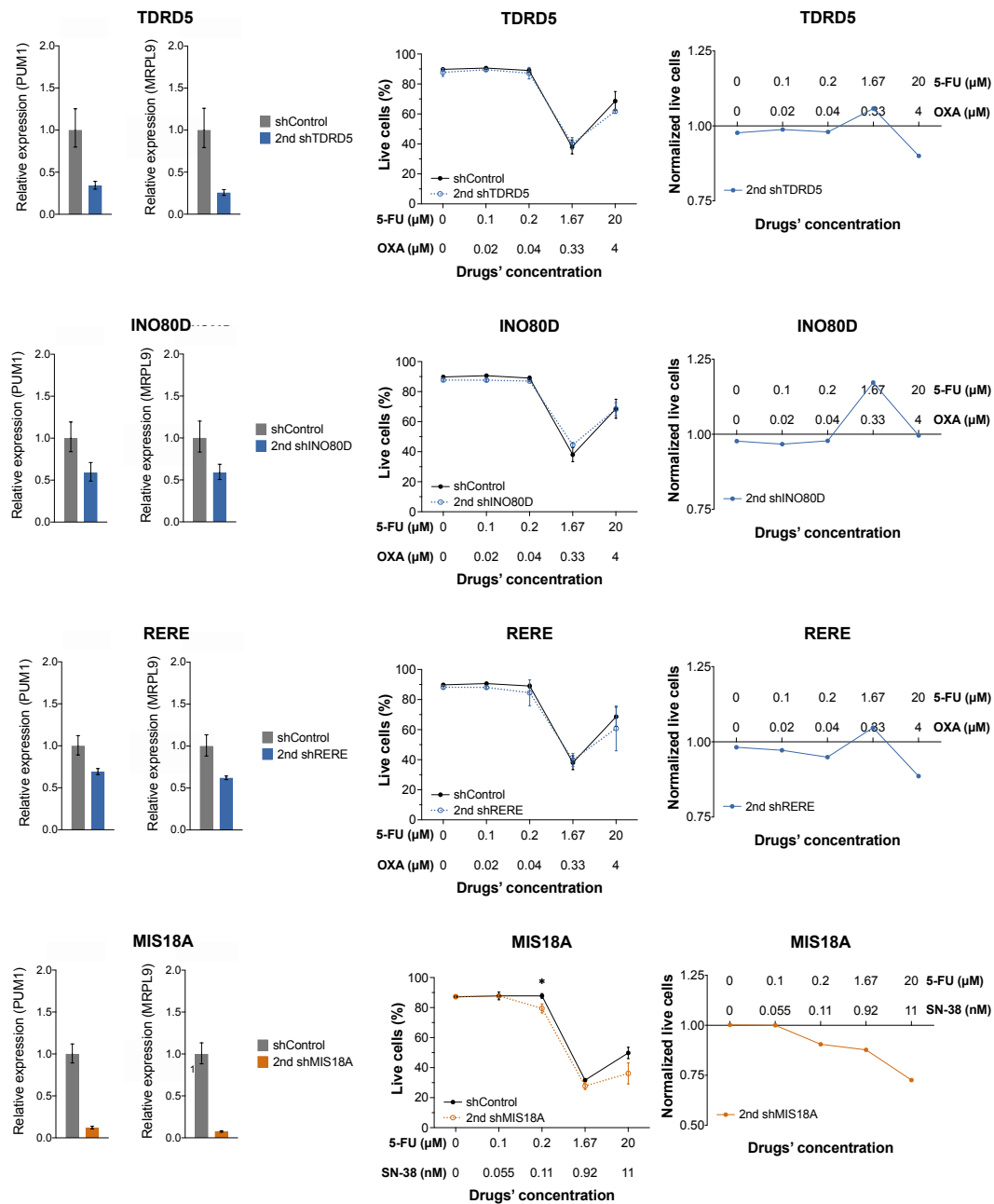
Figure 20. Individual validation of top candidate genes infected with a second shRNA. In blue, genes that arisen when cells were treated with FUOX; in orange, genes that appeared when cells were treated with FUIRI. **(A)** Relative expression measured by RT-qPCR in HT-29 EcoR cells infected with a shControl vs. shRNA of our candidate genes. Two normalizers genes were used: *PUM1* and *MRPL9*. **(B)** Percentage of live cells in HT-29 EcoR cells infected with a shControl vs. shRNA of our candidate gens after FUOX or FUIRI treatment at different doses measured by flow cytometry in three biological replicates. Two-way ANOVA statistic test with multiple comparisons was applied and differences on cell viability within the same treatment condition that were statistically significant are represented as * (p-value < 0.05). **(C)** Normalized graphs of section's B data after dividing live cells' mean of shGene between live cells' mean of shControl cells.

Regarding sensitivity genes, *ACTR5* did not present a clear downregulation and viability results were not consistent, thus, it was discarded; *TDRD5* was correctly downregulated at RNA level (60%) yet viability experiments failed and, therefore, it was discarded as well. The second shRNA against *PHC3* produced a similar effect than the first one: there was not a good downregulation at expression level despite a good mortality tendency can be observed in the cell viability assays with 25% more cell death than control cells;

for this reason, this gene was discarded and considered as a false positive whose shRNA provided off-target effects.

In the case of *MIS18A*, with the second shRNA we obtained a better downregulation, which was over 80%, despite cell viability tendency was not as good as with the first shRNA; it decreased 20% in comparison to control cells, whereas 30% was achieved with the first shRNA; however, as long as we want to assure the maximum downregulation of the gene, we decided to perform future experiments in cells infected with the second shRNA.

Furthermore, regarding resistance genes such as *INO80D* and *RERE*, we correctly validated *INO80D* with a second shRNA as downregulation was under 50% of expression and cells presented around 25% highest survival than shControl in the presence of FUOX. Since downregulation at RNA level of *RERE* was not optimal (30%) and viability assays presented an unclear survival tendency, this gene was discarded.

After performing all the battery of experiments, validation results can be summarized in Table 3. On the top part, the validation results can be observed before trying a second shRNA in all the scenarios; bottom part of the figure shows the results of this validation step. At the end, we validated 15 out of 22 candidate genes coming from the screening as "sensitizer" or "resistant" genes. At the beginning, only 3 genes were directly discarded as false positives arisen from the background of the screenings; but finally 5 genes were discarded after the validation with the second shRNA. Nevertheless, we were able to rescue one of the 5 possible genes with a second shRNA. Altogether, we validated almost 70% of all candidate genes that were selected from the screening.

		SCENARIO 1	SCENARIO 2	SCENARIO 3	SCENARIO 4
BEFORE	NUMBER OF GENES	14 out of 22	3 out of 22	2 out of 22	3 out of 22
BEFORE	PERCENTAGE	63%	14%	9%	14%
AFTER	NUMBER OF GENES	15 out of 22	5 out of 22	1 out of 22	1 out of 22
AFTER	PERCENTAGE	68%	23%	4.5%	4.5%

Table 3. Summary of individual validation results. Top part represents the results before trying a second shRNA in scenarios 3 and 4; bottom part shows the results after the validation also with the second shRNAs experiments.

3.2. Selected validated genes were monitored for protein downregulation

Individual validation of all candidate genes lead us to 15 genes with great potential involvement in the resistance to chemotherapy in CRC. Nevertheless, it was unaffordable to study all of them in detail. After revising available information of each gene in the bibliography and public databases, we decided to focus all our efforts in four

candidate genes: *BRIP1* and *MIS18A,* as sensitizers to chemotherapy and suitable drug targets for combinatorial therapies, and *PBRM1* and *SMARCA4*, which conferred resistance and would present opposite behaviors to sensitizers.

For this reason, we firstly performed RT-qPCR in three biological replicates to confirm the knock-down efficiency in the three replicates used for the viability assays; moreover, WB was also conducted to confirm decreased protein levels in the three biological replicates. In the left, Figure 21A presents the RT-qPCR results with two normalizer genes (*PUM1* and *MRPL9*); it should be remarked the robust downregulation achieved for all four genes in all the biological replicates. WB results are shown in Figure 21B, where it can be appreciated a good downregulation for *BRIP1*, *MIS18A*, *PBRM1* and *SMARCA4* proteins.
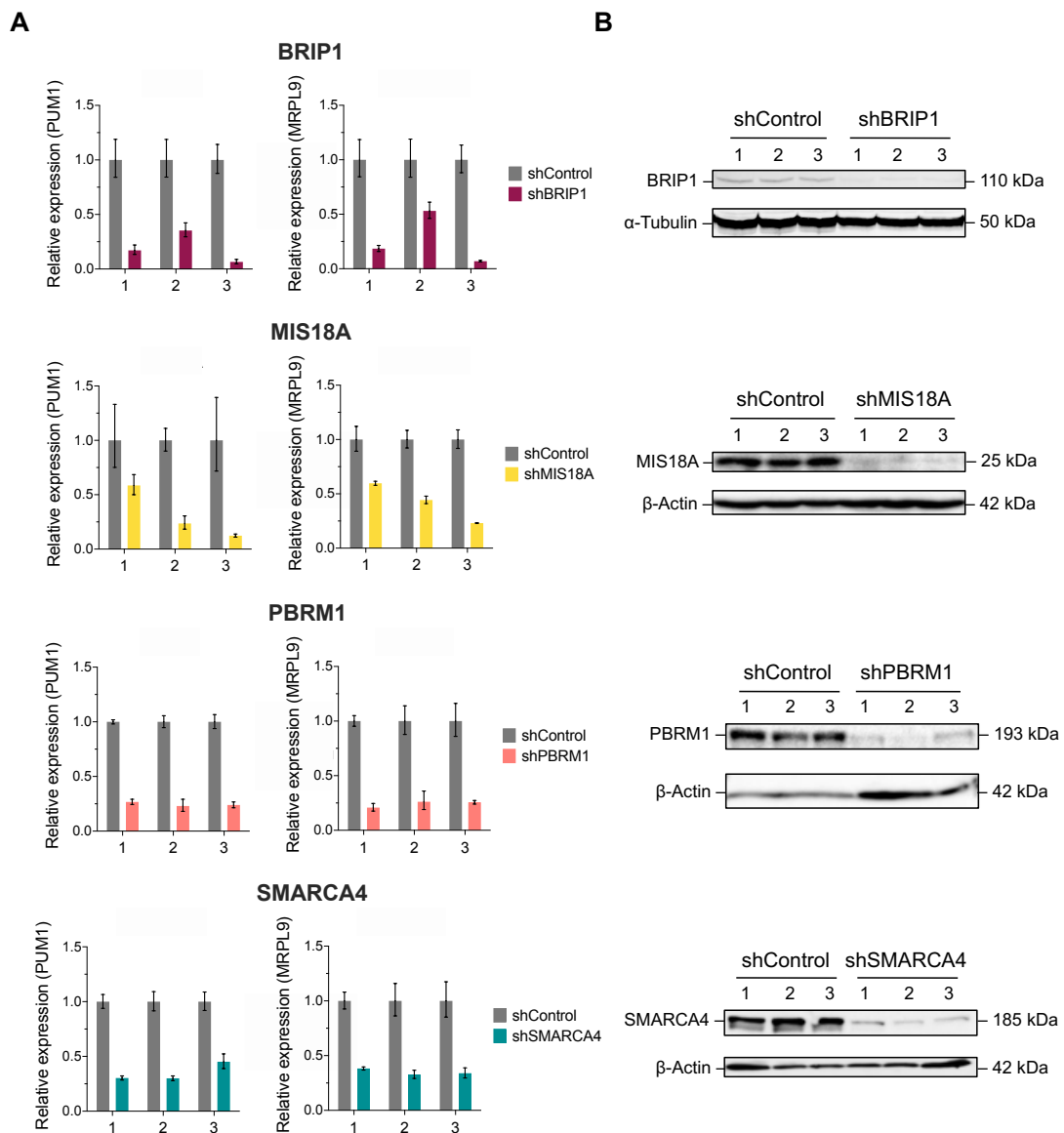
Figure 21. Expanded validation in three biological replicates of *BRIP1, MIS18A, PBRM1* and *SMARCA4*. (A) Relative expression measured by RT-qPCR in HT-29 EcoR cells infected with a shControl vs. shBRIP1, shMIS18A, shPBRM1 or shSMARCA4. **(B)** Protein levels of HT-29 EcoR cells infected with shControl vs. shBRIP1, shMIS18A, shPBRM1 or shSMARCA4 monitored by Western Blot. β-Actin levels were sued as a normalizer.

4. Generation of knock-out cell lines for the final candidate genes of interest by CRISPR-Cas9 technology

To test the clinical relevance of the selected sensitizer and resistant genes, usually researchers look for drugs that specifically inhibit their targets and analyze potential beneficial therapies in combination with standard CT *in vitro* and in mice with transplanted tumors. However, our selection included genes without commercial drugs available, neither they present enzymatic domains that can be targeted straightforward. Moreover, downregulation by shRNA results in fair decreased protein levels yet they are not 100% efficient. For these reasons, we decided to generate knock-out (KO) HT-29 cell lines for *BRIP1*, *MIS18A*, *PBRM1* and *SMARCA4* genes by CRISPR-Cas9 technology. In this manner, we would have a robust absence of these genes, and as such, a crucial tool to study the mechanisms of action where these genes affect cell survival upon CT treatments.

To perform this objective, three different guide RNAs (gRNAs) were selected per gene (Annex III), each one of them targeting different exons. HT-29 CRC cell line was infected with a pool of the three gRNAs and single-cell clones were isolated to obtain homogeneous populations; after clonal expansion, 48 clones per gene were analysed to verify if our genes of interest were correctly edited. In addition, HT-29 cell line was also infected with a non-target gRNA that will serve as the control condition. Figure 22 shows the presence or absence at protein level of *BRIP1*, *MIS18A* and *SMARCA4* in all clones in comparison to the parental HT-29 cell line monitored by WB, which is a rapid technique to obtain a first read-out of all clones; also, two clones of the control gRNA were included. In the case of *PBRM1*, several problems appeared related to the antibody, therefore, after several methodological changes without success, it was decided to postpone its study and continue with the other targets for phenotypic and mechanistic experiments.

In general, it should be remarked that protein levels of *BRIP1*, *MIS18A* and *SMARCA4* were similar between the parental HT-29 cell line and the two clones infected with the control non-target gRNA (C1 and C2), pointing out that all the infection and selection processes suffered by the clones did not affect our preferred genes. In the case of *BRIP1* clones, absolute disappearance of the protein was less clear, although some clones presented a good downregulation. For this reason, we decided to analyse the DNA sequence of clones 11, 12, 27 and 43, which presented a range of almost no BRIP1 expression to approximately 50% of expression in comparison to control gRNAs' clones (C1 and C2), which indeed presented similar expression levels than parental HT-29 cell

Figure 22. Protein levels of *BRIP1*, *MIS18A* and *SMARCA4* KO clones. Protein expression of parental HT-29 cell line, two control gRNA clones (C1 and C2) and all expanded KO clones of *BRIP1*, *MIS18A* and *SMARCA4* genes monitored by Western Blot. β-Actin levels were sued as a normalizer.
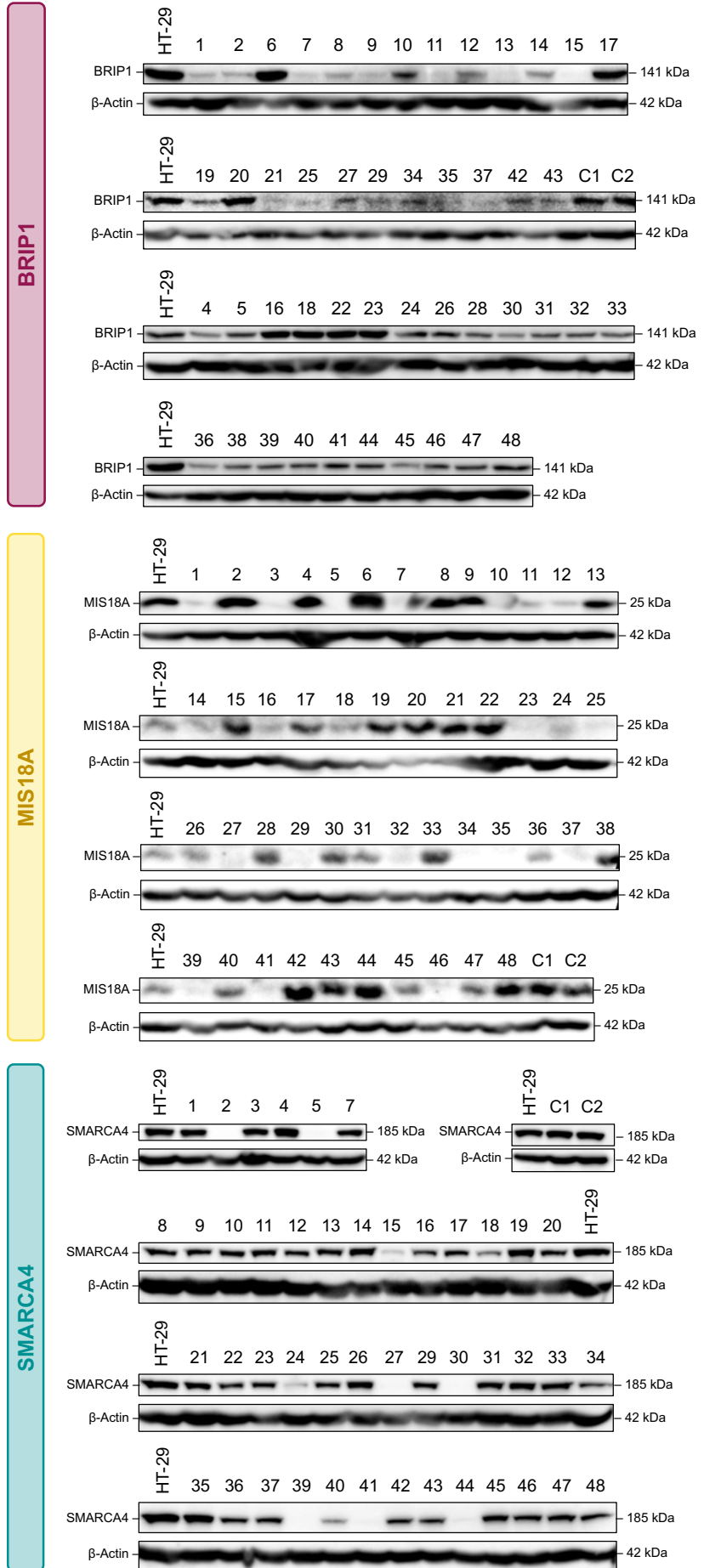
line. In *MIS18A* KO clones, WB showed 14 clones with no expression of this protein and 12 clones with low levels, thus, we decided to sequence and validate further clones 32, 34 and 35, which showed undetectable MIS18A protein levels. For *SMARCA4* gene, the difference between clones was clear as well by WB as some clones presented a complete loss of SMARCA4; here, we decided to sequence clones 2, 5 and 30.

To confirm the genome editing of our target genes, DNA from the best downregulated clones indicated above was sequenced by Sanger. Figure 23 presents an alignment of sequences from parental HT-29 cell line, two control gRNAs clones (C1 and C2), and several clones of *BRIP1* (11, 12, 27 and 43) and *MIS18A* (32, 34 and 35) in the different areas were the gRNAs should act (marked in red squares); in addition, base pairs that differ from the parental HT-29 sequence are marked in yellow. Overall, for all sequences it was confirmed that there were no changes in cells infected with the control gRNA respect to the parental HT-29 cell line since C1 and C2 sequences were always identical to HT-29.

In the case of *BRIP1* clones, it can be observed how, with the gRNA1, clones 11 and 27 were edited: both of them presented a deletion in the area of the targeted sequence, which consequently creates a change on the reading frame of all the sequence that continues after the edited region. Similarly, clones 12, 27 and 43 were edited in the region recognized by gRNA4: although clone 12 had only a punctual insertion of a "T", clones 27 and 43 showed several insertions and deletions in the squared area, which affects all the remaining sequence as well, changing again the reading frame.

However, is also clear that gRNA3 had no effect in any of the clones; this was the same case than gRNA4 of *MIS18A* clones, where there was no change on any of the sequences. Regarding *MIS18A* clones 34 and 35, it can be appreciated a punctual "T" insertion in clone 34 in the target region of gRNA3, whereas clone 35 suffers some deletions in this region that probably affect the reading frame of all the remaining sequence; however, clone 32 did not show a clear edition, thus it was not included in downstream experiments. Regarding the sequences of the region that was targeted by gRNA1, there are not present due to problems on amplifying and sequencing this region that are pending to be solved.

Unfortunately, several problems in amplifying and sequencing the target regions of *SMARCA4* gRNAs need still to be solved as well; for this reason, we cannot confirm the editing of these clones at DNA level. However, since the absence of SMARCA4 protein was extremely clear in some clones (Figure 22), we decided to further analyse clones 2, 5 and 30 in functional assays.
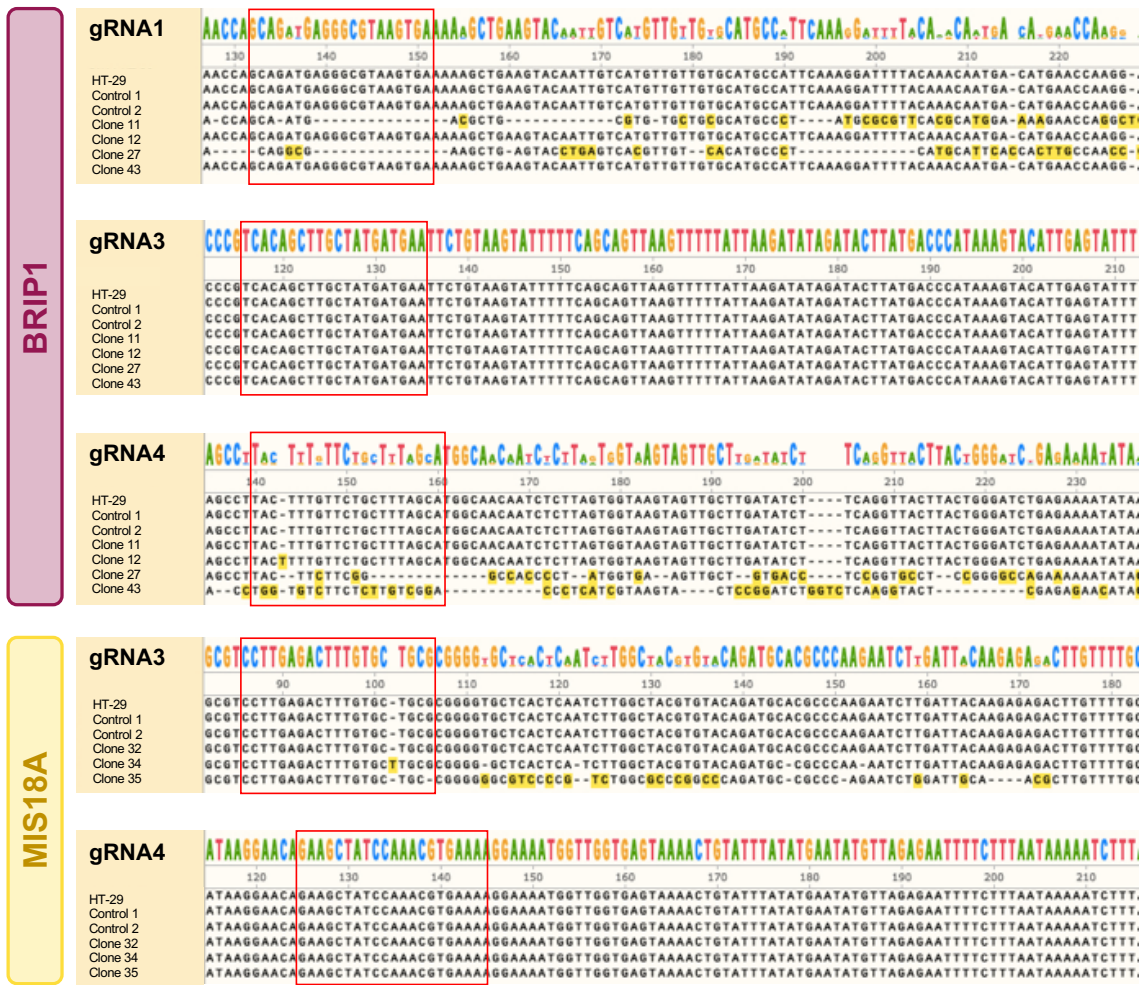
Figure 23. DNA sequences of *BRIP1* and *MIS18A* clones. Sequences' alignment from parental HT-29 cell line, two control gRNA clones (C1 and C2) and several *BRIP1* (11, 12, 27 and 43) and *MIS18A* (32, 34 and 35) clones; sequences were obtained by Sanger sequencing. In red squares are represented the areas were the different gRNAs should act; in yellow, base pairs that differ from the sequence of parental HT-29 cell line are marked.

5. Characterization of several features in *BRIP1*, *MIS18A* and *SMARCA4* knock-out clones

5.1. Phenotype of knock-out cell clones

Figure 24 presents bright-field microscope images (20X) of the different KO clones in culture generated for *BRIP1*, *MIS18A* and *SMARCA4* genes together with the control non-target condition and parental HT-29 cell line. It can be appreciated that *BRIP1* KO cells presented a smaller and rounded shape in comparison to control cells, which are bigger and present some lamellipodium. Moreover, the shape of *BRIP1* KO clones remind to parental HT-29 cell line. However, the shape of *MIS18A* and *SMARCA4* clones are maintained similar to control non-target clones, although they seem to present more lamellipodium.
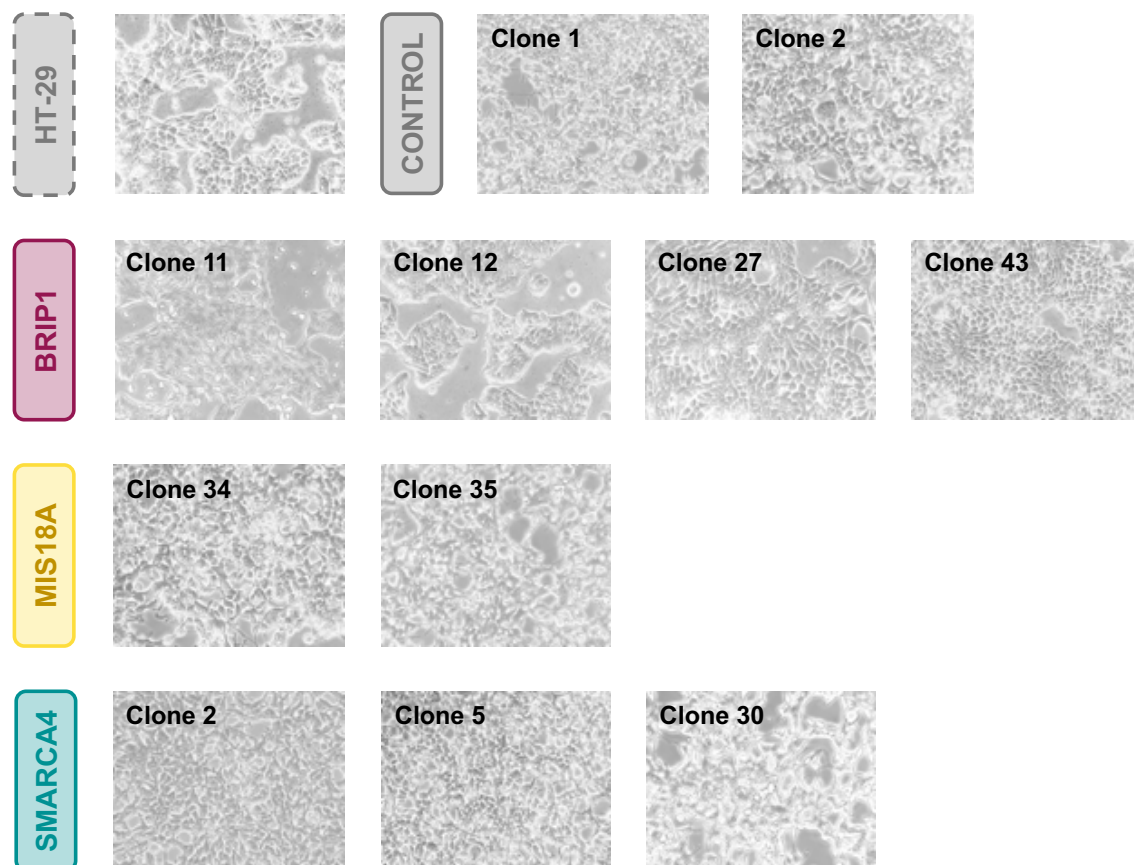
Figure 24. Phenotype of parental HT-29, Control, *BRIP1*, *MIS18A* and *SMARCA4* KO clones. Microscope images (20X) of parental HT-29, Control (1 and 2), *BRIP1* (11, 12, 27 and 43), *MIS18A* (34 and 35) and *SMARCA4* (2, 5 and 30) KO clones.

5.2. Implication in cell viability in combination with FUIRI chemotherapy regimen

To corroborate the role of *BRIP1*, *MIS18A* and *SMARCA4* in sensitizing or increasing cell survival in front of chemotherapeutic agents such as FUIRI, firstly we performed cell viability assays in the same way as in the individual validation phase. Knock-out cells

were treated at different concentrations of FUIRI and live, apoptotic and dead cells were monitored by flow cytometry. FUIRI treatments were chosen to be analyzed first because our cohort of CRC primary tumors belongs to patients treated with FOLFIRI, and as such, we thought it would be more reasonable to better characterize these gene's effects upon this treatment only, at least during a first approach.

Figure 25 presents a comparison of parental HT-29 and cells infected with the control non-target gRNA to assure a similar behavior between these two conditions; in this regard, cell viability was comparable (increased cell death was not significant for the highest CT concentration) and, therefore, control non-target KO cells will be used as the control condition from now on.
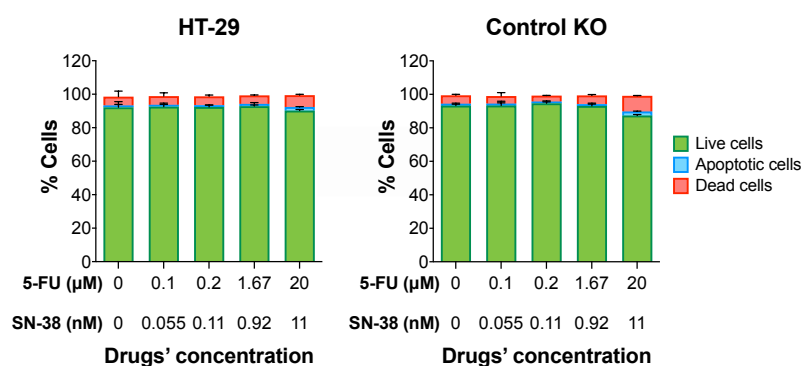
Figure 25. Cell viability assays of parental HT-29 and HT-29 cells infected with a control non-target gRNA. Live, apoptotic and dead cells monitored by flow cytometry at different doses of FUIRI indicates a similar behavior in both conditions.

Regarding *BRIP1* KO clones (11, 12, 27 and 43), Figure 26 presents their results on cell viability assays with two technical replicates. Figure 26A shows the percentage of live (in green), apoptotic (in blue) and dead (in red) cells monitored by flow cytometry at different doses of FUIRI treatment. It can be observed in all four clones a strong decrease on cell viability, especially at higher doses of FUIRI. Figure 26B gives a deeper focus on live cells' behavior: these graphs present live cells' tendency in comparison with control KO cells when treated at the same doses of FUIRI. Here, the variations in cell viability when *BRIP1* is not present were evident in a dose-dependent manner; at higher doses of FUIRI, cell viability strongly decreases, arriving to around 20% more dead and 60% more apoptotic cells in comparison to control cells.

To better understand the magnitude of the differences in cell viability, Figure 26C shows normalized graphs of live cells where *BRIP1* KO mean live percentages were divided by control KO mean on each treatment condition. It can be appreciated how, especially at higher doses of FUIRI, cell viability decreases around 75% in comparison to control KO cells in all *BRIP1* KO clones, which is in agreement with obtained results in the validation step.

Figure 26. Cell viability assays of *BRIP1* KO clones 11, 12, 27 and 43. (A) Live (in green), apoptotic (in blue) and dead (in red) cells monitored by flow cytometry at different doses of FUIRI. **(B)** Percentage of live cells in HT-29 cells infected with a control non-target gRNA (Control KO) vs. *BRIP1* KO cells (clones 11, 12, 27 and 43) monitored by flow cytometry at different doses of FUIRI in two technical replicates. Two-way ANOVA statistic test with multiple comparisons was applied to all clones; differences on cell viability within the same treatment condition that were statistically significant are represented as *** (*p*-value < 0.0005) or **** (*p*-value < 0.0001). **(C)** Normalized graphs of section's B data after dividing live cells' mean of *BRIP1* KO clones between live cells' mean of Control KO cells.

Figure 27 presents cell viability results of *MIS18A* clones 34 and 35 in two technical replicates. *MIS18A* KO did not show clear changes on cell viability between control KO cells and clones 34 and 35, where *MIS18A* gene was completely impaired through CRISPR-Cas9 technology. Even though, there were some statistically significant

differences that appeared in clone 35 by applying a two-way ANOVA statistic test with multiple comparisons (represented as * [p-value < 0.05] or ** [p-value < 0.005]).
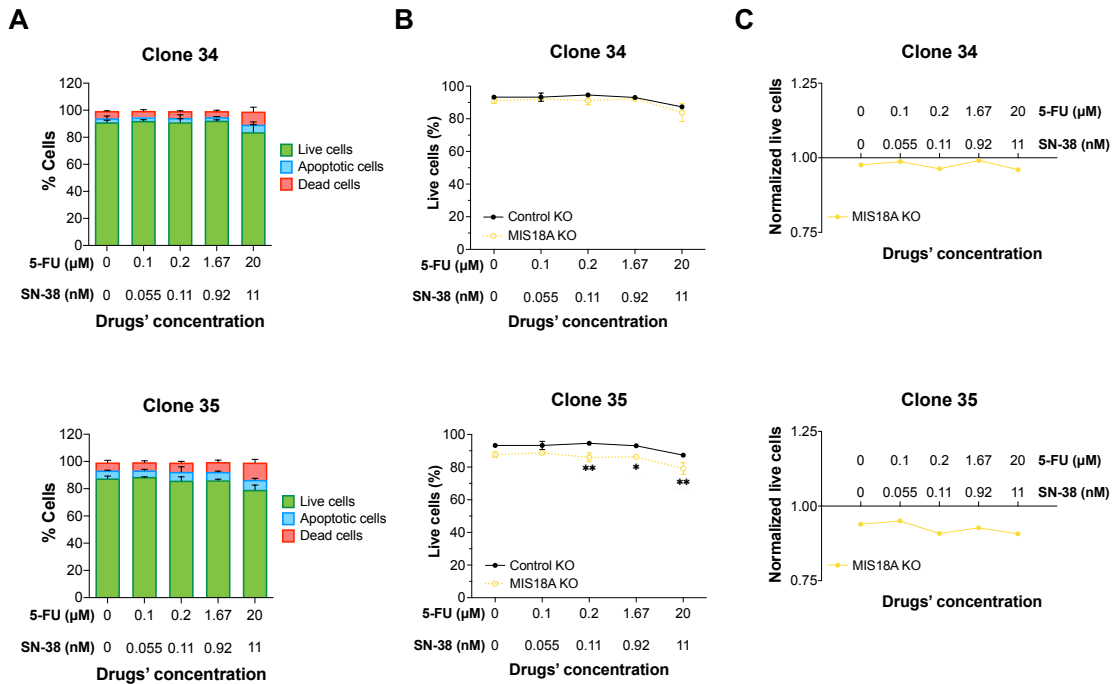
Figure 27. Cell viability assays of *MIS18A* KO clones 34 and 35. (A) Live (in green), apoptotic (in blue) and dead (in red) cells monitored by flow cytometry at different doses of FUIRI. **(B)** Percentage of live cells in HT-29 cells infected with a control non-target gRNA (Control KO) vs. *MIS18A* KO cells (clones 34 and 35) monitored by flow cytometry at different doses of FUIRI in two technical replicates. Two-way ANOVA statistic test with multiple comparisons was applied to all clones; differences on cell viability within the same treatment condition that were statistically significant are represented as * (p-value < 0.05) or ** (p-value < 0.005). **(C)** Normalized graphs of section's B data after dividing live cells' mean of *MIS18A* KO clones between live cells' mean of Control KO cells.

In the case of *SMARCA4*, we were evaluating a possible role on increasing HT-29 cell survival in front of FUIRI, thus, contributing to CRC chemoresistance. As it can be observed on Figure 28, there were no big differences in cell viability comparing *SMARCA4* clones 2, 5 and 30 to control KO cells. Despite we obtained good SMARCA4 knock-out cells, it seems that *SMARCA4* disappearance does not contribute on increasing cell survival *per se* under the pressure of FUIRI chemotherapy at the tested concentrations.
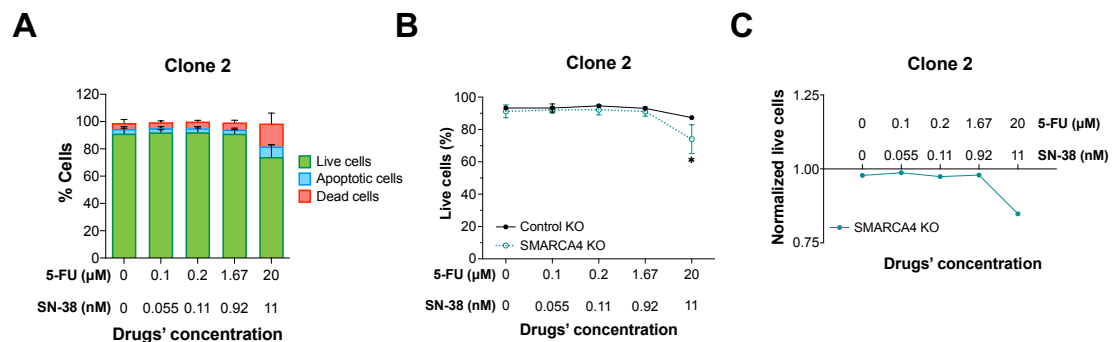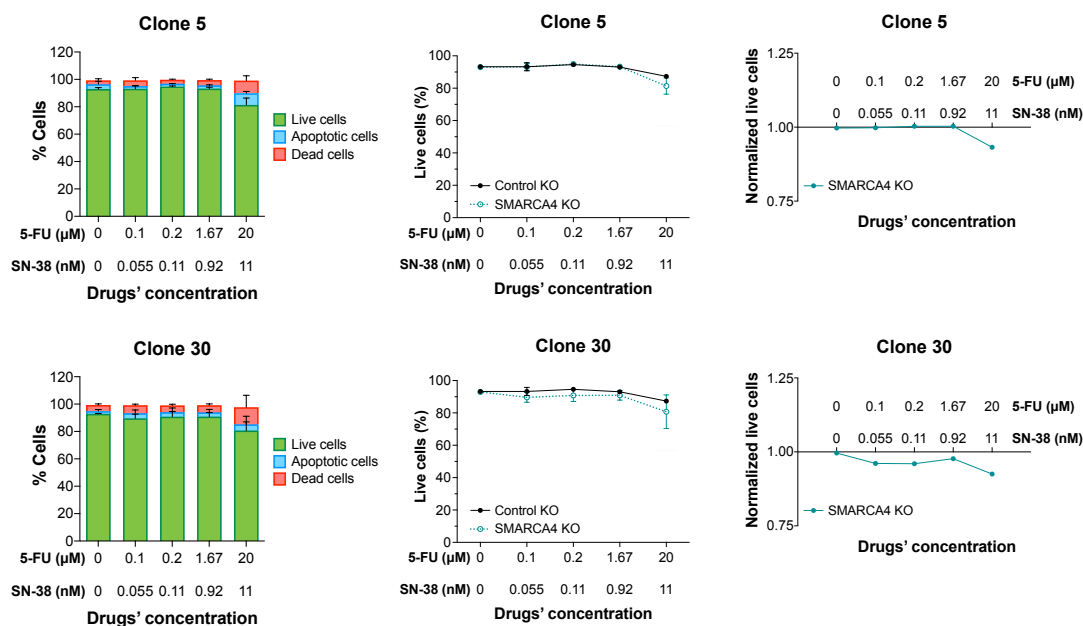
Figure 28. Cell viability assays of *SMARCA4* KO clones 2, 5 and 30. (A) Live (in green), apoptotic (in blue) and dead (in red) cells monitored by flow cytometry at different doses of FUIRI. **(B)** Percentage of live cells in HT-29 cells infected with a control non-target gRNA (Control KO) vs. *SMARCA4* KO cells (clones 2, 5 and 30) monitored by flow cytometry at different doses of FUIRI in two technical replicates. Two-way ANOVA statistic test with multiple comparisons was applied to all clones; differences on cell viability within the same treatment condition that were statistically significant are represented as * (p-value < 0.05). **(C)** Normalized graphs of section's B data after dividing live cells' mean of *SMARCA4* KO clones between live cells' mean of Control KO cells.

5.3. DNA damage quantification in response to FUIRI

In the Introduction of this thesis, the mechanisms by which chemotherapeutic agents, especially FUOX and FUIRI combinations, promote the death of CRC cells were described. The main idea is that these drugs directly impact onto DNA to promote enough DNA damage that leads cells to apoptosis. For this reason, we decided to evaluate and quantify DNA damage promoted on HT-29 KO cells at the FUIRI doses that resulted in viability differences. Figure 29A shows DNA damage monitored by phosphorylated H2A.X (γ-H2A.X) levels, which is a modification that appears when DSBs occur. Figure 29B presents a bar graph of normalized results after band quantification.

It can be appreciated how, in control KO cells, there was not an evident increase on DNA damage, even for the highest concentration of FUIRI, where total H2A.X protein was low; thus, it is difficult to assess whether there was an increase on DNA damage at this sample. However, in the case of *BRIP1* KO clones, we observed two subpopulations: DNA damage was slightly increased on clones 11 and 27 at higher doses of FUIRI, around 1.5 times more than control cells; however, clones 12 and 43 showed a large increase on DNA damage response in a dose-dependent manner, having three times more γ-H2A.X in comparison to control KO cells. Of note, for all clones tested, the highest
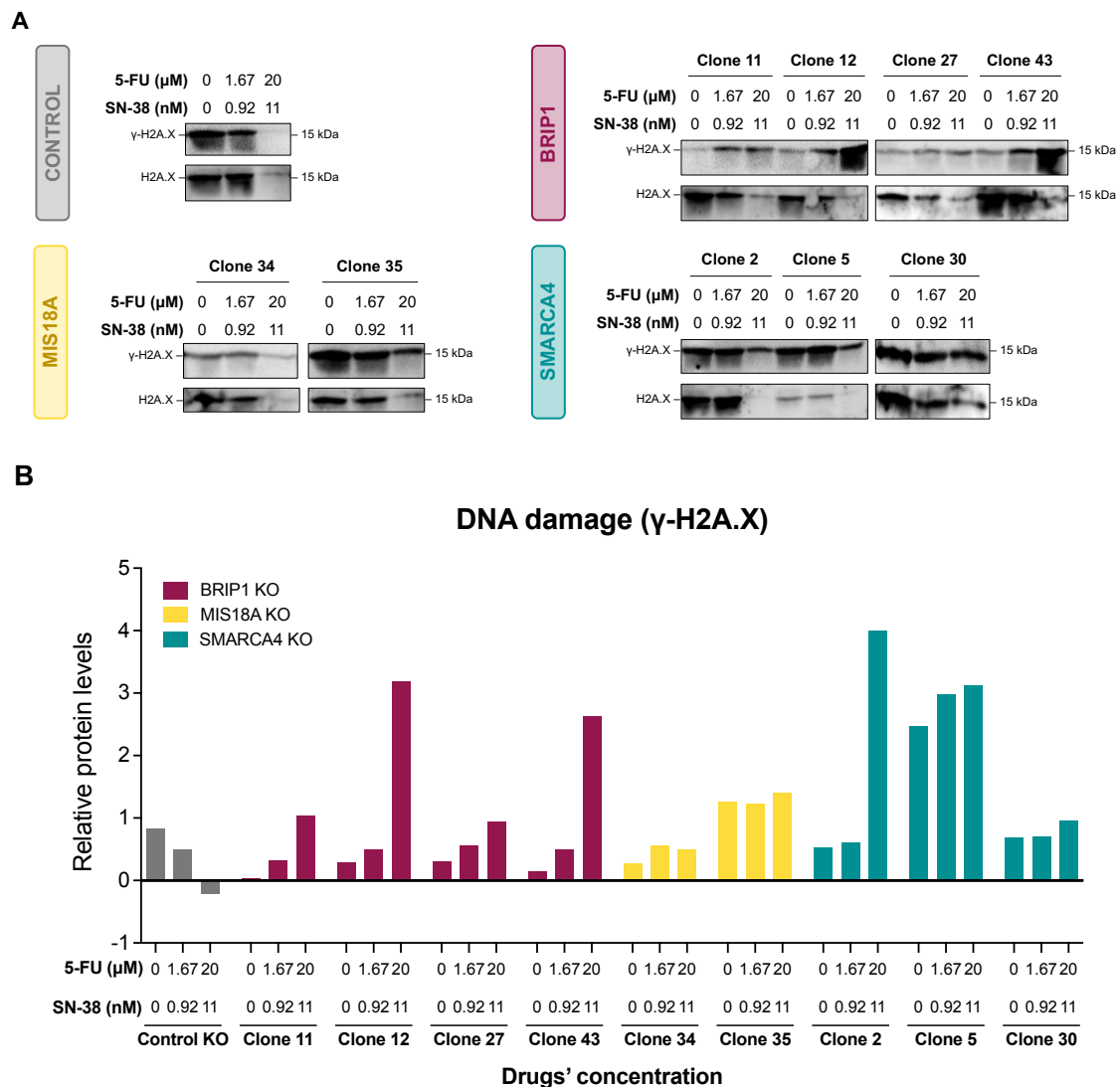
Figure 29. Quantification of DNA damage in Control, *BRIP1*, *MIS18A* and *SMARCA4* KO clones in combination with FUIRI. (A) DNA damage from the clones of Control, *BRIP1* (11, 12, 27 and 43), *MIS18A* (34 and 35) and *SMARCA4* (2, 5 and 30) was measured through H2A.X phosphorylation (γ-H2A.X) by Western Blot after FUIRI treatment at different doses. Total H2A.X was used as a normalizer. **(B)** Relative protein levels were obtained by dividing γ-H2A.X band intensities with their corresponding total H2A.X values, quantified through ImageJ.

concentration of FUIRI treatment showed less total H2A.X. For *SMARCA4* clones, we observed higher levels of DNA damage at higher concentrations of chemotherapy in clone 2, which was four times higher than control KO cells, although again there was less total H2A.X protein than control KO cells as well. Clone 5 showed upregulation of DNA damage in all FUIRI concentrations tested, and did not show a dose-dependent behavior, however one must be cautious interpreting this data because the total protein was really low, which then could result in an over estimation of the bands obtained. Of note, clone 30 did not show a dose-dependent increase of DNA damage. *MIS18A* clones, although seemed to behave differently between them if we look at the normalization, it

seems clear that there was not a dose-dependent increase of DNA damage, and it was not very superior to observed changes in control KO cells.

