G9a/EHMT2 methyltransferase activity controls stem-like identity and tumor-initiating function in human colorectal cancer

Aïcha Zouggar

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Master of Science degree in Cellular and Molecular Medicine
Specialization in Bioinformatics

Department of Cellular and Molecular Medicine
Faculty of Medicine
University of Ottawa
Ottawa, Ontario, Canada

uOttawa

Abstract

Colorectal tumors are hierarchically organized and governed by populations of self-renewing cancer stem cells, representing one of the deadliest types of cancers worldwide. Emergence of a cancer stem-like phenotype depends on epigenetic reprogramming, associated with profound transcriptional changes. As described for pluripotent reprogramming, epigenetic modifiers play a key role in developing and maintaining cancer stem cells by establishing embryonic stem-like transcriptional programs, thus altering the balance between self-renewal and differentiation. Through my work, I have identified overexpression of histone methyltransferase G9a as a risk factor for colorectal cancer, associated with shorter relapse-free survival. Moreover, using human transformed pluripotent cells as a surrogate model for cancer stem cells, I demonstrate that G9a activity is essential for the maintenance of an embryonic stem-like transcriptional signature that is required to promote self-renewal, tumorigenicity and an undifferentiated state. Such a role was also applicable to colorectal cancer, where inhibitors of G9a histone methyltransferase function induced intestinal differentiation while restricting tumor-initiating activity in patient-derived colorectal tumor samples. By integrating transcriptome profiling with G9a/H3K9me2 loci co-occupancy, the canonical Wnt pathway, epithelial-to-mesenchyme transition and extracellular matrix organization were identified as potential targets of such a chromatin regulation mechanism in colorectal cancer stem cells. Considering such novel insights on the role of G9a as a driver of the cancer stem cell phenotype, as well as a promoter of self-renewal, tumorigenicity and an undifferentiated state, I established and executed a multi-step drug screening pipeline to identify new repurposed drugs that selectively alter G9a functions in human CSCs. This pipeline revealed 3 new drug candidates that inhibit H3K9me2 deposition and impair human CSCs in culture. Future in-depth characterization of those candidates will represent an important step toward the development of novel CSC-targeting therapeutics.

Table of Contents

List of Figures

List of abbreviations

ALCAM	Activated leukocyte cell adhesion molecule
ALK	Anaplastic lymphoma receptor tyrosine kinase
AML	Acute myeloid leukemia
APC	Adenomatous polyposis coli
CCSCs	Colorectal cancer stem cells
CICs	Cancer-initiating cells
COAD	Colon adenocarcinoma
CRC	Colorectal cancer
CSC	Cancer stem cell
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DMSO	Dimethyl sulfoxide
DPP4	Dipeptidyl peptidase 4
EC 50	Half maximal effective concentration
EMT	Epithelial to mesenchymal transition
ESCs	Embryonic stem cells
FOLFIRI	Folinic acid/fluorouracil/irinotecan
FOLFOX	Folinic acid/fluorouracil/oxaliplatin
G9a/EHMT2	Euchromatin histone methyl transferase 2
GLP/EHMT1	Euchromatin histone methyl transferase 1
GSEA	Gene set enrichment analysis
H3K9me2	Histone 3 lysine 9 dimethylation
hESCs	Human embryonic stem cells
HMT	Histone methyltransferase
HSCs	Hematopoietic stem cells
IC50	Half maximum inhibitory concentrations
LINE1	Long interspersed element -1
KDM	Lysine demethylase
KRAS	Kirsten rat sarcoma

KRT20	Keratin 20
MAPK	Mitogen activated protein kinase
mRNAsi	Transcriptional stem cell index
OCT-4	Octamer binding transcription factor 4
OSGIN1	Oxidative stress-induced growth inhibitor 1
PDO	Patient derived organoids
PI3K	Phosphoinositide 3 kinase
PKC	Protein kinase C
PTK6	Protein tyrosine kinase 6
RFS	Relapse free survival
RNA	Ribonucleic acid
TCGA	The cancer genome atlas
t-hESCs	Transformed human embryonic stem cells
TP53	Tumor protein 53
Wnt	Wingless related integration sites