5.4. Evaluation of the colony formation capacity of knock-out cells in combination with FUIRI

This type of colony formation assay measures the capacity of isolated cells (seed at very low density to obtain single-cell populations) to divide and expand until a colony is formed; in this manner, is obtained an idea of the ability of a cell to initiate growth, which could be associated to metastatic potential.

Figure 30 presents a colony formation assay of control, *BRIP1*, *MIS18A* and *SMARCA4* KO clones in combination with different concentrations of FUIRI. Figure 30A shows the colony images extracted from 6-well cell culture plates and stained with crystal violet; in a visual manner, it can be appreciated huge differences in the size and number of colonies depending on the KO cell type and throughout the different FUIRI treatments. The different parameters extracted from these images by ImageJ are presented on Figures 30B, 30C and 30D. Figure 30B presents the total number of colonies quantified for *BRIP1* (in purple), *MIS18A* (in yellow) and *SMARCA4* (in green) KO clones. In addition, as the size of the colonies have shown high differences between conditions, we also measured factors such as the area covered by colonies on each well (Figure 30C) and the intensity of these colonies (Figure 30D), which indicates the amount of cells that are in the colonies of each well, thus, the ability of cells to grow densely.

In general terms, there was a tendency indicating that the number of colonies decreases when FUIRI concentration increases; in other words, at highest doses of FUIRI, the number of colonies was reduced. However, this reduction in colony number differed between *BRIP1*, *MIS18A* and *SMARCA4* KO clones: in the highest FUIRI concentration (20 µM 5-FU + 11 nM SN-38), there were almost no colonies in *MIS18A* and *SMARCA4* KO cells whereas *BRIP1* KO clones had more than 100 colonies, but still notably smaller than without FUIRI (Figure 30B).

Additionally, Figure 30C indicates that the area covered by *BRIP1* KO clones was lower at dose 0 compared to control cells; indeed, it decreased compared to controls (in clone 12 in a statistically significant manner) except for clone 11 and for the highest FUIRI concentration, where *BRIP1* KO showed a slightly higher number than control KO. Similar results were observed when looking at the intensity percentage (Figure 30D): at lower doses of chemotherapy *BRIP1* KO clones presented less colony intensity,

A
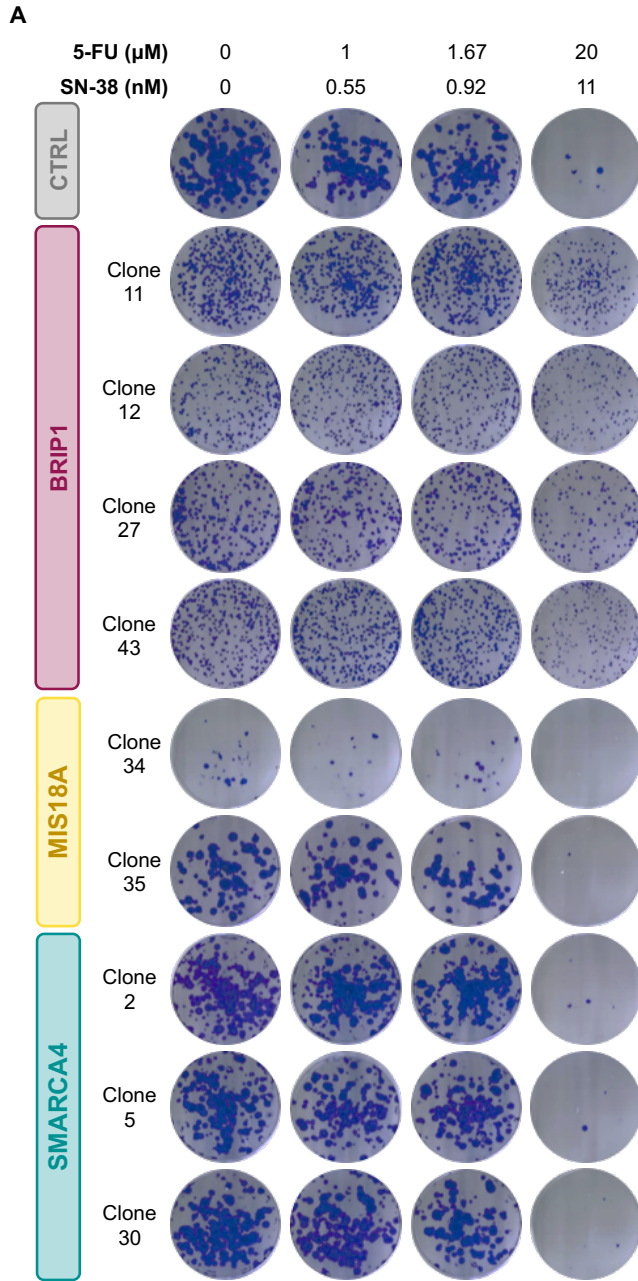
Figure 30. Colony formation assay of Control, *BRIP1*, *MIS18A* and *SMARCA4* KO clones in combination with FUIRI. (A) Colony images stained with crystal violet of Control, *BRIP1* (11, 12, 27 and 43), *MIS18A* (34 and 35) and *SMARCA4* (2, 5 and 30) KO clones at different FUIRI treatments. (B) Quantification by ImageJ of the total number of colonies of *BRIP1* (in purple), *MIS18A* (in yellow) and *SMARCA4* (in green) KO clones at the different FUIRI concentrations. (C) Quantification by ImageJ of the area (%) covered by the total number of colonies on each well for *BRIP1* (in purple), *MIS18A* (in yellow) and *SMARCA4* (in green) KO clones. (D) Quantification by ImageJ of the intensity (%) of the colonies on each well, which indicates the number of cells forming the colonies, thus, the ability of cells to grow densely. Two-way ANOVA statistic test with multiple comparisons was applied to all clones; differences on cell area and intensity within the same treatment condition that were statistically significant are represented as * (p-value < 0.05).

96
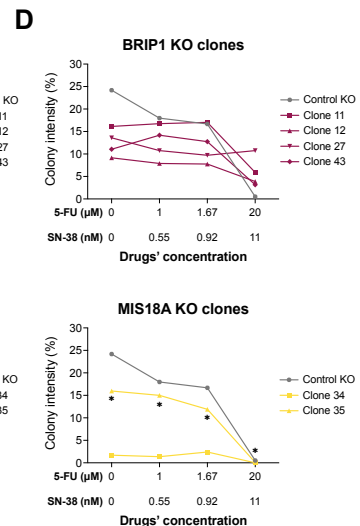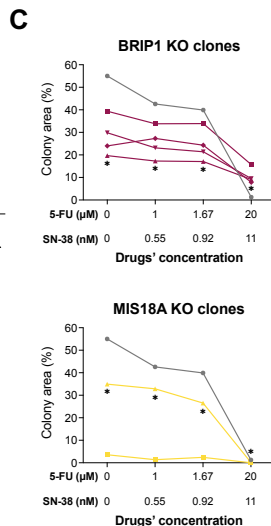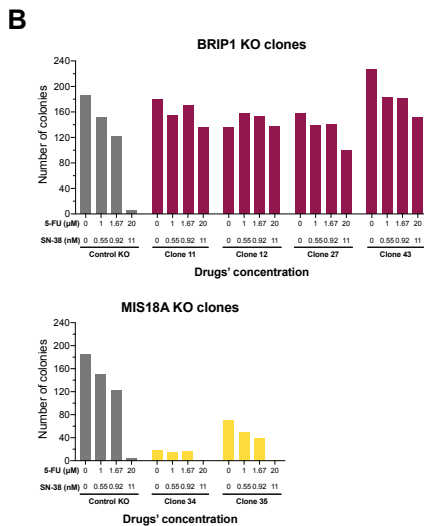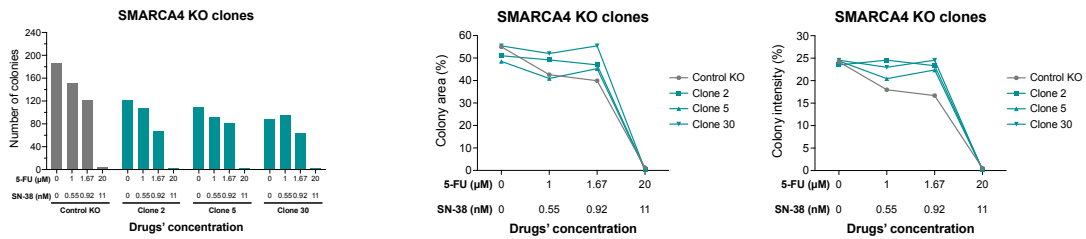
indicating a slower growth, yet these numbers were more maintained in the presence of FUIRI than in control KO cells, which started decreasing the number of cells progressively. In particular, at moderate concentrations of FUIRI, there were no big differences on colony intensity respect to control condition: clones 12 and 27 presented a slight decrease on this parameter whereas clones 11 and 43 showed a little increase in colony intensity. However, at the highest dose of FUIRI (20 µM 5-FU + 11 nM SN-38), all *BRIP1* KO clones (except 27) strongly decreased their colony intensity, although it was always maintained superior to control clone; in the case of clone 27, it slightly increased colony intensity in the highest concentration of FUIRI, being above control clone as well.

Figure 30B (middle) demonstrates that *MIS18A* KO clones grow much less than control clones: in absolute numbers, while we found more than 100 colonies in control conditions (except at the highest dose of FUIRI), *MIS18A* KO clones had less than 30 colonies in clone 34 and less than 60 colonies in clone 35, even in non-treated conditions. Furthermore, these differences between clones 34 and 35 were also translated in colony areas (Figure 30C, middle) and intensities (Figure 30D, middle): in both parameters, clone 35 presented a decrease on colony area and intensity as long as FUIRI concentration is increased, and was always maintained below control condition. Clone 34 showed a flat tendency in both parameters due to the lack of growth of the colonies in all the different conditions, even in the non-treated.

Regarding *SMARCA4* KO clones, although they seemed to appear less colonies than in the control clone (Figure 30B, bottom), intensities and areas were clearly above control in all FUIRI doses except the highest one (Figures 30C and 30D, bottom). This behavior clearly differed from the *BRIP1* and *MIS18A* KOs. Therefore, even though the absolute number of colonies was lower, these colonies were occupying more space and had more cells throughout different FUIRI treatments, indicating a higher capacity to survive and expand than control cells.

6. Analysis of *BRIP1*, *MIS18A*, *PBRM1* and *SMARCA4* in patients' samples

6.1. Analysis in primary tumor samples by NanoString technology

We next monitored candidate's gene expression in 96 primary tumor samples of CRC patients. These patients were subsequently treated with FOLFIRI and their response outcome recorded following the RECIST criteria (209). As such, it was considered as responsive to FOLFIRI when the tumor followed the criteria of complete remission (CR) and partial response (PR), whereas non-responders included those that fell in the category of progressive disease (PD). A customized panel of 25 genes (Annex IV) which included top validated hits coming from the screening and selected genes involved in immunoresponse, was used to analyze RNA expression by NanoString technology. This technique allows to detect small and/or degraded RNA fragments. Therefore, this approach seemed the best option since samples were processed from paraffin-embedded tissue and were quite old, suspecting high RNA degradation, which would prevent a robust analysis by qPCR even if using TaqMan probes.

Figure 31 presents a heat map of expression levels from all the genes that were analyzed based on their *z* score, which ranges from -3 (in blue) to 3 (in brownish). It can be

Patient Samples
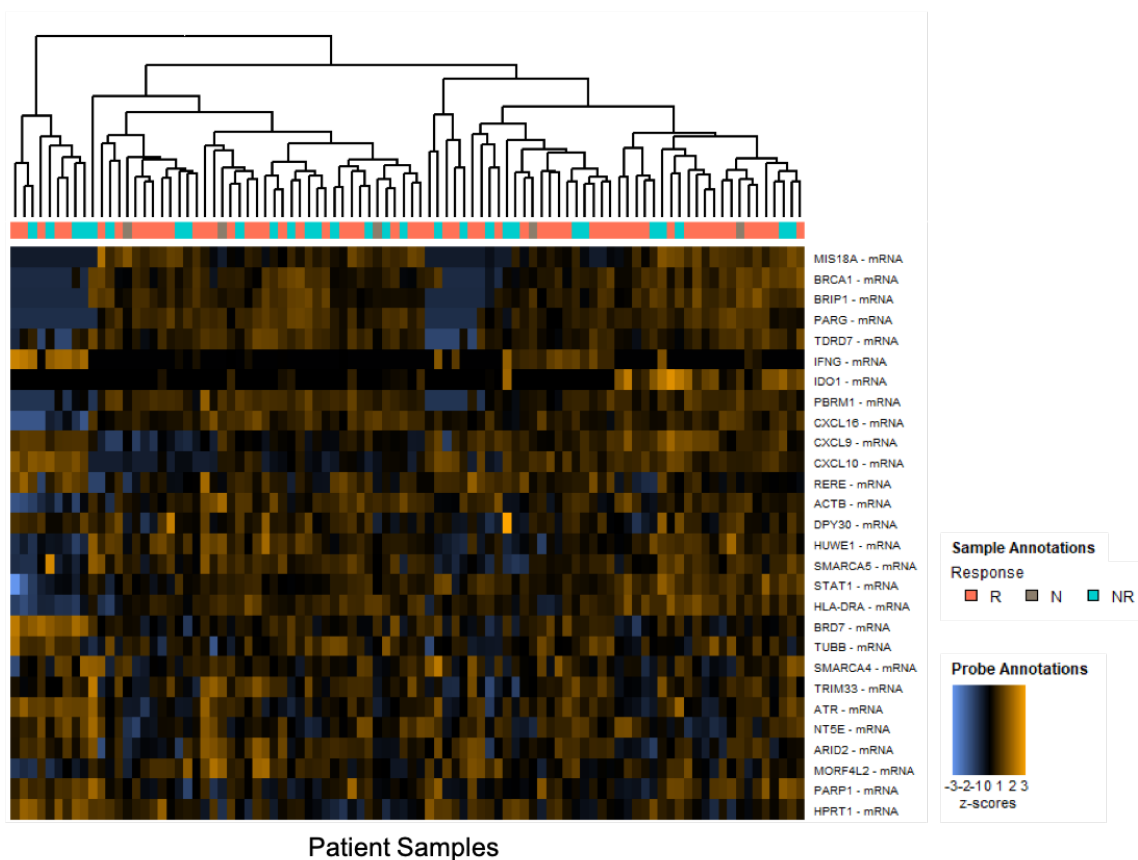
Figure 31. Heat map of expression levels from the 25 genes analyzed by NanoString. In the right, there are specified all the genes analyzed by NanoString. The columns comprise the 96 samples. Expression levels are plotted based on their *z* score, which ranges from -3 (in blue) to 3 (in brownish). IFNG and IDO1 failed on the sequencing. In the upper legend are represented responders ("R"; in orange), non-responders ("NR"; in blue) and in grey samples that finally did not show clear data ("N").

observed that in this unsupervised analysis, there were no clear clusters or groups of genes created based on their expression levels association to treatment responses.

Next, we performed a more detailed analysis for our four genes of interest by monitoring their expression status in this cohort of patients associated to responses, progression free survival and overall survival probabilities. *BRIP1* expression levels were similar in average between patients that responded or not to FOLFIRI (Figure 32A). Figure 32B presents Kaplan-Meier curves of the Progression Free Survival (PFS) from these patients divided by high (in green) or low (in blue) levels of *BRIP1* expression; PFS indicates the time between treatment initiation until disease progression. In this case, there were no differences on PFS according to *BRIP1* expression levels. In Figure 32C is shown the Kaplan-Meier curves of the Overall Survival (OS), which is the time a patient survives since the moment of diagnosis or treatment initiation; again, in this case there were no differences on OS when patients were stratified by *BRIP1* expression.
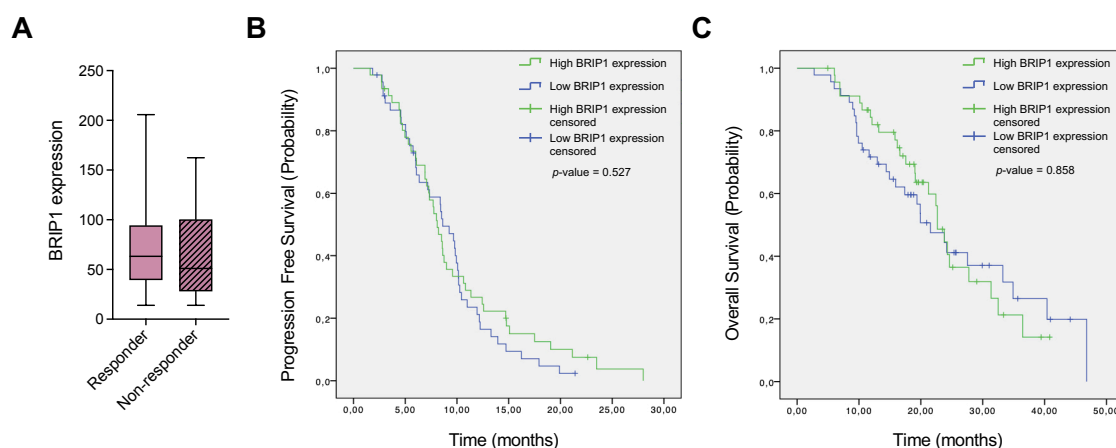
Figure 32. *BRIP1* implication in primary tumor samples. (A) Box plot of *BRIP1* expression levels comparing responders vs. non-responders to FOLFIRI. **(B)** Kaplan-Meier curves of PFS probability comparing high *BRIP1* expression levels (in green) vs. low *BRIP1* expression levels (in blue). **(C)** Kaplan-Meier curves of OS probability comparing high *BRIP1* expression levels (in green) vs. low *BRIP1* expression levels (in blue). Mantel-Cox test was used to compare between the two survival conditions and *p*-values (log-rank values) are plotted on each graph. High expression levels were considered above the median and low expression levels were considered below the median.

Figure 33 presents the results of *MIS18A* expression levels in this patient cohort. *MIS18A* expression levels did not vary between responders and non-responders to FOLFIRI (Figure 33A). Figures 33B and 33C present Kaplan-Meier curves with PFS and OS, respectively; again none of the ratios showed a difference on patients' survival when stratified by *MIS18A* expression. Figures 33D and 33E show PFS and OS Kaplan-Meier curves were *MIS18A* expression has been divided by terciles; in this case, Tercile 1 (in blue) represents samples expressing low *MIS18A* levels, Tercile 2 (in green) are samples expressing middle levels of *MIS18A*, and Tercile 3 (in yellow) present the

samples with the highest levels of *MIS18A* expression. In PFS graph (Figure 33D) it can be observed a tendency where patients expressing the highest levels of *MIS18A* had better survival, although this was not statistically significant; on the other hand, OS graph divided by terciles (Figure 33E) did not indicate a tendency on better outcome according to *MIS18A* expression.
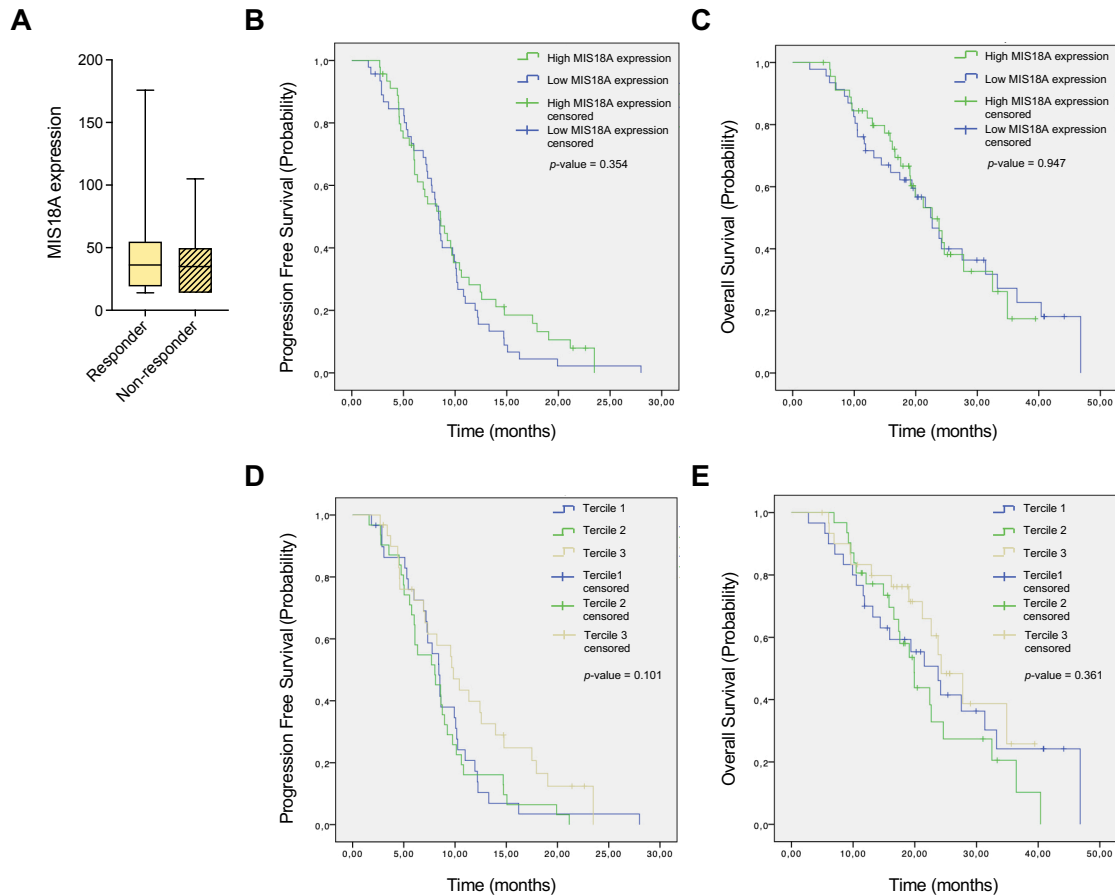
Figure 33. *MIS18A* implication in primary tumor samples. (A) Box plot of *MIS18A* expression levels comparing responders vs. non-responders to FOLFIRI. **(B)** Kaplan-Meier curves of PFS probability comparing high *MIS18A* expression levels (in green) vs. low *MIS18A* expression levels (in blue). **(C)** Kaplan-Meier curves of OS probability comparing high *MIS18A* expression levels (in green) vs. low *MIS18A* expression levels (in blue). **(D)** Kaplan-Meier curves of PFS probabilities comparing *MIS18A* expression levels divided by Tercile 1 (in blue), Tercile 2 (in green) and Tercile 3 (in yellow). **(E)** Kaplan-Meier curves of OS probabilities comparing *MIS18A* expression levels divided by Tercile 1 (in blue), Tercile 2 (in green) and Tercile 3 (in yellow). Mantel-Cox test was used to compare between the survival conditions and *p*-values (log-rank values) are plotted on each graph. High expression levels were considered above the median and low expression levels were considered below the median. Tercile 1 was samples expressing the lowest levels of *MIS18A* (1-33%), Tercile 2 was samples expressing middle levels of *MIS18A* (34-66%), and Tercile 3 was samples expressing the highest levels of *MIS18A* (67-100%).

Regarding the results of *PBRM1* expression, they are plotted on Figure 34. *PBRM1* expression levels were similar on average between patients that responded or not to FOLFIRI (Figure 34A). Regarding PFS and OS probabilities, when patients were
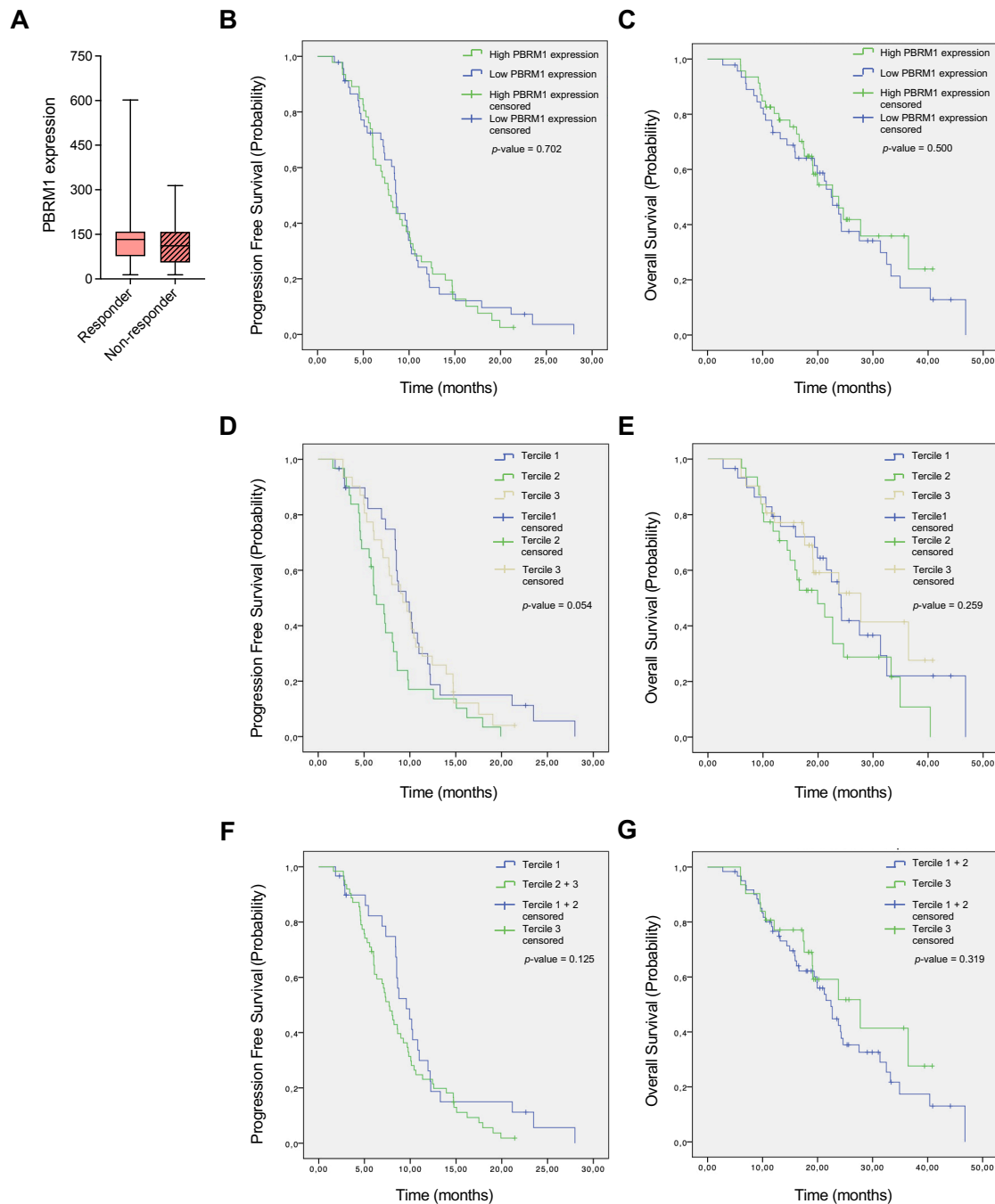
Figure 34. *PBRM1* implication in primary tumor samples. (A) Box plot of *PBRM1* expression levels comparing responders vs. non-responders to FOLFIRI. **(B)** Kaplan-Meier curves of the PFS probability comparing high *PBRM1* expression levels (in green) vs. low *PBRM1* expression levels (in blue). **(C)** Kaplan-Meier curves of the OS probability comparing high *PBRM1* expression levels (in green) vs. low *PBRM1* expression levels (in blue). **(D)** Kaplan-Meier curves of the PFS probabilities comparing *PBRM1* expression levels divided by Tercile 1 (in blue), Tercile 2 (in green) and Tercile 3 (in yellow). **(E)** Kaplan-Meier curves of the OS probabilities comparing *PBRM1* expression levels divided by Tercile 1 (in blue), Tercile 2 (in green) and Tercile 3 (in yellow). **(F)** Kaplan-Meier curves of the PFS probability comparing *PBRM1* expression levels divided by Tercile 1 (in blue) and Tercile 2 + 3 (in green). **(G)** Kaplan-Meier curves of the OS probability comparing *PBRM1* expression levels divided by Tercile 1 + 2 (in blue) and Tercile 3 (in green). Mantel-Cox test was used to compare between the survival conditions and *p*-values (log-rank values) are plotted on each graph. High expression levels were considered above the median and low expression levels were considered below the median. Tercile 1 corresponded to samples expressing the lowest levels of *PBRM1* (1-33%), Tercile 2 to samples expressing middle levels of *PBRM1* (34-66%), and Tercile 3 to samples expressing the highest levels of *PBRM1* (67-100%).

stratified by high (in green) or low (in blue) levels of *PBRM1* expression (Figures 34B and 34C), there were no significant differences between the two groups. If *PBRM1* expression levels were divided by Terciles, in both PFS and OS Kaplan-Meier curves (Figures 34D and 34E) seemed that there was a tendency where patients having middle *PBRM1* expression levels (Tercile 2; in green) had worst progression of the disease, albeit without significant *p*-values. When samples from Terciles 2 and 3 were plotted together (in blue) vs. samples of Tercile 1 (in green), it also appeared that patients with middle-high levels of *PBRM1* had worst progression of the tumor (Figure 34F). However, these associations were not significant, with *p*-values above 0.05. OS of patients expressing the highest levels of *PBRM1* (Tercile 3; in green) was higher than in patients with middle-low *PBRM1* expression (Figure 34G), again without significance.
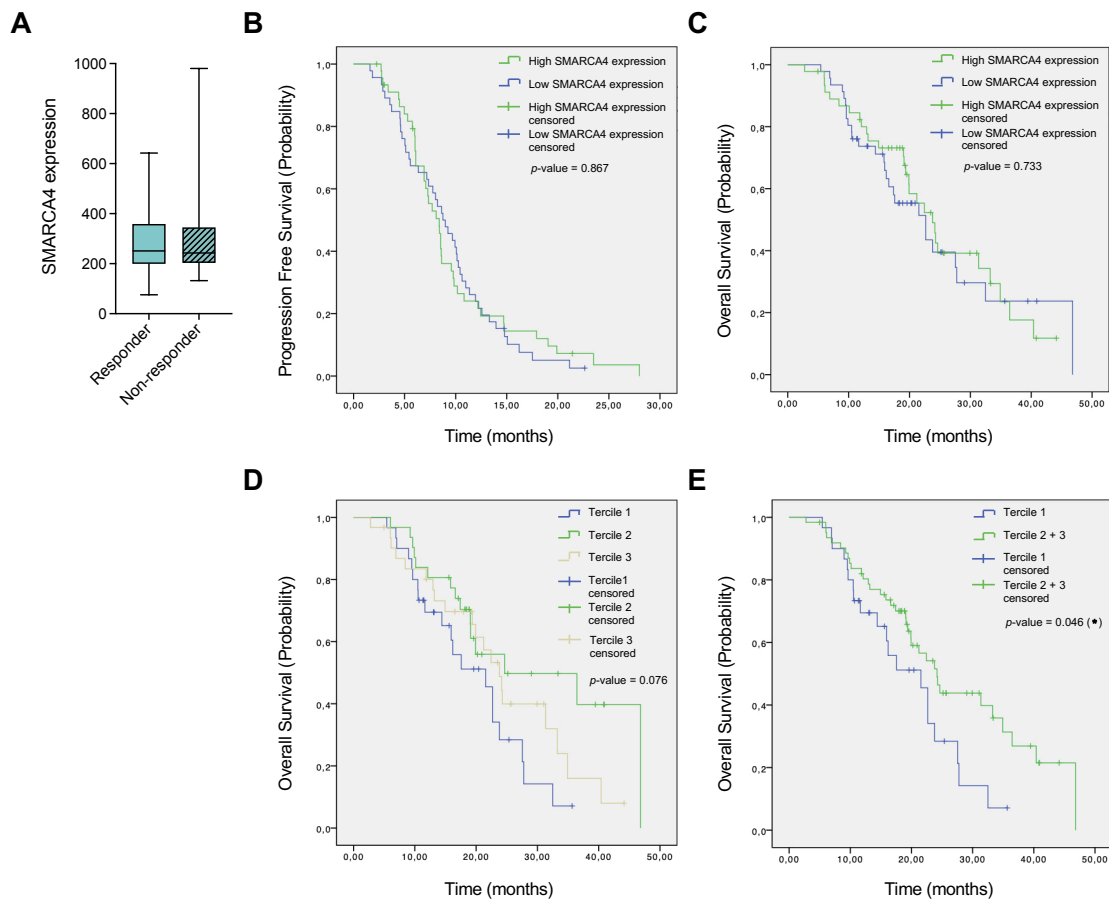
Figure 35. *SMARCA4* **implication in primary tumor samples. (A)** Box plot of *SMARCA4* expression levels comparing responders vs. non-responders to FOLFIRI. **(B)** Kaplan-Meier curves of PFS probability comparing high *SMARCA4* expression levels (in green) vs. low *SMARCA4* expression levels (in blue). **(C)** Kaplan-Meier curves of OS probability comparing high *SMARCA4* expression levels (in green) vs. low *SMARCA4* expression levels (in blue). **(D)** Kaplan-Meier curves of OS probabilities comparing *SMARCA4* expression levels divided by Tercile 1 (in blue), Tercile 2 (in green) and Tercile 3 (in yellow). **(E)** Kaplan-Meier curves of OS probabilities comparing *SMARCA4* expression levels divided by Tercile 1 (in blue) and Tercile 2 + 3 (in green). Mantel-Cox test was used to compare between the survival conditions and *p*-values (log-rank values) are plotted on each graph. High expression levels were considered above the median and low expression levels were considered below the median. Tercile 1 corresponds to samples expressing the lowest levels of *SMARCA4* (1-33%), Tercile 2 samples expressing middle levels of *SMARCA4* (34-66%), and Tercile 3 samples expressing the highest levels of *SMARCA4* (67-100%).

In the case of SMARCA4, expression levels were similar, on average, when responders and non-responders were compared (Figure 35A). Regarding PFS and OS rates, when *SMARCA4* levels were divided by high (in green) or low (in blue) expression levels (Figures 35B and 35C, respectively), there were no significant differences on survival or prognosis in both cases.

However, if *SMARCA4* levels of expression were divided by Terciles, OS Kaplan-Meier curves were separated (Figure 35D), revealing a tendency where patients expressing the lowest *SMARCA4* levels (Tercile 1; in blue) presented the worst survival ratio (35 months vs. 45 months in Tercile 2 and almost 50 months in Tercile 3). Furthermore, when samples from Terciles 2 and 3 were plotted together (in green) vs. samples of Tercile 1 (in blue), there was a statistically significant difference on OS (Figure 35E), where patients with lowest *SMARCA4* expression levels had worst survival (35 months) in contrast to patients expressing middle-high levels of *SMARCA4* (almost 50 months).

6.2. *In silico* analysis in public databases (GSEs)

6.2.1. *In silico* analysis in GSE104645 cohort

GSE104645 is a public available cohort of 193 CRC patients (210) treated with FOLFOX or FOLFIRI where gene expression was analyzed by microarray from the primary tumor tissue. In this database, there is clinicopathological information about tumor stage, number of metastases, type of response to chemotherapy, CMS classification and information about months of PFS and OS.

Figure 36 presents the implication of *BRIP1* in this cohort of patients. Figure 36A presents a box plot of *BRIP1* expression levels at different tumor stages (I, II, III or IV), where *BRIP1* levels were, in average, similar between stages II, III and IV, although the number of cases and deviation per condition was slightly different. Figure 36B shows in a box plot *BRIP1* expression divided by the number of metastases that these patients further developed (1, 2 or 3), where *BRIP1* expression was decreased as long as more metastases appeared, although there was not a significant association. The type of response that these patients had in front of FOLFOX (left) or FOLFIRI (right) according to *BRIP1* expression is shown in Figure 36C: in patients treated with FOLFOX, *BRIP1* expression was slightly elevated in those patients that progress on the disease in comparison to those ones that partially responded to chemotherapy; patients treated with FOLFIRI had a similar *BRIP1* expression in average between those who partially responded and patients that presented an stabilization of the tumor. In Figure 36D is shown a box plot of *BRIP1* levels according to the CMS classification; it can be observed that patients belonging to CMS3, which is the "metabolic" subtype characterized by
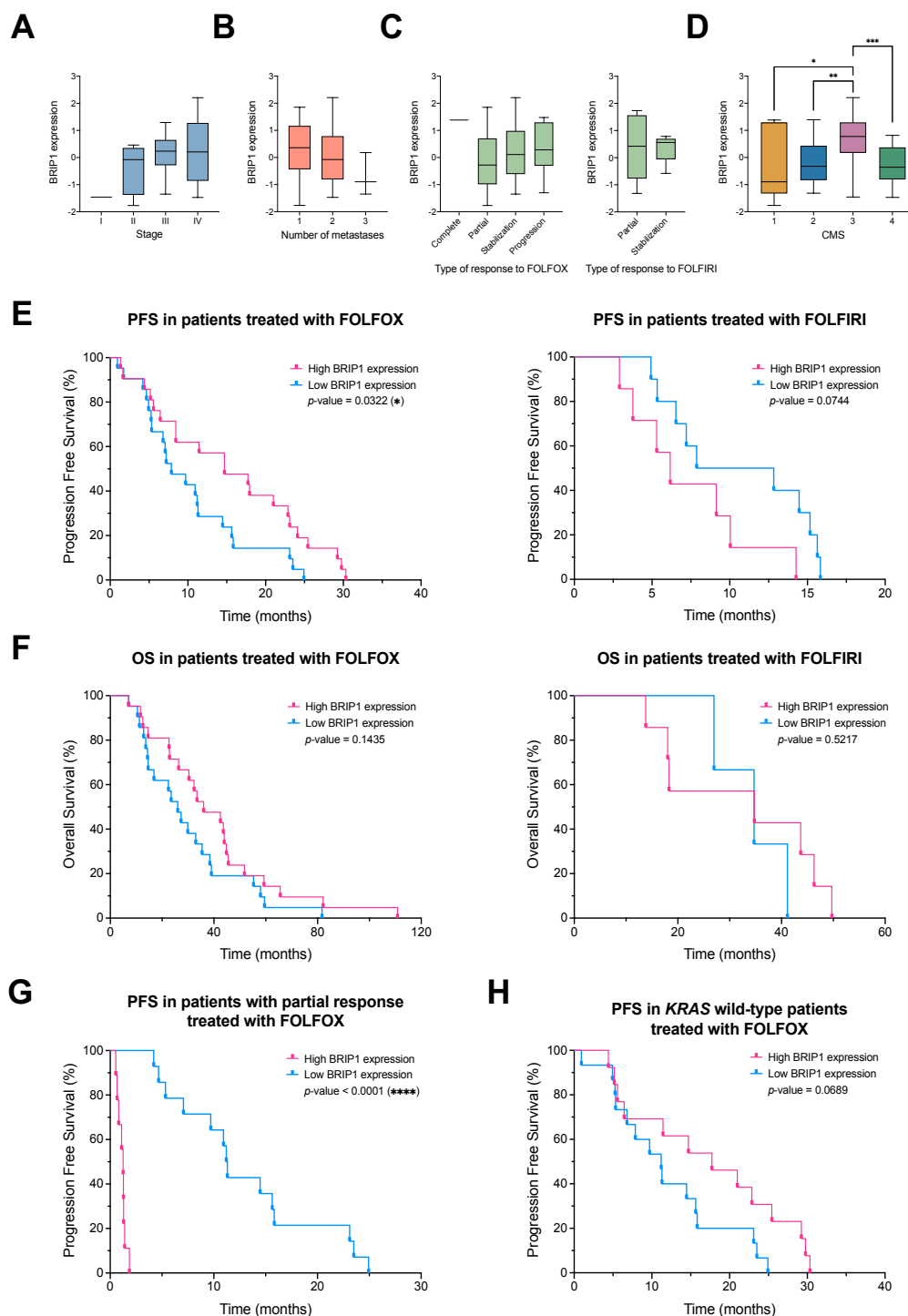
Figure 36. Exploration of *BRIP1* expression in GSE104645 cohort of CRC patients. (A) *BRIP1* expression levels in the different stages of CRC. **(B)** *BRIP1* expression levels divided by the number of metastases. **(C)** *BRIP1* expression levels throughout the different type of responses to FOLFOX (left) or FOLFIRI (right). **(D)** *BRIP1* expression levels according to the different CRC CMSs. **(E)** Kaplan-Meier curves of PFS percentages comparing high *BRIP1* expression levels (in pink) vs. low *BRIP1* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). **(F)** Kaplan-Meier curves of OS percentages comparing high *BRIP1* expression levels (in pink) vs. low *BRIP1* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). **(G)** Kaplan-Meier curves of the PFS comparing high *BRIP1* expression levels (in pink) vs. low *BRIP1* expression levels (in blue) in patients treated with FOLFOX that had a partial response. **(H)** Kaplan-Meier curves of the PFS comparing high *BRIP1* expression levels (in pink) vs. low *BRIP1* expression levels (in blue) in *KRAS* wild-type patients treated with FOLFOX. In graphs A-D, unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (*p*-value < 0.05), ** (*p*-value < 0.005) and *** (*p*-value < 0.0005); in graphs E-H, Mantel-Cox test was used to compare between the two survival conditions and *p*-values (log-rank values) are plotted on each graph. High expression levels were considered as ≥ 0.5, and low expression levels were considered as ≤ -0.5.

metabolic deregulation and mutations on *KRAS*, had the highest expression of *BRIP1*, whereas CMS1 patients, which are MSI, have high immune infiltration and are *BRAF* mutated, presented the lowest levels of *BRIP1*. Figure 36E presents Kaplan-Meier curves of PFS in patients treated with FOLFOX (left) and FOLFIRI (right) where *BRIP1* expression levels have been separated by high (in pink) and low (in blue) levels; in this case, the tendency presented by the two graphs was the opposite: in patients treated with FOLFOX, higher levels of *BRIP1* were significantly correlated with better survival of patients (30 months) with significant value (*p*-value = 0.0322), whereas in patients treated with FOLFIRI higher *BRIP1* levels were correlated with worst progression of CRC (less than 15 months). In Figure 36F are shown OS Kaplan-Meier curves comparing patients with high (in pink) and low (in blue) levels of *BRIP1* expression. In the left, patients that received FOLFOX chemotherapy presented better prognosis when they had higher *BRIP1* levels (110 months) in comparison to patients with low *BRIP1* levels (80 months); patients that received FOLFIRI (in the right) showed an unclear tendency of better outcome based on *BRIP1* expression. These associations were not significant but followed the same tendencies as Kaplan-Meier curves of PFS. Figure 36G shows PFS Kaplan-Meier curves of patients that were treated with FOLFOX and responded partially to it; of note, there was a clear and significant difference on disease progression when patients were stratified by high or low *BRIP1* levels; patients with high *BRIP1* levels presented a PFS of around 2.5 months in contrast to 25 months if *BRIP1* levels were low. Since the CRC cell line used during the thesis (HT-29) was *KRAS* wild-type, it was analysed the PFS of *KRAS* wild-type patients treated with FOLFOX from GSE104645 cohort (Figure 36H); in this regard, higher *BRIP1* levels of expression were correlated with better PFS (30 months) in comparison to the group with low *BRIP1* expression (25 months).