Acknowledgments

First and foremost, I would like to thank Dr. Yannick Benoit for his time, guidance and support during my master's journey. His extensive knowledge and experience as well as his continuous support and enthusiasm for my M.Sc. study have allowed me to gain experience in and be exposed to a new upcoming field. I would also like to thank him for giving me a chance to prove myself in his laboratory. He provides me with exceptional mentorship in developing my skills as a scientist, and I truly appreciate it.

I thank Dr. Marjorie Brand for being my co supervisor. I also thank my Thesis Advisory Committee members Dr. Mario Tiberi and Dr. Theodore Perkins for the time they took out of their busy schedules to train and provide me with assistance and for their guidance and faith in me and in my potential to develop my skills.

Also, I would like to acknowledge the importance of support from my labmates. These amazing people were always an example of motivation and perseverance for me.

Special thanks to my parents who made it possible for me to study and contribute to the science. I would, also, like to thank them for the inspiration, support and encouragement throughout this process.

I hope my scientific path will continue and I will be able to achieve my goal to become a successful scientist.

Introduction

1.1. Colorectal cancer

Cancer is a heterogeneous disease, commonly believed to solely arise from the acquisition of genetic mutations, leading to a loss of functionality of genes that prevent uncontrolled cell growth (tumor suppressor genes), as well as from deregulated activity of genes that promote proliferation (oncogenes)[1]. Colorectal cancer (CRC) is one of the most lethal and prevalent cancers globally, leading to tumor-related morbidity and mortality [2,3]. Therefore, it is important to better understand the molecular mechanisms that contribute to CRC aggressiveness. Mainly, the development of CRC is a complex and multistep process mediated by a variety of factors including the dysregulation of genetic and epigenetic mechanisms under the influence of the microenvironment[2]. These dysfunctions involve alterations in the normal colonic epithelium, leading to the development of colorectal adenomas and invasive adenocarcinomas[3,4]. Traditionally, CRC was proposed to progress according to a sequential acquisition of specific genetic alterations[5,6] (**Figure 1**). However, recent studies demonstrate that genetic mutations play only part of the role in carcinogenesis and that epigenetics is strongly implicated in the initiation and progression of CRC[7-9]. Moreover, new concepts such as cancer cell dormancy, in which cancer cell division stops and cells enter a quiescent-like state, challenge the contemporary CRC carcinogenic cascade[10]. Despite new understandings of the genetic contribution to CRC, the introduction of screening tests and improvements to treatment, colorectal cancer remains a heavy burden on society and health care systems. While the majority of CRC patients are diagnosed with localized neoplasia, for which treatment options mainly include ablative surgery combined with cytotoxic chemotherapy and/or radiotherapy, advanced/late-stage cases remain very difficult to treat[11, 12]. In advanced and/or metastatic stages of CRC, standard chemotherapeutics (*e.g.* FOLFOX and FOLFIRI regimens) can still be effective at debulking tumor masses. Specifically, regimens of this nature consist of multiple drugs, each acting along unique anticancer

axis, ultimately, targeting actively dividing cells through inhibition of mitotic processes[13]. However, these therapies largely fail to eradicate the disease and many patients struggle with important side effects that destroy their quality of life[23].

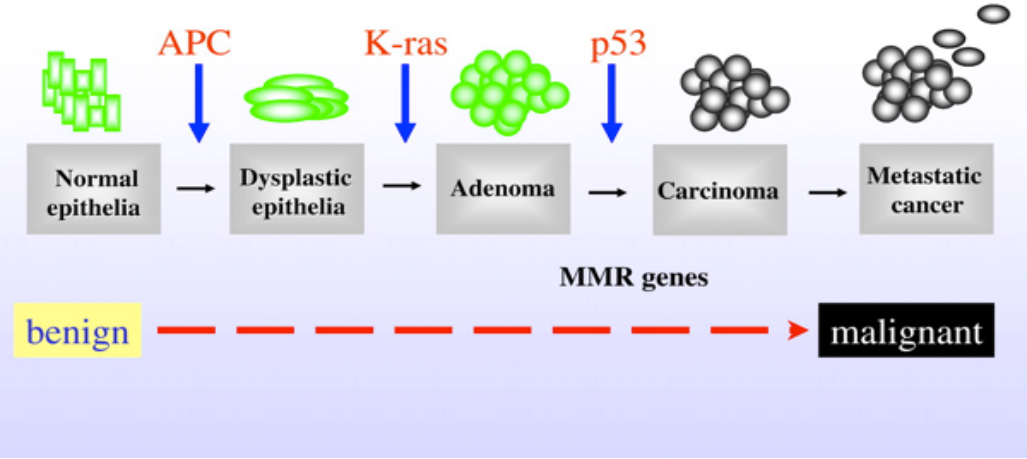


Figure 1. Colorectal cancer progression model.

Colorectal cancer (CRC) originates from hyper proliferative regions in normal colonic mucosa, known as polys, that develop into adenoma and then carcinoma. A genetic model describing the transition from healthy colonic epithelia through dysplastic adenoma to malignant cancer has been historically accepted, implicating sequential mutations in APC, K-ras and p53[4].

1.2. Cancer stem cells and stemness signature

Research advances in cancer biology have provided insights into the phenotypic complexity of cancer cells within a single tumor. This cell-to-cell diversity emphasizes the importance of characterizing tumor heterogeneity, often presenting as a major limitation for the efficacy of targeted therapies[13]. This histological complexity was, at first, defined as the result of clonal evolution, where cells progressively accumulate genetic mutations, gaining growth advantages due to micro-environmentally driven selective pressure[13]. Alternatively, a second model posits that cancer's hierarchical organization is governed by a subpopulation of tumorigenic cells called cancer stem cells (CSCs)[2,14].

CSCs are a small subpopulation of cells possessing the capacity to initiate the growth of a tumor and exhibiting embryonic stem cell (ESC)-like transcriptional characteristics[15]. Specifically, CSCs are defined by their ability to both self-renew and give rise to undifferentiated tumor progenitor cells. CSCs originate from either normal progenitor cells or cancer cells acquiring stem-like properties. Moreover, CSCs were generally shown to be resistant to therapies that target rapidly dividing cells[14,16]. Specifically, this chemotherapy and radiotherapy resistance is mediated by many mechanisms such as acquisition of dormancy, increase of DNA repair and reduction of apoptosis susceptibility.

The idea that tumors are driven by CSCs, maintaining the ability to self-renew, is now widely accepted for many cancers[17] (**Figure 2**). In addition, conventional chemotherapeutics were suggested to enrich CSCs within post-treatment residual tumor cell populations, providing the basis for cancer relapse[18] . As illustrated in **Figure 3**, CSCs remain in the tumor tissue after chemotherapy and expand[18].

The existence of a CSC populations was first studied in the context of human leukemia[19]. Thought to drive tumor growth and disease progression, these putative originators of cancer were first isolated from transformed hematopoietic tissue that manifests disease such as adult acute myeloid leukemia (AML)[19].

It has been shown that these CSC subpopulations, namely leukemic stem cells (LSCs), present a specific cell surface phenotype ($CD34^+$ $CD38^-$). Interestingly, they display a greater capacity for stem cell self-renewal as compared to normal adult bone marrow cells, giving rise to uncontrolled amplification of differentiated cell populations with altered molecular and cellular phenotypes[19,21]. The resulting tumor heterogeneity is explained by clonal evolution due to the genetic diversity and to the hierarchical organization governed by CSCs[20].