In the same way, Figure 37 presents the implication of *MIS18A* expression in this cohort of CRC patients. In Figure 37A is represented *MIS18A* expression throughout the different stages of these patients; there was a non-significant tendency of decreased *MIS18A* expression when the tumor stage was more advanced. Figure 37B shows increasing *MIS18A* expression levels in patients that presented 1, 2 or 3 metastases; *MIS18A* levels classified by the type of response that patients had to FOLFOX (left) and FOLFIRI (right) are shown in Figure 37C; in average, there were no differences on *MIS18A* expression levels. Figure 37D shows *MIS18A* levels according to the different CMS: in this case, patients belonging to CMS3 had lowest *MIS18A* expression levels, while CMS1 patients presented the highest *MIS18A* levels. In Figure 37E are represented Kaplan-Meier curves with the PFS of patients treated with FOLFOX (left) or

FOLFIRI (right) according to their *MIS18A* level of expression; differences between having high (in pink) or low (in blue) levels of *MIS18A* seemed unclear for both chemotherapies, presenting a PFS of 30 months in FOLFOX-treated patients and a PFS of 15 months in FOLFIRI-treated patients. Figure 37F shows OS Kaplan-Meier curves of CRC patients treated with FOLFOX (left) or FOLFIRI (right); again, there were no differences in prognosis if *MIS18A* expression levels were considered.
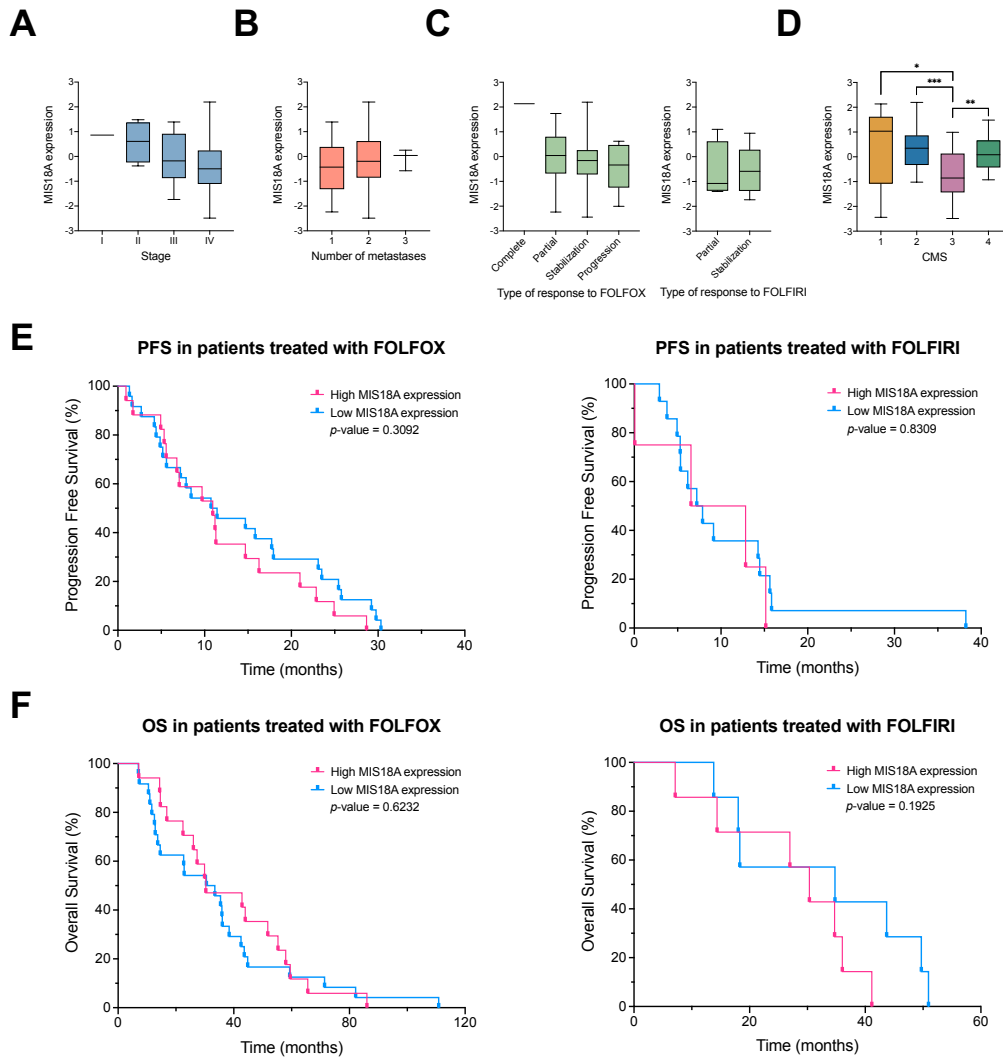
Figure 37. Exploration of *MIS18A* expression in GSE104645 cohort of CRC patients. (A) Box plot of *MIS18A* expression levels in the different stages of CRC. **(B)** Box plot of *MIS18A* expression levels divided by the number of metastases. **(C)** Box plot of *MIS18A* expression levels throughout the different type of responses to FOLFOX (left) or FOLFIRI (right). **(D)** Box plot of *MIS18A* expression levels according to the different CRC CMSs. **(E)** Kaplan-Meier curves of the PFS comparing high *MIS18A* expression levels (in pink) vs. low *MIS18A* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). **(F)** Kaplan-Meier curves of the OS comparing high *MIS18A* expression levels (in pink) vs. low *MIS18A* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). In graphs A-D, unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (p-value < 0.05), ** (p-value < 0.005) and *** (p-value < 0.0005); in graphs E and F, Mantel-Cox test was used to compare between the two survival conditions and p-values (log-rank values) are plotted on each graph. High expression levels were considered as ≥ 0.5, and low expression levels were considered as ≤ -0.5.

Figure 38. Exploration of *PBRM1* expression in GSE104645 cohort of CRC patients. (A) Box plot of *PBRM1* expression levels in the different stages of CRC. **(B)** Box plot of *PBRM1* expression levels divided by the number of metastases. **(C)** Box plot of *PBRM1* expression levels throughout the different type of responses to FOLFOX (left) or FOLFIRI (right). **(D)** Box plot of *PBRM1* expression levels according to the different CRC CMSs. **(E)** Kaplan-Meier curves of the PFS comparing high *PBRM1* expression levels (in pink) vs. low *PBRM1* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). **(F)** Kaplan-Meier curves of the OS comparing high *PBRM1* expression levels (in pink) vs. low *PBRM1* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). In graphs A-D, unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (p-value < 0.05), *** (p-value < 0.0005) and **** (p-value < 0.0001); in graphs E and F, Mantel-Cox test was used to compare between the two survival conditions and p-values (log-rank values) are plotted on each graph. High expression levels were considered as ≥ 0.5, and low expression levels were considered as ≤ -0.5.

Regarding the implication of *PBRM1* in GSE104645 cohort, Figure 38A shows *PBRM1* expression at the different stages of these patients whereas Figure 38B presents *PBRM1* expression in patients that developed 1, 2 or 3 metastases; in general, there were no differences in the average expression of *PBRM1* in none of the two graphs. Figure 38C shows in box plots *PBRM1* expression depending on the type of response that patients did when treated with FOLFOX (left) or FOLFIRI (right); overall, there is a slightly decrease on *PBRM1* expression in patients that suffered a progression of the disease in

comparison to those patients that partially responded to chemotherapy or that stabilized the disease. Figure 38D shows *PBRM1* expression classified by the CMS of each patient; in this regard, *PBRM1* was expressed the highest in patients from CMS3 (metabolic subtype), although differences with patients from CMS1 (immune subtype) were not big and not significant. Here, *PBRM1* showed the lowest expression levels in patients belonging to CMS2, which is the "canonical" subtype characterized by high SCNAs and *MYC* activation. Figure 38E presents the PFS Kaplan-Meier curves of patients treated with FOLFOX (left) or FOLFIRI (right) divided by high (in pink) or low (in blue) *PBRM1* expression; none of the graphs presented big differences on disease progression: in patients treated with FOLFOX, PFS was around 25-30 months while in FOLFIRI-treated patients PFS was around 15 months. Kaplan-Meier curves with OS ratios are represented in Figure 38F; of note, there was a worst OS tendency of patients with low *PBRM1* levels compared to high-expressing ones, while there was unclear tendency in patients treated with FOLFIRI (right). Importantly, when positive vs. negative *PBRM1* levels were plotted (not using cut-offs of higher of 0.5 FDR or lower than -0.5) worst OS with significant *p*-value (0.0102) was obtained for *PBRM1* negative expressing tumors (data not shown).

In the case of *SMARCA4* implication in this cohort of patients, Figure 39A presents a tendency to show decreased *SMARCA4* expression levels in more advanced tumor stages. Regarding *SMARCA4* expression in patients that develop 1, 2 or 3 metastases (Figure 39B), there was a significant increase on *SMARCA4* expression in patients that present 2 metastases in comparison to those ones presenting only 1 (*p*-value = 0.0058). Figure 39C shows *SMARCA4* expression levels depending on the type of response that patients had to FOLFOX (left) or FOLFIRI (right), presenting less *SMARCA4* expression in those patients that progressed on the disease in comparison to those that partially responded to chemotherapy or stabilized the disease. In Figure 39D is clear that patients belonging to CMS3 (metabolic subtype) presented the lowest expression levels of *SMARCA4*. Figure 39E presents the Kaplan-Meier curves of the PFS in patients treated with FOLFOX (left) or FOLFIRI (right) according to high expression levels of *SMARCA4* (in pink) or low *SMARCA4* expression levels (in blue). In FOLFOX-treated patients, although the PFS was of 30 months for both groups, there was a tendency to worst PFS in low expressing *SMARCA4* patients, but not significant, whereas PFS in FOLFIRI-treated patients was around 15 months and there were no differences between high and low *SMARCA4* levels. Figure 39F shows OS Kaplan-Meier curves of patients treated with FOLFOX (left) or FOLFIRI (right) depending on *SMARCA4* expression levels; in FOLFOX-treated patients OS was 110 months when *SMARCA4* was low in comparison

to patients with high *SMARCA4* levels (50 months); in the case of patients treated with FOLFIRI, OS was similar when patients were stratified by low and high *SMARCA4* expression (around 50 months), although it seems there was a worst but not significant tendency in patients with lower levels of *SMARCA4*.
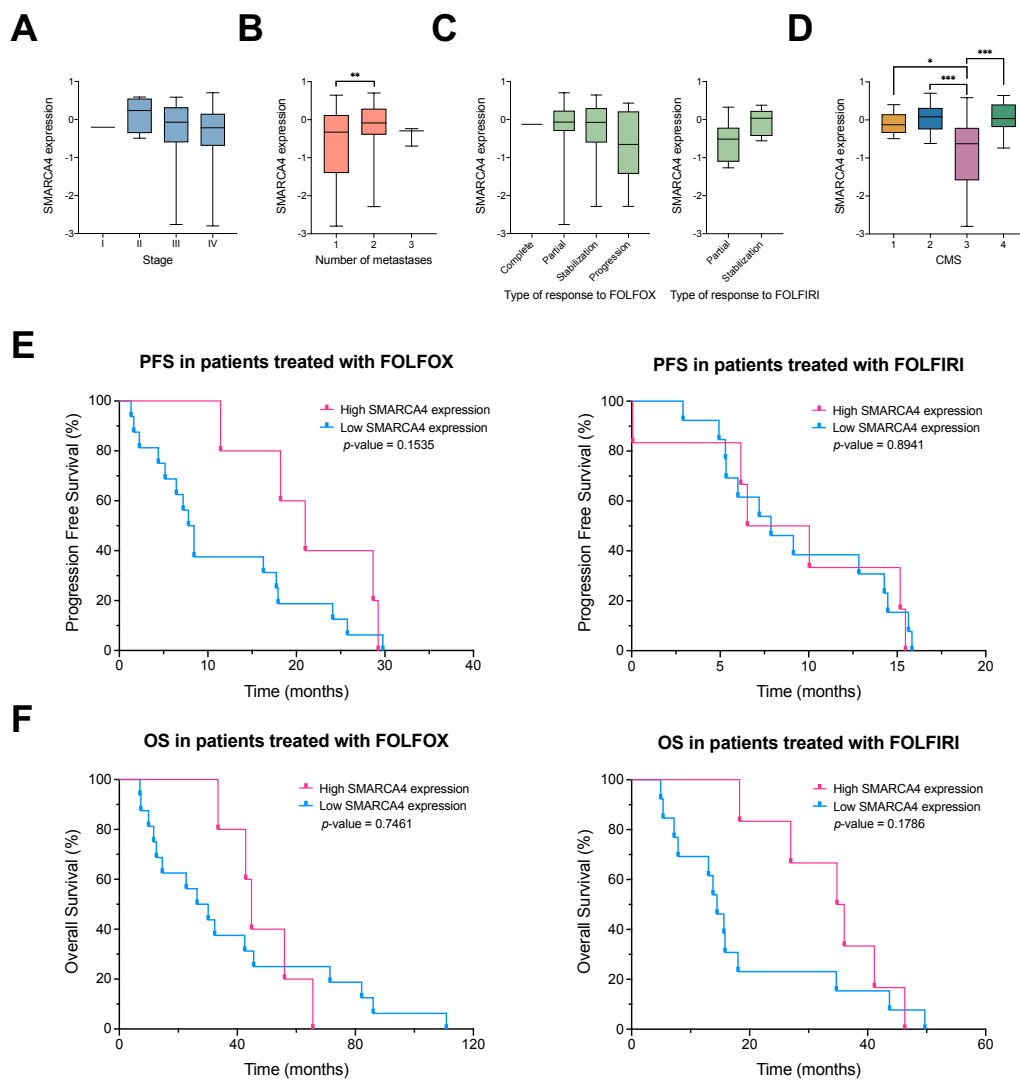
Figure 39. Exploration of *SMARCA4* expression in GSE104645 cohort of CRC patients. (A) Box plot of *SMARCA4* expression levels in the different stages of CRC. **(B)** Box plot of *SMARCA4* expression levels divided by the number of metastases. **(C)** Box plot of *SMARCA4* expression levels throughout the different type of responses to FOLFOX (left) or FOLFIRI (right). **(D)** Box plot of *SMARCA4* expression levels according to the different CRC CMSs. **(E)** Kaplan-Meier curves of the PFS comparing high *SMARCA4* expression levels (in pink) vs. low *SMARCA4* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). **(F)** Kaplan-Meier curves of the OS comparing high *SMARCA4* expression levels (in pink) vs. low *SMARCA4* expression levels (in blue) in patients treated with FOLFOX (left) or FOLFIRI (right). In graphs A-D, unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (p-value < 0.05), ** (p-value < 0.005) and *** (p-value < 0.0005); in graphs E and F, Mantel-Cox test was used to compare between the two survival conditions and p-values (log-rank values) are plotted on each graph. High expression levels were considered as ≥ 0.5 for FOLFOX and ≥ 0 for FOLFIRI, and low expression levels were considered as ≤ -0.5.

6.2.2. *In silico* analysis in GSE62322 cohort

GSE62322 is a public available cohort of advanced CRC patients (211) treated with FOLFIRI where gene expression was analyzed by microarray from the normal colon tissue, primary tumor tissue and liver metastases. In this database, there is information about the response to FOLFIRI (responder vs. non-responder).

Again, a possible predictive role for our candidate genes was explored. Figure 40A shows a before-after plot assessing *BRIP1* expression in matched normal, tumor and liver metastases of each patient. In addition, Figure 40B presents the same data in a box plot to monitor how *BRIP1* expression changes in average; it can be observed a significant increase on *BRIP1* expression in primary tumor samples compared to normal colon. In Figure 40C is presented a box plot of *BRIP1* expression comparing responders vs. non-responders to FOLFIRI throughout the different sample types (normal colon in green, primary tumor in blue, and liver metastases in red). There were no big differences between responders and non-responders inside the same sample type (normal colon, primary tumor or liver metastases); however, there was a significant increase on *BRIP1* expression when normal colon is compared to primary tumor tissue of responders.
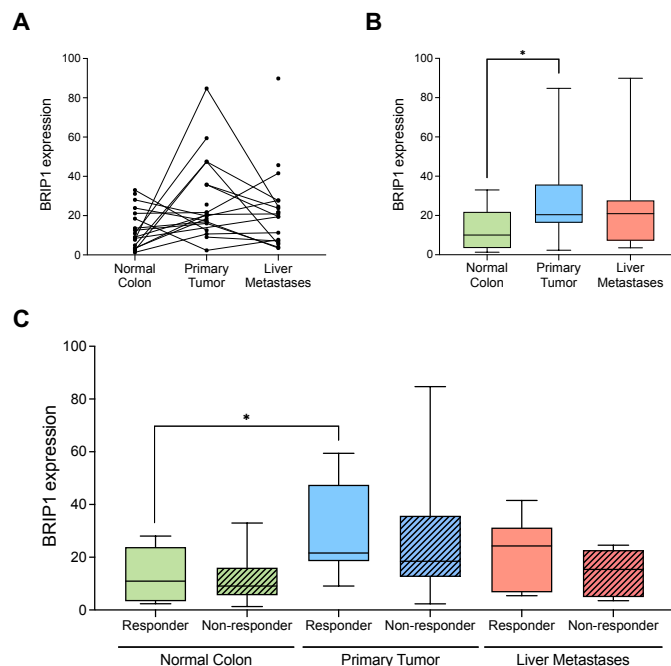
Figure 40. *BRIP1* implication in GSE62322 cohort of CRC patients treated with FOLFIRI. (A) Before-after plot of *BRIP1* expression levels in the different samples throughout the different tissue types (normal colon, primary tumor or liver metastases). **(B)** Box plot of *BRIP1* expression levels divided by normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). **(C)** Box plot of *BRIP1* expression levels comparing Responders vs. Non-responders to FOLFIRI in normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). Unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (*p*-value < 0.05).

Figure 41 presents *MIS18A* implication in this cohort of patients. Figure 41A presents a before-after plot comparing *MIS18A* expression of the same patient throughout the different types of samples (normal colon, primary tumor and liver metastases). Figure 41B presents the same data in a box plot to observe *MIS18A* expression in average; there was a slightly but significant increase on *MIS18A* expression when normal colon

and primary tumor samples were compared. Figure 41C shows a box plot of *MIS18A* expression levels that separates patients responding or not to FOLFIRI in the different sample types (normal colon in green, primary tumor in blue, and liver metastases in red). It seems there is a tendency of increased *MIS18A* expression in primary tumor samples from responder patients in comparison to the normal colonic tissue of responder patients, but with no significance.
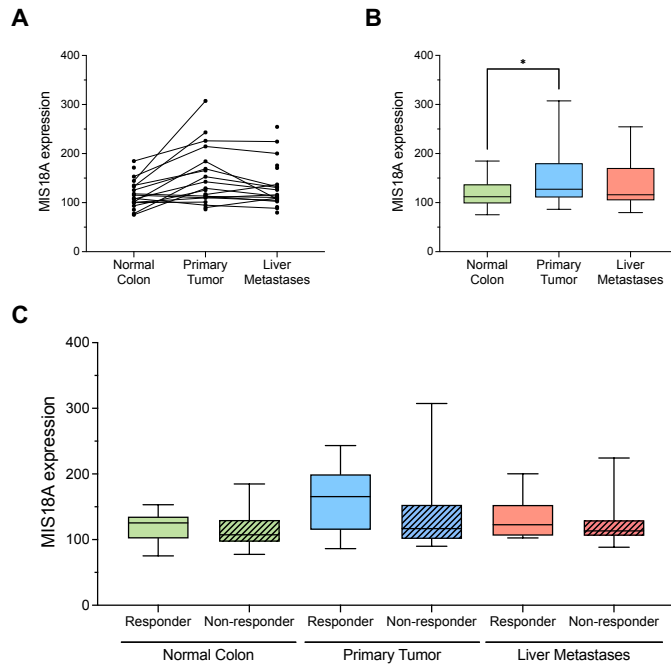
Figure 41. *MIS18A* implication in GSE62322 cohort of CRC patients treated with FOLFIRI. (A) Before-after plot of *MIS18A* expression levels in the different samples throughout the different tissue types (normal colon, primary tumor or liver metastases). **(B)** Box plot of *MIS18A* expression levels divided by normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). **(C)** Box plot of *MIS18A* expression levels comparing Responders vs. Non-responders to FOLFIRI in normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). Unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (*p*-value < 0.05).

Regarding *PBRM1* implication, Figure 42A presents a before-after plot to observe its expression throughout the different sample types (normal colon, primary tumor and liver metastases) of the same patient; it seems there was a tendency of decreased *PBRM1* levels when the sample evolves from normal colonic tissue to liver metastases. In Figure
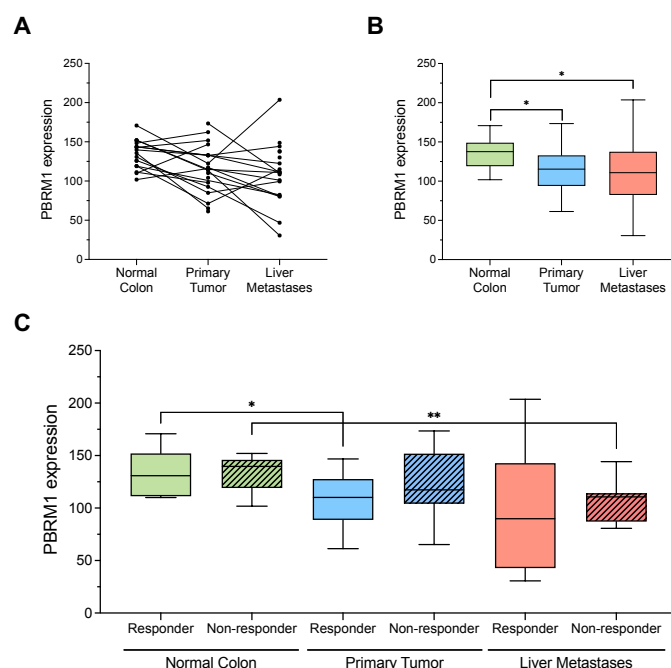
Figure 42. *PBRM1* implication in GSE62322 cohort of CRC patients treated with FOLFIRI. (A) Before-after plot of *PBRM1* expression levels in the different samples throughout the different tissue types (normal colon, primary tumor or liver metastases). **(B)** Box plot of *PBRM1* expression levels divided by normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). **(C)** Box plot of *PBRM1* expression levels comparing Responders vs. Non-responders to FOLFIRI in normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). Unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (*p*-value < 0.05) and ** (*p*-value < 0.005).

42B is presented the same data in a box plot; here, was clear and statistically significant a decrease on *PBRM1* expression levels when the stage of the sample is more advanced. *PBRM1* expression levels in responders vs. non-responders are shown in Figure 42C, where there were also classified by the type of sample (normal colon in green, primary tumor in blue, and liver metastases in red). It should be remarked a significant decrease (*p*-value = 0.0212) on *PBRM1* expression levels where the sample was in a more advanced stage (liver metastases) (Figure 42B). Indeed, Figure 42C presents a statistically significant decrease in *PBRM1* expression levels when normal colon and primary tumor samples were compared between responders (*p*-value = 0.0387); also, when normal colon was compared to liver metastases sample in non-responder patients, there was a significant decrease on *PBRM1* expression (*p*-value = 0.0094).

In the case of *SMARCA4* implication, Figure 43A shows a before-after plot to follow *SMARCA4* expression throughout the different sample types (normal colon, primary tumor and liver metastases) of the same patient; in general, it seems a tendency of increased *SMARCA4* expression when compared primary tumor vs. normal colon samples. In Figure 43B is represented the same data but in a box plot, showing an increase on *SMARCA4* expression in primary tumor samples compared to normal colonic tissue. Figure 43C presents *SMARCA4* expression when responders vs. non-responders are compared for each sample type (normal colon in green, primary tumor in blue, and liver metastases in red). It can be observed a significant increase on *SMARCA4* expression when normal colon vs. primary tissue was compared in responder patients as well as in non-responder patients.
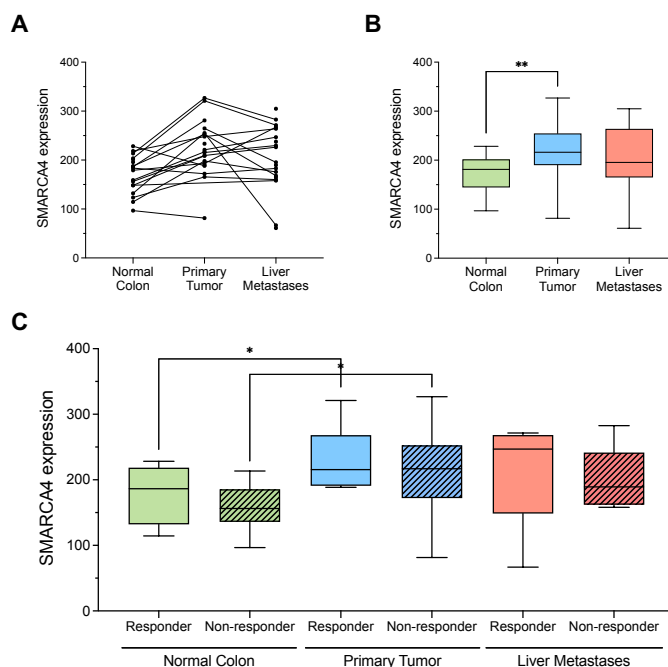
Figure 43. *SMARCA4* implication in GSE62322 cohort of CRC patients treated with FOLFIRI. (A) Before-after plot of *SMARCA4* expression levels in the different samples throughout the different tissue types (normal colon, primary tumor or liver metastases). **(B)** Box plot of *SMARCA4* expression levels divided by normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). **(C)** Box plot of *SMARCA4* expression levels comparing Responders vs. Non-responders to FOLFIRI in normal colonic tissue (in green), primary tumor tissue (in blue) or tissue from liver metastases (in red). Unpaired T-test comparing two subgroups was performed in all conditions and statistically significant differences are represented as * (*p*-value < 0.05) and ** (*p*-value < 0.005).

7. Contributions

This project was originated by sharing data and discussions between Dr. Marcus Buschbeck and Dr. Sonia Forcales. Dr. Forcales was awarded by Fundació Olga Torres with a grant in 2016, which inspired the present "Response Project", conceived by Dr. Buschbeck that expanded the study to elucidate chromatin factors' roles in resistance to chemotherapy in three cancer types: AML/MDS, CRC and lung cancer. The project is multidisciplinary and includes basic research and translational groups from the Can Ruti Campus. Dr. Sonia Forcales, Dr. Eva Martinez-Balibrea, Dr. Cristina Queralt (Dr. Eva Martinez-Balibrea's Lab) and I designed and supervised the realization of this project in CRC disease.

Dr. Cristina Queralt and I performed all the experiments included in the setting of the LOF, the realization of the LOF and the individual validation of candidate genes. DNA extraction for NGS was performed in collaboration with Raquel Casquero (Dr. Marcus Buschbeck Lab); analysis of NGS data following Zuber's lab pipeline was performed by Dr. Jeannine Diesch (Dr. Marcus Buschbeck Lab).

Gene expression data and associated clinicopathological information extracted from two public datasets (GSE104645 and GSE62322) was performed by Dr. Mireia Ramos (Dr. Lorenzo Pasquali Lab) and Dr. Lorenzo Pasquali; the study of the implication of our preferred genes in these CRC cohorts of patients by *in silico* and associative analysis (mostly Kaplan-Meier curves) were performed by myself.

Colorectal cancer patient's cohort analyzed by NanoString was provided by Dr. Eva Martinez-Balibrea's Lab. Dr. Cristina Queralt and I purified and prepared the 96 samples for sequencing; expression data analysis was performed by Dr. Cristina Queralt.

I generated CRISPR-Cas9 KO cell lines and performed the characterization of KO clones.

Discussion

Chemoresistance is one of the main obstacles to cure advanced CRC. Newer treatments, such as immune checkpoint inhibitors or targeted therapies have represented remarkable breakthroughs, for instance, for advanced melanoma. Some CRCs share key mutations with melanoma, such as in *BRAF*. Therefore, oncologists suspected that these innovative therapies would also benefit advanced CRC patients. Unfortunately, this has not been the case, leaving CRC cases with the same treatments that have been the gold-standard for the past decades, which consist on the combination of cytotoxic drugs. This fact contrasts with the huge efforts that have been made to characterize CRC tumors at the genetic, epigenetic and molecular levels, being one of the cancer types where the different layers of alterations have been put together in the consensus molecular subtypes classification.

This categorization, albeit not perfect, highlights that CRCs are highly heterogeneous, and that perhaps we should consider CRC as several different diseases, which require distinct treatments based on the characteristics of each particular tumor. On the other hand, some researchers consider that founder driver alterations should be the ones targeted in synthetic lethality combinations, getting rid of all subsequent clones that complicate downstream treatments, arguing that, otherwise, resistances are more likely to appear.

In any case, CRC needs novel therapeutic strategies, especially for tumors that do not respond to chemotherapy or that become resistant. Our approach has tried to preferentially identify "sensitizers", genes that when targeted could improve responses to two main chemotherapy regimens given in the clinics, such as FOLFOX and FOLIFIRI. Moreover, since the approach allowed to identify factors that when missing provided higher cell survival, we also explored their biomarker value in primary tumors, and in this way attempt to predict better or worst responses to chemotherapy. We were skeptical to find predictive values: identifying downregulated expression of a gene that may be present in a small subclone of a primary tumor is likely to be masked by the rest of the clones that bear normal levels of expression. Nevertheless, that particular small subclone could be the one that survives once chemotherapy selective pressure is applied. In other words, we suspected that finding a biomarker with value in terms of predicting response to treatment in a high heterogenic disease such as CRC would be a complicated task. However, chances were that in some tumors, these alterations could be present in predominant clones or appeared early along the oncogenic process. Therefore, it was worth to explore this possibility because prognostic factors that predict outcomes are major areas of research in clinical oncology. We will discuss all the data gathered in the following sections.

1. A loss-of-function screening to search for novel drug targets

We have performed a loss-of-function screening in combination with chemotherapies to identify novel targets that synergize with chemodrugs and as such, that could result in better responses to treatments. To achieve this objective, we firstly invested our efforts on setting up the different conditions to minimize false positive candidate genes. These types of screenings are noisy by nature; they have a lot of background. In this regard, one of the main concerns was to avoid the entry of more than one shRNA per cell. Therefore, we first dealt with the appropriate dilution at which double positive infections were avoided; however, we could not be 100% sure this was not occurring until we would validate the top hit genes individually.

An additional challenge was to find the correct chemotherapeutic concentrations that mimic the ones given in the clinics, but also that allowed us to find the sensitizers. We were very conscious to apply adequate concentrations of chemotherapy, and reasoned that high concentrations of chemodrugs (IC_{80}, even IC_{50}s) could be suitable to identify genes that when targeted enhance resistance (Figure 44). However, as I mentioned, we were more interested in finding genes that when targeted could synergize with the drugs already given in the clinics, so that the cells die more. Therefore, we agreed a sort of
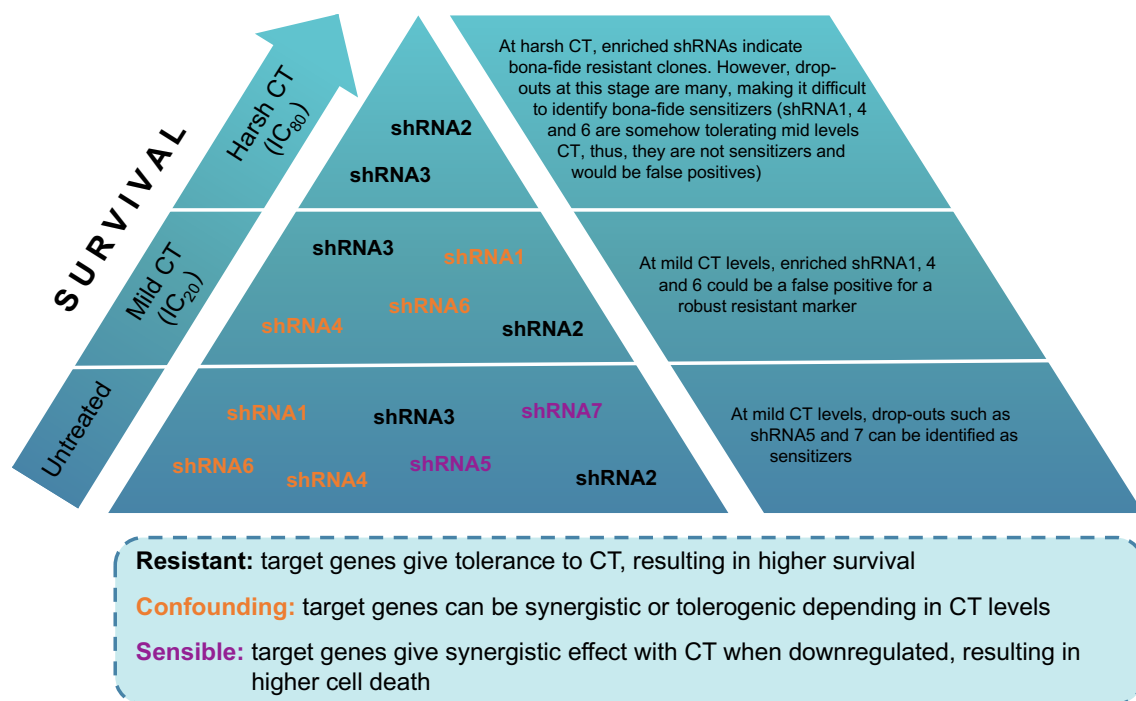
Figure 44. Representation of how chemotherapy doses affect shRNA presence or absence. At mild CT levels, drop-out shRNAs can be identified as sensitizers (in purple); at higher CT doses (IC_{80}), enriched shRNAs (in black) can be correctly identified as resistant genes. However, even at mild doses of CT, confounding shRNAs (in orange) may appear, which indicates false positive genes.

"Goldilocks concentration", which would be around $IC_{20}s$, the one that favors the identification of these sensitizers.

IC_{20} determination was a challenge: combining individual $IC_{20}s$ for 5-Fluorouracil and oxaliplatin or irinotecan resulted in higher cell death than expected by individual calculated $IC_{50}s$. Moreover, since we were trying to mimic the regimes given in the clinics, treating the cells in four consecutive cycles resulted as well in higher toxicity than expected; thus, combined IC_{20} needed to be carefully adjusted in periods of 21 days (4 cycles of chemotherapy). However, achieving IC_{20} was crucial to identify synergistic effects of "disappearing genes"; again, if cells died too much, probably the effect of the shRNA would have not been detected (Figure 44).

After the analysis of genome integrated shRNAs by NGS and obtaining the final lists of top candidate genes, we were moderately optimistic that most of them were reliable. Supporting our rationale, both enriched and drop-out lists include well-described genes involved in chemotherapy response of different cancer types, such as *ATR* in the drop-out and *HDAC4* in the enriched (212–215). This fact, indicated that part of our targets were in agreement with the work of other groups, somehow reassuring us that the screening was performed correctly. Additionally, we tried to implement another screening at IC_{80}, but cells died massively, which resulted in insufficient material to perform three different sequencing experiments as required to reach statistical significance. Nevertheless, we decided to sequence the remaining material, generating one replicate at IC_{80} treated cells. It resulted in enriched shRNAs for several genes, and some of them also appeared as hits for resistance in the IC_{20} screening (enriched shRNAs), such as *SMARCA4*, *PBRM1* or *ARID2*. Again, this concordance of hits in both screenings supported the idea that absence or downregulation of these genes' contributed to chemoresistance. For all these reasons, we argued that less-known genes could also probably be authentic targets implicated in modulating the chemotherapy response in CRC.

In summary, even if screenings need to be carefully planned and set up, we consider that they are excellent tools for discovery purposes, allowing the identification of factors involved in unexpected functions. A "Fishing Expedition" has associated negative connotations, related to science that relies in the analysis of huge amounts of data without pursuing a relevant biological question. However, we combined it with an hypothesis-driven question, the potential role of chromatin factors as facilitators or obstructors for chemotherapy. This hypothesis directed us to perform a LOF screening

focused on chromatin factors and not on the whole genome, therefore, limiting the size of the experiment, which we think benefited the quality of the data obtained. Indeed, even before the validation step, we observed that many top hit genes fell in the category of chromatin remodelers, coinciding with some of our preliminary observations where members of the SWI/SNF family appeared altered at low frequencies. This low frequency alteration rate could signify a passenger (inconsequential) role in CRC pathogenesis, but also they could contribute to respond better or worse to chemotherapy. This fact also gave us some confidence that we were on the right path.

Moreover, the methodology of a pool-approach, where you can evaluate in only one experiment the potential implication of hundreds of genes is timesaving, representing an interesting technique from which loads of information can be extracted if data is rigorously analysed. Therefore, our experience with this type of approach demonstrated its capacity to generate reliable data as we could individually validate numerous top candidate genes.

2. Validation of top candidate genes arisen in the screening

From a list of hundreds of possible candidate genes, we decided to narrow the list to 15 to 20, and individually validate them. This reduced list was delineated after going through several criteria, such as the same behavior of all shRNAs used for any candidate gene, their function in terms of DNA biology, or their implication in resistance known for other cancers. We also considered an attractive candidate when little information was available, since it could uncover a complete unprecedented player. We validated each candidate by individual downregulation and observed the expected mortality tendencies *in vitro*. Despite having promising candidates that had to be discarded during the validation process, as they resulted to be false positive genes, we were able to validate 15 out of 22 candidate genes, which is almost the 70% of the initial selected candidates. These validation numbers are higher than what we expected given that previous data from collaborators showed much lower success rates. We could demonstrate that the previously effect observed in cell viability was the result of the combination of the shRNA + chemotherapy. When there was no chemotherapy (dose 0) cells presented almost no change in viability but together with the shRNA the chemotherapeutic pressure contributed to the mortality or the survival of these cells (Figures 18, 19 and 20).

We had 15 candidate genes validated, but it was unaffordable to study all of them in detail. To select the most interesting genes, we analysed the available bibliography together with an *in silico* analysis of public databases, such as cBioPortal for Cancer Genomics (216) or the The Cancer Genome Atlas program (TCGA) (217), to identify which genes seemed to be most relevant in patients' CRC cohorts. At the end, we decided to focus our efforts on elucidating the role in modulating chemotherapeutic drugs' action of *BRIP1*, *MIS18A*, *PBRM1* and *SMARCA4* genes. This selection includes two genes that when absent or downregulated are involved in sensitizing cells to chemotherapy (*BRIP1* and *MIS18A*) and two genes that when absent or downregulated promote cell survival under the pressure of chemotherapy (*PBRM1* and *SMARCA4*).

3. BRIP1 (BRCA1 Interacting Protein C-Terminal Helicase 1)

BRIP1, also known as BACH1 or FANCJ, is a DNA-dependent ATPase and a 5'-3' helicase required for the maintenance of chromosomal stability; moreover, is implicated in the repair of DNA DSBs by HR when is associated with the BRCA1 complex, thus, appears to be highly involved in DNA damage repair (218,219). In this regard, it has been widely demonstrated that BRCA1 recognizes phosphorylated BRIP1 through BRCT motifs, which are tandem repeats implicated in DNA damage and repair responses; in the case of BRCA1, BRCT motifs seems to present an important role on its tumor suppressor function as they appeared frequently mutated in breast and ovarian cancers (220–224). Missense mutations found in these BRCT motifs have been proven to impair the binding between BRCA1 and BRIP1 (218,225). Furthermore, there is strong evidence supporting that BRIP1 is essential for BRCA1 to correctly repair DNA damage by HR: mutations on the catalytic domain of BRIP1 have been demonstrated to interfere with DBSs repair in a dependent-manner to BRCA1 binding (218,226). Also breast cancer cell lines deficient for BRIP1 failed in repairing these DSBs through HR (227). Moreover, it has been reported that BRCA1 binding to BRIP1 regulates the choice of the DNA damage repair mechanism: when the binding between BRCA1 and BRIP1 is depleted, DNA damage repair by HR is blocked and replaced by a polη-dependent bypass, which might contribute to resistance to some drugs through an interaction with the MMR protein MLH1 (227).

Nevertheless, BRIP1 has been mainly studied in breast and ovarian cancer due to its link with BRCA1. Therefore, the role of BRIP1 in CRC remains mostly unexplored. There is a study where BRIP1 is linked to hereditary nonpolyposis colon cancer (HNPCC) through MLH1, which presents several mutations in the binding region to BRIP1 that resulted in alterations on the MMR signalling pathways and apoptotic responses. However, the MMR pathway seems not to be depleted but just delayed; with this delay, the MMR-independent methylation reversal by MGMT has time to enhance DNA methylation resistance. Whether this delay in MMR signalling pathway may constitute a mechanism linked to cancer or chemoresistance needs still to be addressed (227).

In the present thesis, *BRIP1* was one of the candidate genes arisen in the screening for both FUOX and FUIRI treatments. Since the beginning, BRIP1 appeared as one of the most interesting candidates due to its clear implication in DNA damage repair and its links to BRCA1 complex, which is one of the main tumor suppressor genes involved in breast and ovarian cancers. In addition, the little evidence reported in CRC suggested a novelty aspect that we were also looking for when the screening was planned.

When *BRIP1* had to be individually validated by shRNA to confirm it was a good candidate, it turned out to be one of the genes with best validation ratios presenting more than 70% downregulation by RT-qPCR, which was translated in strong decreases on the cell viability when cells were treated at different FUIRI concentrations. Since *BRIP1* was extracted from the drop-out lists, this means that theoretically cells are more sensible to chemotherapy when is absent, thus, *BRIP1* could have a potential role as a sensitizer to FUOX and FUIRI. In this regard, at the highest doses of chemotherapy, there was a decrease on cell viability of 25-30% in comparison with the control condition; these results were accompanied by fair downregulation of BRIP1 at protein level as well. These consistent results converted *BRIP1* in an interesting hit for further analysis.

Nevertheless, targeting BRIP1 with specific drugs to perform several functional assays was impossible because there are no commercially available compounds that inhibit BRIP1. This issue led us to find techniques more robust than shRNAs to expand the study on the mechanism of action of BRIP1 in CRC; for this reason, generating a HT-29 knock-out cell line for *BRIP1* by CRISPR-Cas9 technology seemed the best option.

We successfully edited several clones for *BRIP1* in HT-29 cell line, which provided us with a crucial tool to study the implication of this gene in chemoresistance. It should be remarked that in a first WB (Annex V) a complete absence of BRIP1 protein was not observed in any clone, although clones with higher than 50% of downregulated expression showed gRNA target regions clearly edited. A subsequent WB with improved conditions indicated that perhaps a complete KO was achieved in clone 15. This clone will be further analysed in the future. Nevertheless, these selected edited clones, single-cell derived, although presenting a normal growth, showed higher residual mortality than control targeted cells or even the parental HT-29, which led us to think that a complete loss of BRIP1 may trigger apoptosis. Thus, cells able to be maintained in culture might present *BRIP1* edited in heterozygosis. However, this reasoning is not supported by *in vivo* models, since KO mice for *BRIP1* in homozygosis are not embryonically lethal according to data showed by the international mouse phenotyping consortium (228). Nevertheless, we suspected that complete absence of *BRIP1* was not well tolerated. In the future, we will also sequence clone 15 and probably use it as a bona-fide KO.

Once we obtained the confirmation of *BRIP1* clones correctly edited, the main feature that we wanted to analyse was if the mortality tendencies that we observed in the screening and validation steps were maintained (synergy with chemodrugs). For this reason, viability assays were performed in 4 edited clones and results were stunning: in

a clearly dose-dependent manner, HT-29 *BRIP1* KO cells treated with FUIRI presented around 80% less viability than control cells; clone 27 seems to present the lowest decrease in cell viability but still shows around 50% increased cell death than control cells. Thus, it seems clear that BRIP1 absence promotes a higher sensitivity to FUIRI in HT-29 cell line, which is in concordance to what we observed in the validation step when *BRIP1* was downregulated by shRNA.