Numerous studies have confirmed the existence of CSCs in most types of malignancies, including ovarian, pancreatic, brain, lung and colon cancers[18,21-23]. The identification of a CSC population in human colon cancer was first described in 2007, demonstrating the existence of self-renewing cells enriched in frequency within the CD133-positive tumor fraction[16]. CD133 is a surface protein that has also been shown to be expressed by normal intestinal stem cells[24]. CD133-positive cells were shown to be capable of initiating tumor growth in murine xenograft models[21]. They are also shown to be responsible for driving tumor growth following chemotherapy, explaining the inability of chemotherapeutic agents to improve overall survival despite tumor regression[25].

Tumor recurrence following treatment remains a major clinical challenge. Evidence from xenograft models and human trials indicates selective enrichment of cancer-initiating cells (CICs) in tumors that survive therapy[16]. These studies predict that targeting self-renewal, the key 'stemness' property of CSCs, may represent a new paradigm in cancer therapy[26]. For example, the molecular analyses of functionally defined leukemic stem cell (LSC) populations from acute myeloid leukemia (AML) patients led to the generation of an LSC signature that largely reflects a self-renewal or stemness signature[27,28]. This signature was found to be a strong predictor of poor prognosis, with the implication that it may be possible to identify patients at highest risk[14]. Moreover, substantial studies have identified gene signatures in malignant tumors that are associated with human ESCs, suggesting a molecular relationship between aggressive cancers and

pluripotency[29]. These findings suggest that targeting self-renewal pathways may represent one of the most effective strategies for eradicating CSCs. Indeed, the ESC-like transcriptional program is activated in diverse human epithelial cancers and strongly predicts metastasis and death[30].

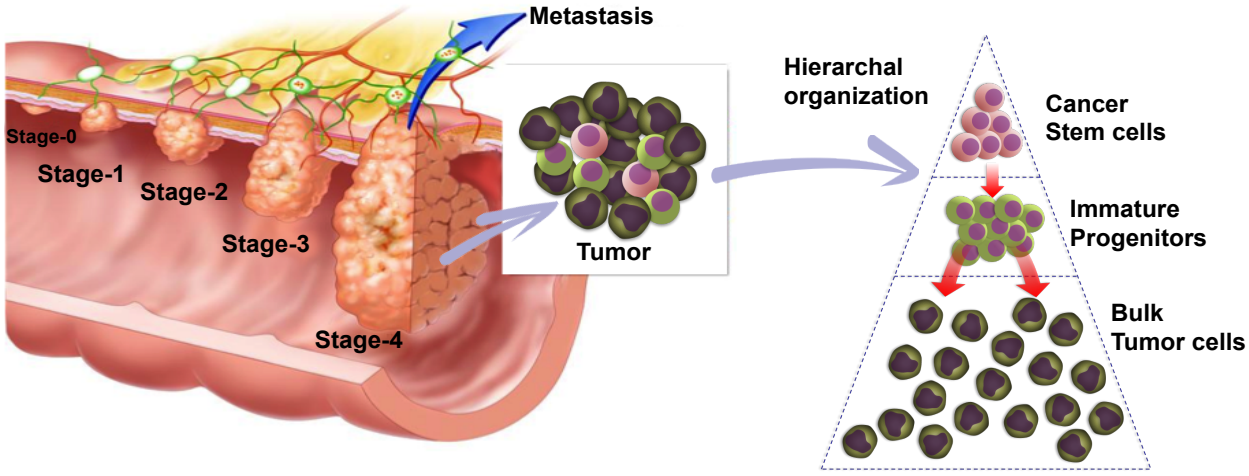


Figure 2. Colorectal tumor heterogeneity.

Colorectal tumors are heterogeneous and organized according to a cellular hierarchy where cancer stem cells (CSCs) at the top are responsible for tumor initiation and supporting tumor growth.

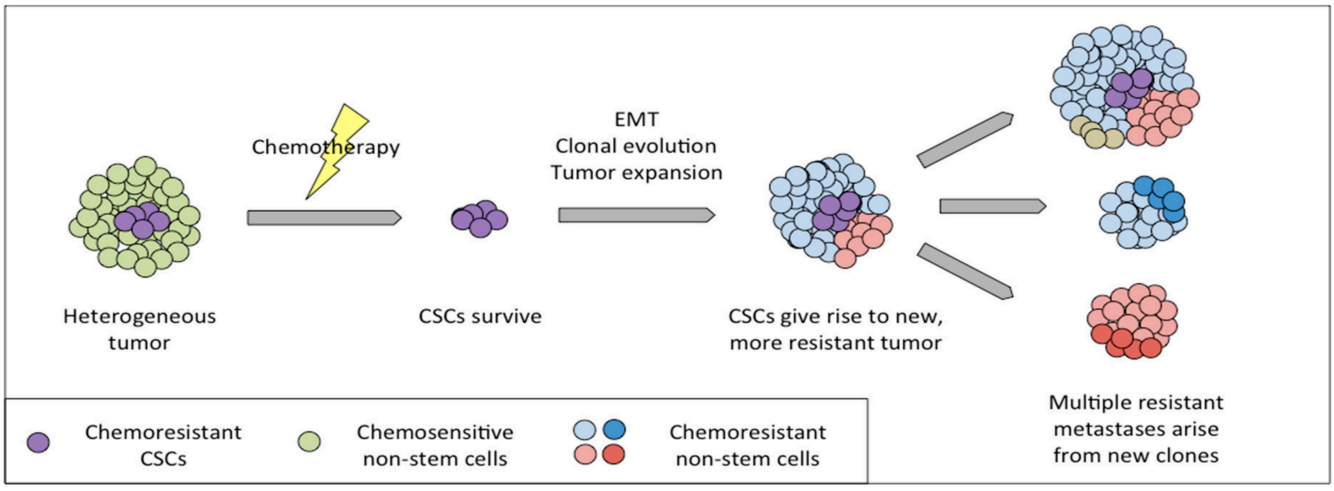


Figure 3. Cancer stem cell (CSC) contributes to treatment resistance and metastasis.

Cancer stem cells (CSCs), undifferentiated cancer cells with self-renewal and tumorigenic activity (purple), are resistant to conventional chemotherapy. Post-treatment, surviving CSCs initiate recurring tumor growth and re-establish a chemo-resistant heterogeneity (dark red and dark blue), which can eventually lead to metastasis[18].

Human ESCs can acquire genetic and epigenetic changes that promote enhanced self-renewal, proliferation and tumor-initiating cell capacity, making them vulnerable to transformation[29, 31]. These observations suggest that the regulatory networks controlling the function of stem cells may also be active in certain tumors[32].

In the longer term, detailed characterization of the stem-cell regulatory networks active in cancer is likely to identify powerful diagnostic and prognostic markers and, quite possibly, attractive targets for therapeutic intervention[32]. Clearly, targeting CSCs is predicted to have widespread clinical implications, since it is assumed that these cells are involved in tumour progression and metastatic dissemination[5]. We speculate that for most tumor types, it will still prove necessary to test novel anti-CSC therapies in combination with tumor debulking (non-CSC specific) therapies.

Taken together, herein summarized discoveries that are helping to clarify colorectal CSC (CCSC) biology and enhance our ability to isolate the subpopulations of metastasising, tumor-initiating and chemo-resistant cells for more reliable and in-depth interrogation. Further understanding of CSCs contribution to CRC heterogeneity, and the mechanisms behind acquisition of therapeutic resistance may lead to novel CSC-directed therapy.