As explained above, BRIP1 is involved in DNA damage repairing pathways, thus, it was of interest the analysis of DNA damage in KO clones treated at different concentrations of FUIRI. An easy approach to assess DNA damage is to measure phosphorylated levels of H2A.X variant by Western Blot. The results of this experiment in control cells were a little bit confusing: it seems to be more DNA damage in non-treated cells than in treated conditions; we believe this might be a combination of a non-efficient extraction of all histone proteins combined with the fact that cells treated with chemotherapy, perhaps because they are already compromised, resulted in lower extractions of total histones. However it may be, this issue complicates the comparison of obtained results to a control condition. Nevertheless, all BRIP1 clones showed a tendency to increase DNA damage in a FUIRI dose-dependent manner, especially at higher doses of chemotherapy.

When *BRIP1* KOs were more carefully analysed we observed that there were two subgroups. On one hand, clones 11 and 27 presented a consistent increment on DNA damage when FUIRI concentration increased. On the other hand, clones 12 and 43 presented an increase as well on DNA damage in a dose-dependent manner, and a dramatic increase at the highest FUIRI concentration in comparison with non-treated cells. These results reinforce the role of DNA damage repair already described for BRIP1 but in a colorectal cancer cell line, which has not been previously described; altogether, it appears that BRIP1 absence is correlated with an increment on DNA damage in HT-29 cells when chemotherapeutic pressure is applied.

Furthermore, we also evaluated the capacity of KO clones to grow in a colony formation assay, where cells are seeded at low density, in order to be isolated. This approach gives a hint of the possible ability of these cells to start metastatic disease by colonizing new tissues. In this regard, at simple sight there was a clear difference on the type of growth of *BRIP1* KO clones in comparison with the control condition: *BRIP1* KO clones had small colonies whereas control colonies were bigger and wider. This could be due to a faster growth of control colonies, with some of the bigger colonies containing several smaller colonies; in this way, we would be underestimating the number of control

colonies. However, to overcome this issue, we analysed not only the absolute number of colonies, but also the area covered by colonies and their intensity. With these parameters, we were able to discriminate areas covered by *BRIP1* KO clones in comparison to control clones, indicating a slower growth of *BRIP1* KO colonies. Additionally, colony intensity indicates the amount of cells forming the colony, with *BRIP1* KO clones showing fewer cells than controls. However, if we take into account when colonies were treated at different concentrations of FUIRI, surprising results appeared: the number of control colonies almost disappeared at the highest dose of FUIRI whereas the four *BRIP1* KO clones had more than 100 colonies in this condition. Moreover, the colony area and intensity, at the highest dose of FUIRI, were maintained above control ratios in all *BRIP1* KO clones.

These results are somehow puzzling since BRIP1 seemed to have a clear role on sensitizing cells to chemotherapy as supported by viability data, thus, we expected less colonies than the controls. It should be taken into account that experimental conditions are completely different between this approach and cell viability or DNA damage assays: in a colony formation assay, cells are practically isolated at the beginning of the experiment, whereas in cell viability assays cells are seeded at a confluence of around 50%. These differences on seeding conditions imply that several aspects, such as the cell-to-cell contact or the chemokines and signals that are released to the media, might be completely different when cells are single or in confluence. For this reason, we hypothesized that the absence of BRIP1 might deregulate different pathways depending on the external signals or stimuli that cancer cells receive, which could be in close relation with the fact of how cells were seeded in the different approaches tested; we are planning to check senescent markers as well as EMT markers, to evaluate whether BRIP1 may play a role in these pathways as well.

Nevertheless, the analysis of the BRIP1 KO features needs further studies to unravel whether it has mild, average, or strong synergizing effects with chemotherapy. Probably *in vivo* assays in transplanted immunodeficient mice could help us to evaluate the combinatorial effects at different timings and with different doses. For instance, we could also evaluate whether BRIP1 KO alone impairs tumor growth *in vivo* as *in vitro* colony assays indicate. If so, it could indicate *BRIP1* as a target *per se*, without necessarily be considered a combinatorial target with FOLFOX and FOLFIRI chemotherapies. Moreover, regarding the DNA damage data, aside from the control values that were unexpected, earlier time points should be also evaluated.

Regarding a possible biomarker role for *BRIP1*, our data is unclear. NanoString data for responders and non-responders to treatment showed no differences in *BRIP1* expression between these two groups. On one hand, in the cohort of FOLFIRI-treated patients analysed by NanoString we did not find any tendency on PFS or OS indicating that stratifying patients by *BRIP1* levels would not imply a better or worst outcome. However, in the GSE104645 cohort that we analysed (Figure 36), it appears a division on PFS and OS whereby patients having low *BRIP1* levels treated with FOLFOX would have a worst progression of the tumor. This is somehow the opposite tendency that we would expect since, in our hands, low *BRIP1* levels would give an advantage to FOLFOX treated patients in responding to chemotherapy indicated by survival assays on shRNAs.

Of note, in the case of patients treated with FOLFIRI, Kaplan-Meier survival curves gave somehow more expected tendencies where patients with low *BRIP1* levels have a better PFS. For patients treated with FOLFOX, if we only consider those that partially responded to chemotherapy and analysed the PFS, the obtained ratios were remarkable: patients expressing high levels of *BRIP1* showed a PFS of 2.5 months while patients with low *BRIP1* levels had a PFS of 25 months, which is ten times higher; this would be in agreement with low *BRIP1* being a good response factor, but with little value for prediction since these are all partial responders, which cannot be a priori identified. Nevertheless, this data supports what we found in the screening: that *BRIP1* low levels could synergize with chemotherapy at least for a subgroup of patients. At this point we do not know what other particularities the subgroup of better responders to FOLFOX have in common that could help identify these two groups in the general patient population. Moreover, since HT-29 cell line is *KRAS* wild-type, we analysed PFS in *KRAS* wild-type patients treated with FOLFOX: the results resembled those of PFS in FOLFOX-treated patients.

Furthermore, data from GSE62322 cohort shows that the tumors from patients responding to FOLFIRI had higher levels of *BRIP1* than the normal colon tissues (Figure 40C), which is the opposite to what is shown in the GSE104645 cohort for FOLFIRI (Figure 36E), whereby patients with higher *BRIP1* levels had worst PFS; moreover, as we were evaluating a sensitizer role of *BRIP1*, we would expect that patients responding to chemotherapy would be the ones with lower *BRIP1* expression.

Given the unclear data obtained, we cannot assess a biomarker value for *BRIP1* in terms of prognostic (prediction) to better or worst response to FOLFOX and FOLFIRI. Nevertheless, even if *BRIP1* could not be considered a good biomarker of response, it

does not mean that it could not be a good target for therapy. Preliminary data suggests that it actually might; *BRIP1* KO cells are more sensitive to DNA damage and their viability is also impaired. However, colony formation assays suggest that at higher doses of chemotherapy, the combination of targeting BRIP1 may not be as beneficial; this is important when considering combinatorial treatments, which probably need to be fine-tuned in terms of concentrations alongside monitorizations of the tumor evolution. *In vivo* experiments would probably help to answer this and other questions, such as when it would be the best moment to give the combinatorial drug treatment. In this regard, it would be better to have an inducible KO system or drugs that specifically target BRIP1, to evaluate in detail all these issues. Moreover, whether *BRIP1* could also be considered a good sensitizer for *KRAS* mutated cells, remains to be addressed.

4. MIS18A (MIS18 Kinetochore Protein A)

MIS18A is required together with MIS18B and MIS18BP1 for the recruitment of CENP-A, a histone H3 variant only present in centromeres, thus, is involved in chromosomal segregation during mitosis (229). Furthermore, it has been well stablished that MIS18A has to form an heterotetramer with MIS18B, so that MIS18BP is able to bind the complex; in this manner, this complex is active and recruits the chaperone HJURP, which is in charge of finally deposit CENP-A nucleosomes (230,231). To maintain the homeostasis in this CENP-A deposition process across cell cycle, the active complex formed by MIS18A, MIS18B and MIS18BP1 needs to be carefully regulated. In this regard, CDK1 has arisen as one of the main regulators since it has been reported to phosphorylate several MIS18BP1 residues in order to avoid the assembly of the functional complex MIS18A-MIS18B-MIS18BP1 in early G1 phase, thus restricting temporally CENP-A deposition (232,233).

More recently, other mechanisms of regulation, such as deSUMOylation, have appeared: absence of SENP6, which is responsible of SUMO chains' depolymerization, leads to poly-SUMOylation of MIS18A and MIS18BP1 and, consequently, CENP-A accumulation at centromeres is reduced (234,235). Overall, despite MIS18A seems not to play a direct role on damage repair, it has a crucial role on avoiding chromosomal instability, which has been described as one of the hallmarks of cancer (16,236,237). Nevertheless, there are few studies where the direct implication of MIS18A in cancer is assessed. There are a couple of recently published studies where MIS18A could be deregulated in MSI CRC tumors (238,239). On one hand, it was observed in deficient *Mlh1* mice where CRC was not developed, that there was a subsequent downregulation of a group of genes, especially *Mis18a*; furthermore, *Mis18a* downregulation was accompanied by *Cenpa* downregulation, which resulted in improper chromosome segregation, suggesting that these changes might be a signal of carcinogenesis in normal colon mucosa (238). On the other hand, Sun et al. reported *in vitro* that ATG5, which is a cytoplasmic factor that can be aberrantly translocated into the nucleus in the presence of DNA damaging agents such as 5-FU, would bind MIS18A in the nucleus to hypermethylate MLH1 promoter, therefore, favoring MMR deficiency in MSI CRC (239).

MIS18A arose as a drop-out gene from FUIRI-treated cells in the screening. In the validation phase we achieved a good downregulation of around 50% by RT-qPCR, which was translated in 30% less cell viability at higher doses of FUIRI in comparison to control cells; downregulation at protein levels was also satisfactory. Altogether, these data indicated that *MIS18A* could be a good potential candidate to further validate and study

its role as a sensitizer to chemotherapy; additionally, the fact that little to nothing was known about its implication in CRC represented a challenge, but would give also the novelty aspect sought. Finally, MIS18A has a role in chromosome segregation through CENPA binding at centromeric regions, and our group is also interested in studying a deregulated macrosatellite repeat in cancer, recently mapped to centromeric areas. For all these reasons, it was decided to include *MIS18A* as one of the selected genes for a deeper study; again, since compounds targeting MIS18A were not available, we generated knock-out HT-29 cell lines for *MIS18A* by CRISPR-Cas9 technology.

We successfully obtained several clones that presented clear edited characteristics: protein levels completely disappeared in at least 2 clones; and we could corroborate the editing of the targeted areas by Sanger sequencing in clones 34 and 35.

Cell viability assays were surprising: on the one hand, clone 34 presented a viability tendency similar to control cells, which means that *MIS18A* KO in this clone was not promoting more cell death in the presence of FUIRI; nevertheless, these cells presented an extremely slower growth in culture in comparison to clone 35 or controls. We hypothesized that this growth impairment might be related to a senescence/quiescent state of these cells, which has been proven to be a protective mechanism against damaging agents such as chemotherapy (240,241). On the other hand, clone 35 presented around 10% less cell viability in comparison to control cells, although it was not observed any clear dose dependent effect.

When DNA damage was measured by γ-H2A.X we obtained slightly different results for clones 34 and 35: both clones showed stable DNA damage levels that did not seem to increase with higher doses of FUIRI; however, clone 35 showed higher γ-H2A.X levels than clone 34. Although MIS18A is not directly involved in DNA damage repair pathways, it is involved in chromosomal instability, which is tightly related to genomic instability and DNA damage (242,243). The data showed that DNA damage did not increase in a dose-dependent manner when combined with FUIRI, which is different from what happens with BRIP1. Although BRIP1 and MIS18A are potentially involved in synergistic effects with chemotherapy, the mechanisms of action seem to be quite different, which is in agreement with the different known functions of these proteins.

Regarding the colony formation assay that we performed with *MIS18A* KO clones, there was an evident growth difference between clone 34 and 35; the number of colonies in clone 34 was lower than in clone 35, even at non-treated condition. Furthermore, the

different parameters analysed in this assay (number of colonies, colony area and colony intensity) clearly indicated that there were less colonies in *MIS18A* KO clones than in control cells at all time points (without and with chemotherapy). In the case of clone 34, colony area and intensity parameters presented a flat tendency near 0, probably due to the slower growth that we observed in culture. However, in clone 35 there were less colonies since colony areas and intensities were also maintained below control parameters and decreased in a dose-dependent manner. Since *MIS18A* was also a gene selected for its potential role on sensitizing cells to chemotherapy, these results are in agreement with the screening: a decreased capacity to grow and resist FUIRI when MIS18A is absent. Of note, both clones showed lower levels of colony growth even without chemotherapy, again supporting a possible individual target value independent of FOLFIRI.

When *MIS18A* implication was analysed in the cohort of CRC patients sequenced by NanoString, it appeared that no differences of *MIS18A* expression were present in responders vs. non-responders. Moreover, better PFS and OS ratios associated with extremely high *MIS18A* levels (Tercile 3) in comparison with patients expressing the lowest levels of *MIS18A* (Tercile 1), although it was not statistically significant. These results were somehow not expected according to our data of the screening for *MIS18A*, where it clearly presented a role on sensitizing cells to chemotherapy when downregulated, and as such, we would had expected better outcomes for patients with low levels of *MIS18A*.

However, as we stated at the beginning of the Discussion, finding biomarker values relevant for a subclone in a heterogeneous primary tumor such as CRC, is improbable. The contradictory data obtained along with the non-significance, indicates that this may be indeed the case. In this same line, the results of *MIS18A* analysis in the GSE104645 cohort did not give clear survival tendencies for neither PFS nor OS parameters. Of note, in this cohort, *MIS18A* expression was highest in the CMS1 subtype, which is characterized by a MSI status, high immune infiltration and *BRAF* mutations. It should be remarked that HT-29 cell line (used to perform the screening) is *BRAF*-mutated and the role of MIS18A in chromosomal instability has been widely mentioned. Both aspects are in concordance with the characteristics of CMS1, which may indicate a potential role of this gene in this subset of patients. Of note, CMS1 has been proven to present worst survival after relapse; for this reason, whether *MIS18A* could contribute to better sensitize this subset of patients under chemotherapy should be deeper explored.

5. PBRM1 (Protein Polybromo-1)

PBRM1 is an accessory subunit of the PBAF complex required for its stability, thus, is also involved in regulating transcription activation or repression through chromatin remodelling, as it was commented in the Introduction. Of note, *PBRM1* is not only strongly associated with clear cell renal cell carcinoma (ccRCC) when mutated, but also low levels have been correlated with increased expression of proinflammatory chemokines in patients with Crohn's disease (97,99,100,244). In fact, the Catalogue Of Somatic Mutations In Cancer (COSMIC) (245) considers *PBRM1* as a hallmark cancer genes since it appears truncated in several cancer types.

PBRM1 appeared in the screening as one of the enriched genes in cells treated with FUIRI at IC_{20} and IC_{80} concentrations, while it appeared in FUOX treated cells at IC_{20}. Therefore, its down-regulation may contribute to cell survival in the presence of chemotherapy, having perhaps a potential role as a biomarker for predicting response to treatment. *PBRM1* validation yielded great results: around 80% of downregulation was efficiently achieved at the RNA level together with an increment on cell viability of 25% in comparison to control cells; in addition, there was a good downregulation at protein level as well. Although PBRM1 has been well-described and characterized, there are still no drugs available to block this PBAF complex member. Nevertheless, since many SWI/SNF subunits arose from the screening, a further study of PBRM1 seemed worthwhile; for this reason, we also decided to create a *PBRM1* KO HT-29 cell line by CRISPR-Cas9 technology. As previously done with the other genes, all the process to edit *PBRM1* seemed to be done correctly; however, when the different clones had to be analysed by Western Blot to observe which ones were edited, it was impossible to successfully analyse them. This issue was not solved when other antibodies were tested and several conditions were changed, therefore, due to the lack of time, we decided to leave apart these clones temporally until this problem could be properly assessed.

Although we postponed the analysis of the function of *PBRM1* in KO clones, it was possible to perform the analysis of its involvement in patients' cohorts. Since *PBRM1* was identified as a "resistant" gene, it would be expected that lower *PBRM1* expression levels could associate with worst outcomes in patients, as these low *PBRM1* tumors would be resistant to chemotherapy. In the cohort of patients analysed by NanoString (Figure 34), this is only the case when OS was analysed by comparing expression levels of Terciles 1 + 2 vs. Tercile 3; here, a clear tendency was observed whereby patients expressing the highest *PBRM1* levels had better survival to FOLFIRI than patients expressing middle-low *PBRM1* levels, although *p*-values were above 0.05. However, the

OS of FOLFOX-treated patients from the GSE104645 cohort (Figure 38F), where *PBRM1* expression levels were stratified by high and low, did present a clear difference between patients with higher *PBRM1* levels with an OS of 110 months while low *PBRM1* expressing patients showed an OS of less than 60 months. This association was close to be statistically significant (p-value = 0.0583). Of note, this p-value becomes significant when *PBRM1* expression levels are divided in positive or negative values (p-value = 0.0102). Similarly, OS in FOLFIRI-treated patients presented exactly the same tendency; however, in this subtype of patients there were too few cases, which made this association far from being statistically significant. Interestingly, regarding patients treated with FOLFIRI in the GSE62322 cohort (Figures 42A and 42B), there was a clear tendency to decrease *PBRM1* expression levels along tumor progression (as the normal colon evolves to a malignant state, first primary tumor and, afterwards, metastases); indeed, this association was significant in non-responder patients when comparing the expression in normal vs. liver metastases (Figure 42C). There was also an association showing a decrease on *PBRM1* expression in responder patients comparing normal vs. primary tumor samples.

Taken all together these results suggest a potential biomarker role of *PBRM1*; patients with low expressing *PBRM1* tumors could have a worst OS if treated with FOLFOX but not with FOLFIRI. Additional analyses increasing the number of samples should be performed to further consolidate this data.

Nevertheless, we still need to elucidate what happens with KO clones and whether this possible value for predicting chemoresistance could be further corroborated in our model system.

6. SMARCA4 (SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4)

SMARCA4 was also introduced in the present thesis as one of the ATPases of SWI/SNF complexes, with a critical role on chromatin remodelling to modulate DNA structure and transcription, among other important functions (97–100,246); indeed, *SMARCA4* is the most frequently mutated chromatin-remodelling ATPase component in cancer, such as in small cell ovarian cancer (247,248) or in a subset of NSCLC patients (249).

SMARCA4 appeared as a resistant gene in FUIRI-treated samples in the screening. Due to the expertise in our laboratory working on SWI/SNF members, *SMARCA4* seemed an attractive candidate to further study. It was successfully validated individually, with a downregulation of 50% at RNA level as well as at protein level, which resulted in an increment on cell viability of 40% respect to control cells. Thus, it seemed clear that *SMARCA4* could act as a resistance marker when absent since cells survived more in the presence of FUIRI.

Surprisingly, despite *SMARCA4* is one of the two ATPases of the SWI/SNF complex, currently there are not commercially available specific inhibitors to target it. Foghorn Therapeutics has developed FHD-286, which inhibits both SWI/SNF ATPases (BRG1 and BRM); moreover, clinical trials are underway for uveal melanoma and AML. For this reason, with *SMARCA4* it was also decided to create a KO cell line by CRISPR-Cas9 in order to generate a tool to explore *SMARCA4* implication in resistance to CT. When *SMARCA4* KO clones had to be analysed to assure the correct edition of the gene, Western Blot images clearly revealed the clones that were knocked-out; many showed a complete disappearance of *SMARCA4* protein (BRG1). However, sequencing of DNA target regions to confirm genome editing was an issue impossible to solve. Several primers were tested for all the regions and several conditions were changed in order to amplify the DNA region of interest, but none of them worked. Nevertheless, as Western Blot images were robust, we decided to further study three KO clones (2, 5 and 30) and leave the confirmation of the edition changes for later.

Cell viability assays performed in *SMARCA4* KO clones revealed surprising results: the behavior of all three clones was similar to the control; cell viability did not decrease, which is what we would expect taking into account the increased cell survival data in *SMARCA4* shRNA analyses. Also, there was not an increment on cell survival respect to control cells. Therefore, the KO clones were not growing faster than the control condition, but they were not growing slower either; they had a similar viability tendency

as control KO. Recent studies are starting to suggest that BRG1 and BRM are mutually exclusive, yet they might replace each other function when one is absent (250). Additionally, analysis from the Connectivity Map (251) revealed an upregulation of BRM in HT-29 cell line when *SMARCA4* gene is knocked-out (data not shown); in this regard, we are planning to study BRM protein levels in our *SMARCA4* KO clones to see whether these association is corroborated in our model system. Altogether, these results may be indicating a regulatory role of BRG1 inhibiting BRM expression in normal conditions, although if BRG1 is completely absent, then BRM might be supplementing BRG1 function more efficiently than when BRG1 is just downregulated. This hypothesis could explain our results since a partial absence of BRG1 (downregulation by shRNA) gives a resistant capacity to cells whereas a complete BRG1 absence (knock-out by CRISPR-Cas9) promotes a replacement of its functions by BRM. Furthermore, this theory would also explain the results of *SMARCA4* KO clones in the colony formation assay: although *SMARCA4* KO clones present a lower number of colonies in all conditions respect to control, the reduction of this number of colonies is dose-dependent, similarly to the behavior of control clones. Intriguingly, the area covered by *SMARCA4* KO clones was higher than the control condition, which was accompanied by a higher number of cells forming each colony (colony intensity), which goes in favor of a resistant contribution when BRG1 is absent. Regarding the evaluation of DNA damage, results were unclear: in general terms, there was a moderate increase on DNA damage in a dose-response manner. However, normalization with total protein of these samples was inconclusive, with no consistent results when different clones were compared. Probably DNA damage experiments should be reconsidered in terms of time points.

Despite these shortcomings, regarding the implication of *SMARCA4* in patients' cohorts, in the NanoString analysis, patients presenting lower levels of *SMARCA4* have poorer OS ratios than patients expressing middle-high levels of *SMARCA4*; moreover, in a statistically significant association, patients with low *SMARCA4* levels had an OS of around 35 months whereas patients with middle-high *SMARCA4* levels showed an OS of almost 50 months (Figure 35E). Thus, the possible predictive role of *SMARCA4* to chemoresistance in CRC was reinforced. Furthermore, in patients from the GSE104645 cohort, despite not being statistically significant, there was a clear tendency of worst progression and outcome (PFS and OS) when *SMARCA4* expression levels were low, which agrees to observed results in our cohort of patients and with the screening findings. However, results obtained from the analysis in the GSE62322 cohort were ambiguous: there was an increase on *SMARCA4* expression in primary tumors when compared to the counterpart normal colonic tissue in both responders and non-

responders groups of patients. According to our working hypothesis, we could expect an increment in expression in responding patients, since cells should be resistant when *SMARCA4* is lower expressed; however, the fact that *SMARCA4* expression levels increased as well in the non-responding group of patients is at variance with all our previously presented results. Whether other parameters could further classify or differentiate between these subsets of patients that could explain this conflicting behavior is unclear, but it could be the case given that CRC is characterized by its heterogeneity.

7. Future perspectives

We have generated crucial tools to further study the role of our selected genes. On one hand, several functional *in vitro* assays, such as wound-healing or invasion assays, are still pending to be performed to correctly characterize KO clones. KO cells can be evaluated through RNA-seq and ChIP-seq to elucidate which partners are altered in the absence of our preferred genes. Furthermore, since these KO cells were prepared to be used in *in vivo* assays, several experiments on mice can be designed to evaluate the role of KO genes in tumor growth and metastasis in combination with chemotherapy administration.

One of the limitations that this thesis had was that all the assays, including the screening, were performed in only one CRC cell line. We are conscious that the results obtained here may not be extrapolated to other CRC cell lines, with a different mutational background. However, due to all the time required to set up the screening, together with having several types of chemotherapy to be administered and the huge amount of cells that needed to be cultured, it became practically impossible to perform this screening in more than one cell line. In the near future we are planning to validate our results in the Colo205 CRC cell line, which presents a similar background to HT-29 (MSI –, CIMP +, *KRAS* wild-type and *BRAF* mutated); moreover, we would like to test our candidate genes in additional CRC cell lines that present different characteristics such as SW48 (MSI +, CIMP +, *KRAS* and *TP53* wild-type). Of note, we have a SW48 cell line where *KRAS* has been mutated; this is a great tool to investigate the role of *KRAS* status in the same cell background when our favourite genes will be targeted and challenged with CT.

Another field that should be explored is finding drugs against these genes. KO cells are incredible tools to study the mechanism of action of our selected genes; however, this technique cannot be applied to the clinics. Since this project aimed to be translational, discovery of compounds able to target our preferred genes would represent a leap from bench to bedside. In this regard, several approaches can be considered: on one hand, we are contemplating to collaborate with 3D-modelling laboratories that could search for the best domains of our proteins to be targeted, thus, a process to design and develop a drug could start there. On the other hand, we recently discovered an interesting tool on Connectivity Map (251), which is called Repurposing (252), where identifies the best domains of proteins and compares them with a library of already available compounds to see whether one of these drugs could target your protein of n. In this manner, we could also start testing these already available compounds in our CRC cell lines to corroborate the potential role of our genes as biomarkers or sensitizers to chemotherapy.

Furthermore, this tool can compare the expression profile of HT-29 cell line with several KO genes (*BRIP1* among others), with the expression profile generated by the above mentioned library of commercial drugs; therefore, it allows to find drugs that promote the same expression profile than the KO of the preferred target, which can be tested as potential drugs that mimic the effect of having, for instance, BRIP1 absent.

Conclusions

The main conclusions that can be extracted from the present thesis are:

1. By performing a loss-of-function screening with two types of chemotherapies, FUOX and FUIRI, targeting 912 chromatin factors using a pool-approach method with an improved retroviral shRNAs library, we learned that:
    a. The use of mild concentrations of chemotherapy, such as IC_{20}s, was crucial to favor the identification of sensitizers.
    b. Appropriate shRNA controls that set upper and lower viability ratios were essential to delineate the window of possible outcome responses.

2. We achieved an individual validation of 70% of the selected candidate genes arisen from the loss-of-function screening. This efficiency is unusually high for these type of screenings since efficacies of gene silencing and off-target effects continue to create significant limitations to RNA interference-based approaches.

3. Functional assays with downregulated *BRIP1* and *MIS18A* levels indicate they can be promising targets for individual or combined therapies in CRC, but a deeper understanding of the mechanisms that play needs to be further explored, in combination with *in vivo* experiments. However, *BRIP1* and *MIS18A* do not seem to be good biomarkers for predicting response to FOLFOX or FOLFIRI, at least in the cohorts examined.

4. *PBRM1* and *SMARCA4* low levels seem to be crucial to potentiate resistant behaviors in CRC, therefore therapeutic approaches undertaken in other cancer types to individually block these SWI/SNF subunits may not be adequate for CRC.

5. A clear biomarker value for predicting response to FOILFIRI in CRC was not found in any of the genes evaluated by NanoString.

6. *PBRM1* shows a potential biomarker value for worst response to FOLFOX, suggesting that patients with low *PBRM1* expression in primary CRC tumors should not be treated with FOLFOX.

Materials and Methods

1. Cell culture

All the different cell lines that were used during the development of the present thesis with their culture conditions are shown in the following table:

Cell line	Type	Culture media	Incubator's conditions
HT-29	Human colorectal cancer cell line	McCoy's 5A medium (GlutaMAX™) + 10% FBS + 1% Penicillin-Streptomycin-Amphotericin B	37°C 5% CO_2
HT-29 EcoR	Human colorectal cancer cell line with an ecotropic receptor		
293T	Human lentiviral packaging cell line	DMEM/F12 medium + 10% FBS + 1% L-Glutamine + 1% Penicillin-Streptomycin-Amphotericin B	
Platinum-E (Plat-E)	Ecotropic retroviral packaging cell line		
C2C12	Murine myoblast cell line		

Table 4. Cell lines. Type and cell culture characteristics of all the cell lines used in this thesis.

Based on Berg et al. (253), the main characteristics of the CRC cell line that was used to perform all the experiments are:

Cell line	MSI	CIMP	*KRAS*	*BRAF*	*TP53*	*PIK3CA*	*PTEN*	CMS
HT-29	–	+	wild-type	V600E	pR273H	wild-type	wild-type	3

2. Lentiviral production and infection

2.1. Production of lentiviral supernatants

293T packaging cells were plated the day before transfection at a seeding density of 3 millions cells per 10 cm dish to obtain a 70% density at the moment of transfection. Two types of lentiviral vectors were used in this thesis: one to transduce the Eco receptor, present in a lentiviral-derived vector named pWPXLd-rtTA3-IRES-EcoRec-PGK-Puro; and another one to transduce the gRNAs and Cas9 protein, present in a lentiviral-derived vector that contains a resistance cassette to puromycin (pLentiCRISPR v2; GenScript®) (Annex VI). Both were co-transfected with the packaging vectors psPAX2 and pMD2.G in 293T cells using Lipofectamine™ 2000 or 3000 Transfection Reagents (ThermoFisher Scientific) following manufacturer's protocols. Briefly, 8 µg of each vector were transfected (Moore 1:1:1 ratio); 48 hours post-transfection, the supernatant containing lentiviruses was harvested and filtered (0.45 µm; Merck-Millipore). HT-29 cells were titrated with a range of 0.1 to 2 µg/mL of puromycin, and 0.35 µg/mL resulted as the minimum concentration at which non-transduced cells die.

2.2. Infection with EcoR lentiviral supernatant

HT-29 cells were plated the day before infection at a seeding density of 2.9 millions cells per 10 cm dish to obtain a 70% density at the moment of infection. The day after, 4 mL of supernatant containing EcoR lentiviral particles were added in the presence of 8 μg/μL of Polybrene (Merck-Sigma Aldrich). After 4 hours, fresh media up to 10 mL was added. 72 hours post-infection, HT-29 EcoR-transduced cells were placed in selection media containing 0.35 μg/mL of puromycin. Cells were selected for one week and resistant cells were further amplified in selection media.

2.3. Infection with guide RNAs / CRISPR-Cas9 lentiviral supernatants

HT-29 cells were plated the day before infection in 6-well plates at a seeding density of 500.000 cells/well to obtain a 70% density at the moment of infection. The day after, 1 mL of lentiviral supernatant diluted 40 times was added in the presence of 8 μg/μL of Polybrene (Merck-Sigma Aldrich); in this case, we infected with a pool of the three gRNAs from the same gene mixed in a 1:1:1 proportion (8.3 μL of each virus). After 4 hours, 1 mL of fresh media was added; 24 hours later, another 1 mL of fresh media was added. 72 hours post-infection, HT-29 transduced cells were placed in selection media containing 0.35 μg/mL of puromycin. Cells were selected for 10 days and resistant cells were further amplified in selection media.

3. Retroviral production and infection

17.5 μg of pMSCV-LEPG, pMSCV-LENC, pMSCV-LENC-Rpa3, pMSCV-LENC-Myc, pMSCV-LENC-Ren, pMSCV-LENC cloned with individual shRNAs vectors (Annex VI) and the retroviral hEpi9 library were separately transfected into the retroviral packaging cell line Plat-E. Plat-E cells were seed the day before transfection at a density of 2.9 million cells per 10 cm dish to obtain a 70% of confluence at the moment of transfection. Transfection was performed using Lipofectamine™ 3000 Transfection Reagent (for pMSCV-LEPG and pMSCV-LENC) or Lipofectamine™ 2000 Transfection Reagent (for pMSCV-LENC cloned with individual shRNAs and hEpi9 library) (ThermoFisher Scientific) following manufacturer's protocol. However, instead of using Opti-Mem™ Medium, it was used DMEM/F12 Medium without adding any extra compound. 60 hours post-transfection, viral supernatant was collected and filtered with 0.45 μm filters (Merck-Millipore).

For pMSCV-LEPG and pMSCV-LENC vectors, HT-29, HT-29 EcoR and C2C12 cells were seed in 6-well plates at a density of 500.000 cells/well to obtain a confluence of 70% at the moment of infection. The day after, 1 mL of viral supernatant per well was added in the presence of 8 μg/μL of Polybrene (Merck-Sigma Aldrich). After 4 hours,

fresh media up to 3 mL per well was added. 72 hours post-infection, HT-29 EcoR-infected cells were analysed by flow cytometry to monitor the percentages of GFP and mCherry-positive cells (section 15). For pMSCV-LENC vector cloned with the individual shRNAs, HT-29 EcoR cells were seed at a density of 2.9 million cells in 10 cm dishes to obtain a confluence of 70% at the moment of infection. The day after, 3 mL per plate of viral supernatant were added in the presence of 8 µg/µL of Polybrene (Merck-Sigma Aldrich). After 4 hours, fresh media up to 10 mL per plate was added. 72 hours post-infection, part of HT-29 EcoR-infected cells were analysed by flow cytometry to monitor the percentage of mCherry-positive cells (section 15). The remaining HT-29 EcoR-infected cells were seed again in 10 cm dishes and selected during 10 days by changing the media every 2-3 days and adding Geneticin (G418) at a final concentration of 600 µg/mL until at least 60% of cells were mCherry positive. At this point, dry pellets were collected for RNA and protein extraction, part of the cells were frozen in Recovery™ Cell Culture Freezing Medium (ThermoFisher Scientific), and the remaining cells were used to perform cell viability assays. In the case of infected HT-29 EcoR cells with hEpi9 library, our previous data demonstrated us that mCherry vectors usually infect HT-29 EcoR cells at a percentage of 1%. Taking into account that we had 7.300 different shRNAs in the library and we wanted to achieve a 1.000X representativeness of each one, 730 million HT-29 EcoR cells should be infected in order to obtain 7.3 million positive cells with, at least, 1.000 times each shRNA. For this reason, 440 million HT-29 EcoR cells were seed in twenty-two 15 cm dishes at a density of 20 million cells per plate. They day after, 8 mL of viral supernatant per plate were added in the presence of 8 µg/µL of Polybrene (Merck-Sigma Aldrich). After 4 hours, fresh media up to 20 mL per plate was added. 72 hours post-infection, each plate was split in three parts:

a) 1/3 of HT-29 EcoR-infected cells were analysed by flow cytometry to monitor the percentage of mCherry-positive cells (section 15). This time, the shRNA library infected at a percentage of 8.7% instead of 1%, therefore we had to assume that the probability of double-infected cells or multiple-infected cells is higher.

b) 1/3 of HT-29 EcoR-infected cells were frozen in FBS + 10% DMSO and stored at -80°C.

c) The remaining 1/3 of HT-29 EcoR-infected cells were seed again in 15 cm dishes and selected during 10 days by changing the media every 2-3 days and adding G418 at a final concentration of 600 µg/mL until at least 70% of cells were mCherry positive.

4. RNA extraction

RNA was extracted from HT-29 EcoR cells (selected and non-selected with puromycin) to synthetized afterwards the cDNA to monitor the presence of EcoR receptor in HT-29 cell line; in addition RNA was extracted from HT-29 EcoR cells infected with the different shRNAs (including a control shRNA: shRenilla) to perform afterwards the cDNA synthesis and verify the knock-down efficiency of the shRNAs by RT-qPCR. RNA extraction was performed using the Maxwell® 16 LEV simplyRNA Cells Kit (Promega), which is an automatic procedure performed by a machine that uses paramagnetic particles (PMPs) to expose samples to different purification solutions to isolate the RNA. The only change in the protocol is that 10 µL of DNAse I were added instead of 5 µL.

5. DNA extraction

DNA was extracted from selected HT-29 KO clones to further sequence the edited regions. DNA extraction was performed using the Maxwell® RSC Cultured Cells DNA Kit (Promega), which is an automatic procedure performed by a machine that uses PMPs to expose samples to different purification solutions to isolate the DNA.

6. cDNA synthesis

cDNA synthesis was performed with the SuperScript™ IV First-Strand Synthesis System of ThermoFisher Scientific. It was done following manufacturer's protocol, except in step two, on which it was not added the Ribonuclease Inhibitor; and in step four, where the reaction was incubated 10 minutes at 52°C.

7. Real-time PCR (RT-qPCR)

To monitor EcoR expression, synthetized cDNA's from HT-29 EcoR cells were amplified by real time qPCR. Reaction mix was set up following the datasheet of SYBR® Green master mix (Roche). Briefly, it included per sample 5 µL of SYBR® Green, 0.5 µL of forward and reverse primers (10 µM), 1 µL of cDNA, and 3 µL of H_2O (V_{final} = 10 µL). Two different sets of primers were used: EcoR and PUM1 (Annex VII). The PCR reaction was: 1 cycle of pre-incubation (95°C during 10 minutes), 40 cycles of amplification divided on denaturalization (95°C during 15 seconds), annealing (60°C during 25 seconds) and extension (72°C during 30 seconds), 1 cycle of melting curve (95°C during 15 seconds and 65°C during 1 minute) and the cooling (40°C during 30 seconds). To visualize the RT-qPCR products, 10 µL of each sample were run in a 1% agarose gel at 100V during 35 minutes. The correct size bands were monitored by UV light in the Molecular Imager® Gel Doc™ XR+ System with Image Lab™ Software (Bio-Rad).

For cDNA's coming from HT-29 EcoR cells infected with the different shRNAs, reaction mix was also set up following the datasheet of SYBR® Green master mix (Roche). Briefly, in the case of normalizer genes (*PUM1* and *MRPL9*), it included per sample 5 µL of SYBR® Green, 0.5 µL of forward and reverse primers (10 µM), 1 µL of cDNA, and 3 µL of H_2O. To monitor candidate genes' expression, 1 µL of pre-designed assays (10 µM) from IDT™ company was used for each reaction (Annex VII). The PCR reaction was: 1 cycle of pre-incubation (95°C during 10 minutes), 45 cycles of amplification divided on denaturalization (95°C during 15 seconds), annealing (55°C during 25 seconds) and extension (72°C during 30 seconds), 1 cycle of melting curve (95°C during 15 seconds and 65°C during 1 minute) and the cooling (40°C during 30 seconds).

8. PCR and Sanger sequencing

DNA was amplified by PCR using Phusion™ High-Fidelity DNA Polymerase (2 U/µL) (ThermoFisher Scientific) following manufacturer's protocol. Briefly, reaction mix included per sample 50 ng of DNA, 4 µL GC Buffer (5X) (except for MIS18A gRNA4, where HF Buffer (5X) was used), 0.4 µL dNTPs (10 mM), 0.5 µL of forward primer and reverse primers (10 µM) (Annex VII), 0.4 µL Phusion™ High-Fidelity DNA Polymerase (2 U/µL), and H_2O until V_{final} of 20 µL. The PCR reaction was: 1 cycle of pre-incubation (98°C during 3 minutes), 30 cycles of amplification divided on denaturalization (98°C during 10 seconds), annealing (20 seconds) and extension (72°C), 1 cycle of final extension (72°C during 5 minutes) and the hold (4°C). Annealing temperatures and extension times are specified on Table 5.

Gene	gRNA	Annealing Temperature	Extension Time
BRIP1	1	60°C	10 seconds
	3	58°C	20 seconds
	4	61°C	10 seconds
MIS18A	3	62°C	11 seconds
	4	57°C	8 seconds

Table 5. PCR conditions. Annealing temperatures and extension times used for DNA amplification of the different gRNAs target areas of *BRPI1* and *MIS18A* genes.

5 µL of PCR products correctly amplified were cleaned using ExoSAP-IT™ PCR Product Cleanup Reagent (ThermoFisher Scientific) following manufacturer's protocol. Afterwards, 1 µL of 5 µM forward primer (depending on the type of gene knocked-out) was added and samples were dried at 80°C during 20 minutes. Samples were sequenced by Sanger sequencing in the genomic unit of CCiTUB.

9. Co-infection assay of pMSCV-LEPG and pMSCV-LENC vectors

Co-infection assays of pMSCV-LEPG and pMSCV-LENC vectors were performed to observe when DP cells disappear. The transfection in Plat-E cells was performed as explained in section 3. HT-29 EcoR cells were seed in 6-well plates at a density of 500.000 cells/well to obtain a confluence of 70% at the moment of infection. The day after, cells were infected with the individual vectors and in combination (1:1 proportion) at different concentrations:

- 2X: contained the double amount of viral supernatant than a normal well, therefore we added 1 mL of viral supernatant of pMSCV-LEPG vector and 1 mL of viral supernatant of pMSCV-LENC vector.
- 1X: contained 500 µL of pMSCV-LEPG viral supernatant and 500 µL of pMSCV-LENC viral supernatant (V_{final} = 1 mL).

Also, viral supernatants of pMSCV-LEPG and pMSCV-LENC vectors were mixed in a 1:3 proportion to try to equilibrate them, as it was always observed that mCherry vector infected at lower percentage than vector containing GFP. From this stock mix, several dilutions were made to infect HT-29 EcoR cells (V_{final} = 1 mL/well):

- 1X: contained 1 mL of vector's mix in a 1:3 proportion.
- 2/3 dilution: contained 667 µL of vectors' mix + 333 µL of media.
- 1/3 dilution: contained 333 µL of vectors' mix + 667 µL of media.
- 1/6 dilution: contained 167 µL of vectors' mix + 833 µL of media.
- 1/15 dilution: contained 67 µL of vectors' mix + 933 µL of media.

In all wells was also added Polybrene (Merck-Sigma Aldrich) at a concentration of 8 µg/µL. After 4 hours, fresh media up to 3 mL was added on each well. 72 hours post-infection, cells were analysed by flow cytometry to monitor GFP and mCherry levels as explained on section 15.

10. Kill curve assays of control vectors

Control vectors that will be used in LOF screening (pMSCV-LENC-Rpa3, pMSCV-LENC-Myc and pMSCV-LENC-Ren) were tested performing kill curve assays. HT-29 EcoR cells were seed in 6-well plates at a density of 500.000 cells/well to obtain a confluence of 70% at the moment of infection. The day after, 1 mL of viral supernatant per well (produced as explained on section 3) was added in the presence of 8 µg/µL of Polybrene (Merck-Sigma Aldrich). After 4 hours, fresh media up to 3 mL per well was added. 72 hours post-infection, part of HT-29 EcoR-infected cells were analysed by flow cytometry to monitor the percentage of mCherry-positive cells; the rest were maintained in culture. Cells were split every 2-3 days in a 1/3 dilution during 17 days in total; in every split, part

of infected-cells were analysed by flow cytometry to monitor the evolution of mCherry levels (section 15).

11. Individual IC$_{50}$s of 5-Fluorouracil, Oxaliplatin and Irinotecan

Individual IC$_{50}$s of 5-FU (provided by the pharmacy of Germans Trias i Pujol Hospital), OXA (Sanofi) and SN-38 (Merck) in HT-29 EcoR cell line were established by XTT method (Roche). Briefly, HT-29 EcoR cells were seed in 96-well plates at a density of 1.500 cells/well. The day after, cells were treated with a wide range of doses from each chemotherapeutic drug (5-FU, OXA or SN-38). 24 hours post-treatment, media was changed to remove the drugs. 72 hours later, XTT reagent was added following manufacturer's protocol. After 6 hours, plates were read at a λ = 492 nm in the spectrophotometer SPECTROstar® *Nano* (BMG Labtech).

12. Combined IC$_{20}$ and IC$_{80}$ of FUOX and FUIRI treatments

To achieve the most accurate IC$_{20}$ and IC$_{80}$ of the combination of 5-FU + OXA (FUOX) and 5-FU + SN-38 (FUIRI) in HT-29 EcoR cell line, cells infected with the shRNA library were seed at a density of 12 million cells per 15 cm dish. The day after, cells were treated at different dilutions with the combination of the individual IC$_{50}$s of FUOX and FUIRI:

Chemotherapy \ Dilution		1/6	1/12	1/24	1/48	1/75	1/100
FUOX	5-Fluorouracil (µM)	1.67	0.83	0.42	0.21	0.13	0.1
	Oxaliplatin (µM)	0.33	0.17	0.083	0.042	0.026	0.02
FUIRI	5-Fluorouracil (µM)	1.67	0.83	0.42	0.21	0.13	0.1
	Irinotecan (nM)	0.92	0.46	0.23	0.11	0.073	0.055

Table 6. Determination of FUOX and FUIRI IC$_{20}$ and IC$_{80}$. Concentration of FUOX and FUIRI treatments given at different dilutions to HT-29 EcoR cells to set up IC$_{20}$ and IC$_{80}$ for the screening.

24 hours post-treatment, media was changed to remove drugs. 72 hours later, part of the cells were analysed by flow cytometry to measure cell viability (section 13) and, thus, establish the IC$_{20}$ and the IC$_{80}$. Remaining cells were seed again in a density of 12 million per 15 cm dish to start the process explained above again. Four consecutive treatments were made trying to mimic chemotherapeutic regimes given in the clinics to patients. IC$_{20}$ and IC$_{80}$ doses were determined after these four rounds of treatments (21 days in total).

13. Cell viability assays

To perform cell viability assays, HT-29 EcoR cells infected with shRNAs or HT-29 KO clones were seed in 6-well plates at a density of 500.000 cells/well. The day after, they were treated at different doses with the combination of the individual IC$_{50}$s of FUOX and

FUIRI (KO clones were only treated with FUIRI). Table 7 shows the chemotherapeutic doses for HT-29 EcoR cells:

Chemotherapy / Dilution		2X	1/6	1/50	1/100
FUOX	5-Fluorouracil (µM)	20	1.67	0.2	0.1
	Oxaliplatin (µM)	4	0.33	0.04	0.02
FUIRI	5-Fluorouracil (µM)	20	1.67	0.2	0.1
	Irinotecan (nM)	11	0.92	0.11	0.055

Table 7. Concentration of FUOX and FUIRI treatments given at different dilutions to HT-29 EcoR and HT-29 KO cell lines to perform cell viability assays.

24 hours post-treatment, media was changed to remove drugs. 72 hours later, cells were analysed by flow cytometry to measure cell viability: live cells (stained with DiOC), dead cells (stained with DAPI) and apoptotic cells (stained with both dyes). Briefly, trypsinized cells (live cells) were centrifuged at 1.200 rpm during 5 minutes together with the supernatant (where dead cells remain). The pellet was resuspended in 1 mL of PBS (1X) and live cells were stained with 5 µL of DiOC 10 µM (ThermoFisher Scientific) for 30 minutes at 37°C; then, cells were washed twice with PBS (1X) and dead cells were stained with DAPI (Merck-Sigma Aldrich) at a final concentration per sample of 3 µM in 1 mL of PBS (1X). Fluorescent levels were measured by flow cytometry as explained on section 15.

14. Chemotherapeutic treatments with FUOX and FUIRI in the screening

HT-29 EcoR cells infected with the shRNA library were divided in triplicates into three different conditions: untreated, treated with FUOX, and treated with FUIRI. Cells were seed in 15 cm dishes at a density of 12 million cells per plate and were biological and technically independent from now on. Also, control 10 cm dishes were seed (4 million cells per dish) to monitor live and dead cells by flow cytometry without manipulating the 15 cm plates that should be used for Next Generation Sequencing (NGS). The remaining cells were all frozen in Recovery™ Cell Culture Freezing Medium (ThermoFisher Scientific) and stored at -80°C. All conditions were treated four consecutive times during three weeks following this scheme: 24 hours after seeding the cells, they were treated with FUOX or FUIRI; 24 hours post-treatment, media was changed to remove the drugs and cells were left recovering for 72 hours. At this point, the process starts again by counting and seeding 12 million cells per 15 cm dish; remaining cells of each condition at the different time points were frozen in Recovery™ Cell Culture Freezing Medium (ThermoFisher Scientific) and stored at -80°C. The drug doses administered were 0.1 µM of 5-FU and 0.02 µM of OXA for FUOX treatment whereas FUIRI treatment had a

concentration of 0.1 µM of 5-FU and 0.055 nM of SN-38. Every time that a cycle of treatment was finished, control 10 cm dishes were analysed by flow cytometry to monitor live and dead cells, stained with DiOC and DAPI respectively (section 15), to assure that the treatments were acting as expected.

15. Cytometry analysis

Fluorescent levels of HT-29 EcoR cells were measured in LSRFortessa SORP Flow Cytometry (BD Biosciences) at the IGTP facilities. GFP fluorophore of pMSCV-LEPG vector was excited at a $\lambda = 488$ nm and emitted at a $\lambda = 530$ nm ±10; mCherry fluorophore of pMSCV-LENC vector was excited at a $\lambda = 532$ nm and emitted at a $\lambda = 616$ nm ±11.5. Fluorescent levels of DiOC and DAPI from the cell viability assays were measured in LSRFortessa SORP Flow Cytometry (BD Biosciences) or in the cytometry platform of CCiTUB by using the Gallios™ Flow Cytometer (Beckman Coulter). DiOC was excited at a $\lambda = 488$ nm and emitted at a $\lambda = 530$ nm, and DAPI was excited at a $\lambda = 405$ nm and emitted at a $\lambda = 450$ nm.

16. DNA extraction and preparation for Next Generation Sequencing by Solexa Technology

DNA extraction and sample preparation for NGS of the samples from the screening was performed following a confidential protocol provided by Johannes Zuber's laboratory, as samples were sequenced there. First of all, genomic DNA was isolated by a phenol extraction. Briefly, each cell pellet was resuspended in 400 µL of DNA Extraction Buffer (10 mM Tris-HCl (pH = 8), 150 mM NaCl and 10 mM EDTA). 4 µL of 10% SDS and 4 µL of Proteinase K (20 mg/mL) were added per sample and the mixture was incubated overnight at 55°C. The next day, after centrifuging at maximum speed for 3 minutes at RT, 400 µL of phenol were added per sample. Mixture was centrifuged at maximum speed for 8 minutes at RT and top 300 µL were removed from the water phase. Then, DNA precipitation was performed by adding 1 volume of NaAc 3M (pH = 5.2) and 3 volumes of 100% ethanol (at -20°C), and mixture was left 1 hour at -80°C. Samples were centrifuged at maximum speed during 30 minutes at 4°C and supernatant was immediately removed; pellets were washed with 200 µL of 70% ethanol and centrifuged again 5 minutes at 4°C. Supernatant was removed immediately and pellets were air-dried for 5 minutes; then, resuspended in 100 µL of Elution Buffer. To facilitate pellet resuspension, samples were left overnight at -80°C. The day after, 10 cycles of Freeze-Thawing were done to better resuspend the DNA, passing samples from -80°C to 55°C.

Afterwards, DNA concentration was quantified at NanoDrop™ 1000 Spectrophotometer (ThermoFisher Scientific), and final concentrations were adjusted to 0.5 µg/µL.

To verify that samples would amplify correctly and to add a barcode to each sample (untreated 1, 2 and 3; treated with FUOX 1, 2 and 3; treated with FUIRI 1, 2 and 3), a PCR test was performed. Briefly, reaction mix included per sample 5 µL of Buffer II 10X, 0.5 µL of dNTPs (25 mM), 4 µL MgCl$_2$ (25 mM), 1.5 µL of barcoded forward primer p7+Loop (10 µM), 1.5 µL of reverse primer p5+PGK (10 µM) (Annex VII), 0.5 µL of AmpliTaq Gold™ DNA Polymerase (ThermoFisher Scientific), 1 µL of template DNA (0.5 µg/µL) and 36 µL of H$_2$O (V$_{final}$ = 50 µL). The PCR reaction was: 1 cycle of pre-incubation (95°C during 10 minutes), 31 cycles of amplification divided on denaturalization (95°C during 30 seconds), annealing (52°C during 45 seconds) and extension (72°C during 1 minute), 1 cycle to finish the extension (72°C during 7 minutes) and the hold (4°C). To visualize the PCR products, 5 µL of each sample were run in a 1% agarose gel at 120V during 30 minutes. The correct size bands (340 bp) were monitored by UV light in the Molecular Imager® Gel Doc™ XR+ System with Image Lab™ Software (Bio-Rad).

To amplify the shRNA library that contains 7.300 shRNAs in order to have a representativeness of 1.000 times each shRNA, and taking into account that ideally we had 1 shRNA/cell, we needed to have on each sample 7.3 million genomes. As 1 cell contains 6 pg of DNA, we needed 43.8 µg of DNA in total; as our samples were at a concentration of 0.5 µg/µL and we could only add 1 µL of template DNA per PCR reaction, 88 PCR reactions needed to be performed in total for each sample to amplify the whole shRNA library and reach the representativeness that we want. In this way, each sample reaction was done in a 96-well plate and the reaction mix included per well the same reagents and amounts of previous step. The PCR reaction was also the same of previous step. Once the PCR reactions finished, all the PCR products from the same plate (same sample) were collected and half of them was frozen at -80°C as a back-up. The rest was column-purified with the FavorPrep™ Gel/PCR Purification Mini Kit (Favorgen Biotech Corp.) following manufacturer's protocol, except in step five, on which centrifugation time was 4.5 minutes instead of 3. Purified products were run in an agarose gel at 80V for 2 hours to assure that the shRNA library was correctly amplified. The correct size bands (340 bp) were monitored by UV light in the Molecular Imager® Gel Doc™ XR+ System with Image Lab™ Software (Bio-Rad), and they were all quickly cut. Afterwards, they were purified with the FavorPrep™ Gel/PCR Purification Mini Kit (Favorgen Biotech Corp.) following manufacturer's protocol, except in step seven, where the wash step was done twice. At the end, DNA concentration of all samples was quantified by NanoDrop™ 1000 Spectrophotometer (ThermoFisher Scientific). As DNA concentration of all samples was correct, they were deeply analysed by 2100

Bioanalyzer Instrument. It appeared the correct peak at 370 bp in all samples, therefore, they were sent to sequence by NGS to Johannes Zuber's laboratory in Vienna.

17. Analysis of NGS data

Sequenced DNA FASTA files were processed using bash commands. Firstly, read sequences were trimmed to contain the guide strand and barcode sequences, and counted; the reverse complement sequences of counted reads were aligned to the hEpi9 shRNA library hairpins. Only counts with at least one read count per million (CPM) were took into account, which is equivalent to a log-CPM value of 0. shRNA hairpin log-CPM values were normalized by the trimmed mean of M-values method (254); normalized factors were afterwards used as a scaling factor for the library sizes. At the end, R programme edgeR was used to calculate log-fold change (logFC) values, counts per millions and significance values, p-values and false discovery rates (FDR). Only genes with at least 6 out of 8 hairpins behaving in the same direction were selected and ranked according to mean, p-value and FDR values.

18. Cloning of the individual shRNAs into a retroviral backbone

After selecting the two best shRNAs that target each gene based on their p-value (Annex VIII), they were individually ordered to IDT™ company as 97 bp ultramers. To clone the individual shRNAs into the retroviral backbone pMSCV-LENC, firstly restriction sites of EcoRI and XhoI were introduced to the ultramers by PCR. Reaction mix was set up following the datasheet of Phusion™ High-Fidelity DNA Polymerase (ThermoFisher Scientific). Briefly, it included per sample 10 µL of Buffer GC (5X), 1 µL of dNTPs (10 mM), 2.5 µL of forward primer 5' miRE-XhoI (10 µM), 2.5 µL of reverse primer 5' miRE-EcoRI (10 µM) (Annex VII), 0.5 µL of Phusion™ High-Fidelity DNA Polymerase (2 U/µL), 1 µg of the 97 bp ultramer and H$_2$O until final volume (V_{final} = 50 µL). PCR reaction was: 1 cycle of pre-incubation (98°C during 30 seconds), 30 cycles of amplification divided on denaturalization (98°C during 10 seconds), annealing (72°C during 20 seconds) and extension (72°C during 10 seconds), 1 cycle to finish the extension (72°C during 5 minutes) and the hold (4°C). Successful oligo amplification was confirmed by running 2 µL of the PCR product on 2% agarose gel, which showed a single band at 131 bp. PCR products were purified with the FavorPrep™ Gel/PCR Purification Mini Kit (Favorgen Biotech Corp.) following manufacturer's protocol, except in step four, where the wash step was done twice, and in step five, where the centrifugation time was 6 minutes instead of 3. Afterwards, they were digested with EcoRI and XhoI restriction enzymes (both from ThermoFisher Scientific) during 1 hour at 37°C. Briefly, reaction mix included per sample 11 µL of Tango Buffer (2X), 0.2 µL of EcoRI (10 U/µL), 0.6 µL of XhoI (10

U/µL), 40 µL of purified PCR product and 3.2 µL of H_2O (V_{final} = 55 µL). Digested shRNAs were run in a 2% agarose gel and 110 bp bands were obtained. These bands were cut and gel purified with the FavorPrep™ Gel/PCR Purification Mini Kit (Favorgen Biotech Corp.) following manufacturer's protocol, except in step seven, where the wash step was done twice, and in step eight, where the centrifugation time was 6 minutes instead of 3. At the end, DNA concentration was measured at NanoDrop™ 1000 Spectrophotometer (ThermoFisher Scientific).

In parallel, pMSCV-LENC vector was digested with EcoRI and XhoI restriction enzymes during 1 hour at 37°C. Briefly, reaction mix included 6 µL of Tango Buffer (2X), 0.6 µL of EcoRI (10 U/µL), 1.8 µL of XhoI (10 U/µL), 3 µg of vector and H_2O until final volume (V_{final} = 30 µL). To avoid re-ligation, pMSCV-LENC vector was treated afterwards with a Calf Intestinal Phosphatase (CIP; New England Biolabs) during 30 minutes at 37°C. The reaction mix included 4 µL of CutSmart Buffer (10X), 6 U of CIP (1 µL) and 5 µL of H_2O. Digested vector was run in a 2% agarose gel and a 8 kb band was obtained. This band was cut and gel purified with the FavorPrep™ Gel/PCR Purification Mini Kit (Favorgen Biotech Corp.) following manufacturer's protocol, except in step seven, where the wash step was done twice, and in step eight, where the centrifugation time was 6 minutes instead of 3. At the end, DNA concentration was measured at NanoDrop™ 1000 Spectrophotometer (ThermoFisher Scientific).

Ligation of each individual shRNA into the backbone was done with a T4 DNA Ligase (ThermoFisher Scientific) following manufacturer's protocol. Briefly, reaction mix included per sample 2 µL of Buffer (10X), 100 ng of pMSCV-LENC vector, 5.25 µL of insert, 0.5 µL of T4 DNA Ligase (5 U/µL) and H_2O until final volume (V_{final} = 20 µL). The reactions were incubated overnight at 16°C. The next day, 5 µL of each ligation reaction were transformed into 50 µL of Stbl3 bacteria by a heat-shock method. Briefly, mixtures were incubated on ice for 30 minutes, 45 seconds at 42°C and back on ice for 2 minutes. Then, 1 mL of pre-warmed SOC media was added per sample and bacteria were left recovering for 1 hour at 37°C with shacking (300 rpm). Afterwards, 200 µL of each reaction mixture were spread in LB plates with ampicillin, that were growing overnight at 37°C. The day after, 10 colonies per each shRNA were individually pick up and put to grow in culture tubes with 5 mL of LB media + 5 µL of ampicillin (100 µg/mL) overnight at 37°C. 12 hours later, minipreps were performed to purify the plasmids using the NucleoSpin® Plasmid DNA Purification EasyPure Kit (Macherey-Nagel) following manufacturer's protocol. To verify that the shRNAs were correctly cloned into pMSCV-LENC vector, all minipreps were sent to sequence to GATC Services (Eurofins Genomics). Sequencing reactions included per sample 500 ng of plasmid, 2.5 µL of ZUB-

SEQ-SH primer at 10 µM (Annex VII) and H$_2$O until final volume (V$_{final}$ = 10 µL). The insertions were checked by CLC Sequence Viewer 6 software.

19. Total protein extraction and quantification

For HT-29 EcoR cells infected with shRNAs, total protein was extracted from dry cell pellets stored at -20°C. Pellets were resuspended in 100 µL approximately (depending on pellet size) of RIPA buffer (for a V$_{final}$ of 5 mL contained 3.88 mL PBS (1X), 500 µL Sodium deoxycholate (0.5%), 400 µL PMSF protease inhibitors (25X), 100 µL Sodium Fluoride (NaF; 2.5 M), 50 µL SDS (10%), 50 µL NP-40, 10 µL EDTA (0.5 M) and 10 µL Sodium Orthovanadate (Na$_3$VO$_4$; 2.5 M)) and homogenized with a low-intensity sonicator. For KO clones, pellets were resuspended in 60 µL approximately (depending on the size) of another type of RIPA buffer (for a V$_{final}$ of 50 mL contained 40 mL milliQ H$_2$O, 5 mL Sodium deoxycholate (10%), 500 µL SDS (10%), 500 µL NP-40, 2.5 mL HEPES (1 M), 1.5 mL NaCl (5 M), PMSF (250X; 1/250 dilution), Leupeptin (10 mg/mL; 1/1.000 dilution), Aprotinin (10 mg/mL; 1/1.000 dilution) and Iodoacetamide (86 mg/mL; 1/1.000 dilution)) and homogenized with a low-intensity sonicator.

In both cases, samples were incubated with RIPA on ice (4°C) for 15 minutes and centrifuged afterwards at highest speed (13.000 rpm) during 15 minutes. Total proteins are present in the supernatant that appears after centrifugation. The quantification of the protein concentration of these samples was done using Pierce™ BCA Protein Assay Kit (ThermoFisher Scientific) following manufacturer's protocol.

20. Western Blot

During this thesis, two type of protocols were used to perform Western Blot. On one hand, a protocol based on the NuPAGE™ system (ThermoFisher Scientific) for the electrophoresis and on the LI-COR Odissey Imaging System to block, incubate with secondary antibody and reveal the membranes. Here, protein samples were prepared by mixing 11.25 µL NuPAGE™ LDS Sample Buffer (4X), 50 µg of protein, 4.5 µL of Reductor Agent Novex™ NuPAGE™ (10X) and H$_2$O (V$_{final}$ = 45 µL), and denatured at 95°C during 5 minutes. Denatured samples were charged in NuPAGE™ 8 or 10% Bis-Tris Midi Protein Gels (1.0 mm). Electrophoresis tray has to be filled with 1X MOPS Buffer (pH = 7.7). Electrophoresis was performed at 200V during 1 hour. Dry transference was performed in PVDF membranes using the iBlot 2 Dry Blotting System (ThermoFisher Scientific) following manufacturer's protocol. Blocking of the membranes was done during 1 hour at RT in Intercept® (TBS) Blocking Buffer (LI-COR). Incubation of the membranes with the primary antibody (Table 8) was done overnight at 4°C diluting the antibody in Intercept® (TBS) Blocking Buffer (LI-COR). Membranes were washed 3

times of 5 minutes each one in Intercept® (TBS) Blocking Buffer (LI-COR), and then incubated with the secondary antibody (Table 8) at RT during 45 minutes in the dark. Membranes were washed 3 times of 5 minutes each one in Intercept® (TBS) Blocking Buffer (LI-COR), and then revealed in the LI-COR Odyssey 9120 Digital Imaging System.

On the other hand, the second protocol used was the classic. Briefly, protein samples were prepared by mixing 8.33 µL Laemmli Buffer (3X), 50 µg of protein and RIPA buffer (V_{final} = 20 µL), and denatured at 95°C during 5 minutes; in the case of histone proteins, samples were prepared by mixing 80 µL Laemmli Buffer (3X) with 160 µL of histone proteins (V_{final} = 240 µL), and denatured at 95°C during 5 minutes. Denatured samples were charged in 8 or 10% Acrylamide Gels (1.5 mm) for KO clones and 50µL/well of histone proteins' samples were charged in 12% Acrylamide Gels (1.5 mm). Electrophoresis tray has to be filled with 1X SDS-PAGE Buffer; electrophoresis was performed at 145V during 1.5 hours. Wet transference was performed in PVDF membranes during 75 minutes at 100V; the tray has to be filled with 1X Transfer Buffer. Blocking of the membranes was done during 30 minutes at RT in StartingBlock™ (TBS) Blocking Buffer (ThermoFisher Scientific). Incubation of the membranes with the primary antibody (Table 8) was done overnight at 4°C diluting the antibody in TBS-T (0.05%). Membranes were washed 3 times of 5 minutes each one in TBS-T (0.05%), and then incubated with the secondary antibody (Table 8) at RT during 1.5 hours. Membranes were washed 3 times of 5 minutes each one in TBS-T (0.05%) and then revealed in the Amersham™ Imager 680 system using ECL™ Western Blotting Reagents (Amersham™ – Sigma Aldrich (Merck)). For histone proteins, to normalize the result, total H2A.X levels had to be evaluated, which is a protein of the same size as γ-H2A.X; for this reason, membranes were stripped using the ReBlot Plus Mild Antibody Stripping Solution 10X (Merck-Millipore) following manufacturer's protocol. Membranes were blocked again during 30 minutes at RT in StartingBlock™ (TBS) Blocking Buffer (ThermoFisher Scientific) and incubated with the primary antibody (Table 8) overnight at 4°C diluting the antibody in TBS-T (0.05%). Membranes were washed 3 times of 5 minutes each one in TBS-T (0.05%), and then incubated with the secondary antibody (Table 8) at RT during 1.5 hours. Membranes were washed 3 times of 5 minutes each one in TBS-T (0.05%) and then revealed in the Amersham™ Imager 680 system using ECL™ Western Blotting Reagents (Amersham™ – Sigma Aldrich (Merck)).

To be able to plot WB results in graphs, the bands were quantified by ImageJ. Briefly, in the pipeline used, images were transformed to 8-bit (black & white). Then, a rectangle that contained the biggest band of the image was made and saved; the intensity mean

of all the bands was measured by putting this square over the band and pressing "⌘+M". This process was repeated for the normalizer images and also was made in the background part of each image, where there was no band. All the values were saved in an Excel file where firstly was created the "inverted values" column by subtracting "255 – Mean value"; afterwards, a "net values" column was created by subtracting the values of our bands of interest minus the values of the background from the same image. Lastly, net values of our bands of interest were divided into net values of normalizer bands to obtain a relative protein level quantification.

All the antibodies used in WBs are specified in Table 8:

Gene	Brand	Reference	Type	Dilution	Protocol
α–Tubulin	Sigma Aldrich (Merck)	T6074 Clone B-5-1-2	Primary	1/20.000	NuPAGE™ & LI-COR
β–Actin	Sigma Aldrich (Merck)	A3854	Primary	1/20.000	Classic
γ–H2A.X	Merck (Millipore)	O5-636 Clone JBW301	Primary	1/1.000	Classic
Anti-mouse	LI-COR	926-32210	Secondary	1/10.000	NuPAGE™ & LI-COR
Anti-rabbit	LI-COR	926-68071	Secondary	1/7.500	NuPAGE™ & LI-COR
Anti-rabbit	Bio-Rad	1721019	Secondary	1/10.000	Classic
BRIP1	Abcam	ab180853	Primary	1/1.000	Both
H2A.X	Abcam	ab11175	Primary	1/5.000	Classic
MIS18A	Cell Signaling Technology	69625	Primary	1/1.000	Both
PBRM1	Cell Signaling Technology	91894	Primary	1/1.000	NuPAGE™ & LI-COR
SMARCA4	Cell Signaling Technology	49360	Primary	1/1.000	Both

Table 8. Antibodies. Characteristics of all the antibodies used in this thesis.

21. Amplification of the CRISPR-Cas9 gRNAs

The three different DNA of gRNAs from *BRIP1*, *MIS18A*, *PBRM1* and *SMARCA4* genes, together with the control non-target gRNA (Annex III), were ordered to GenScript® company and arrived at a concentration of 0.2 µg/µL. As it was not enough to perform the lentiviral transfection (section 2), each gRNA was transformed into competent

bacteria to expand them. Briefly, 50 ng of each gRNA were transformed into 50 µL of competent bacteria by a heat-shock method: mixtures were incubated on ice for 30 minutes, 45 seconds at 42°C and back on ice for 2 minutes. Then, 1 mL of pre-warmed SOC media was added per sample and bacteria were left recovering for 1 hour at 37°C with shacking (300 rpm). Afterwards, 100 µL of each reaction mixture were spread in LB plates with ampicillin, that were growing overnight at 37°C. The day after, 2 colonies per each gRNA were individually pick up and put to grow in culture tubes with 5 mL of LB media + 5 µL of ampicillin (100 µg/mL) overnight at 37°C. 12 hours later, minipreps were performed to purify the plasmids using the QIAprep Spin Miniprep Kit (QIAGEN) following manufacturer's protocol.

22. Single-cell sorting and clonal expansion of HT-29–transduced KO cells

HT-29–transduced KO cells selected with puromycin were trypsinized to perform a single-cell sorting in the cytometry platform of CCiTUB by using the MoFlo Astrios EQ Sorter (Beckman Coulter). Three 96-well plates were seed for each gene knock-out condition; after 4 hours post-sorting, all wells were analysed in the microscope and around 50% of wells contained only one cell. To expand single-cell populations of the different KOs, media was regularly changed until it was observed an 80% of confluence by microscope. At this point, 48 clones of each KO condition were randomly selected, trypsinized and seed into 24-well plates; again, media was regularly changed until it was observed an 80% of confluence. Then, all clones (90% of them survived) were trypsinized and seed into 6-well plates. Cells were maintained in 6-well plates culture until cell pellets were collected to further extract DNA and protein; remaining cells of all clones were frozen in Recovery™ Cell Culture Freezing Medium (ThermoFisher Scientific) and stored at -80°C.

23. Treatment of KO cells with DNA damaging agents

In our case, since chemotherapy is a DNA damaging agent, we treated HT-29 KO cells at different concentrations of FUIRI to further analyse the phosphorylation levels of H2A.X. Briefly, infected HT-29 KO cells were seed in 6-well plates at a density of 500.000 cells/well. The day after, they were treated at different doses with the combination of the individual IC_{50}s of FUIRI. Table 9 shows the chemotherapeutic doses for HT-29 KO cells:

Chemotherapy / Dilution		2X	1/6
FUIRI	5-Fluorouracil (µM)	20	1.67
	Irinotecan (nM)	11	0.92

Table 9. Concentration of FUIRI given at different dilutions to HT-29 KO cells to induce DNA damage.

24 hours post-treatment, media was changed to remove drugs. 72 hours later, cells were collected in pellets to perform the acid extraction of histone proteins (section 24).

24. Acid extraction of histone proteins

To evaluate DNA damage through phosphorylated H2A.X, histone proteins have to be extracted by an acid method. Firstly, total protein extraction was performed as explained on section 19. Total proteins present in the supernatant that appeared after centrifugation were kept at -80°C; histone proteins remained in the pellet, that was resuspended in 100 µL of HCl (0.2 M) and incubated in rotation at 4°C during 15 minutes. Afterwards, samples were neutralized by adding 60 µL of Tris-HCl (1 M; pH = 8).

25. Colony formation assay

Since colony formation assays aimed to identify the capacity of isolated KO clones to expand, they have to be seed at really low density by counting cells and making several dilutions in order to accurate the number of cells seed. In this way, infected HT-29 KO cells were seed in 6-well plates at a density of 500 cells/well. The day after, they were treated at different doses with the combination of the individual IC_{50}s of FUIRI. Table 10 shows the chemotherapeutic doses for HT-29 KO cells:

Chemotherapy / Dilution		2X	1/6	1/10
FUIRI	5-Fluorouracil (µM)	20	1.67	1
	Irinotecan (nM)	11	0.92	0.55

Table 10. Concentration of FUIRI given at different dilutions to HT-29 KO cells in the colony formation assay.

24 hours post-treatment, media was changed to remove drugs. 72 hours later, media was changed again and from now on media was regularly changed every 2-3 days during 10 days. Afterwards, media was removed and colonies were washed with PBS (1X). Colonies were fixed by adding 2 mL/well of a methanol and acetic acid solution (3:1 proportion) during 10 minutes; afterwards, cells were washed with PBS (1X). Staining of colonies was performed by adding 2 mL/well of a 0.5% crystal violet solution during 10 minutes; then, cells were washed 3 times with PBS (1X).

Stained plates were scanned and colonies quantified through ImageJ. The pipeline followed included transforming the image to 8-bit (black and white). Afterwards, a circle was created surrounding the desired well and was cut by pressing "Edit → Clear

Outside"; then, the threshold between real colonies and background needed to be adjusted by pressing "Image → Adjust → Threshold". Afterwards, individual colonies were detected by pressing "Process → Binary → Watershed"; finally the number of colonies was obtained by pressing "Analyze → Analyze Particles". Colony area and intensity percentages were quantified in ImageJ using the plugin ColonyArea and following the pipeline developed by Guzmán et al. (255). Briefly, wells were selected and cropped from the image and converted to 8-bit images; then, threshold of each well was detected to eliminate background. Afterwards, colony area and intensity were measured on each well.

26. NanoString expression data analysis

Gene expression raw data from a customized panel of 25 genes (Annex IV) generated by NanoString technology in 96 CRC primary tumor samples was firstly analysed with their own software (nSolver 4.0 Analysis Software). It was followed their pipeline to assure the quality of the data and normalize it according to internal and external controls (housekeeping genes) included in the sequencing. Final expression levels of each patient were analysed by PASW® Statistics (18.0 version) through SPSS to obtain Kaplan-Meier curves of PFS and OS with their log-rank values.

Annexes

I. hEpi9 library

GENES' NAME										
ACTL6A	BRD7	CTSL2	DPF3	H2AFZ	KDM1A	MIER2	PADI1	PRDM8	SAP30L	SUPV3L1	UBE2N

GENES' NAME										
ACTL6A	BRD7	CTSL2	DPF3	H2AFZ	KDM1A	MIER2	PADI1	PRDM8	SAP30L	SUPV3L1
ACTL6B	BRD8	CUX1	DPY30	HAT1	KDM1B	MIER3	PADI2	PRDM9	SAP130	SUV39H1
ACTR5	BRD9	CXXC1	DQX1	HCFC1	KDM2A	MINA	PADI3	PRDM10	SATB1	SUV39H2
ADA	BRDT	CXXC4	DTX3L	HDAC1	KDM2B	MIS12	PADI4	PRDM11	SBNO1	SUV420H1
ADAR	BRE	CXXC5	DZIP3	HDAC2	KDM3A	MIS18A	PADI6	PRDM12	SCMH1	SUV420H2
ADPRH	BRIP1	DAXX	ECE2	HDAC3	KDM3B	MIS18BP1	PAG1	PRDM13	SCML2	SUZ12
ADPRHL2	BRPF1	DCTD	EDF1	HDAC4	KDM4A	MKRN1	PARG	PRDM14	SCML4	SVEP1
AEBP2	BRPF3	DDB1	EED	HDAC5	KDM4B	MLH3	PARP1	PRDM15	SETD1A	TADA1
AFF1	BRWD1	DDX1	EHMT1	HDAC6	KDM4C	MLL	PARP2	PRDM16	SETD1B	TADA2A
AFF4	BRWD3	DDX3X	EHMT2	HDAC7	KDM4D	MLL2	PARP3	PRKAA1	SETD2	TADA2B
AICDA	BTAF1	DDX3Y	EID1	HDAC8	KDM4E	MLL3	PARP4	PRKAA2	SETD3	TADA3
AIRE	BUB3	DDX4	EIF4A1	HDAC9	KDM5A	MLL4	PARP6	PRKCB	SETD4	TAF1
AKAP1	C6orf130	DDX5	EIF4A2	HDAC10	KDM5B	MLL5	PARP8	PRKCD	SETD5	TAF1L
AKAP8	C11orf30	DDX6	EIF4A3	HDAC11	KDM5C	MLLT6	PARP9	PRKDC	SETD6	TAF3
ALKBH2	C14orf43	DDX10	ELP2	HDGF	KDM5D	MLLT10	PARP10	PRMT1	SETD7	TAF5L
ALKBH3	C14orf93	DDX11	ELP3	HDGFL1	KDM6A	MORC1	PARP11	PRMT2	SETD8	TAF6L
ALKBH8	C14orf169	DDX17	ELP4	HDGFRP2	KDM6B	MORC2	PARP12	PRMT3	SETDB1	TAF9
APBB1	C16orf53	DDX18	EMG1	HDGFRP3	KDM8	MORC3	PARP14	PRMT5	SETDB2	TAF9B
APOBEC2	C17orf49	DDX19A	ENY2	HELLS	KEAP1	MORC4	PARP15	PRMT6	SETMAR	TAF10
APOBEC3B	C17orf96	DDX19B	EP300	HELQ	KIAA1045	MORF4L1	PARP16	PRMT7	SF3B1	TAF12
APOBEC4	C20orf20	DDX20	EP400	HFM1	KIAA1958	MORF4L2	PAX5	PRMT8	SFMBT1	TBL1X
ARID1A	CALCOCO1	DDX21	EPC1	HIF1AN	KIAA2026	MPHOSPH8	PAX6	PRMT10	SFMBT2	TBL1XR1
ARID1B	CAMK4	DDX23	EPC2	HIRA	KIF22	MSH6	PAX9	PSD3	SHPRH	TBL1Y

Additional final column (UBE2N):

UBE2N
UBE2N
UBE2Q2
UBE3A
UBE4A
UBR1
UBR2
UBR3
UBR4
UBR5
UBR7
UCHL5
UHRF1
UHRF2
UIMC1
USF2
USP3
USP7
USP12
USP13
USP16
USP17L5
USP21
USP22

ARID2	CAMKMT	DDX24	ERCC2	HJURP	KLF1	MSL2	PAXIP1	PSIP1	SIAH2	TCEA1	USP27X
ARID4A	CARM1	DDX25	ERCC3	HLTF	KLHDC3	MSL3	PBRM1	PWWP2B	SIK1	TCF19	USP35
ARID4B	CBL	DDX27	ERCC5	HMG20A	L3MBTL1	MTA1	PBX1	PYGO1	SIN3A	TCF20	USP46
ARIH2	CBX1	DDX28	ERCC6	HMG20B	L3MBTL2	MTA2	PCGF1	PYGO2	SIN3B	TCP1	USP51
ART1	CBX2	DDX31	ERCC6L	HMGA1	L3MBTL3	MTA3	PCGF2	RAD18	SIRT1	TDG	UTP3
AS3MT	CBX3	DDX39A	ERCC6L2	HMGA2	L3MBTL4	MTERF	PCGF5	RAD54L2	SIRT2	TDRD1	UTY
ASAP1	CBX4	DDX39B	ERG	HMGN1	LBR	MTF2	PCGF6	RAG1	SIRT3	TDRD3	VPS72
ASCC3	CBX5	DDX41	ESCO1	HMGN2	LCOR	MTR	PCMT1	RAG2	SIRT4	TDRD5	VRK2
ASF1A	CBX6	DDX42	ESCO2	HMGN3	LMNA	MUM1	PHC1	RAI1	SIRT5	TDRD6	WBSCR22
ASF1B	CBX7	DDX43	EYA1	HMGN4	LMNB1	MUM1L1	PHC2	RB1	SIRT6	TDRD7	WBSCR27
ASH1L	CBX8	DDX46	EYA2	HMGN5	LMNB2	MYB	PHC3	RBBP4	SIRT7	TDRD9	WDR5
ASH2L	CCDC79	DDX47	EZH1	HNRNPA0	LNX1	MYBL1	PHF1	RBBP5	SKIV2L	TDRD10	WDR61
ASMT	CCDC101	DDX49	EZH2	HNRNPA1	LRWD1	MYBL2	PHF2	RBBP7	SKIV2L2	TDRD12	WDR77
ASXL1	CDA	DDX50	FABP1	HNRNPA2B1	MACROD1	MYPOP	PHF3	RBL1	SLK	TDRKH	WDR82
ASXL2	CDC5L	DDX51	FAM175A	HOMEZ	MACROD2	MYSM1	PHF5A	RBL2	SMARCA1	TERF1	WHSC1
ASXL3	CDC6	DDX52	FAM175B	HOXC11	MAEL	NAA10	PHF6	RC3H1	SMARCA2	TET1	WHSC1L1
ATAD2	CDC34	DDX53	FANCM	HPRT1	MAP3K12	NAA15	PHF7	RCBTB1	SMARCA4	TET2	WIZ
ATAD2B	CDCA7L	DDX54	FBXL19	HR	MARCH5	NAA60	PHF8	RCOR1	SMARCA5	TET3	WNT5A
ATAT1	CDK17	DDX55	FBXO10	HSPBAP1	MBD1	NAP1L1	PHF10	RCOR2	SMARCAD1	TFPT	WRN
ATF2	CDK2AP1	DDX56	FBXO11	HUWE1	MBD2	NAP1L2	PHF11	RCOR3	SMARCAL1	TIPARP	YBX2
ATM	CDK3	DDX58	FBXO17	IFIH1	MBD3	NAP1L3	PHF12	RECQL	SMARCB1	TLE1	YEATS4
ATOH1	CDK5	DDX59	FBXO44	IGHMBP2	MBD4	NAT10	PHF13	RECQL4	SMARCC1	TLE4	YTHDC2
ATR	CDK9	DDX60	FBXW9	IKBKAP	MBD5	NAT6	PHF14	RECQL5	SMARCC2	TLK1	YWHAQ
ATRX	CDY1	DDX60L	FEN1	ING1	MBD6	NBN	PHF15	RERE	SMARCD1	TLK2	YY1
ATXN3	CDY1B	DHX8	FIZ1	ING2	MBTD1	NCOA1	PHF16	REST	SMARCD2	TMPO	ZAR1

ATXN7	CDY2A	DHX9	FKBP1A	ING3	MCM2	NCOA3	PHF17	RFWD2	SMARCD3	TNKS	ZBED5
ATXN7L3	CDY2B	DHX15	FKBP2	ING4	MCM3	NCOA6	PHF19	RFX4	SMARCE1	TNKS2	ZBTB9
AURKA	CDYL	DHX16	FKBP5	ING5	MCM4	NCOR1	PHF20	RFX5	SMN1	TNRC18	ZBTB12
AURKB	CDYL2	DHX29	FLYWCH1	INMT	MCM5	NCOR2	PHF20L1	RFXANK	SMN2	TOX4	ZBTB33
BABAM1	CECR2	DHX30	FMR1	INO80	MCM6	NDUFAF5	PHF21A	RFXAP	SMNDC1	TP53BP1	ZBTB40
BAHCC1	CHAF1A	DHX32	FOXA1	INO80B	MCM7	NEDD4L	PHF21B	RING1	SMYD1	TRDMT1	ZC3HAV1
BAHD1	CHAF1B	DHX33	FTSJ3	INO80C	MCRS1	NEK6	PHF23	RNF2	SMYD2	TRERF1	ZCWPW1
BANP	CHD1	DHX34	FUBP3	INO80D	MDC1	NEK9	PHIP	RNF8	SMYD3	TRIM5	ZCWPW2
BAP1	CHD1L	DHX35	FUS	INO80E	MEAF6	NFRKB	PHRF1	RNF17	SMYD4	TRIM13	ZFP57
BAZ1A	CHD2	DHX36	FXR1	INTS12	MECOM	NNMT	PIAS1	RNF20	SMYD5	TRIM24	ZGPAT
BAZ1B	CHD3	DHX37	FXR2	IWS1	MECP2	NOC2L	PIAS2	RNF40	SNAPC4	TRIM28	ZHX1
BAZ2A	CHD4	DHX38	G2E3	JAK2	MED1	NONO	PKN1	RNF168	SND1	TRIM33	ZHX2
BAZ2B	CHD5	DHX40	GADD45A	JARID2	MEN1	NOP2	PKNOX1	RNF217	SNRNP200	TRIM37	ZHX3
BCL6	CHD6	DHX57	GADD45B	JHDM1D	METTL2A	NR0B2	PNMT	RPA1	SOX10	TRIM66	ZKSCAN2
BCOR	CHD7	DHX58	GATAD2A	JMJD1C	METTL2B	NR2C1	POGZ	RPA3	SOX15	TRRAP	ZMYM2
BCORL1	CHD8	DIDO1	GATAD2B	JMJD4	METTL6	NSD1	POLE3	RPH3A	SP100	TSPYL2	ZMYM3
BDP1	CHD9	DMAP1	GDAP2	JMJD6	METTL7A	NSUN2	POLQ	RPS19BP1	SP110	TSSK6	ZMYND11
BLM	CHRAC1	DMTF1	GLI1	JMJD7-PLA2G4B	METTL7B	NSUN3	POLR1B	RPS6KA3	SP140	TTF1	ZMYND8
BMI1	CHUK	DNA2	GLI3	JMJD8	METTL8	NSUN4	POLR2B	RPS6KA4	SP140L	TTF2	ZNF295
BOP1	CLOCK	DNAJC1	GLYR1	JUNB	METTL10	NSUN5	PPARGC1A	RPS6KA5	SRC	UBA1	ZNF451
BPTF	COMMD3-BMI1	DNAJC2	GON4L	KANSL1	METTL11B	NSUN6	PPIB	RSF1	SRCAP	UBA7	ZNF519
BRCA1	COQ3	DNMT1	GPS2	KAT2A	METTL12	NSUN7	PPP4C	RTEL1	SSRP1	UBE2A	ZNF541
BRCA2	COQ5	DNMT3A	GSG2	KAT2B	METTL13	NTMT1	PRDM1	RUNX1	STAT5B	UBE2B	ZNF740
BRCC3	CRAMP1L	DNMT3B	GTF2B	KAT5	METTL20	NUFIP1	PRDM2	RUVBL1	STK31	UBE2C	ZNF787
BRD1	CREBBP	DNMT3L	GTF2F1	KAT6A	METTL21D	OGT	PRDM4	RUVBL2	SUB1	UBE2E1	ZRANB3

BRD2	CSNK2A1	DOT1L	GTF2H1	KAT6B	MGEA5	OIP5	PRDM5	SAP18	SUDS3	UBE2E3	ZSCAN20
BRD3	CSNK2B	DPF1	GTF3C4	KAT7	MGMT	ORC1	PRDM6	SAP25	SUPT16H	UBE2I	ZSCAN21
BRD4	CSTL1	DPF2	GTF3C5	KAT8	MIER1	ORC2	PRDM7	SAP30	SUPT7L	UBE2M	ZZZ3

Table 11. Name of 912 chromatin factor genes included in the hEpi9 shRNA library.

II. Enriched and drop-out shRNAs of the candidate genes arisen from the screening after FUOX or FUIRI treatment

GENE	logFC-sh1	logFC-sh2	logFC-sh3	logFC-sh4	logFC-sh5	logFC-sh6	logFC-sh7	logFC-sh8	mean (>0)	FDR	*p*-value mixed
				ENRICHED GENES IN FUOX SCREENING							
KDM6B	-0,42381	0,28763	-0,23698	1,48231	0,38746	0,66734	0,42997	-0,26243	0,65	0,8666	0,345
KEAP1	-0,16297	0,01132	0,68949	0,82907	0,40518	1,36189	0,52855	-0,19929	0,64	0,6052	0,003
SMARCB1	0,57347	0,83381	0,69914	0,51925	0,03444	0,19632	1,21009	0,95224	0,63	0,5546	0,053
CECR2	0,27957	-0,04805	0,22236	-0,11276	1,41146	0,28062	0,93903	-0,11362	0,63	0,6052	0,258
SMARCA4	0,54926	-0,43764	0,03271	1,34995	0,36243	-0,06512	0,50898	0,91557	0,62	0,6052	0,038
NOP2	0,37859	-0,40847	0,34841	0,15313	0,82086	-0,30799	1,43648	0,57935	0,62	0,3691	0,066
HDAC4	-0,26158	1,23513	0,13557	-0,06152	0,04883	1,63561	0,03774	-0,09809	0,62	0,8898	0,784
HPRT1	0,28522	0,61738	0,28681	1,47056	0,60763	0,22375	-1,00951	0,8376	0,62	0,5546	0,001
BUB3	-0,27636	0,38047	0,44294	0,30279	-0,16725	-0,2736	1,42813	0,52185	0,62	0,8898	0,463
UBE2M	0,4872	0,4596	0,29701	0,28229	-0,01627	1,20568	0,0381	1,25908	0,58	0,2280	0,149
KAT2A	0,74613	0,87649	0,23152	0,00516	0,08182	-0,16688	-0,0144	1,5063	0,57	0,7953	0,393
ZNF451	0,67556	0,59778	0,20069	-0,5333	0,21714	0,28771	1,02289	0,95692	0,57	0,8044	0,105
ARID1A	0,43828	0,76522	0,34902	-0,07882	0,7352	0,27521	0,09481	1,26152	0,56	0,6052	0,034
PADI4	0,59614	1,52319	0,40922	0,51037	0,5287	0,1843	0,14877	-0,19064	0,56	0,6840	0,007
PBRM1	0,49319	0,13628	0,60516	0,8983	0,30398	0,63983	0,60901	0,73338	0,55	0,3600	0,011
ARID2	0,60894	0,72365	-0,24242	0,80118	0,46187	0,13462	0,20872	0,83196	0,54	0,5546	0,006
HDAC5	1,51904	0,03566	0,42753	0,04724	-0,25702	-0,06874	0,65477	-0,40857	0,54	0,8615	0,302
RUVBL2	-0,91405	-0,31798	1,4845	0,08199	-0,07685	0,65343	0,30842	0,13654	0,53	0,8898	0,24
G2E3	1,05218	0,60098	0,80528	-0,22027	0,03361	0,27717	0,37334	-0,41975	0,52	0,3600	0,013
BAP1	-0,03349	0,20932	0,99709	-0,39057	0,07457	0,44523	0,8794	-0,47378	0,52	0,8898	0,18
TAF9B	0,35351	-0,40748	0,3543	-0,17188	0,10036	0,96269	-0,02552	0,82975	0,52	0,7811	0,123
PARP9	-1,06976	0,62664	-0,22052	0,33945	0,40864	0,53743	0,66279	-0,05704	0,51	0,8506	0,026

HDGFL1	0,87564	-0,59912	0,52638	0,20109	0,5689	0,54375	-0,39743	0,32505	0,51	0,8044	0,04
ORC1	0,36265	1,02501	0,47952	0,32811	-1,07329	0,31641	-0,25752	-0,00862	0,50	0,7973	0,045
PRDM14	0,81085	-0,0232	-0,69317	0,55416	0,0044	0,89205	-0,01369	0,24459	0,50	0,8044	0,357
RCOR3	0,46832	-0,43903	1,03949	0,37989	1,04312	0,0973	0,13534	0,2742	0,49	0,5546	0,024
DPF2	-0,35211	0,62155	-0,15893	0,0472	-0,58924	0,21782	0,60979	0,90923	0,48	0,8898	0,274
KIAA1045	0,34251	0,55715	0,13332	1,02874	0,48739	0,33164	-0,14436	-0,08808	0,48	0,6042	0,234
FTSJ3	0,31143	0,95367	0,36407	0,71034	0,0835	0,44168	-1,27988	-0,0376	0,48	0,7973	0,013
KAT7	0,20518	-0,00326	-0,34869	0,08759	0,73845	0,16097	1,18828	-0,46489	0,48	0,8898	0,529
BAHCC1	0,56121	0,71985	0,51013	0,28678	0,65143	-0,23856	-0,06671	0,10823	0,47	0,5546	0,389
RING1	-1,32622	0,18804	0,26463	0,50451	0,45989	0,0574	0,70053	1,12217	0,47	0,7973	0,002
DDX51	-0,29412	0,64835	0,10082	0,66047	0,82003	-0,32697	-0,12726	0,12373	0,47	0,8898	0,218
KIAA2026	-0,48106	0,767	-0,28826	-0,12834	0,503	0,26566	0,3075	0,49889	0,47	0,7811	0,006
SIRT6	-0,04646	0,06552	0,23713	-0,75921	-0,0196	0,50475	0,88084	0,65123	0,47	0,7953	0,213
RNF168	0,48713	0,8044	0,14864	0,44298	0,38611	0,45497	0,53073	-0,08516	0,46	0,3600	0,001
L3MBTL2	0,40342	0,1864	0,00085	0,12864	-0,14671	1,70134	-0,03093	0,35284	0,46	0,8044	0,644
INO80E	0,60735	-0,41988	0,47883	0,06009	-0,20175	0,73767	-0,05822	0,4206	0,46	0,8805	0,149
TADA2B	-0,14795	0,15517	-0,49846	0,80263	0,12225	-1,05488	0,3115	0,89452	0,46	0,8044	0,046
PIAS2	-0,81045	0,57998	0,71201	0,27276	0,13002	-0,06776	0,57651	-0,08216	0,45	0,8898	0,356
IFIH1	0,30366	-0,38814	-0,48859	0,72716	0,16373	-0,01017	0,13198	0,91313	0,45	0,7973	0,534
PHRF1	-0,17588	-0,88748	0,27483	0,8662	0,05851	0,42037	0,60139	-0,6108	0,44	0,9335	0,04
UBE2I	0,3032	-0,03655	-0,31591	0,88827	0,04455	-0,69593	0,8825	0,09359	0,44	0,8924	0,26
METTL20	0,40623	-0,06533	0,42892	0,20547	0,68298	-0,08878	0,4665	-0,56778	0,44	0,8805	0,114
OIP5	-0,0019	1,13519	0,32749	0,01844	0,57706	0,10736	-0,15371	-0,5294	0,43	0,8898	0,579
UBE3A	0,57821	0,13915	0,73838	0,46822	-0,19634	0,55816	-0,32572	0,11579	0,43	0,8613	0,049
BRD1	0,64935	0,25716	-0,25844	0,62499	0,31276	0,43699	-0,18547	0,30987	0,43	0,7811	0,07

STAT5B	-0,52719	0,91284	0,09144	0,248	0,33262	-0,03849	-0,07139	0,57283	0,43	0,7811	0,189
TRIM5	0,07308	0,54121	0,03688	-0,12912	0,0189	0,16497	1,14961	1,01677	0,43	0,3600	0,19
DDX17	-0,74429	0,05618	0,36778	-0,40258	0,27365	1,18335	0,25843	-0,13984	0,43	0,8898	0,22

Table 12. Top 50 enriched genes found in the screening after treating the cells with FUOX. Table presenting top 50 enriched genes with logFC values of the 8 different shRNAs, *p*-value mixed of all the shRNAs, FDR values above 0.5 and ordered by the mean.