1.3. Epigenetic mechanisms affecting colorectal CSC maintenance

Colon CSCs are distinct from normal stem cells by the way they organize their chromatin and regulate gene expression, conferring them with CSC-specific oncogenic properties[33]. Although mutated genes unequivocally contribute to the initiation and progression of cancer, epigenetic-based mechanisms have been shown to play a key role in the formation and the regulation of tumor-initiating cell subpopulations (CSCs)[34]. Functionally, epigenetic mechanisms refer to the regulation of gene expression

that does not involve changing the sequence of DNA[35]. These epigenetic events are essential to regulate the condensation state of chromatin and, hence, to regulate gene accessibility. This is specifically mediated by covalent histone modifications, DNA methylation and expression of non-coding RNAs[36,37].

Thus, the epigenetic machinery controls the expression of genes by attaching or removing chemical groups to DNA and histones, consequently modifying the structure and accessibility of DNA inside the nucleus. The most documented epigenetic modifications in the literature are currently methylation, phosphorylation, and acetylation[38].

1) Methylation involves the addition of a methyl group (CH3), catalyzed by an enzyme called methyltransferase, on DNA or specific lysine or arginine residues of histones. Methyl groups can be removed by a variety of demethylase enzymes. DNA methylation is generally associated with chromatin condensation and transcriptional repression. However, histone methylation can be linked with both active or silenced chromatin, depending on which residues are methylated and to what extent methyl groups are added (mono (me1), di (me2), or tri-methylation (me3))[39,40]. For instance, methylation of histone H3 lysine 4 (H3K4) is generally associated with transcriptionally active genes, whereas methylation of H3K9 and H3K27 are generally hallmarks of condensed chromatin at silent loci (**Figure 4**)[36-41].

2) Phosphorylation refers to kinases adding a phosphoryl group (PO3) on targeted histone residues, mainly leading to chromatin loosening and enhanced accessibility for transcriptional machinery. Phosphoryl groups can be removed by phosphatases.

3) Histone acetylation, catalyzed by histone acetyltransferases (HAT), adds an acetyl group (CH3CO) to specific lysine residues to reduce the positive charge of histones, therefore, decreasing its affinity for negatively-charged DNA. Histone acetylation typically opens up the chromatin and promotes gene transcription. Acetyl groups can be removed by deacetylases[3,33].

Histones can undergo different post-translational modifications in an interrelated way, in which one post-translational modification could either promote or impede another modification. This interrelation also applies to DNA methylation. For instance, G9a catalyses the methylation of H3K9, a key marker for the recruitment of HP1, a protein involved in the DNA methylation[37] (**Figure 6**).

Epigenetic mechanisms are important during normal mammalian development and are ascribed to govern normal cell diversification[42]. Hence, dysfunction of histone modification patterns has been correlated with the etiology of a variety of human diseases including allergic diseases, multiple sclerosis, as well as cancer[36]. In colorectal cancer, patterns of histone methylation have been found to be severely altered resulting from both gains and losses of histone methylation. Both play pivotal roles in the formation of colorectal cancer by activating oncogenes and/or silencing tumor suppressor genes. DNA methylation has also been associated with CRC[20-22]. Specifically, DNA methylation is catalysed by DNA methyltransferases (DNMTs) that mediate the addition of methyl groups to the 5th position of the cytosine rings of CpG dinucleotides (C-phosphodiester-G bond). DNA methylation abnormalities such as DNA hypermethylation is believed to influence CRC development by affecting the chromatin structure[43].

Thus, epigenetic marks represent important molecular hallmarks of cancer, as they occur very early in disease pathogenesis and are involved in the modulation of several key cancer-associated pathways. Importantly, histone modifications can be used as clinically-relevant disease biomarkers for diagnosis, prognostication and prediction of treatment response[3]. Due to the reversible nature of the epigenetic marks, tumors that show a pro-oncogenic epigenome could be targeted by small molecules altering the activity of key chromatin modifying enzymes [33,36,44]. Moreover, chromatin regulators are now acknowledged as key drivers of oncogenesis, including colorectal tumors[44-47]. For example, the histone methyltransferase EZH2 was reported to be overexpressed in colorectal tumors vs. normal intestine and associated with the

acquisition of tumor-initiating functions[48,49]. Loss of H3K27 methylation by inactivation of Polycomb repressive complex-2 (PRC 2) was also shown to induce apoptosis in colon tumor-initiating cells [45]. In addition, direct inhibitors of EZH2 were recently confirmed as a powerful tool to suppress CRC tumor-initiating function *in vivo*[48,49].