DROP-OUT GENES IN FUOX SCREENING											
GENE	**logFC-sh1**	**logFC-sh2**	**logFC-sh3**	**logFC-sh4**	**logFC-sh5**	**logFC-sh6**	**logFC-sh7**	**logFC-sh8**	**mean (<0)**	**FDR**	**p-value mixed**
SIAH2	0,48231	0,42213	0,24115	-4,04107	-0,35227	-0,08729	-0,11664	-0,44355	-1,01	0,7811	0,012
ATR	-0,62199	-0,83476	-0,60357	-1,96207	-0,32012	-0,62229	0,7062	-0,69883	-0,81	0,3600	0,006
TRIM28	-0,09921	-0,83994	-0,24347	0,63611	0,06931	-1,04175	-1,2473	-1,11314	-0,76	0,7973	0,066
SETD5	-1,19564	0,06391	-0,18855	0,48197	0,12616	-0,99514	-0,76926	-0,57768	-0,75	0,6052	0,007
WBSCR27	-0,72697	-0,88425	0,10003	-1,10899	-0,09225	0,2468	0,24412	-0,79771	-0,72	0,3600	0,039
FBXL19	0,47358	-1,31121	0,1416	-1,00186	-0,28318	-0,90277	0,16751	-0,01058	-0,70	0,5546	0,113
MLLT6	-0,91667	1,23707	-0,43858	0,02748	-0,89915	-0,72841	-0,22036	0,37821	-0,64	0,9128	0,223
EDF1	-0,33263	-0,71594	-1,11052	-0,07088	-0,21579	-0,44346	-1,55756	0,20795	-0,64	0,5546	0,088
GLYR1	0,11491	-1,43639	0,27486	-0,6999	-0,26357	-0,0187	-0,83081	-0,40322	-0,61	0,7811	0,021
USP22	0,29923	-1,01827	-0,07741	-0,19348	0,94304	-0,90627	0,15229	-0,83384	-0,61	0,8898	0,002
FUS	-0,40093	0,21563	-1,54016	-0,10762	0,29395	-0,86574	-0,08547	0,05976	-0,60	0,7564	0,062
PIAS1	0,49581	-0,85565	0,26115	-1,22482	-0,10393	-0,294	-0,44499	0,66281	-0,58	0,8898	0,001
HCFC1	-0,14672	0,67571	-0,49944	-0,0101	-0,51345	0,08562	-1,68444	0,16926	-0,57	0,8658	0,101
ASF1A	-0,54177	-0,70228	-0,27652	-0,91717	0,19981	0,051	0,29822	-0,40054	-0,57	0,3600	0,186
METTL12	0,57918	0,6355	-1,68459	-0,47645	-0,75216	-0,23958	-0,13403	-0,0691	-0,56	0,8898	0,071
INO80	0,00604	0,13095	-0,37092	-0,79839	0,21963	-0,84963	-0,47655	-0,2784	-0,55	0,7953	0,149
PSIP1	0,78509	-0,59868	-0,53808	0,24783	-0,68628	-0,0714	0,02613	-0,84026	-0,55	0,3420	0,045
PARG	-0,542	-0,16038	0,19256	-0,27272	-1,01835	-0,9712	-0,37698	-0,43393	-0,54	0,5729	0,013

	-0,1062	-0,1601	0,24508	0,16602	-2,12884	-0,43101	-0,13	-0,20806	-0,53	0,7542	0,093
DHX29	-0,1062	-0,1601	0,24508	0,16602	-2,12884	-0,43101	-0,13	-0,20806	-0,53	0,7542	0,093
BRIP1	-0,47745	0,04469	-0,22362	-0,43981	-0,32318	-1,23599	0,77494	-0,44896	-0,52	0,5640	0,029
ELP4	0,34147	0,41399	-0,08872	0,01326	-1,33593	-0,74585	-0,06637	-0,37925	-0,52	0,8898	0,585
BRD4	-0,5447	-0,18508	-0,82862	-1,00074	-0,67313	-0,44485	-0,05357	-0,42986	-0,52	0,8044	0,201
MTA2	-0,40017	0,08525	-0,52248	-1,45764	-0,17653	-0,08821	0,3649	-0,4441	-0,51	0,7811	0,299
MCM6	-0,64526	-0,03855	-0,2283	0,16807	-1,1232	-0,47923	-0,55293	0,36521	-0,51	0,8044	0,243
KDM5C	0,25026	0,38657	-0,40571	-0,33101	-0,26557	-0,15548	-1,55183	-0,35418	-0,51	0,7010	0,289
DDX52	-0,64351	-0,05294	-0,17726	-0,1143	0,04284	0,38326	-1,31327	-0,76046	-0,51	0,7971	0,059
DHX8	0,20902	-0,21759	0,95901	0,11788	-0,1005	-0,58598	-0,51965	-1,12158	-0,51	0,8898	0,095
SAP130	-0,58946	-0,90733	0,03616	-0,28561	-0,60272	-0,15366	0,26359	0,56166	-0,51	0,8842	0,095
UBA1	-0,25854	-0,19373	-0,81453	-0,57757	-0,40727	0,13258	-0,78122	0,2961	-0,51	0,6052	0,12
PHF23	-0,07089	-0,9593	-0,24204	-0,99247	-0,17933	0,02303	1,0288	-0,53937	-0,50	0,6052	0,053
UBR4	0,37111	0,35436	0,21086	-0,28488	-0,01965	-0,14969	-1,3142	-0,71103	-0,50	0,8898	0,265
CCDC101	-0,84029	-0,631	-0,31716	-0,46786	0,44848	0,18875	-0,22086	1,04468	-0,50	0,8506	0,009
SETDB2	-0,21935	-0,79899	-0,42988	0,09501	-0,40097	0,06592	0,09999	-0,62699	-0,50	0,6054	0,068
DDX3Y	0,43746	-0,15306	-0,30393	-0,2108	0,00446	0,52998	-0,74679	-1,05732	-0,49	0,8805	0,093
GTF3C5	0,91238	-1,28768	-0,24317	-0,44883	-0,08708	0,35569	0,28162	-0,37371	-0,49	0,8969	0,001
CHAF1A	-0,60341	-0,16538	-0,78268	-0,42114	-0,23521	-0,62758	-0,43102	-0,6276	-0,49	0,3600	0,251
TBL1XR1	-0,41803	-0,69189	-0,26005	-0,30627	0,27729	-0,75746	-0,48507	0,28382	-0,49	0,8044	0,004
MKRN1	0,43528	-0,96587	-0,23939	-0,2408	-0,6565	0,08729	-0,32916	0,01083	-0,49	0,7811	0,252
PHF5A	0,11758	-0,73709	0,2055	0,00766	-0,05021	-0,21445	-0,93931	0	-0,49	0,8898	0,988
TRIM33	-0,39368	-0,47022	0,02395	-0,57113	0,07908	-0,81924	0,03284	-0,14679	-0,48	0,7973	0,141
PPP4C	0,28726	-0,64126	-0,57416	-0,35186	0,0574	0,08981	-0,40275	-0,41641	-0,48	0,8898	0,525
MIS18A	-0,67328	0,6313	-0,52817	-0,55299	-0,05249	-1,02061	0,11866	-0,01082	-0,47	0,8044	0,142
HIF1AN	-0,15481	-0,61041	0,05585	0,4073	-0,99963	-0,12893	0,13259	-0,45208	-0,47	0,8898	0,543

	logFC-sh1	logFC-sh2	logFC-sh3	logFC-sh4	logFC-sh5	logFC-sh6	logFC-sh7	logFC-sh8	mean	FDR	p-value
BRCA1	-0,53385	-0,6695	-0,42594	-0,28039	-0,81183	0,41769	0,18782	-0,04551	-0,46	0,3600	0,005
SF3B1	-0,58334	-0,43883	0,12915	-0,17001	0,09031	-0,63474	-0,59958	-0,33879	-0,46	0,8506	0,426
PRKDC	-0,53993	-0,39963	-0,4641	0,0576	-0,14786	0,39948	-0,3437	-0,85242	-0,46	0,8044	0,114
PARP15	-0,99671	-0,62889	-0,21496	-0,2196	0,12774	0,19647	0,63528	-0,2245	-0,46	0,8898	0,128
SND1	-1,10212	-0,20082	0,11408	-0,91461	-0,05408	-0,4443	0,10315	-0,0221	-0,46	0,8898	0,518
MLL4	0,38014	-0,57244	-0,98116	0,78359	-0,28351	-0,05508	-0,38219	0,38768	-0,45	0,9066	0,077
DDX24	-0,77254	0,32691	-0,13167	0,08026	-0,8596	-0,4691	-0,46771	-0,02007	-0,45	0,8044	0,376

Table 13. Top 50 drop-out genes found in the screening after treating the cells with FUOX. Table presenting top 50 drop-out genes with logFC values of the 8 different shRNAs, *p*-value mixed of all the shRNAs, FDR values above 0.5 and ordered by the mean.

ENRICHED GENES IN FUIRI SCREENING											
GENE	logFC-sh1	logFC-sh2	logFC-sh3	logFC-sh4	logFC-sh5	logFC-sh6	logFC-sh7	logFC-sh8	mean (>0)	FDR	*p*-value mixed
SMYD1	0,54537	4,18721	-0,29419	0,29527	0,69024	0,57676	-0,20123	-1,86043	1,26	0,8952	0,116
SMARCB1	-0,01001	1,09321	0,95368	0,74904	0,34951	0,4354	1,23557	1,0412	0,84	0,2470	0,007
RCOR3	0,69508	-0,51686	0,8228	0,97197	0,78351	-0,19789	-0,00049	0,2275	0,70	0,4915	0,027
SUV39H1	0,99028	0,06023	0,19197	0,1239	1,70381	-0,37229	0,83549	-0,24925	0,65	0,4002	0,05
PBRM1	0,44708	0,16429	0,79799	1,27098	0,58958	0,9391	0,11729	0,85457	0,65	0,2470	0,008
BRWD3	0,54844	0,45541	0,65187	0,90316	0,66813	-0,15569	-0,21964	-0,14256	0,65	0,4745	0,009
BAZ2A	0,66083	-0,2041	0,28194	-0,38018	-0,22229	0,50189	0,64654	0,96568	0,61	0,6988	0,082
HPRT1	-0,09874	0,67171	0,2649	0,91698	-0,03973	0,05011	-0,32514	1,12964	0,61	0,5908	0,32
KEAP1	-0,0322	0,126	0,63611	0,90184	0,37507	1,25155	0,27643	-0,12169	0,59	0,5908	0,106
HDAC4	-0,0006	1,72757	0,5758	-0,33115	0,37825	0,43481	0,02288	0,42137	0,59	0,3448	0,273
ATAD2	1,06041	0,09476	0,66198	-0,03225	0,79642	0,39905	0,38063	-0,2725	0,57	0,2533	0,157
SMARCD3	1,14074	-0,27446	0,40311	0,51192	0,46643	-0,15346	0,24061	-1,32434	0,55	0,8917	0,077
SOX15	0,50467	0,36196	1,10124	0,26861	1,24806	0,21213	0,14503	-0,02953	0,55	0,2470	0,063
SMARCA4	0,14917	-0,55394	0,53572	1,30897	0,07628	0,62466	0,99783	0,12448	0,55	0,4560	0,011

ING5	1,8912	0,51474	-0,51559	0,07	0,04498	-0,57622	0,14581	-0,44266	0,53	0,9127	0,091
KDM4A	0,84248	0,44931	0,86211	0,44532	0,4094	-0,23794	-0,01862	0,16079	0,53	0,4307	0,029
ARID1A	0,29543	0,68388	0,53806	-0,09468	0,70198	0,53138	0,03724	0,84462	0,52	0,3784	0,046
CDC34	-0,53244	0,04279	0,8732	0,51476	0,32221	0,17597	0,08449	1,60579	0,52	0,6978	0,417
PRMT2	0,59381	-0,43451	0,48153	0,53046	-0,6041	0,10257	0,87093	-0,9573	0,52	0,9771	0,087
ASMT	0,49645	0,1279	-0,2127	-0,39129	0,62003	-0,21941	0,92369	0,40614	0,51	0,6069	0,281
KIAA2026	-0,85507	0,45972	-0,16239	-0,41026	0,20273	0,60497	0,29636	0,98717	0,51	0,7198	0,02
NCOA3	0,26601	0,3997	-0,73609	-0,59049	0,90433	0,30202	-0,24828	0,67861	0,51	0,8363	0,085
ARID2	0,14429	0,46907	-0,51536	0,72674	0,46539	0,13998	0,69589	0,92108	0,51	0,3085	0,004
TAF6L	0,10866	0,27028	0,36022	1,37376	-0,41927	0,2461	0,63245	-0,3677	0,50	0,6945	0,282
BTAF1	-0,18325	0,33606	-0,57324	0,71766	0,28788	0,61314	-0,13302	0,53488	0,50	0,1140	0,007
KAT2A	-0,69472	0,54576	0,64337	0,15393	-0,1974	0,08327	0,10007	1,45793	0,50	0,7179	0,063
LMNB2	0,12119	1,1844	0,35848	-1,31558	0,28455	0,09119	0,10247	1,30877	0,49	0,5908	0,012
NCOR2	0,40766	1,0405	-0,75659	0,45983	0,29253	-0,027	0,26208	-0,2647	0,49	0,8896	0,145
HDAC6	0,07859	-0,01069	0,11425	0,61943	1,03762	0,29907	0,63332	0,65318	0,49	0,5939	0,164
NR2C1	-0,45944	0,19635	-0,10312	0,17368	0,06659	-0,66066	1,00833	1,00314	0,49	0,6022	0,118
CHD8	-0,21507	-0,09901	0,01786	0,2479	0,57403	1,03819	0,56688	-0,55901	0,49	0,8479	0,327
TRIM5	0,34873	0,30436	-0,13312	-0,13727	0,31519	-0,07314	1,05501	0,39784	0,48	0,5908	0,138
SMARCE1	0,30318	0,70219	-0,57589	1,06567	0,22936	-0,02298	0,36729	0,23017	0,48	0,4202	0,088
FKBP1A	0,16493	0,0347	0,10663	-0,16208	-0,27669	1,14495	-0,3239	0,94596	0,48	0,5908	0,353
KDM3B	1,31859	-0,39728	0,36881	0,37848	0,08372	-0,60424	0,2084	-0,0681	0,47	0,8708	0,066
ZC3HAV1	0,65433	0,28025	0,46658	0,47353	-0,07706	-0,28257	0,47276	-0,18279	0,47	0,6908	0,012
DDX46	0,79876	0,26951	-0,02656	0,43783	-0,4485	0,60635	0,23323	-0,44631	0,47	0,7179	0,176
JMJD4	-0,14528	-0,11316	0,58007	0,6698	0,35812	0,51418	0,22122	-0,57713	0,47	0,7271	0,277
CECR2	0,06808	0,27866	-0,01851	0,1804	0,87841	0,21795	1,18796	-0,05069	0,47	0,5908	0,437

BRD7	-0,61944	-0,12625	0,63987	0,1683	0,691	0,0015	0,55746	0,74998	0,47	0,6018	0,003
MIER2	0,36801	0,72413	0,26768	0,62395	-0,99703	0,28565	0,69443	0,30877	0,47	0,7156	0,064
KAT8	-0,05985	0,42404	-0,4352	-0,42686	0,53272	0,82509	0,37799	0,17331	0,47	0,9404	0,497
USP21	0,38628	0,68764	0,41922	0,43106	-0,31252	0,28508	-0,33574	0,57291	0,46	0,7179	0,242
PADI4	0,83342	0,35172	0,51201	0,47497	0,40548	0,37318	0,27312	-0,35472	0,46	0,2128	0,023
FLYWCH1	-0,12953	0,33709	0,2107	0,28271	0,88231	-0,19457	0,58989	-0,58689	0,46	0,5855	0,193
PARP9	-0,4489	0,47482	-0,42924	0,21749	0,06574	0,67307	0,81635	0,48873	0,46	0,7652	0,035
ZMYM3	0,62416	0,06526	0,91703	0,48554	-0,10694	0,34412	-0,00797	0,24992	0,45	0,1629	0,193
APOBEC4	-0,02639	0,32663	-0,17676	-0,56049	0,63144	0,40618	0,36105	0,50149	0,45	0,7036	0,177
VPS72	-0,71612	0,14508	-0,41738	0,18831	-0,52481	1,2278	0,49294	0,16007	0,44	0,9911	0,025
L3MBTL2	0,22965	0,52192	0,41462	0,06371	-0,18029	1,32361	-0,08104	0,09312	0,44	0,4202	0,336

Table 14. Top 50 enriched genes found in the screening after treating the cells with FUIRI. Table presenting top 50 enriched genes with logFC values of the 8 different shRNAs, *p*-value mixed of all the shRNAs, FDR values above 0.5 and ordered by the mean.

DROP-OUT GENES IN FUIRI SCREENING											
GENE	logFC-sh1	logFC-sh2	logFC-sh3	logFC-sh4	logFC-sh5	logFC-sh6	logFC-sh7	logFC-sh8	mean (<0)	FDR	*p*-value mixed
RUVBL1	-0,54572	0,15268	-0,81213	-0,73014	0,01377	0,17879	-1,56035	-1,28368	-0,99	0,4002	0,055
ATR	-0,66768	-0,76732	-1,06359	-1,92453	-0,53786	-1,05964	0,03692	-0,63681	-0,95	0,1140	0,001
SMARCA5	-0,57366	-0,89312	1,09496	-0,48137	-1,64498	0,04952	-0,5676	-0,64675	-0,80	0,1140	0,027
ATM	0,00932	-0,70773	-1,14586	0,02828	-0,5906	-0,28765	-1,07715	-0,37553	-0,70	0,3448	0,038
ATXN7L3	-0,74469	0,27246	-0,75215	-0,06497	-0,36468	0,14303	-1,22191	-0,96038	-0,68	0,2533	0,004
CARM1	-0,36296	-0,96898	0,22531	0,54814	-0,53432	0,3563	-0,51761	-0,91057	-0,66	0,3448	0,081
SAP130	-1,60208	-0,65508	-0,31908	0,2088	-0,08972	-0,5561	0,19169	0,04346	-0,64	0,4560	0,12
USP22	0,12212	-0,96032	-0,17763	-0,84159	0,37895	-0,34828	0,4494	-0,81359	-0,63	0,7198	0,024
PARP10	0,03347	-0,3918	-0,06277	-0,26546	-0,82033	-1,59889	0,1664	1,04961	-0,63	0,7028	0,022
BRWD1	-0,67061	0,03696	0,21569	0,28043	-0,00312	-0,09919	-0,29016	-1,8763	-0,59	0,8480	0,364
MBTD1	0,06487	-0,05767	-0,14675	-0,55959	-1,33037	-0,56235	-0,8718	-0,5694	-0,59	0,1629	0,108
TOX4	-0,20763	-0,91053	0,28243	-0,23968	-0,34873	-0,91066	-0,87805	0,21596	-0,58	0,4745	0,031
BRD4	-0,4506	-0,16608	-1,09211	-0,84234	-0,52142	-0,6125	-0,28127	0,27932	-0,57	0,1629	0,026
DHX38	-0,23135	0,5434	-0,32752	0,42819	-0,40029	-1,32664	-0,63491	-0,458	-0,56	0,8363	0,134

ING3	-0,55656	-1,10186	-0,73112	-0,41053	-0,16526	0,41962	-0,35445	0,20687	-0,55	0,6908	0,071
SIK1	-0,17747	-0,43451	0,06212	-0,01529	-0,02469	-2,10868	0,37814	0,29647	-0,55	0,7028	0,296
METTL7B	-0,59865	-0,87838	0,42301	-0,28909	0,31513	-0,85929	0,31	-0,1009	-0,55	0,6501	0,019
EDF1	-0,7258	-0,49866	-0,90875	0,14324	-0,14363	-0,25536	-0,72224	0,43929	-0,54	0,7304	0,363
MIS18A	-1,15195	0,02951	-0,36164	-0,50585	-0,51323	-0,95893	-0,26656	-0,00488	-0,54	0,5908	0,082
SCMH1	-0,14565	0,3653	-0,16006	-0,85828	-0,92892	0,68543	-0,30127	-0,79803	-0,53	0,3448	0,026
RECQL4	-0,43249	0,17118	-0,11682	0,04511	-1,00416	-1,02854	0,46152	-0,00189	-0,52	0,5415	0,251
ZHX3	-0,91669	0,91092	-0,36672	0,09997	-0,42638	-0,14005	-0,72173	0,13548	-0,51	0,8224	0,051
DDX21	-0,79132	-0,58911	-0,65507	0,01422	-0,40422	-0,44428	-0,62821	-0,05286	-0,51	0,6908	0,502
TBL1XR1	-0,044	-0,48919	-0,42122	-0,35323	-0,37307	-0,6397	-1,23961	0,32383	-0,51	0,5908	0,071
TADA3	-0,22155	-0,89086	1,42499	-0,49377	0,34475	-0,56502	-0,11682	-0,74936	-0,51	0,3712	0,021
PARG	-0,62877	-0,44986	-0,07213	0,08178	-0,47144	-0,21914	-0,63502	-0,97668	-0,49	0,4745	0,07
DHX15	-0,5887	-0,52299	-0,54263	0,22315	-0,28718	0,0109	0,67668	-0,51759	-0,49	0,7179	0,052
MIER3	-0,10666	0,05555	-0,43714	0,01575	-0,81432	-0,94116	0,15327	-0,15914	-0,49	0,6945	0,085
DHX8	-0,16011	-0,66477	1,01503	0,00009	-0,56208	-0,24114	-0,09341	-1,22859	-0,49	0,7028	0,147
GTF3C5	0,4664	-1,03764	-0,60301	-0,69044	-0,03414	0,10764	0,09877	-0,084	-0,49	0,6018	0,235
TRIM28	0,21726	0,27809	-0,53895	0,37345	-0,36694	-0,01795	-0,94901	-0,56647	-0,49	0,8363	0,119
GLI1	-0,48038	0,35613	0,09126	0,36775	-0,75643	-0,10259	-0,17965	-0,90605	-0,49	0,7198	0,121
TCP1	-0,26959	0,48615	-0,39581	-0,58471	-0,64239	-0,51055	0,02902	0,08604	-0,48	0,5966	0,323
WDR61	-0,47537	-0,29267	0,03298	0,04957	-0,1198	-0,52672	-0,96286	-0,49326	-0,48	0,7028	0,329
PHF8	-0,78558	0,26222	0,128	0,57831	-1,03387	-0,26002	-0,21856	-0,09225	-0,48	0,9198	0,336
DDX6	-0,49046	0,06014	-0,99768	-0,0255	-0,68754	-0,06765	-0,54361	-0,49536	-0,47	0,2533	0,136
LRWD1	-0,04684	-0,28206	0,0724	0,07058	-0,37359	0,08557	-1,36299	-0,2946	-0,47	0,7179	0,879
BAZ1A	-0,38754	0,08112	-0,04053	-0,93621	-0,55538	0,23182	0,39927	-0,42407	-0,47	0,6945	0,027
PRMT1	-0,47319	0,2898	-1,02038	-0,05326	-0,51741	-0,36433	-0,36627	0,4213	-0,47	0,8540	0,361
ASF1A	-0,5287	-0,36595	-0,11356	-1,2002	0,19568	0,28522	-0,40663	-0,15966	-0,46	0,5908	0,313
SIN3A	-0,86672	0,13111	-0,40422	-0,64244	0,10546	-0,49577	-0,05798	-0,29826	-0,46	0,6690	0,67
SETD8	-0,57151	-0,24884	0,95204	0,29961	-1,33522	-0,11627	0,97024	-0,03103	-0,46	0,9900	0,019
GTF2F1	-0,21059	-0,4051	-0,55181	-0,77496	0,6884	0,41166	0,03837	-0,33033	-0,45	0,7149	0,216
PARP8	0,06158	-0,40029	-0,00709	-0,6965	-0,61301	-1,14217	-0,21509	-0,10151	-0,45	0,6945	0,396
SF3B1	-0,96887	-0,16704	0,12996	-0,25639	0,20915	-0,33651	-0,61202	-0,37406	-0,45	0,6069	0,465
FBXL19	0,51417	-0,18562	-0,39943	-0,56126	-0,17465	-0,79948	0,40069	-0,58569	-0,45	0,6945	0,084
PRKDC	-0,44376	-0,27155	-0,24706	-0,25078	-0,72949	-0,89772	-0,30483	0,13013	-0,45	0,7028	0,296

ASH2L	-0,72244	-0,70944	-0,33372	-0,03575	-0,2098	0,00122	-0,6818	0,21896	-0,45	0,7028	0,153
HNRNPA2B1	-0,23405	-0,46721	-0,12363	-0,41857	-0,42087	0,30725	-1,1969	-0,27078	-0,45	0,1629	0,118
DTX3L	-0,4311	0,12838	-0,519	-0,33354	-0,57292	-0,33842	0,19425	0,17639	-0,44	0,6177	0,083

Table 15. Top 50 drop-out genes found in the screening after treating the cells with FUIRI. Table presenting top 50 drop-out genes with logFC values of the 8 different shRNAs, *p*-value mixed of all the shRNAs, FDR values above 0.5 and ordered by the mean.

ENRICHED SELECTED GENES IN FUOX SCREENING		
GENE	**hits (>0)**	**mean (>0)**
PBRM1	8	0,55
RNF168	7	0,46
UHRF2	6	0,40
STK31	7	0,33
G2E3	6	0,52
NOP2	6	0,62
BRD7	7	0,40
HPRT1	7	0,62
SMARCB1	8	0,63
HDAC9	6	0,41
ARID2	7	0,54
DDX43	6	0,39
RCOR3	7	0,49
ARID1A	7	0,56
SMARCA4	6	0,62
KEAP1	6	0,64
PADI4	7	0,56
ERCC6	6	0,38
FMR1	8	0,42
BRD1	6	0,43
RING1	7	0,47
FTSJ3	6	0,48
HDGFL1	6	0,51
HUWE1	7	0,37
UBE3A	6	0,43

DROP-OUT SELECTED GENES IN FUOX SCREENING		
GENE	**hits (<0)**	**mean (<0)**
ATR	7	-0,81
TRIM28	6	-0,76
EDF1	7	-0,64
GLYR1	6	-0,61
METTL12	6	-0,56
PARG	7	-0,54
DHX29	6	-0,53
BRIP1	6	-0,52
DDX52	6	-0,51
PHF23	6	-0,50
TBL1XR1	6	-0,49
BRCA1	6	-0,46
SMYD4	6	-0,44
AICDA	6	-0,44
RPS6KA5	7	-0,42
SETD2	7	-0,41
EP400	7	-0,39
PRKCD	7	-0,36
NCOR1	6	-0,36
APOBEC2	7	-0,34
FANCM	6	-0,33
SMARCA5	7	-0,33
ACTR5	6	-0,32
TOX4	6	-0,32
TDRD5	6	-0,31
PHC3	6	-0,30
SMN2	7	-0,30
PAX9	6	-0,29
MORF4L2	6	-0,29
METTL21D	6	-0,23

Table 16. Selected enriched (left) and drop-out (right) genes found in the screening after treating the cells with FUOX. Table presenting the final lists of selected genes after filtering the number of shRNAs that behave in the same direction (equal or over 6) and ordered by mean.

ENRICHED SELECTED GENES IN FUIRI SCREENING		
GENE	hits (>0)	mean (>0)
SMARCB1	7	0,84
SUV39H1	6	0,65
PBRM1	8	0,65
SOX15	7	0,55
SMARCA4	7	0,55
KDM4A	6	0,53
ARID1A	7	0,52
ARID2	7	0,51
KAT2A	6	0,50
LMNB2	7	0,49
SMARCE1	6	0,48
BRD7	6	0,47
MIER2	7	0,47
PADI4	7	0,46
PARP9	6	0,46
UBE3A	6	0,44
EYA2	6	0,44
DDX1	6	0,43
PHRF1	6	0,43
TSSK6	6	0,43
FBXO10	8	0,40
GADD45A	6	0,40
BANP	7	0,40
ERCC6	6	0,37
UBE2N	6	0,37
RNF217	6	0,34
UBR1	6	0,30
PHF14	7	0,28
SP140L	6	0,28
ZCWPW2	6	0,26

DROP-OUT SELECTED GENES IN FUIRI SCREENING		
GENE	hits (<0)	mean (<0)
ATR	7	-0,95
SMARCA5	6	-0,80
ATM	6	-0,70
ATXN7L3	6	-0,68
TOX4	6	-0,58
BRD4	7	-0,57
ING3	6	-0,55
MIS18A	7	-0,54
SCMH1	6	-0,53
TBL1XR1	7	-0,51
TADA3	6	-0,51
PARG	7	-0,49
FBXL19	6	-0,45
KDM5A	6	-0,43
PRKAA1	7	-0,42
BRCA1	6	-0,42
GLYR1	7	-0,41
DHX35	7	-0,41
PARP14	6	-0,37
CDK2AP1	6	-0,37
TRIM33	6	-0,36
MECOM	6	-0,35
BRIP1	6	-0,34
INTS12	6	-0,31
AFF4	7	-0,28
MORF4L2	6	-0,23

Table 17. Selected enriched (left) and drop-out (right) genes found in the screening after treating the cells with FUIRI. Table presenting the final lists of selected genes after filtering the number of shRNAs that behave in the same direction (equal or over 6) and ordered by mean.

ENRICHED GENES TREATED WITH FUOX AT IC₈₀

GENE	logFC-sh1	logFC-sh2	logFC-sh3	logFC-sh4	logFC-sh5	logFC-sh6	logFC-sh7	logFC-sh8	hits (>0)	mean (>0)
ARID2	5,63145	6,26952	1,8248	2,03965	0,51613	6,23487	4,03795	0,65026	8	3,40
HJURP	5,26069	4,39095	1,27668	-1,44002	3,65507	0,43953	5,31618	2,81898	7	3,31
RERE	0,64545	6,15528	2,2581	3,21122	4,32647	2,07856	1,4697	-0,80068	7	2,88
RFWD2	9,32707	3,94812	1,03768	2,14405	1,30529	-0,276	0,81019	0,77975	7	2,76
GATAD2A	0,00086	8,82174	0,42244	6,14854	1,5353	1,79632	0,39798	-0,87191	7	2,73
GON4L	0,97591	5,59363	1,35651	0,97314	3,5378	0,85462	3,81929	-5,27219	7	2,44
CBX3	4,39556	3,00437	0,058	1,78635	0,14265	0,95816	4,56848	-5,62795	7	2,13
RPS19BP1	0,84596	0,25581	0,43068	2,29244	1,75426	1,57809	3,44114	4,07344	8	1,83
FUS	2,66652	-10,86076	0,14133	4,90356	0,08248	0,66403	1,18557	2,34201	7	1,71
RPH3A	3,529	2,22737	-13,34752	2,58057	0,336	0,72256	0,09813	0,01747	7	1,36
INO80D	1,8613	0,21681	0,71661	0,09712	0,92661	1,51362	1,16131	1,92403	8	1,05

Table 18. Top enriched genes found in the screening after treating the cells with FUOX at IC₈₀. Table presenting top enriched genes with logFC values of the 8 different shRNAs, filtered by the number of shRNAs that behave in the same direction (equal or over 7) and ordered by the mean.

ENRICHED GENES TREATED WITH FUIRI AT IC80

GENE	logFC-sh1	logFC-sh2	logFC-sh3	logFC-sh4	logFC-sh5	logFC-sh6	logFC-sh7	logFC-sh8	hits (>0)	mean (>0)
SMARCA4	2,72802	4,02708	4,34175	5,79115	4,82301	0,86776	-7,58801	5,73796	7	4,05
PBRM1	5,0215	5,2503	8,33445	0,33733	-3,2072	3,2568	0,80717	5,27877	7	4,04
CBX5	3,35827	2,05548	3,97887	6,93062	-7,35581	2,65325	3,80839	4,15952	7	3,85
ARID2	2,03141	5,08165	8,4936	1,00714	7,39326	2,55822	-0,82371	0,19182	7	3,82
DPY30	-7,45579	6,05272	1,8525	1,44876	0,70635	3,88894	6,03422	4,66591	7	3,52
HUWE1	1,85465	-12,55334	4,07462	5,05236	3,11204	3,20146	2,24401	2,52664	7	3,15
TDRD7	0,96936	5,88191	1,39352	9,50717	2,01108	1,83407	0,30351	-0,06659	7	3,13
BRD7	3,50134	1,02872	4,04122	1,34442	-0,30332	3,42443	3,27541	4,47796	7	3,01

HMGN5	3,45137	2,63257	0,06938	-0,46822	2,12879	4,76368	2,55172	3,82179	7	2,77
TTF1	0,84679	-12,93675	4,91219	0,92844	2,36096	3,79339	1,9426	0,75837	7	2,22
WDR77	1,20231	1,04617	0,54798	2,41338	3,75232	1,55593	4,29117	-5,58097	7	2,12
BPTF	1,2762	0,83794	0,59038	4,91342	4,1874	-6,24106	2,0283	0,48993	7	2,05

Table 19. Top enriched genes found in the screening after treating the cells with FUIRI at IC$_{80}$. Table presenting top enriched genes with logFC values of the 8 different shRNAs, filtered by the number of shRNAs that behave in the same direction (equal or over 7) and ordered by the mean.

III. Sequences of the CRISPR-Cas9 guide RNAs

Gene	gRNA Number (GenScript®)	gRNA Sequence	gRNA Location
Control	NonTargetingControl GuideForHuman_0001	ACGGAGGCTAAGCGTCGCAA	-
BRIP1	gRNA1 / crRNA1	3'-TCACTTACGCCCTCATCTGC-5'	Exon 4
	gRNA3 / crRNA3	3'-TTCATCATAGCAAGCTGTGA-5'	Exon 2
	gRNA4 / crRNA4	3'-TGCTAAAGCAGAACAAAGTA-5'	Exon 3
MIS18A	gRNA1 / crRNA1	3'-CACCGACGCGTCTTCGCTCA-5'	Exon 1
	gRNA3 / crRNA3	5'-CCTTGAGACTTTGTGCTGCG-3'	Exon 3
	gRNA4 / crRNA4	5'-GAAGCTATCCAAACGTGAAA-3'	Exon 2
PBRM1	gRNA1 / crRNA1	5'-TAATACCATCCGAGACTATA-3'	Exon 3
	gRNA2 / crRNA2	5'-CAAACTCATTTCTTGTTCGA-3'	Exon 5
	gRNA3 / crRNA3	3'-GAAACCACTTCATAATAGTC-5'	Exon 4
SMARCA4	gRNA1 / crRNA1	3'-CTGGCCGAGGAGTTCCGCCC-5'	Exon 1
	gRNA2 / crRNA2	3'-CATCCCGGGGGGCACGCCCG-5'	Exon 4
	gRNA3 / crRNA3	3'-CCTGTTGCGGACACCGAGGG-5'	Exon 3

Table 20. CRISPR-Cas9 gRNAs. Sequences of all the gRNAs used for developing KO clones by CRISPR-Cas9 technology with the exon location on the genome.

IV. Customized panel of genes analyzed by NanoString

Type of gene	Name
Sensitizer Common	ATR
	BRCA1
	BRIP1
	MORF4L2
	PARG
Sensitizer FUIRI	MIS18A
	SMARCA5
	TRIM33
Resistant Common	ARID2
Resistant FUOX	RERE
Resistant FUIRI	BRD7
	DPY30
	HUWE1
	PBRM1
	SMARCA4
	TDRD7
Immune Response	CXCL9
	CXCL10
	CXCL16
	HLA-DRA
	IDO1
	IFNG
	NTE5
	STAT1
Others	PARP1
Housekeeping	ACTB
	HPRT1
	TUBB

Table 21. **Genes of the customized panel analyzed by NanoString.** In this panel were included sensitizer candidate genes common for both chemotherapies and only specific of FUIRI, resistant candidate genes common in both chemotherapies and only specific of FUOX and FUIRI, several genes related with the immune response, and PARP1 since its link with PARG. Three housekeeping genes were included: *ACTB*, *HPRT1* and *TUBB*. In total, 28 genes were analyzed.

V. Western Blot of BRIP1 KO clones.
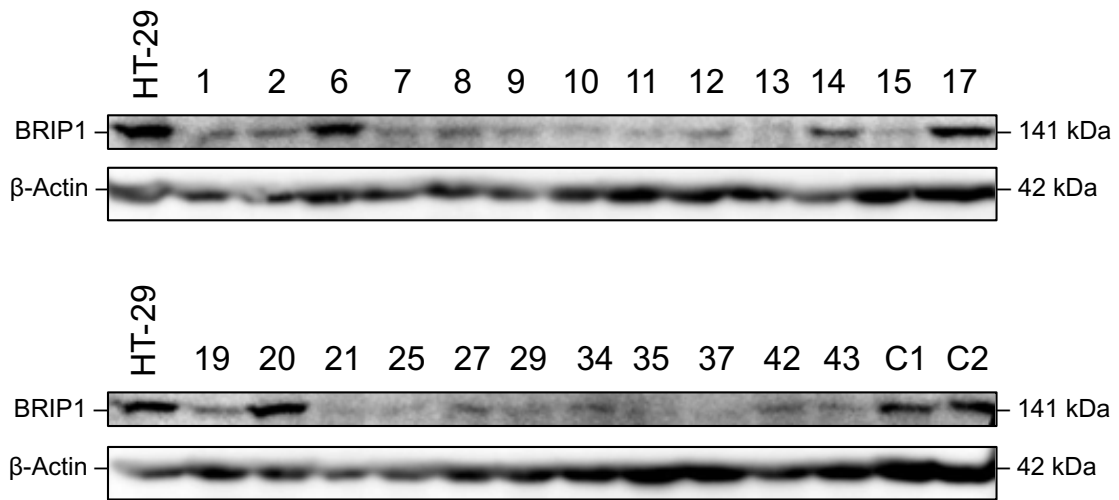
Figure 45. Western Blot of some BRIP1 KO clones. Protein expression of parental HT-29 cell line, two control gRNAs (C1 and C2) and some expanded *BRIP1* KO clones.

VI. Vectors.
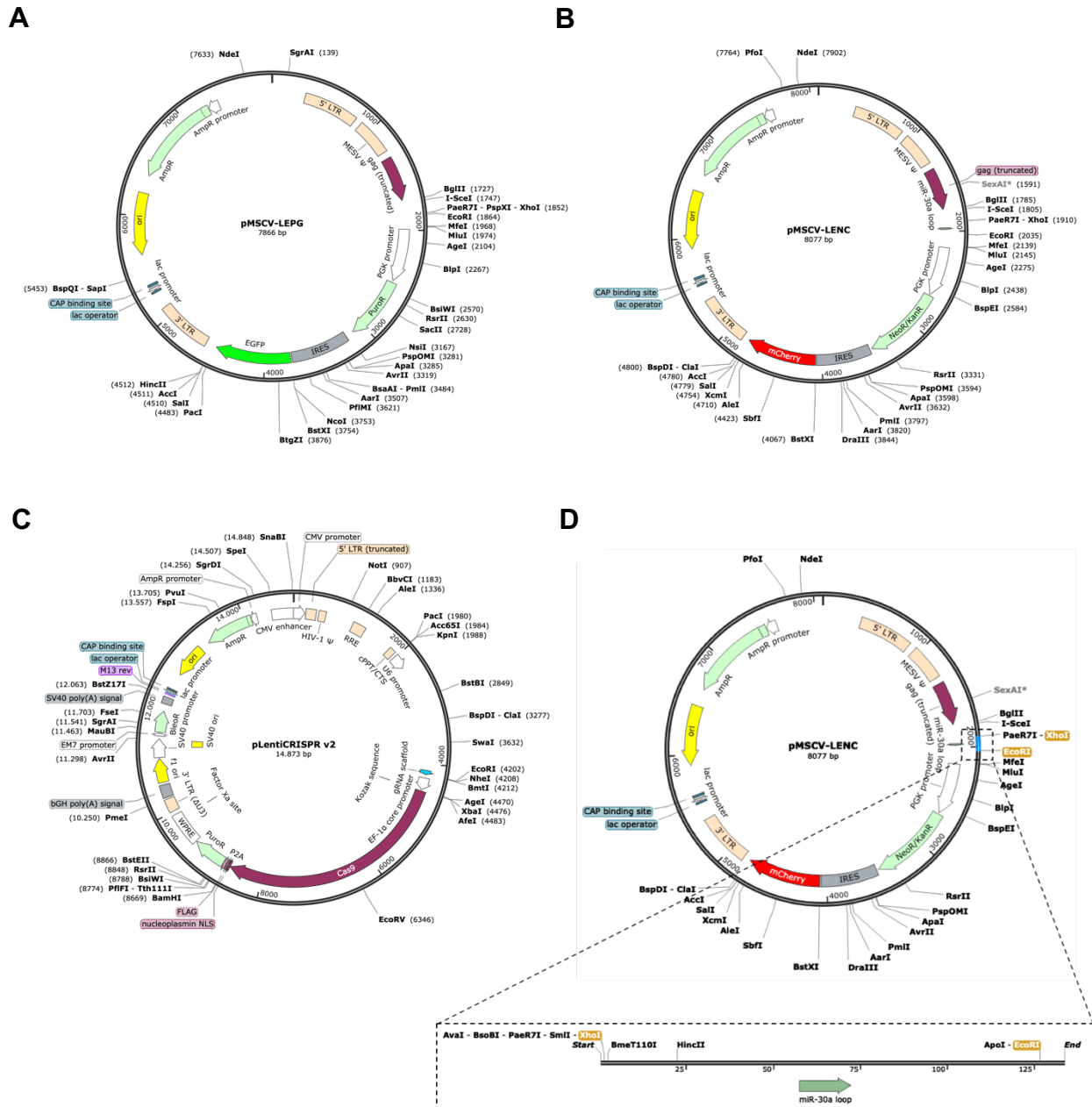
Figure 46. SnapGene circular maps of all the vectors used in this thesis. (A) pMSCV-LEPG vector used to monitor EcoR functionality and in co-infection assays. **(B)** pMSCV-LENC vector used to monitor EcoR functionality, in co-infection assays, and in kill curve assays. **(C)** pLentiCRISPR v2 backbone where all the different gRNAs were cloned; used to generate KO cell lines. **(D)** pMSCV-LENC vector indicating the area were all the individual shRNAs were cloned in the validation step; also, shRNAs of control vectors were cloned in this area (pMSCV-LENC-Rpa3, pMSCV-LENC-Myc and pMSCV-LENC-Ren).

VII. Sequences of primers

EcoR RECEPTOR		
Name	**Sense**	**Sequence (5' - 3')**
EcoR receptor	Forward	GGG TTT ATG CCC TTT GGA TT
	Reverse	CAC GCC AAA GTA CGC TAT GA

CLONING OF INDIVIDUAL shRNAs		
Name	**Sense**	**Sequence (5' - 3')**
miRE-XhoI	Forward	TAC AAT ACT CGA GAA GGT ATA TTG CTG TTG ACA GTG AGC G
miRE-EcoRI	Reverse	TTA GAT GAA TTC TAG CCC CTT GAA GTC CGA GGC AGT AGG CA
ZUB-SEQ-SH	Forward	TGT TTG AAT GAG GCT TCA GTA C

DNA PREPARATION FOR NGS		
Name	**Sense**	**Sequence (5' - 3')**
p7+Loop+ Untreated 1	Forward	CAA GCA GAA GAC GGC ATA CGA CAG ATA GTG AAG CCA CAG ATG T
p7+Loop+ Untreated 2	Forward	CAA GCA GAA GAC GGC ATA CGA ATC CTA GTG AAG CCA CAG ATG T
p7+Loop+ Untreated 3	Forward	CAA GCA GAA GAC GGC ATA CGA AAT GTA GTG AAG CCA CAG ATG T
p7+Loop+ FUOX 1	Forward	CAA GCA GAA GAC GGC ATA CGA ACT CTA GTG AAG CCA CAG ATG T
p7+Loop+ FUOX 2	Forward	CAA GCA GAA GAC GGC ATA CGA TGT TTA GTG AAG CCA CAG ATG T
p7+Loop+ FUOX 3	Forward	CAA GCA GAA GAC GGC ATA CGA TCG ATA GTG AAG CCA CAG ATG T
p7+Loop+ FUIRI 1	Forward	CAA GCA GAA GAC GGC ATA CGA GAG TTA GTG AAG CCA CAG ATG T
p7+Loop+ FUIRI 2	Forward	CAA GCA GAA GAC GGC ATA CGA ATA GTA GTG AAG CCA CAG ATG T
p7+Loop+ FUIRI 3	Forward	CAA GCA GAA GAC GGC ATA CGA CTC ATA GTG AAG CCA CAG ATG T
p5+PGK	Reverse	AAT GAT ACG GCG ACC ACC GAT GGA TGT GGA ATG TGT GCG AGG

RT-qPCR OF INDIVIDUAL shRNAs		
Name	**Sense**	**Sequence (5' - 3')**
ACTR5 IDT™ assay	Forward	CTG TTG GAG ATG AGA CCC TTC
	Reverse	TTC ATA CTC TTT CCT GGT GAT CC
ARID2 IDT™ assay	Forward	ACT AAA CAC ATC CGA CTA ACA GC
	Reverse	CAT GTT ACT AAT GGC TAG CAC TG
ATR IDT™ assay	Forward	GAA GAT GAT GAC CAC ACT GAG A
	Reverse	CCC AGA CAA GCA TGA TCC AG
BRD7 IDT™ assay	Forward	TGA AGG AAT CTG GAG GAA AGC
	Reverse	CAG GAT GGA GAA GTC CCA AC

BRIP1 IDT™ assay	Forward	CTC CGC TTT ATT TGC TCT CAG A
	Reverse	TGC TTT CCT GTT TAT TTC AGA TTC C
CDK2AP1 IDT™ assay	Forward	AAG GAG ATC AGA CCC ACG TA
	Reverse	AAC AAG GCA GCT AGG ATC TG
DPY30 IDT™ assay	Forward	TGC CTA CCT GGA TCA GAC A
	Reverse	CGA TCT TCA AAC TGT GCC TTG
INO80D IDT™ assay	Forward	CCA AGT ATA ACA GCC AAC GC
	Reverse	AGC CAA GTA CCT GCA AGT G
MIS18A IDT™ assay	Forward	AGC CAG GAG GAC ACC AA
	Reverse	ACG CAA CCA TTT TCC TTT TCA C
MORF4L2 IDT™ assay	Forward	GTG CGT ATT TGC CTG AAG AAG
	Reverse	TCC TCA CTA TAC AGC AAA CTT AGC
PARG IDT™ assay	Forward	TGC TGA GAC ATA TCG TTG GTC
	Reverse	GAG GTA GCG TCT GAA GTG AA
PARP14 IDT™ assay	Forward	ACT TGA ACA CAT ACA CTG CCA
	Reverse	TTC TGC TGC TTC ATA TCA CTC C
PAX9 IDT™ assay	Forward	GGA AGC CAA GTA CGG TCA G
	Reverse	GTC CAG CAA CAT AAC CAG AAG
PBRM1 IDT™ assay	Forward	CTG AAG GTT GGC GAC TGT
	Reverse	AGC TCC ATC TCG AAC CCA TA
PHC3 IDT™ assay	Forward	TGC CAT TTC AGA CTC TTC CTC
	Reverse	CTT CTT CAC ACA CAT CTT CTA CCT
RERE IDT™ assay	Forward	CAA CAT GAG GAA CTG GTC TGG
	Reverse	TCT GTA GAG CCT CCA TCA CA
SETD2 IDT™ assay	Forward	CAG AGT CAG CAT CAG AGC AG
	Reverse	ACC CTC ACC ATT TTC CAT CA
SMARCA4 IDT™ assay	Forward	CAT CAC TGA GAA GCT GGA GAA G
	Reverse	GCT GGA GAA TGC TAT TGA GGT
SMARCA5 IDT™ assay	Forward	GTG GTC TTG GCA TCA ATC TTG
	Reverse	AAA GCG GAA CAC TCT GAC TG
TDRD5 IDT™ assay	Forward	GCT TCC AGC TCA GGC TAT C
	Reverse	CCA TCT ACA TAT TCA TCC ACT ACC C
TRIM28 IDT™ assay	Forward	GTC CTG GCA CTA ACT CAA CAG
	Reverse	TTG AGT AGG GAT CAT CTC CTG A
TRIM33 IDT™ assay	Forward	TCC CAA CAC TAC CAA ATC CC
	Reverse	CTG CCT GAG CTA CTT CTG AAT C

RT-qPCR OF KNOCK-OUT CLONES		
Name	**Sense**	**Sequence (5' - 3')**
BRIP1 gRNA1	Forward	TGA CCA GCT TGT CCA GGT TT
	Reverse	TTC AGA AGG TGG TGT GCT TG

BRIP1 gRNA3	Forward	GGT TTT CTT TGT AAG GCG TGT C
	Reverse	TCT TCC AAG TGA ACC CAG AAA
BRIP1 gRNA4	Forward	TGC AAT CTC ACT TTT CCT TGC
	Reverse	ACA GCA TGG CTG AAC CAG TC
MIS18A gRNA3	Forward	CCT AGT GTG GTC AGA ATA GCA T
	Reverse	GTT CTG ACA ATT CCG TGC CT
MIS18A gRNA4	Forward	TGT GTG TTT GAC TTT GGG CT
	Reverse	ACT GAT ATA TGC GAA CGA CTG A

HOUSEKEEPING GENES		
Name	**Sense**	**Sequence (5' - 3')**
GAPDH	Forward	TCG ACA GTC AGC CGC ATC T
	Reverse	CTA GCC TCC CGG GTT TCT CT
MRPL9	Forward	CAG TTT CTG GGG ATT TGC AT
	Reverse	TAT TCA GGA AGG CA TCT CG
PUM1	Forward	CGG TCG TCC TGA GGA TAA AA
	Reverse	CGT ACG TGA GGC GTG AGT AA

Table 22. Primers. Sequences of all the primers used in this thesis.

VIII. Sequences of ultramers

GENE	HAIRPIN ID	97 bp ULTRAMER SEQUENCE
ACTR5	ACTR5.1651	TGCTGTTGACAGTGAGCGAAAACCACCTAGATGATAATGAAGATAGTGAAGCCACAGATGTATTCATTATCATCTAGGTGGTTCTGCCTACTGCCTCGGA
	ACTR5.1036	TGCTGTTGACAGTGAGCGACAGATGGATCAGTTTCACAAATAGTGAAGCCACAGATGTATTTGTGAAACTGATCCATCTGGTGCCTACTGCCTCGGA
ARID2	ARID2.3102 (common)	TGCTGTTGACAGTGAGCGCCAAGTACAGATGCAAGTTCAATAGTGAAGCCACAGATGTATTGAACTTGCATCTGTTGCCTACTGCCTCGGA
	ARID2.3108 (FUOX)	TGCTGTTGACAGTGAGCGCCAGATGCAAGTTCAACCTCAATAGTGAAGCCACAGATGTATTGAGGTTGAACTTGCATCTGTTGCCTACTGCCTCGGA
	ARID2.3950 (FUIRI)	TGCTGTTGACAGTGAGCGCTCAGTGCTCACTAATCAGTAATAGTGAAGCCACAGATGTATTACTGATTAGTGAGCACTGATTGCCTACTGCCTCGGA
ATR	ATR.2186 (common)	TGCTGTTGACAGTGAGCGCCCCAAGATTCTTATAGATAAATAGTGAAGCCACAGATGTATTTATCTATAAGAATCTTGGGATGCCTACTGCCTCGGA
	ATR.812 (FUOX)	TGCTGTTGACAGTGAGCGACAGAGATAGGTTGTGTTCTGCTATAGTGAAGCCACAGATGTATAGCAGAACACAACCTATCTGCTGCCTACTGCCTCGGA
	ATR.3749 (FUIRI)	TGCTGTTGACAGTGAGCGATCCCTTCTCAGTCATGTAATATAGTGAAGCCACAGATGTATATTACATGACTGAGAAGGGAGTGCCTACTGCCTCGGA
BRD7	BRD7.732	TGCTGTTGACAGTGAGCGCTCAGTCCATAGAAGAACTAAATAGTGAAGCCACAGATGTATTTAGTTCTTCTATGGACTGATTGCCTACTGCCTCGGA
	BRD7.1426	TGCTGTTGACAGTGAGCGAAGCAAGGATGATTCTGATTTATAGTGAAGCCACAGATGTATAAATCAGAATCATCCTTGCTGTGCTACTGCCTCGGA
BRIP1	BRIP1.2009 (FUOX)	TGCTGTTGACAGTGAGCGCTGGGGTTGTTGGTTCTACCAAATAGTGAAGCCACAGATGTATTTGGTAGAACCAACAACCCATTGCCTACTGCCTCGGA
	BRIP1.954 (FUOX)	TGCTGTTGACAGTGAGCGATGCTGTTGTTCTACTAAACAATAGTGAAGCCACAGATGTATTGTTTAGTAGAACAACAGCACTGCCTACTGCCTCGGA
	BRIP1.7787 (FUIRI)	TGCTGTTGACAGTGAGCGCAAGGTAGAGATTGTTTATCAATAGTGAAGCCACAGATGTATTGATAAACAATCTCTACCTTATGCCTACTGCCTCGGA
	BRIP1.1986 (FUIRI)	TGCTGTTGACAGTGAGCGCCAGATTGATATTTCAGACAAATAGTGAAGCCACAGATGTATTTGTCTGAAATATCAATCTGATGCCTACTGCCTCGGA
CDK2AP1	CDK2AP1.434	TGCTGTTGACAGTGAGCGCACCTACAGTCTCAAAGATTGATAGTGAAGCCACAGATGTATCAATCTTTGAGACTGTAGGTTTGCCTACTGCCTCGGA
	CDK2AP1.375	TGCTGTTGACAGTGAGCGCCCCTTGTTTTCAAAATGGTTATAGTGAAGCCACAGATGTATAACCATTTTGAAAACAAGGGTTGCCTACTGCCTCGGA

Gene	Construct	Sequence
DPY30	DPY30.294	TGCTGTTGACAGTGAGCGAAACGTTGAGAGAATAGTAGAAATAGTGAAGCCACAGATGTATTCTACTATTCTCTCAACGTTGTGCCTACTGCCTCGGA
	DPY30.598	TGCTGTTGACAGTGAGCGGCTCCCTCTTTTGGATGTATAATAGTGAAGCCACAGATGTATTATACATCCAAAAAGAGGGAATGCCTACTGCCTCGGA
INO80D	INO80D.6487	TGCTGTTGACAGTGAGCGCAAGGCCAGTTAGATTGTGTTATAGTGAAGCCACAGATGTATAACAATCTAACTGGCCTTTGCCTACTGCCTCGGA
	INO80D.13166	TGCTGTTGACAGTGAGCGAAAGGTAGATGGACAATCCTAATAGTGAAGCCACAGATGTATTAGGATTGTCCATCTACCTTCTGCCTACTGCCTCGGA
MIS18A	MIS18A.1204	TGCTGTTGACAGTGAGCGCTGGGACAACATTTATTATAAATAGTGAAGCCACAGATGTATTTATAATAAATGTTGTCCCAATGCCTACTGCCTCGGA
	MIS18A.564	TGCTGTTGACAGTGAGCGCGCCATTGAAAGTTATGTTTTATAGTGAAGCCACAGATGTATAAAACATAACTTTCAATGGCTTGCCTACTGCCTCGGA
MORF4L2	MORF4L2.673	TGCTGTTGACAGTGAGCGGAAGGAATAAAAGAATATTTCAATAGTGAAGCCACAGATGTATTGAAATATTCTTTATTCCTGTGCCTACTGCCTCGGA
	MORF4L2.166	TGCTGTTGACAGTGAGCGGATGCAGAAGAAGAGAACTTCAATAGTGAAGCCACAGATGTATTGAAGTTCTCTTCTTCTGCAGTGCCTACTGCCTCGGA
PARG	PARG.2424 (FUOX)	TGCTGTTGACAGTGAGCGCTACGAAGGTACCATAGAAGAAATAGTGAAGCCACAGATGTATTCTTCTATGGTACCTTCGTAATGCCTACTGCCTCGGA
	PARG.3047 (FUOX)	TGCTGTTGACAGTGAGCGCTAAGCTGTTGCTACGATACTATAGTGAAGCCACAGATGTATAGTAGTATCGTAGCAACAGCTTATTGCCTACTGCCTCGGA
	PARG.2189 (FUIRI)	TGCTGTTGACAGTGAGCGAACGACGAAATGCTAAGATGAATAGTGAAGCCACAGATGTATTCATCTTAGCATTTCGTCGTGTGCCTACTGCCTCGGA
	PARG.846 (FUIRI)	TGCTGTTGACAGTGAGCGGCCACAGTGACACAGATAGTGAATAGTGAAGCCACAGATGTATTCACTATCTGTGTCACTGTGATGCCTACTGCCTCGGA
PARP14	PARP14.386	TGCTGTTGACAGTGAGCGGAAAGACCAAAGAAGATGTTAAATAGTGAAGCCACAGATGTATTTAACATCTTCTTTGGTCTTGTGCCTACTGCCTCGGA
	PARP14.214	TGCTGTTGACAGTGAGCGACCGCTTCCTGGTGTTCTTCTATAGTGAAGCCACAGATGTATAGAAGAACACCAGGAAGCGGGTGCCTACTGCCTCGGA
PAX9	PAX9.2902	TGCTGTTGACAGTGAGCGACCAAATAATGTTGTAATTTAATAGTGAAGCCACAGATGTATTAAATTACAACATTATTTGGGTGCCTACTGCCTCGGA
	PAX9.986	TGCTGTTGACAGTGAGCGCCCGGACCTACAAGCAGAGAGATAGTGAAGCCACAGATGTATCTCTCGCTTGTAGGTCCGGATGCCTACTGCCTCGGA
PBRM1	PBRM1.3093	TGCTGTTGACAGTGAGCGCAAGAGTGACTATTACAACAAATAGTGAAGCCACAGATGTATTTGTTGTAATAGTCACTCTTATGCCTACTGCCTCGGA
	PBRM1.924	TGCTGTTGACAGTGAGCGGAAAGGATGCAAATTCAATTAAATAGTGAAGCCACAGATGTATTTAATTGAATTTGCATCCTTGTGCCTACTGCCTCGGA
PHC3	PHC3.703	TGCTGTTGACAGTGAGCGCCAGGTTCAGAATTTAACATTATAGTGAAGCCACAGATGTATAAATGTTAAATTCTGAACCTGATGCCTACTGCCTCGGA
	PHC3.4933	TGCTGTTGACAGTGAGCGCAAGGCTCATGTTCCAGACCTATAGTGAAGCCACAGATGTATAGGTCTGGAACATGAGCCTTATGCCTACTGCCTCGGA
RERE	RERE.3444	TGCTGTTGACAGTGAGCGCACCAAGTCAGGAAGATTATTATAGTGAAGCCACAGATGTATAATAATCTTCCTGACTTGGTATGCCTACTGCCTCGGA
	RERE.396	TGCTGTTGACAGTGAGCGATCGCATCCCACTTCAAGAAAATATAGTGAAGCCACAGATGTATATTTCTTGAAGTGGATGCGACTGCCTACTGCCTCGGA
SETD2	SETD2.2549	TGCTGTTGACAGTGAGCGCCCAGTTATATGTGATAGTAGAGATAGTGAAGCCACAGATGTATCTACTACTATCACATATAACTGGTTGCCTACTGCCTCGGA

		Sequence
	SETD2.7459	TGCTGTTGACAGTGAGCGCCAAGCAAAGAAGTATTCAGAAATAGTGAAGCCACAGATGTATTTCTGAATACTTCTTTGCTTTTGCCTACTGCCTCGGA
SMARCA4	SMARCA4.2582	TGCTGTTGACAGTGAGCGCCGAGTACATCATCAAAGACAATAGTGAAGCCACAGATGTATTGTCTTTGATGATGTACTCGTTGCCTACTGCCTCGGA
	SMARCA4.4419	TGCTGTTGACAGTGAGCGACCCGTGGACTTCAAGAAGATATAGTGAAGCCACAGATGTATATCTTGAAGTCCACGGGCTGCCTACTGCCTCGGA
SMARCA5	SMARCA5.3184	TGCTGTTGACAGTGAGCGAAGGCGAGAATTCAAAGAAGAATAGTGAAGCCACAGATGTATTCTTCTTTGAATTCTCGCCTCTGCCTACTGCCTCGGA
	SMARCA5.1443	TGCTGTTGACAGTGAGCGAAAGACTACAAAATAGACTATTATAGTGAAGCCACAGATGTATAATAGTCTATTTGTGAGTCTTGTGCCTACTGCCTCGGA
TDRD5	TDRD5.59	TGCTGTTGACAGTGAGCGCCCAGTTCAGAACTAAAACATAATAGTGAAGCCACAGATGTATTATGTTTAGTTCTGAACTGATGCCTACTGCCTCGGA
	TDRD5.2274	TGCTGTTGACAGTGAGCGCCACAGTCTACTTTGATGTGTAATAGTGAAGCCACAGATGTATTACACATCAAAGTAGACTGTTTGCCTACTGCCTCGGA
TRIM28	TRIM28.1134	TGCTGTTGACAGTGAGCGACCCAGGTGTCTGACGTACAGAATAGTGAAGCCACAGATGTATTCTGTACGTCAGACACCTGGCTGCCTGCCTCGGA
	TRIM28.2586	TGCTGTTGACAGTGAGCGACCGCATGTTCAAGCAATTCAATAGTGAAGCCACAGATGTATTGAATTGCTTGAACATGCGGGCTGCCTACTGCCTCGGA
TRIM33	TRIM33.6875	TGCTGTTGACAGTGAGCGCTCGGATCAGTGAGTAGTTTTATAGTGAAGCCACAGATGTATAAAACTACTCACTGATCCGATTGCCTACTGCCTCGGA
	TRIM33.3893	TGCTGTTGACAGTGAGCGCCACCCAAGTAGCTCATTCTATAGTGAAGCCACAGATGTATAGAATCTGAGCTACTTGGGTTTGCCTACTGCCTCGGA

Table 23. Ultramers. 97 bp ultramer sequences of the two top targeting shRNAs of each gene ordered by *p*-value. In the case of genes that appeared as common in cells treated with FUOX and FUIRI, top targeting shRNAs can be common for both treatments or different depending on the type of chemotherapy.

Bibliography

1.   Global Cancer Observatory. https://gco.iarc.fr/.

2.   Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, Van De Velde CJH & Watanabe T. Colorectal cancer. *Nat. Rev. Dis. Prim.* **1**: 1–25 (2015).

3.   American Cancer Society | Information and Resources about for Cancer: Breast, Colon, Lung, Prostate, Skin. https://www.cancer.org/.

4.   Mattiuzzi C, Sanchis-Gomar F & Lippi G. Concise update on colorectal cancer epidemiology. *Ann. Transl. Med.* **7**(21): 609–609 (2019).

5.   Ionov Y, Peinado MA, Malkhosyan S, Shibata D & Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **363**(6429): 558–561 (1993).

6.   Boland CR & Lynch HT. The History of Lynch Syndrome. *Fam. Cancer* **12**(2): 145–157 (2013).

7.   Brenner H, Kloor M & Pox CP. Colorectal cancer. *Lancet* **383**(9927): 1490–1502 (2014).

8.   Vasen HFA, Tomlinson I & Castells A. Clinical management of hereditary colorectal cancer syndromes. *Nat. Rev. Gastroenterol. Hepatol.* **12**(2): 88–97 (2015).

9.   Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AMM & Bos JL. Genetic Alterations During Colorectal-Tumor Development. *N. Engl. J. Med.* **319**(9): 525–532 (1988).

10.   Fearon ER & Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* **61**(5): 759–767 (1990).

11.   Barrow TM & Michels KB. Epigenetic epidemiology of cancer. *Biochem. Biophys. Res. Commun.* **455**(1–2): 70–83 (2014).

12.   Lee KH, Lee JS, Nam JH, Choi C, Lee MC, Park CS, Juhng SW & Lee JH. Promoter methylation status of hMLH1, hMSH2, and MGMT genes in colorectal cancer associated with adenoma-carcinoma sequence. *Langenbeck's Arch. Surg.* **396**(7): 1017–1026 (2011).

13.   Psofaki V, Kalogera C, Tzambouras N, Stephanou D, Tsianos E, Seferiadis K & Kolios G. Promoter methylation status of hMLH1, MGMT, and CDKN2A/p16 in colorectal adenomas. *World J. Gastroenterol.* **16**(28): 3553–3560 (2010).

14.   Toyota M, Ahuja N, Ohe-Toyota M, Herman J, Baylin SB & Issa JPJ. CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. U. S. A.* **96**(15): 8681–8686 (1999).

15.   Hanahan D & Weinberg RA. The hallmarks of cancer. *Cell* **100**(1): 57–70 (2000).

16.   Hanahan D & Weinberg RA. Hallmarks of cancer: the next generation. *Cell* **144**(5): 646–674 (2011).

17.   Lao VV & Grady WM. Epigenetics and colorectal cancer. *Nat. Rev. Gastroenterol. Hepatol.* **8**(12): 686–700 (2011).

18.   Feinberg AP & Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem. Biophys. Res. Commun.* **111**(1): 47–54 (1983).

19.   Feinberg AP & Vogelstein B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**(5895): 89–92 (1983).

20.   Ehrlich M. DNA hypomethylation in cancer cells. *Epigenomics* **1**(2): 239–259 (2009).

21.   Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, Fiala E, Ehrlich M & Laird PW. Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res.* **33**(21): 6823–6836 (2005).

22.   Eden A, Gaudet F, Waghmare A & Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* **300**(5618): 455 (2003).

23.   Worthley DL, Whitehall VLJ, Buttenshaw RL, Irahara N, Greco SA, Ramsnes I, Mallitt K-A, Le Leu RK, Winter J, Hu Y, Ogino S, Young GP & Leggett BA. DNA methylation within the normal colorectal mucosa is associated with pathway-specific predisposition to cancer. *Oncogene* **29**(11): 1653–1662

(2009).

24. Daskalos A, Nikolaidis G, Xinarianos G, Savvari P, Cassidy A, Zakopoulou R, Kotsinas A, Gorgoulis V, Field JK & Liloglou T. Hypomethylation of retrotransposable elements correlates with genomic instability in non-small cell lung cancer. *Int. J. Cancer* **124**(1): 81–87 (2009).

25. Anwar SL, Wulaningsih W & Lehmann U. Transposable elements in human cancer: Causes and consequences of deregulation. *Int. J. Mol. Sci.* **18**(5): 974 (2017).

26. Zauber AG, Winawer SJ, O MJ, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey BF, Shi W, Bond JH, Schapiro M, Panish JF, Stewart ET & Waye JD. Colonoscopic Polypectomy and Long-Term Prevention of Colorectal-Cancer Deaths. *N. Engl. J. Med.* **366**(8): 687–696 (2012).

27. Nishihara R, Wu K, Lochhead P, Morikawa T, Liao X, Qian ZR, Inamura K, Kim SA, Kuchiba A, Yamauchi M, Imamura Y, Willett WC, Rosner BA, Fuchs CS, Giovannucci E, Ogino S & Chan AT. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N. Engl. J. Med.* **369**(12): 1095–1105 (2013).

28. Hurwitz H, Fehrenbacher L, Novotny W, Cartwright T, Hainsworth J, Heim W, Berlin J, Baron A, Griffing S, Holmgren E, Ferrara N, Fyfe G, Rogers B, Ross R & Kabbinavar F. Bevacizumab plus Irinotecan, Fluorouracil, and Leucovorin for Metastatic Colorectal Cancer. *N. Engl. J. Med.* **23**(3): 2335–2377 (2004).

29. Giantonio BJ, Catalano PJ, Meropol NJ, O'Dwyer PJ, Mitchell EP, Alberts SR, Schwartz MA & Benson AB. Bevacizumab in combination with oxaliplatin, fluorouracil, and leucovorin (FOLFOX4) for previously treated metastatic colorectal cancer: Results from the Eastern Cooperative Oncology Group Study E3200. *J. Clin. Oncol.* **25**(12): 1539–1544 (2007).

30. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH & Diaz LA. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat. Rev. Gastroenterol. Hepatol.* **16**(6): 361–375 (2019).

31. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B & Diaz LA. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**(26): 2509–2520 (2015).

32. Le DT, Uram JN, Wang H, Bartlett B, Kemberling H, Eyring A, Azad NS, Laheru D, Donehower RC, Crocenzi TS, Goldberg RM, Fisher GA, Lee JJ, Greten TF, Koshiji M, Kang SP, Anders RA, Eshleman JR, Vogelstein B & Diaz LA. Programmed death-1 blockade in mismatch repair deficient colorectal cancer. *J. Clin. Oncol.* **34**(15_suppl): 103–103 (2016).

33. Overman MJ, McDermott R, Leach JL, Lonardi S, Lenz H-J, Morse MA, Desai J, Hill A, Axelson M, Moss RA, Goldberg M V, Cao ZA, Ledeine J-M, Maglinte GA, Kopetz S & André T. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet Oncol.* **18**(9): 1182–1191 (2017).

34. Overman MJ, Lonardi S, Wong KYM, Lenz H-J, Gelsomino F, Aglietta M, Morse MA, Van Cutsem E, McDermott R, Hill A, Sawyer MB, Hendlisz A, Neyns B, Svrcek M, Moss RA, Ledeine J-M, Cao ZA, Kamble S, Kopetz S & André T. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. *J. Clin. Oncol.* **36**(8): 773–779 (2018).

35. Guinney J, Dienstmann R, Wang X, De Reyniès A, Schlicker A, Soneson C, Marisa L, Roepman P,

Nyamundanda G, Angelino P, Bot BM, Morris JS, Simon IM, Gerster S, Fessler E, De Sousa .E Melo F, Missiaglia E, Ramay H, Barras D, Homicsko K, Maru D, Manyam GC, Broom B, Boige V, Perez-Villamil B, Laderas T, Salazar R, Gray JW, Hanahan D, Tabernero J, Bernards R, Friend SH, Laurent-Puig P, Medema JP, Sadanandam A, Wessels L, Delorenzi M, Kopetz S, Vermeulen L & Tejpar S. The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**(11): 1350–1356 (2015).

36. Llosa NJ, Cruise M, Tam A, Wicks EC, Hechenbleikner EM, Taube JM, Blosser RL, Fan H, Wang H, Luber BS, Zhang M, Papadopoulos N, Kinzler KW, Vogelstein B, Sears CL, Anders RA, Pardoll DM & Housseau F. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov.* **5**(1): 43–51 (2015).

37. Tran B, Kopetz S, Tie J, Gibbs P, Jiang Z-Q, Lieu CH, Agarwal A, Maru DM, Sieber O & Desai J. Impact of BRAF mutation and microsatellite instability on the pattern of metastatic spread and prognosis in metastatic colorectal cancer. *Cancer* **117**(20): 4623–4632 (2011).

38. Popovici V, Budinska E, Bosman FT, Tejpar S, Roth AD & Delorenzi M. Context-dependent interpretation of the prognostic value of BRAF and KRAS mutations in colorectal cancer. *BMC Cancer* **13**: 439–445 (2013).

39. Brunelli L, Caiola E, Marabese M, Broggini M & Pastorelli R. Capturing the metabolomic diversity of KRAS mutants in non-small-cell lung cancer cells. *Oncotarget* **5**(13): 4722–4731 (2014).

40. Son J, Lyssiotis CA, Ying H, Wang X, Hua S, Ligorio M, Perera RM, Ferrone CR, Mullarky E, Shyh-Chang N, Kang Y, Fleming JB, Bardeesy N, Asara JM, Haigis MC, Depinho RA, Cantley LC & Kimmelman AC. Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496**(7443): 101–105 (2013).

41. Ying H, Kimmelman AC, Lyssiotis CA, Hua S, Chu GC, Fletcher-Sananikone E, Locasale JW, Son J, Zhang H, Coloff JL, Yan H, Wang W, Chen S, Viale A, Zheng H, Paik J, Lim C, Guimaraes AR, Martin ES, Chang J, Hezel AF, Perry SR, Hu J, Gan B, Xiao Y, Asara JM, Weissleder R, Wang YA, Chin L, Cantley LC & DePinho RA. Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149**(3): 656–670 (2012).

42. Chin LJ, Ratner E, Leng S, Zhai R, Nallur S, Babar I, Muller RU, Straka E, Su L, Burki EA, Crowell RE, Patel R, Kulkarni T, Homer R, Zelterman D, Kidd KK, Zhu Y, Christiani DC, Belinsky SA, Slack FJ & Weidhaas JB. A SNP in a let-7 microRNA complementary site in the KRAS 3′ untranslated region increases non-small cell lung cancer risk. *Cancer Res.* **68**(20): 8535–8540 (2008).

43. Du XY, Hu YY, Xie C, Deng CY, Liu CY, Luo ZG, Niu YM & Shen M. Significant association between Let-7-KRAS rs712 G > T polymorphism and cancer risk in the Chinese population: A meta-analysis. *Oncotarget* **8**(8): 13863–13871 (2017).

44. Park SM, Gaur AB, Lengyel E & Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* **22**(7): 894–907 (2008).

45. Carmona FJ, Davalos V, Vidal E, Gomez A, Heyn H, Hashimoto Y, Vizoso M, Martinez-Cardus A, Sayols S, Ferreira HJ, Sánchez-Mut J V., Morán S, Margelí M, Castella E, Berdasco M, Stefansson OA, Eyfjord JE, Gonzalez-Suarez E, Dopazo J, Orozco M, Gut IG & Esteller M. A comprehensive DNA methylation profile of epithelial-to-mesenchymal transition. *Cancer Res.* **74**(19): 5608–5619 (2014).

46. Dawson MA & Kouzarides T. Cancer epigenetics: From mechanism to therapy. *Cell* **150**(1): 12–27 (2012).

47. Kumar R, Li DQ, Müller S & Knapp S. Epigenomic regulation of oncogenesis by chromatin

remodeling. *Oncogene* **35**(34): 4423–4436 (2016).

48. Saksouk N, Simboeck E & Déjardin J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**: 3 (2015).

49. Waddington CH. The epigenotype. *Int. J. Epidemiol.* **41**(1): 10–13 (2012).

50. Esteller M. Epigenetics in Cancer. *N. Engl. J. Med.* **358**(11): 1148–1159 (2008).

51. Okugawa Y, Grady WM & Goel A. Epigenetic Alterations in Colorectal Cancer: Emerging Biomarkers. *Gastroenterology* **149**(5): 1204–1225 (2015).

52. Jung G, Hernández-Illán E, Moreira L, Balaguer F & Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat. Rev. Gastroenterol. Hepatol.* **17**(2): 111–130 (2020).

53. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science* **338**(6113): 1435–1439 (2012).

54. Shen H & Laird PW. Interplay between the cancer genome and epigenome. *Cell* **153**(1): 38–55 (2013).

55. Sharrard RM, Royds JA, Rogers S & Shorthouse AJ. Patterns of methylation of the c-myc gene in human colorectal cancer progression. *Br. Joumal Cancer* **65**(5): 667–672 (1992).

56. Luo J, Li Y-N, Wang F, Zhang W-M & Geng X. S-Adenosylmethionine Inhibits the Growth of Cancer Cells by Reversing the Hypomethylation Status of c-myc and H-ras in Human Gastric Cancer and Colon Cancer. *Int. J. Biol. Sci.* **6**(7): 784–795 (2010).

57. Mukherjee K, Sur D, Singh A, Rai S, Das N, Sekar R, Narindi S, Dhingra VK, Jat B, Balraam KVV, Agarwal SP & Mandal PK. Robust expression of LINE-1 retrotransposon encoded proteins in oral squamous cell carcinoma. *BMC Cancer* **21**(1): 1–15 (2021).

58. Suzuki K, Suzuki I, Leodolter A, Alonso S, Horiuchi S, Yamashita K & Perucho M. Global DNA demethylation in gastrointestinal cancer is age dependent and precedes genomic damage. *Cancer Cell* **9**(3): 199–207 (2006).

59. Samuelsson JK, Dumbovic G, Polo C, Moreta C, Alibés A, Ruiz-Larroya T, Giménez-Bonafé P, Alonso S, Forcales S-V & Manuel P. Helicase Lymphoid-Specific Enzyme Contributes to the Maintenance of Methylation of SST1 Pericentromeric Repeats That Are Frequently Demethylated in Colon Cancer and Associate with Genomic Damage. *Epigenomes* **1**(1): 2 (2017).

60. Dumbovic G, Forcales S-V & Perucho M. Emerging roles of macrosatellite repeats in genome organization and disease development. *Epigenetics* **12**(7): 515–526 (2017).

61. Wang W, Kandimalla R, Huang H, Zhu L, Li Y, Gao F, Goel A & Wang X. Molecular subtyping of colorectal cancer: Recent progress, new challenges and emerging opportunities. *Semin. Cancer Biol.* **55**: 37–52 (2019).

62. Baubec T & Schübeler D. Genomic patterns and context specific interpretation of DNA methylation. *Curr. Opin. Genet. Dev.* **25**(1): 85–92 (2014).

63. Neri F, Krepelova A, Incarnato D, Maldotti M, Parlato C, Galvagni F, Matarese F, Stunnenberg HG & Oliviero S. Dnmt3L Antagonizes DNA Methylation at Bivalent Promoters and Favors DNA Methylation at Gene Bodies in ESCs. *Cell* **155**(1): 121–134 (2013).

64. Suetake I, Shinozaki F, Miyagawa J, Takeshima H & Tajima S. DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a and Dnmt3b through a Direct Interaction. *J. Biol. Chem.* **279**(26): 27816–27823 (2004).

65. Gong Z, Morales-Ruiz T, Ariza RR, Roldán-Arjona T, David L & Zhu JK. ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase. *Cell* **111**(6): 803–814 (2002).

66. Wu H & Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev.* **25**(23): 2436–2452 (2011).

67. Finley A & Copeland RA. Small molecule control of chromatin remodeling. *Chem. Biol.* **21**(9): 1196–1210 (2014).

68. Qin J, Wen B, Liang Y, Yu W & Li H. Histone Modifications and their Role in Colorectal Cancer (Review). *Pathol. Oncol. Res.* **26**(4): 2023–2033 (2020).

69. Ashktorab H, Belgrave K, Hosseinkhah F, Brim H, Nouraie M, Takkikto M, Hewitt S, Lee EL, Dashwood RH & Smoot D. Global histone H4 acetylation and HDAC2 expression in colon adenoma and carcinoma. *Dig. Dis. Sci.* **54**(10): 2109–2117 (2009).

70. Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Pérez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MÁ, Ahn N, Imhof A, Caldas C, Jenuwein T & Esteller M. Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**(4): 391–400 (2005).

71. Ingen H van, Schaik FMA van, Wienk H, Ballering J, Rehmann H, Dechesne AC, Kruijzer JAW, Liskamp RMJ, Timmers HTM & Boelens R. Structural Insight into the Recognition of the H3K4me3 Mark by the TFIID Subunit TAF3. *Structure* **16**(8): 1245–1256 (2008).

72. Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH & Roeder RG. H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. *Cell* **152**(5): 1021–1036 (2013).

73. Salz T, Li G, Kaye F, Zhou L, Qiu Y & Huang S. hSETD1A Regulates Wnt Target Genes and Controls Tumor Growth of Colorectal Cancer Cells. *Cancer Res.* **74**(3): 775–786 (2014).

74. Margueron R & Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**(7330): 343–349 (2011).

75. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS & Zhang Y. Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science* **298**(5595): 1039–1043 (2002).

76. Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, De Launoit Y & Fuks F. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**(7078): 871–874 (2006).

77. McCabe MT, Lee EK & Vertino PM. A Multifactorial Signature of DNA Sequence and Polycomb Binding Predicts Aberrant CpG Island Methylation. *Cancer Res.* **69**(1): 282–291 (2009).

78. Benard A, Goossens-Beumer IJ, Hoesel AQ van, Horati H, Putter H, Zeestraten ECM, Velde CJH van de & Kuppen PJK. Prognostic Value of Polycomb Proteins EZH2, BMI1 and SUZ12 and Histone Modification H3K27me3 in Colorectal Cancer. *PLoS One* **9**(9): e108265 (2014).

79. ENCODE: Encyclopedia of DNA Elements. https://www.encodeproject.org/.

80. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome ENCODE Encyclopedia of DNA Elements. *Nature* **489**(7414): 57–74 (2012).

81. Motoyama K, Inoube H, Takatsuno Y, Tanaka F, Mimori K, Uetake H, Sugihara K & Mori M. Over- and under-expressed microRNAs in human colorectal cancer. *Int. J. Oncol.* **34**(4): 1069–1075 (2009).

82. Nagel R, Le Sage C, Diosdado B, Van Der Waal M, Oude Vrielink JAF, Bolijn A, Meijer GA & Agami R. Regulation of the Adenomatous Polyposis Coli Gene by the miR-135 Family in Colorectal Cancer. *Cancer Res.* **68**(14): 5795–5802 (2008).

83. Arndt GM, Dossey L, Cullen LM, Lai A, Druker R, Eisbacher M, Zhang C, Tran N, Fan H, Retzlaff K, Bittner A & Raponi M. Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer. *BMC Cancer* **9**(374): (2009).

84. Dong Y, Zhao J, Wu C-W, Zhang L, Liu X, Kang W, Leung W-W, Zhang N, Chan FKL, Sung JJY, Ng SSM & Yu J. Tumor Suppressor Functions of miR-133a in Colorectal Cancer. *Mol. Cancer Res.* **11**(9): 1051–1060 (2013).

85. Vogt M, Munding J, Grüner M, Liffers ST, Verdoodt B, Hauk J, Steinstraesser L, Tannapfel A & Hermeking H. Frequent concomitant inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal, pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft tissue sarcomas. *Virchows Arch.* **458**(3): 313–322 (2011).

86. Tazawa H, Tsuchiya N, Izumiya M & Nakagama H. Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **104**(39): 15427–15477 (2007).

87. Pagliuca A, Valvo C, Fabrizi E, Di Martino S, Biffoni M, Runci D, Forte S, De Maria R & Ricci-Vitiani L. Analysis of the combined action of miR-143 and miR-145 on oncogenic pathways in colorectal cancer cells reveals a coordinate program of gene repression. *Oncogene* **32**(40): 4806–4813 (2013).

88. Luo J, Qu J, Wu DK, Lu ZL, Sun YS & Qu Q. Long non-coding RNAs: A rising biotarget in colorectal cancer. *Oncotarget* **8**(13): 22187–22202 (2017).

89. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, Tanaka F, Shibata K, Suzuki A, Komune S, Miyano S & Mori M. Long Noncoding RNA HOTAIR Regulates Polycomb-Dependent Chromatin Modification and Is Associated with Poor Prognosis in Colorectal Cancers. *Cancer Res.* **71**(20): 6320–6326 (2011).

90. Wang KC & Chang HY. Molecular Mechanisms of Long Noncoding RNAs. *Mol. Cell* **43**(6): 904–914 (2011).

91. Valencia AM & Kadoch C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat. Cell Biol.* **21**(2): 152–161 (2019).

92. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, Shmulevich I, Sander C & Stuart JM. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **45**(10): 1113–1120 (2013).

93. Beck S, Bernstein BE, Campbell RM, Costello JF, Dhanak D, Ecker JR, Greally JM, Issa J-P, Laird PW, Polyak K, Tycko B & Jones PA. A Blueprint for an International Cancer Epigenome Consortium. A Report from the AACR Cancer Epigenome Task Force. *Cancer Res.* **72**(24): 6319–6324 (2012).

94. Sahu RK, Singh S & Tomar RS. The mechanisms of action of chromatin remodelers and implications in development and disease. *Biochem. Pharmacol.* **180**: 114200 (2020).

95. Clapier CR, Iwasa J, Cairns BR & Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18**(7): 407–422 (2017).

96. Bracken AP, Brien GL & Verrijzer CP. Dangerous liaisons: Interplay between SWI/SNF, NURD, and polycomb in chromatin regulation and cancer. *Genes Dev.* **33**(15–16): 936–959 (2019).

97. Cenik BK & Shilatifard A. COMPASS and SWI/SNF complexes in development and disease. *Nat. Rev. Genet.* **22**(1): 38–58 (2021).

98. Harrod A, Lane KA & Downs JA. The role of the SWI/SNF chromatin remodelling complex in the response to DNA double strand breaks. *DNA Repair (Amst).* **93**: 102919 (2020).

99. Mittal P & Roberts CWM. The SWI/SNF complex in cancer — biology, biomarkers and therapy. *Nat. Rev. Clin. Oncol.* **17**(7): 435–448 (2020).

100. Centore RC, Sandoval GJ, Soares LMM, Kadoch C & Chan HM. Mammalian SWI/SNF Chromatin Remodeling Complexes: Emerging Mechanisms and Therapeutic Strategies. *Trends Genet.* **36**(12): 936–950 (2020).

101. Ogiwara H, Ui A, Otsuka A, Satoh H, Yokomi I, Nakajima S, Yasui A, Yokota J & Kohno T. Histone

acetylation by CBP and p300 at double-strand break sites facilitates SWI/SNF chromatin remodeling and the recruitment of non-homologous end joining factors. *Oncogene* **30**(18): 2135–2146 (2011).

102. Watanabe R, Ui A, Kanno S, Ogiwara H, Nagase T, Kohno T & Yasui A. SWI/SNF Factors Required for Cellular Resistance to DNA Damage Include ARID1A and ARID1B and Show Interdependent Protein Stability. *Cancer Res.* **74**(9): 2465–2475 (2014).

103. Brownlee PM, Chambers AL, Cloney R, Bianchi A & Downs JA. BAF180 Promotes Cohesion and Prevents Genome Instability and Aneuploidy. *Cell Rep.* **6**(6): 973–981 (2014).

104. Kakarougkas A, Ismail A, Chambers AL, Riballo E, Herbert AD, Kü J, Lö M, Jeggo PA & Downs JA. Molecular Cell Article Requirement for PBAF in Transcriptional Repression and Repair at DNA Breaks in Actively Transcribed Regions of Chromatin. *Mol. Cell* **55**(5): 723–732 (2014).

105. Michel BC, D'Avino AR, Cassel SH, Mashtalir N, McKenzie ZM, McBride MJ, Valencia AM, Zhou Q, Bocker M, Soares LMM, Pan J, Remillard DI, Lareau CA, Zullow HJ, Fortoul N, Gray NS, Bradner JE, Chan HM & Kadoch C. A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat. Cell Biol.* **20**(12): 1410–1420 (2018).

106. Alpsoy A & Dykhuizen EC. Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J. Biol. Chem.* **293**(11): 3892–3903 (2018).

107. Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J & Crabtree GR. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**(6): 592–601 (2013).

108. Dykhuizen EC, Hargreaves DC, Miller EL, Cui K, Korshunov A, Kool M, Pfister S, Cho Y-J, Zhao K & Crabtree GR. BAF complexes facilitate decatenation of DNA by topoisomerase IIα. *Nature* **497**(7451): 624–627 (2013).

109. Corona DFV & Tamkun JW. Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1677**(1–3): 113–119 (2004).

110. Dirscherl SS & Krebs JE. Functional diversity of ISWI complexes. *Biochem. Cell Biol.* **82**(4): 482–489 (2004).

111. Collins N, Poot RA, Kukimoto I, García-Jiménez C, Dellaire G & Varga-Weisz PD. An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. *Nat. Genet.* **32**(4): 627–632 (2002).

112. Neve B, Jonckheere N, Vincent A & Van Seuningen I. Long non-coding RNAs: the tentacles of chromatin remodeler complexes. *Cell. Mol. Life Sci.* **78**(4): 1139–1161 (2021).

113. Lee S-K, Park E-J, Lee H-S, Lee YS & Kwon J. Genome-Wide Screen of Human Bromodomain-Containing Proteins Identifies Cecr2 as a Novel DNA Damage Response Protein. *Mol. Cells* **34**(1): 85–91 (2012).

114. Ji J-H, Min S, Chae S, Ha G-H, Kim Y, Park Y-J, Lee C-W & Cho H. De novo phosphorylation of H2AX by WSTF regulates transcription-coupled homologous recombination repair. *Nucleic Acids Res.* **47**(12): 6299–6314 (2019).

115. Torchy MP, Hamiche A & Klaholz BP. Structure and function insights into the NuRD chromatin remodeling complex. *Cell. Mol. Life Sci.* **72**(13): 2491–2507 (2015).

116. Basta J & Rauchman M. The Nucleosome Remodeling and Deacetylase Complex in Development and Disease. *Transl. Epigenetics to Clin.* **165**(1): 37–72 (2017).

117. Gururaj AE, Singh RR, Rayala SK, Holm C, Den Hollander P, Zhang H, Balasenthil S, Talukder AH, Landberg G & Kumar R. MTA1, a transcriptional activator of breast cancer amplified sequence 3. *Proc. Natl. Acad. Sci. U. S. A.* **103**(17): 6670–6675 (2006).

118. Fujita N, Jaye DL, Geigerman C, Akyildiz A, Mooney MR, Boss JM & Wade PA. MTA3 and the Mi-2/NuRD complex regulate cell fate during B lymphocyte differentiation. *Cell* **119**(1): 75–86 (2004).

119. Morrison AJ. Genome maintenance functions of the INO80 chromatin remodeller. *Philos. Trans. R. Soc. B Biol. Sci.* **372**(1731): 20160289 (2017).

120. Poli J, Gasser SM & Papamichos-Chronakis M. The INO80 remodeller in transcription, replication and repair. *Philos. Trans. R. Soc. B Biol. Sci.* **372**(1731): 20160290 (2017).

121. Willhoft O & Wigley DB. INO80 and SWR1 complexes: the non-identical twins of chromatin remodelling. *Curr. Opin. Struct. Biol.* **61**: 50–58 (2020).

122. Papamichos-Chronakis M, Krebs JE & Peterson CL. Interplay between Ino80 and Swr1 chromatin remodeling enzymes regulates cell cycle checkpoint adaptationin response to DNA damage. *Genes Dev.* **20**(17): 2437–2449 (2006).

123. Chambers AL, Ormerod G, Durley SC, Sing TL, Brown GW, Kent NA & Downs JA. The INO80 chromatin remodeling complex prevents polyploidy and maintains normal chromatin structure at centromeres. *Genes Dev.* **26**(23): 2590–2603 (2012).

124. Zhou B, Wang L, Zhang S, Bennett BD, He F, Zhang Y, Xiong C, Han L, Diao L, Li P, Fargo DC, Cox AD & Hu G. INO80 governs superenhancer-mediated oncogenic transcription and tumor growth in melanoma. *Genes Dev.* **30**(12): 1440–1453 (2016).

125. Ahuja N, Sharma AR & Baylin SB. Epigenetic Therapeutics: A New Weapon in the War Against Cancer. *Annu. Rev. Med.* **67**: 73–89 (2016).

126. Hashimoto Y, Zumwalt TJ & Goel A. DNA methylation patterns as noninvasive biomarkers and targets of epigenetic therapies in colorectal cancer. *Epigenomics* **8**(5): 685–703 (2016).

127. Tsai H-C, Li H, Van Neste L, Cai Y, Robert C, Rassool F V, Shin JJ, Harbom KM, Beaty R, Pappou E, Harris J, Yen R-WC, Ahuja N, Brock M V, Stearns V, Feller-Kopman D, Yarmus LB, Lin Y-C, Welm AL, Issa J-P, Minn I, Matsui W, Jang Y-Y, Sharkis SJ, Baylin SB & Zahnow CA. Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells. *Cancer Cell* **21**(3): 430–446 (2012).

128. Jones PA, Issa JPJ & Baylin S. Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.* **17**(10): 630–641 (2016).

129. McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, Liu Y, Pietra A Della, LaFrance L V., Mellinger M, Duquenne C, Tian X, Kruger RG, McHugh CF, Miller WH, Dhanak D, Verma SK, Tummino PJ, Creasy CL, Graves AP & Diaz E. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**(7427): 108–112 (2012).

130. Stein EM. IDH2 inhibition in AML: Finally progress? *Best Pract. Res. Clin. Haematol.* **28**(2–3): 112–115 (2015).

131. Pfister SX & Ashworth A. Marked for death: Targeting epigenetic changes in cancer. *Nat. Rev. Drug Discov.* **16**(4): 241–263 (2017).

132. Flis S, Gnyszka A & Flis K. DNA Methyltransferase Inhibitors Improve the Effect of Chemotherapeutic Agents in SW48 and HT-29 Colorectal Cancer Cells. *PLoS One* **9**(3): e92305 (2014).

133. Overman MJ, Morris V, Moinova H, Manyam G, Ensor J, Lee MS, Eng C, Kee B, Fogelman D, Shroff RT, LaFramboise T, Mazard T, Feng T, Hamilton S, Broom B, Lutterbaugh J, Issa J-P, Markowitz SD, Kopetz S, Overman MJ, Morris V, Moinova H, Manyam G, Ensor J, Lee MS, Eng C, Kee B, Fogelman D, Shroff RT, LaFramboise T, Mazard T, Feng T, Hamilton S, Broom B, Lutterbaugh J, Issa J-P, Markowitz SD & Kopetz S. Phase I/II study of azacitidine and capecitabine/oxaliplatin (CAPOX) in refractory CIMP-high metastatic colorectal cancer: evaluation of circulating methylated

vimentin. *Oncotarget* **7**(41): 67495–67506 (2016).

134. Patnaik S & Anupriya. Drugs targeting epigenetic modifications and plausible therapeutic strategies against colorectal cancer. *Front. Pharmacol.* **10**: 1–15 (2019).

135. Rajendran P, Kidane AI, Yu T-W, Dashwood W-M, Bisson WH, Löhr C V, Ho E, Williams DE & Dashwood RH. HDAC turnover, CtIP acetylation and dysregulated DNA damage signaling in colon cancer cells treated with sulforaphane and related dietary isothiocyanates. *Epigenetics* **8**(6): 612–623 (2013).

136. Robert C & Rassool F V. HDAC Inhibitors: Roles of DNA damage and repair. *Adv. Cancer Res.* **116**: 87–129 (2012).

137. Kim JC, Shin ES, Kim CW, Roh SA, Cho DH, Na YS, Kim TW, Kim MB, Hyun YL, Ro S, Kim SY & Kim YS. In vitro evaluation of histone deacetylase inhibitors as combination agents for colorectal cancer. *Anticancer Res.* **29**(8): 3027–3034 (2009).

138. Fakih MG, Groman A, McMahon J, Wilding G & Muindi JR. A randomized phase II study of two doses of vorinostat in combination with 5-FU/LV in patients with refractory colorectal cancer. *Cancer Chemother. Pharmacol.* **69**(3): 743–751 (2012).

139. Reuter S, Gupta SC, Park B, Goel A & Aggarwal BB. Epigenetic changes induced by curcumin and other natural compounds. *Genes Nutr.* **6**(2): 93–108 (2011).

140. Liu H, Chen Y, Cui G & Zhou J. Curcumin, a potent anti-tumor reagent, is a novel histone deacetylase inhibitor regulating B-NHL cell line Raji proliferation. *Acta Pharmacol. Sin.* **26**(5): 603–609 (2005).

141. Chen Y, Shu W, Chen W, Wu Q, Liu H & Cui G. Curcumin, both histone deacetylase and p300/CBP-specific inhibitor, represses the activity of nuclear factor kappa B and Notch 1 in Raji cells. *Basic Clin. Pharmacol. Toxicol.* **101**(6): 427–433 (2007).

142. Bora-Tatar G, Dayangaç-Erden D, Demir AS, Dalkara S, Yelekçi K & Erdem-Yurter H. Molecular modifications on carboxylic acid derivatives as potent histone deacetylase inhibitors: Activity and docking studies. *Bioorg. Med. Chem.* **17**(14): 5219–5228 (2009).

143. Balasubramanyam K, Varier RA, Altaf M, Swaminathan V, Siddappa NB, Ranga U & Kundu TK. Curcumin, a Novel p300/CREB-binding Protein-specific Inhibitor of Acetyltransferase, Represses the Acetylation of Histone/Nonhistone Proteins and Histone Acetyltransferase-dependent Chromatin Transcription. *J. Biol. Chem.* **279**(49): 51163–51171 (2004).

144. Neckers L, Trepel J, Lee S, Chung E-J, Lee M-J, Jung Y-J & Marcu M. Curcumin is an Inhibitor of p300 Histone Acetylatransferase. *Med. Chem.* **2**(2): 169–174 (2006).

145. Link A, Balaguer F, Shen Y, Lozano JJ, Leung H-CE, Boland CR & Goel A. Curcumin Modulates DNA Methylation in Colorectal Cancer Cells. *PLoS One* **8**(2): e57709 (2013).

146. Filippakopoulos P, Qi J, Picaud S, Shen Y, Smith WB, Fedorov O, Morse EM, Keates T, Hickman TT, Felletar I, Philpott M, Munro S, McKeown MR, Wang Y, Christie AL, West N, Cameron MJ, Schwartz B, Heightman TD, La Thangue N, French CA, Wiest O, Kung AL, Knapp S & Bradner JE. Selective inhibition of BET bromodomains. *Nature* **468**(7327): 1067–1073 (2010).

147. McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, Liu Y, Pietra A Della, LaFrance L V., Mellinger M, Duquenne C, Tian X, Kruger RG, McHugh CF, Miller WH, Dhanak D, Verma SK, Tummino PJ, Creasy CL, Graves AP & Diaz E. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**(7427): 108–112 (2012).

148. Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, Johnston LD, Scott MP, Smith JJ, Xiao Y, Jin L, Kuntz KW, Chesworth R, Moyer MP, Bernt KM, Tseng J-C, Kung AL, Armstrong SA, Copeland RA, Richon VM & Pollock RM. Cancer Cell Article Selective Killing of Mixed

Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. *Cancer Cell* **20**(1): 53–65 (2011).

149. Kim KH, Kim W, Howard TP, Vazquez F, Tsherniak A, Wu JN, Wang W, Haswell JR, Walensky LD, Hahn WC, Orkin SH & Roberts CWM. SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of EZH2. *Nat. Med.* **21**(12): 1491–1496 (2015).

150. Cameron EE, Bachman KE, Myöhänen S, Herman JG & Baylin SB. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat. Genet.* **21**(1): 103–107 (1999).

151. Gore SD, Jiemjit A, Silverman LB, Aucott T, Baylin S, Carraway H, Dauses T, Fandy T, Herman J, Karp JE, Licht JD, Murgo AJ, Odchimar-Reissig R, Smith BD, Zwiebel JA & Sugar E. Combined Methyltransferase/Histone Deacetylase Inhibition with 5-Azacitidine and MS-275 in Patients with MDS, CMMoL and AML: Clinical Response, Histone Acetylation and DNA Damage. *Blood* **108**(11): 517–517 (2006).

152. Prebet T, Sun Z, Figueroa ME, Ketterling R, Melnick A, Greenberg PL, Herman J, Juckett M, Wang ES, Smith MR, Malick L, Paietta E, Czader M, Litzow M, Gabrilove J, Erba HP, Gore SD & Tallman MS. Prolonged Administration of Azacitidine With or Without Entinostat for Myelodysplastic Syndrome and Acute Myeloid Leukemia With Myelodysplasia-Related Changes: Results of the US Leukemia Intergroup Trial E1905. *J. Clin. Oncol.* **32**(12): 1242–1248 (2014).

153. Maslak P, Chanel S, Camacho LH, Soignet S, Pandolfi PP, Guernah I, Warrell R & Nimer S. Pilot study of combination transcriptional modulation therapy with sodium phenylbutyrate and 5-azacytidine in patients with acute myeloid leukemia or myelodysplastic syndrome. *Leukemia* **20**(2): 212–217 (2006).

154. Ramalingam SS, Maitland ML, Frankel P, Argiris AE, Koczywas M, Gitlitz B, Thomas S, Espinoza-Delgado I, Vokes EE, Gandara DR & Belani CP. Carboplatin and Paclitaxel in Combination With Either Vorinostat or Placebo for First-Line Therapy of Advanced Non–Small-Cell Lung Cancer. *J. Clin. Oncol.* **28**(1): 56–62 (2009).

155. Matei D, Fang F, Shen C, Schilder J, Arnold A, Zeng Y, Berry WA, Huang T & Nephew KP. Epigenetic Resensitization to Platinum in Ovarian Cancer. *Cancer Res.* **72**(9): 2197–2205 (2012).

156. Fang F, Zuo Q, Pilrose J, Wang Y, Shen C, Li M, Wulfridge P, Matei D, Nephew KP, Fang F, Zuo Q, Pilrose J, Wang Y, Shen C, Li M, Wulfridge P, Matei D & Nephew KP. Decitabine reactivated pathways in platinum resistant ovarian cancer. *Oncotarget* **5**(11): 3579–3589 (2014).

157. Fang F, Munck J, Tang J, Taverna P, Wang Y, Miller DFB, Pilrose J, Choy G, Azab M, Pawelczak KS, VanderVere-Carozza P, Wagner M, Lyons J, Matei D, Turchi JJ & Nephew KP. The Novel, Small-Molecule DNA Methylation Inhibitor SGI-110 as an Ovarian Cancer Chemosensitizer. *Clin. Cancer Res.* **20**(24): 6504–6516 (2014).

158. Danese E & Montagnana M. Epigenetics of colorectal cancer: Emerging circulating diagnostic and prognostic biomarkers. *Ann. Transl. Med.* **5**(13): 279 (2017).

159. Coppedè F, Lopomo A, Spisni R & Migliore L. Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer. *World J. Gastroenterol.* **20**(4): 943–956 (2014).

160. Nagai Y, Sunami E, Yamamoto Y, Hata K, Okada S, Murono K, Yasuda K, Otani K, Nishikawa T, Tanaka T, Kiyomatsu T, Kawai K, Nozawa H, Ishihara S, Hoon DSB, Watanabe T, Nagai Y, Sunami E, Yamamoto Y, Hata K, Okada S, Murono K, Yasuda K, Otani K, Nishikawa T, Tanaka T, Kiyomatsu T, Kawai K, Nozawa H, Ishihara S, Hoon DSB & Watanabe T. LINE-1 hypomethylation status of circulating cell-free DNA in plasma as a biomarker for colorectal cancer. *Oncotarget* **8**(7): 11906–11916 (2017).

161. Ahn JB, Chung WB, Maeda O, Shin SJ, Kim HS, Chung HC, Kim NK & Issa J-PJ. DNA Methylation Predicts Recurrence From Resected Stage III Proximal Colon Cancer. *Cancer* **117**(9): 1847–1854 (2011).

162. Antelo M, Balaguer F, Shia J, Shen Y, Hur K, Moreira L, Cuatrecasa M, Bujanda L, Giraldez MD, Takahashi M, Cabanne A, Barugel ME, Arnold M, Roca EL, Andreu M, Castellvi-Bel S, Llor X, Jover R, Castells A, Boland CR & Goel A. A High Degree of LINE-1 Hypomethylation Is a Unique Feature of Early-Onset Colorectal Cancer. *PLoS One* **7**(9): e45357 (2012).

163. Karczmarski J, Rubel T, Paziewska A, Mikula M, Bujko M, Kober P, Dadlez M & Ostrowski J. Histone H3 lysine 27 acetylation is altered in colon cancer. *Clin. Proteomics* **11**(1): 24 (2014).

164. Nakazawa T, Kondo T, Ma D, Niu D, Mochizuki K, Kawasaki T, Yamane T, Iino H, Fujii H & Katoh R. Global histone modification of histone H3 in colorectal cancer and its precursor lesions. *Hum. Pathol.* **43**(6): 834–842 (2012).

165. Gezer U, Yörüker EE, Keskin M, Burak Kulle C, Dharuman Y & Holdenrieder S. Histone Methylation Marks on Circulating Nucleosomes as Novel Blood-Based Biomarker in Colorectal Cancer. *Int. J. Mol. Sci.* **16**(12): 29654–29662 (2015).

166. Peng Q, Zhang X, Min M, Zou L, Shen P & Zhu Y. The clinical role of microRNA-21 as a promising biomarker in the diagnosis and prognosis of colorectal cancer: a systematic review and meta-analysis. *Oncotarget* **8**(27): 44893–44909 (2017).

167. Toiyama Y, Takahashi M, Hur K, Nagasaka T, Tanaka K, Inoue Y, Kusunoki M, Boland CR & Goel A. Serum mir-21 as a Diagnostic and Prognostic Biomarker in colorectal cancer. *J. Natl. Cancer Inst.* **105**(12): 849–859 (2013).

168. Mima K, Nishihara R, Yang J, Dou R, Masugi Y, Shi Y, Silva A da, Cao Y, Song M, Nowak J, Gu M, Li W, Morikawa T, Zhang X, Wu K, Baba H, Giovannucci EL, Meyerhardt JA, Chan AT, Fuchs CS, Qian ZR & Ogino S. MicroRNA MIR21 (miR-21) and PTGS2 Expression in Colorectal Cancer and Patient Survival. *Clin. Cancer Res.* **22**(15): 3841–3848 (2016).

169. Sun G, Cheng Y-W, Lai L, Huang T-C, Wang J, Wu X, Wang Y, Huang Y, Wang J, Zhang K, Hu S, Yang J-R & Yen Y. Signature miRNAs in colorectal cancers were revealed using a bias reduction small RNA deep sequencing protocol. *Oncotarget* **7**(4): 3857–3872 (2015).

170. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, Yuen ST, Chan TL, Kwong DLW, Au GKH, Liu C-G, Calin GA, Croce CM & Harris CC. MicroRNA Expression Profiles Associated With Prognosis and Therapeutic Outcome in Colon Adenocarcinoma. *JAMA* **299**(4): 425–436 (2008).

171. Heidelberger C, Chaudhuri NK, Danneberg P, Mooren D, Griesbach L, Duschinsky R, Schnitzer RJ, Pleven E & Scheiner J. Fluorinated pyrimidines, a new class of tumour-inhibitory compounds. *Nature* **179**(4561): 663–666 (1957).

172. Longley DB, Harkin DP & Johnston PG. 5-Fluorouracil: Mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**(5): 330–338 (2003).

173. Crea F, Nobili S, Paolicchi E, Perrone G, Napoli C, Landini I, Danesi R & Mini E. Epigenetics and chemoresistance in colorectal cancer: An opportunity for treatment tailoring and novel therapeutic strategies. *Drug Resist. Updat.* **14**(6): 280–296 (2011).

174. Mader RM, Müller M & Steger GG. Resistance to 5-Fluorouracil. *Gen. Pharmacol. Vasc. Syst.* **31**(5): 661–666 (1998).

175. Escalante PI, Quiñones LA & Contreras HR. Epithelial-Mesenchymal Transition and MicroRNAs in Colorectal Cancer Chemoresistance to FOLFOX. *Pharmaceutics* **13**(75): 1–19 (2021).

176. De Mattia E, Cecchin E & Toffoli G. Pharmacogenomics of intrinsic and acquired

pharmacoresistance in colorectal cancer: Toward targeted personalized therapy. *Drug Resist. Updat.* **20**: 39–62 (2015).

177. De Angelis PM, Svendsrud DH, Kravik KL & Stokke T. Cellular response to 5-fluorouracil (5-FU) in 5-FU-resistant colon cancer cell lines during treatment and recovery. *Mol. Cancer* **5**: 20 (2006).

178. Zhou J, Kang Y, Chen L, Wang H, Liu J, Zeng S & Yu L. The Drug-Resistance Mechanisms of Five Platinum-Based Antitumor Agents. *Front. Pharmacol.* **11**: 343 (2020).

179. Dilruba S & Kalayda G V. Platinum-based drugs: past, present and future. *Cancer Chemother. Pharmacol.* **77**(6): 1103–1124 (2016).

180. Harada K, Okamoto W, Mimaki S, Kawamoto Y, Bando H, Yamashita R, Yuki S, Yoshino T, Komatsu Y, Ohtsu A, Sakamoto N & Tsuchihara K. Comparative sequence analysis of patient-matched primary colorectal cancer, metastatic, and recurrent metastatic tumors after adjuvant FOLFOX chemotherapy. *BMC Cancer* **19**(1): 1–11 (2019).

181. Rabik CA & Dolan ME. Molecular mechanisms of resistance and toxicity associated with platinating agents. *Cancer Treat. Rev.* **33**(1): 9–23 (2007).

182. Martinez-Balibrea E, Martínez-Cardus A, Gines A, Ruiz De Porras V, Moutinho C, Layos L, Manzano JL, Buges C, Bystrup S, Esteller M & Abad A. Tumor-related molecular mechanisms of oxaliplatin resistance. *Mol. Cancer Ther.* **14**(8): 1767–1776 (2015).

183. Song I-S, Savaraj N, Siddik ZH, Liu P, Wei Y, Wu CJ & Kuo MT. Role of human copper transporter Ctr1 in the transport of platinum-based antitumor agents in cisplatin-sensitive and cisplatin-resistant cells. *Mol. Cancer Ther.* **3**(12): 1543–1549 (2004).

184. Samimi G, Katano K, Holzer AK, Safaei R & Howell SB. Modulation of the Cellular Pharmacology of Cisplatin and Its Analogs by the Copper Exporters ATP7A and ATP7B. *Mol. Pharmacol.* **66**(1): 25–32 (2004).

185. Martinez-Balibrea E, Martínez-Cardús A, Musulén E, Ginés A, Manzano JL, Aranda E, Plasencia C, Neamati N & Abad A. Increased levels of copper efflux transporter ATP7B are associated with poor outcome in colorectal cancer patients receiving oxaliplatin-based chemotherapy. *Int. J. Cancer* **124**(12): 2905–2910 (2009).

186. Gherman A, Balacescu L, Gheorghe-Cetean S, Vlad C, Balacescu O, Irimie A & Lisencu C. Current and New Predictors for Treatment Response in Metastatic Colorectal Cancer. The Role of Circulating miRNAs as Biomarkers. *Int. J. Mol. Sci.* **21**(6): 2089 (2020).

187. Martinez-Balibrea E, Martínez-Cardus A, Gines A, Ruiz De Porras V, Moutinho C, Layos L, Manzano JL, Buges C, Bystrup S, Esteller M & Abad A. Tumor-related molecular mechanisms of oxaliplatin resistance. *Mol. Cancer Ther.* **14**(8): 1767–1776 (2015).

188. Hatch SB, Swift LP, Caporali S, Carter R, Hill EJ, MacGregor TP, Middleton MR, McHugh PJ & Sharma RA. XPF protein levels determine sensitivity of malignant melanoma cells to oxaliplatin chemotherapy: Suitability as a biomarker for patient selection. *Int. J. Cancer* **134**(6): 1495–1503 (2014).

189. Boyer J, Mclean EG, Aroori S, Wilson P, Mcculla A, Declan Carey P, Longley DB & Johnston PG. Characterization of p53 Wild-Type and Null Isogenic Colorectal Cancer Cell Lines Resistant to 5-Fluorouracil, Oxaliplatin, and Irinotecan. *Clin. Cancer Res.* **10**(6): 2158–2167 (2004).

190. Souglakos J, Androulakis N, Syrigos K, Polyzos A, Ziras N, Athanasiadis A, Kakolyris S, Tsousis S, Kouroussis C, Vamvakas L, Kalykaki A, Samonis G, Mavroudis D & Georgoulias V. FOLFOXIRI (folinic acid, 5-fluorouracil, oxaliplatin and irinotecan) vs FOLFIRI (folinic acid, 5-fluorouracil and irinotecan) as first-line treatment in metastatic colorectal cancer (MCC): a multicentre randomised phase III trial from the Hellenic Oncolog. *Br. J. Cancer* **94**(6): 798–805 (2006).

191. Moutinho C, Martinez-Cardús A, Santos C, Navarro-Pérez V, Martínez-Balibrea E, Musulen E, Carmona FJ, Sartore-Bianchi A, Cassingena A, Siena S, Elez E, Tabernero J, Salazar R, Abad A & Esteller M. Epigenetic inactivation of the BRCA1 interactor SRBC and resistance to oxaliplatin in colorectal cancer. *J. Natl. Cancer Inst.* **106**(1): djt322 (2014).

192. Smith NF, Figg WD & Sparreboom A. Pharmacogenetics of irinotecan metabolism and transport: An update. *Toxicol. Vitr.* **20**(2): 163–175 (2006).

193. Kciuk M, Marciniak B & Kontek R. Irinotecan—still an important player in cancer chemotherapy: A comprehensive overview. *Int. J. Mol. Sci.* **21**(14): 1–21 (2020).

194. Meisenberg C, Ashour ME, El-Shafie L, Liao C, Hodgson A, Pilborough A, Khurram SA, Downs JA, Ward SE & El-Khamisy SF. Epigenetic changes in histone acetylation underpin resistance to the topoisomerase I inhibitor irinotecan. *Nucleic Acids Res.* **45**(3): 1159–1176 (2016).

195. Xu Y & Villalona-Calero MA. Irinotecan: mechanisms of tumor resistance and novel strategies for modulating its activity. *Ann. Oncol.* **13**(12): 1841–1851 (2002).

196. Bao X, Wu J, Kim S, LoRusso P & Li J. Pharmacometabolomics Reveals Irinotecan Mechanism of Action in Cancer Patients. *J. Clin. Pharmacol.* **59**(1): 20–34 (2019).

197. Martinez-Balibrea E, Abad A, Martínez-Cards A, Ginés A, Valladares M, Navarro M, Aranda E, Marcuello E, Benavides M, Massutí B, Carrato A, Layos L, Manzano JL & Moreno V. UGT1A and TYMS genetic variants predict toxicity and response of colorectal cancer patients treated with first-line irinotecan and fluorouracil combination therapy. *Br. J. Cancer* **103**(4): 581–589 (2010).

198. Van Ark-Offe J, Kedde MA, Van Der Vijgh W, Dingemans A-M, Jansen W, Pinedo HM, Boven E & Giaccone G. Determinants of CPT-11 and SN-38 activities in human lung cancer cells. *Br. Joumal Cancer* **77**(12): 2171–2176 (1998).

199. Wu MH, Yan B, Humerickhouse R & Dolan ME. Irinotecan Activation by Human Carboxylesterases in Colorectal Adenocarcinoma Cells. *Clin. Cancer Res.* **8**(8): 2696–2700 (2002).

200. Liu LF, Desai SD, Li T-K, Mao Y, Sun M & Sim S-P. Mechanism of Action of Camptothecin. *Ann. N. Y. Acad. Sci.* **922**(1): 1–10 (2000).

201. Takeba Y, Kumai T, Matsumoto N, Nakaya S, Tsuzuki Y, Yanagida Y & Kobayashi S. Irinotecan Activates p53 With Its Active Metabolite, Resulting in Human Hepatocellular Carcinoma Apoptosis. *J. Pharmacol. Sci.* **104**(3): 232–242 (2007).

202. Terme J-M, Millán-Ariño L, Mayor R, Luque N, Izquierdo-Bouldstridge A, Bustillos A, Sampaio C, Canes J, Font I, Sima N, Sancho M, Torrente L, Forcales S, Roque A, Suau P & Jordan A. Dynamics and dispensability of variant-specific histone H1 Lys-26/Ser-27 and Thr-165 post-translational modifications. *FEBS Lett.* **588**(14): 2353–2362 (2014).

203. Christophorou MA, Castelo-Branco G, Halley-Stott RP, Oliveira CS, Loos R, Radzisheuskaya A, Mowen KA, Bertone P, Silva JCR, Zernicka-Goetz M, Nielsen ML, Gurdon JB & Kouzarides T. Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature* **507**(7490): 104–108 (2014).

204. Cantariño N, Musulén E, Valero V, Peinado MA, Perucho M, Moreno V, Forcales S-V, Douet J & Buschbeck M. Downregulation of the Deiminase PADI2 Is an Early Event in Colorectal Carcinogenesis and Indicates Poor Prognosis. *Mol. Cancer Res.* **14**(9): 841–848 (2016).

205. Fellmann C, Hoffmann T, Sridhar V, Hopfgartner B, Muhar M, Roth M, Lai DY, Barbosa IAM, Kwon JS, Guan Y, Sinha N & Zuber J. An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**(6): 1704–1713 (2013).

206. Ulukaya E, Colakogullari M & Wood EJ. Interference by anti-cancer chemotherapeutic agents in the MTT-tumor chemosensitivity assay. *Chemotherapy* **50**(1): 43–50 (2004).

207. Luis C, Castaño-Guerrero Y, Soares R, Sales G & Fernandes R. Avoiding the interference of doxorubicin with mtt measurements on the mcf-7 breast cancer cell line. *Methods Protoc.* **2**(2): 1–5 (2019).

208. Kumar N, Afjei R, Massoud TF & Paulmurugan R. Comparison of cell-based assays to quantify treatment effects of anticancer drugs identifies a new application for Bodipy-L-cystine to measure apoptosis. *Sci. Rep.* **8**(1): 1–11 (2018).

209. Padhani AR & Ollivier L. The RECIST criteria: implications for diagnostic radiologists. *Br. J. Radiol.* **74**(887): 983–986 (2001).

210. Okita A, Takahashi S, Ouchi K, Inoue M, Watanabe M, Endo M, Honda H, Yamada Y & Ishioka C. CMS classification of CRC as a predictive factor for chemotherapeutic efficacy against metastatic CRC. *Oncotarget* **9**(27): 18698–18711 (2018).

211. Del Rio M, Molina F, Bascoul-Mollevi C, Copois V, Bibeau F, Chalbos P, Bareil C, Kramar A, Salvetat N, Fraslon C, Conseiller E, Granci V, Leblanc B, Pau B, Martineau P & Ychou M. Gene Expression Signature in Advanced Colorectal Cancer Patients Select Drugs and Response for the Use of Leucovorin, Fluorouracil, and Irinotecan. *J. Clin. Oncol.* **25**(7): 773–780 (2007).

212. Combes E, Andrade AF, Tosi D, Michaud HA, Coquel F, Garambois V, Desigaud D, Jarlier M, Coquelle A, Pasero P, Bonnefoy N, Moreaux J, Martineau P, Rio M Del, Beijersbergen RL, Vezzio-Vie N & Gongora C. Inhibition of ataxia-telangiectasia mutated and Rad3-related (ATR) overcomes oxaliplatin resistance and promotes antitumor immunity in colorectal cancer. *Cancer Res.* **79**(11): 2933–2946 (2019).

213. Job A, Schmitt LM, von Wenserski L, Lankat-Buttgereit B, Gress TM, Buchholz M & Gallmeier E. Inactivation of PRIM1 Function Sensitizes Cancer Cells to ATR and CHK1 Inhibitors. *Neoplasia* **20**(11): 1135–1143 (2018).

214. Stronach EA, Alfraidi A, Rama N, Datler C, Studd JB, Agarwal R, Guney TG, Gourley C, Hennessy BT, Mills GB, Mai A, Brown R, Dina R & Gabra H. HDAC4-Regulated STAT1 Activation Mediates Platinum Resistance in Ovarian Cancer. *Cancer Res.* **71**(13): 4412–4422 (2011).

215. Kaewpiboon C, Srisuttee R, Malilas W, Moon J, Oh S, Jeong HG, Johnston RN, Assavalapsakul W & Chung Y. Upregulation of Stat1-HDAC4 confers resistance to etoposide through enhanced multidrug resistance 1 expression in human A549 lung cancer cells. *Mol. Med. Rep.* **11**(3): 2315–2321 (2015).

216. cBioPortal for Cancer Genomics. https://www.cbioportal.org/.

217. The Cancer Genome Atlas - Using TCGA - National Cancer Institute. https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/using-tcga.

218. Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA & Livingston DM. BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* **105**(1): 149–160 (2001).

219. Cantor S, Drapkin R, Zhang F, Lin Y, Han J, Pamidi S & Livingston DM. The BRCA1-associated protein BACH1 is a DNA helicase targeted by clinically relevant inactivating mutations. *Proc. Natl. Acad. Sci. U. S. A.* **101**(8): 2357–2362 (2004).

220. Friedman LS, Ostermeyer EA, Szabo CI, Dowd P, Lynch ED, Rowell SE & King M-C. Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. *Nat. Genet.* **8**(4): 399–404 (1994).

221. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, Bell R, Rosenthal J, Hussey C, Tran T, McClure M, Frye C, Hattier T, Phelps R, Haugen-Strano A, Katcher H, Yakumo K, Gholami Z, Shaffer D, Stone S, Bayer S, Wray C,

Bogden R, Dayananth P, Ward J, Tonin P, Narod S, Bristow PK, Norris FH, Helvering L, Morrison P, Rosteck P, Lai M, Barrett JC, Lewis C, Neuhausen S, Cannon-Albright L, Goldgar D, Wiseman R, Kamb A & Skolnick MH. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**(5182): 66–71 (1994).

222.   Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J, Miki Y, Eddington K, McClure M, Frye C, Weaver-Feldhaus J, Ding W, Gholami Z, Soderkvist P, Terry L, Jhanwar S, Berchuck A, Iglehart JD, Marks J, Ballinger DG, Barrett JC, Skolnick MH, Kamb A & Wiseman R. BRCA1 mutations in primary breast and ovarian carcinomas. *Science* **266**(5182): 120–122 (1994).

223.   Gayther SA, Harrington P, Russell P, Kharkevich G, Garkavtseva RF & Ponder BAJ. Rapid detection of regionally clustered germ-line BRCA1 mutations by multiplex heteroduplex analysis. UKCCCR Familial Ovarian Cancer Study Group. *Am. J. Hum. Genet.* **58**(3): 451–456 (1996).

224.   Moyer CL, Ivanovich J, Gillespie JL, Doberstein R, Radke MR, Richardson ME, Kaufmann SH, Swisher EM & Goodfellow PJ. Rare BRIP1 Missense Alleles Confer Risk for Ovarian and Breast Cancer. *Cancer Res.* **80**(4): 857–867 (2020).

225.   Shiozaki EN, Gu L, Yan N & Shi Y. Structure of the BRCT Repeats of BRCA1 Bound to a BACH1 Phosphopeptide: Implications for Signaling. *Mol. Cell* **14**(3): 405–412 (2004).

226.   Peng M, Litman R, Jin Z, Fong G & Cantor SB. BACH1 is a DNA repair protein supporting BRCA1 damage response. *Oncogene* **25**(15): 2245–2253 (2006).

227.   Xie J, Litman R, Wang S, Peng M, Guillemette S, Rooney T & Cantor SB. Targeting the FANCJ-BRCA1 interaction promotes a switch from recombination to polη-dependent bypass. *Oncogene* **29**(17): 2499–2508 (2010).

228.   International Mouse Phenotyping Consortium. https://www.mousephenotype.org/data/genes/MGI:2442836#phenotypesTab.

229.   Fujita Y, Hayashi T, Kiyomitsu T, Toyoda Y, Kokubu A, Obuse C & Yanagida M. Priming of Centromere for CENP-A Recruitment by Human hMis18α, hMis18β, and M18BP1. *Dev. Cell* **12**(1): 17–30 (2007).

230.   Nardi IK, Zasadzińska E, Stellfox ME, Knippler CM & Foltz DR. Licensing of Centromeric Chromatin Assembly through the Mis18α-Mis18β Heterotetramer. *Mol. Cell* **61**(5): 774–787 (2016).

231.   Stellfox ME, Nardi IK, Knippler CM & Foltz DR. Differential Binding Partners of the Mis18α/β YIPPEE Domains Regulate Mis18 Complex Recruitment to Centromeres. *Cell Rep.* **15**(10): 2127–2135 (2016).

232.   Pan D, Klare K, Petrovic A, Take A, Walstein K, Singh P, Rondelet A, Bird AW & Musacchio A. CDK-regulated dimerization of M18BP1 on a Mis18 hexamer is necessary for CENP-A loading. *Elife* **6**: e23352 (2017).

233.   Spiller F, Medina-Pritchard B, Abad MA, Wear MA, Molina O, Earnshaw WC & Jeyaprakash AA. Molecular basis for Cdk1-regulated timing of Mis18 complex assembly and CENP-A deposition. *EMBO Rep.* **18**(6): 894–905 (2017).

234.   Liebelt F, Jansen NS, Kumar S, Gracheva E, Claessens LA, Verlaan-de Vries M, Willemstein E & Vertegaal ACO. The poly-SUMO2/3 protease SENP6 enables assembly of the constitutive centromere-associated network by group deSUMOylation. *Nat. Commun.* **10**(1): 1–18 (2019).

235.   Fu H, Liu N, Dong Q, Ma C, Yang J, Xiong J, Zhang Z, Qi X, Huang C & Zhu B. SENP6-mediated M18BP1 deSUMOylation regulates CENP-A centromeric localization. *Cell Res.* **29**(3): 254–257 (2019).

236.   Bolhaqueiro ACF, Ponsioen B, Bakker B, Klaasen SJ, Kucukkose E, van Jaarsveld RH, Vivié J,

Verlaan-Klink I, Hami N, Spierings DCJ, Sasaki N, Dutta D, Boj SF, Vries RGJ, Lansdorp PM, van de Wetering M, van Oudenaarden A, Clevers H, Kranenburg O, Foijer F, Snippert HJG & Kops GJPL. Ongoing chromosomal instability and karyotype evolution in human colorectal cancer organoids. *Nat. Genet.* **51**(5): 824–834 (2019).

237. Burrell RA, McClelland SE, Endesfelder D, Groth P, Weller M-C, Shaikh N, Domingo E, Kanu N, Dewhurst SM, Gronroos E, Chew SK, Rowan AJ, Schenk A, Sheffer M, Howell M, Kschischo M, Behrens A, Helleday T, Bartek J, Tomlinson IP & Swanton C. Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**(7438): 492–496 (2013).

238. Pussila M, Törönen P, Einarsdottir E, Katayama S, Krjutškov K, Holm L, Kere J, Peltomäki P, Mäkinen MJ, Linden J & Nyström M. Mlh1 deficiency in normal mouse colon mucosa associates with chromosomally unstable colon cancer. *Carcinogenesis* **39**(6): 788–797 (2018).

239. Sun SY, Hu XT, Yu XF, Zhang YY, Liu XH, Liu YH, Wu SH, Li YY, Cui SX & Qu XJ. Nuclear translocation of ATG5 induces DNA mismatch repair deficiency (MMR-D)/microsatellite instability (MSI) via interacting with Mis18α in colorectal cancer. *Br. J. Pharmacol.* **178**(11): 2351–2369 (2021).

240. Gordon RR & Nelson PS. Cellular senescence and cancer chemotherapy resistance. *Drug Resist. Updat.* **15**(1–2): 123–131 (2012).

241. Guillon J, Petit C, Toutain B, Guette C, Lelièvre E & Coqueret O. Chemotherapy-induced senescence, an adaptive mechanism driving resistance and tumor heterogeneity. *Cell Cycle* **18**(19): 2385–2397 (2019).

242. Jelluma N & Kops GJPL. Collateral Genome Instability by DNA Damage in Mitosis. *Cancer Discov.* **4**(11): 1256–1258 (2014).

243. Wilhelm T, Said M & Naim V. DNA Replication Stress and Chromosomal Instability: Dangerous Liaisons. *Genes (Basel).* **11**(6): 642 (2020).

244. Shu X, Zhao Y, Sun Y, Zhong L, Cheng Y, Zhang Y, Ning K, Tao Q, Wang Y & Ying Y. The epigenetic modifier PBRM1 restricts the basal activity of the innate immune system by repressing retinoic acid-inducible gene-I-like receptor signalling and is a potential prognostic biomarker for colon cancer. *J. Pathol.* **244**(1): 36–48 (2018).

245. COSMIC | Catalogue of Somatic Mutations in Cancer. https://cancer.sanger.ac.uk/cosmic.

246. Kadoch C & Crabtree GR. Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**(5): e1500447 (2015).

247. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, Grynspan D, Saloustros E, Nadaf J, Rivera B, Gilpin C, Castellsagué E, Silva-Smith R, Plourde F, Wu M, Saskin A, Arseneault M, Karabakhtsian RG, Reilly EA, Ueland FR, Margiolaki A, Pavlakis K, Castellino SM, Lamovec J, Mackay HJ, Roth LM, Ulbright TM, Bender TA, Georgoulias V, Longy M, Berchuck A, Tischkowitz M, Nagel I, Siebert R, Stewart CJR, Arseneau J, McCluggage WG, Clarke BA, Riazalhosseini Y, Hasselblatt M, Majewski J & Foulkes WD. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. *Nat. Genet.* **46**(5): 438–443 (2014).

248. Ramos P, Karnezis AN, Craig DW, Sekulic A, Russell ML, Hendricks WPD, Corneveaux JJ, Barrett MT, Shumansky K, Yang Y, Shah SP, Prentice LM, Marra MA, Kiefer J, Zismann VL, McEachron TA, Salhia B, Prat J, D'Angelo E, Clarke BA, Pressey JG, Farley JH, Anthony SP, Roden RBS, Cunliffe HE, Huntsman DG & Trent JM. Small cell carcinoma of the ovary, hypercalcemic type, displays frequent inactivating germline and somatic mutations in SMARCA4. *Nat. Genet.* **46**(5): 427–429 (2014).

249. Xue Y, Meehan B, Fu Z, Wang XQD, Fiset PO, Rieker R, Levins C, Kong T, Zhu X, Morin G, Skerritt

L, Herpel E, Venneti S, Martinez D, Judkins AR, Jung S, Camilleri-Broet S, Gonzalez AV, Guiot M-C, Lockwood WW, Spicer JD, Agaimy A, Pastor WA, Dostie J, Rak J, Foulkes WD & Huang S. SMARCA4 loss is synthetic lethal with CDK4/6 inhibition in non-small cell lung cancer. *Nat. Commun.* **10**(1): 1–13 (2019).

250. Raab JR, Runge JS, Spear CC & Magnuson T. Co-regulation of transcription by BRG1 and BRM, two mutually exclusive SWI/SNF ATPase subunits. *Epigenetics Chromatin* **10**(1): 62 (2017).

251. Connectivity Map (CMAP) | Broad Institute. https://www.broadinstitute.org/connectivity-map-cmap.

252. REPURPOSING [clue.io]. https://clue.io/repurposing-app.

253. Berg KCG, Eide PW, Eilertsen IA, Johannessen B, Bruun J, Danielsen SA, Bjørnslett M, Meza-Zepeda LA, Eknæs M, Lind GE, Myklebost O, Skotheim RI, Sveen A & Lothe RA. Multi-omics of 34 colorectal cancer cell lines - a resource for biomedical studies. *Mol. Cancer* **16**(1): 1–16 (2017).

254. Robinson MD & Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**(3): 1–9 (2010).

255. Guzmán C, Bagga M, Kaur A, Westermarck J & Abankwa D. ColonyArea: An ImageJ Plugin to Automatically Quantify Colony Formation in Clonogenic Assays. *PLoS One* **9**(3): e92444 (2014).