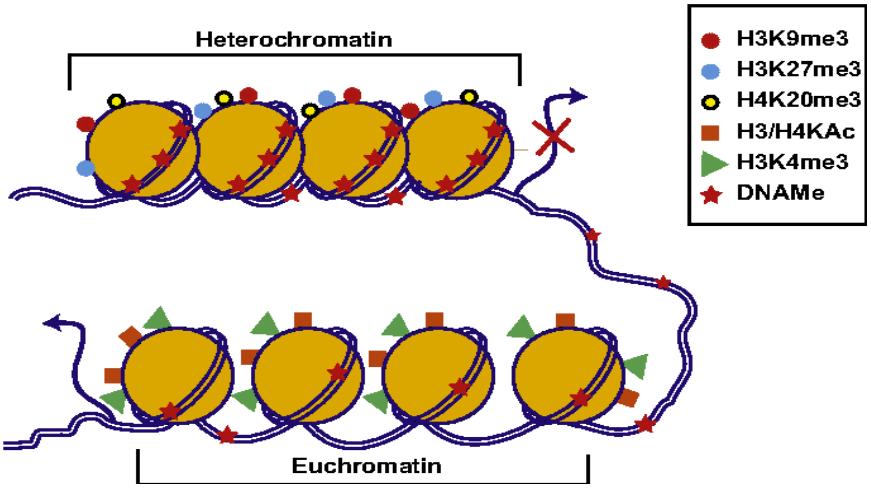


Figure 4. Histone modifications influence chromatin structure and activity.

DNA is wrapped around octamers of histones into nucleosomes. Histone tails undergo post-translational modifications that lead to a change in histone affinity for the DNA, causing the chromatin to shift between open (active) and a closed (suppressive) states[39].

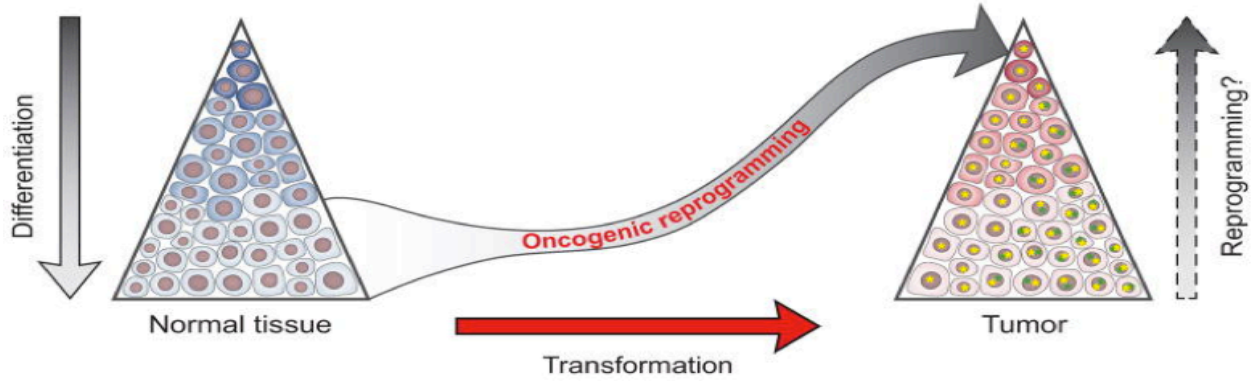


Figure 5. Alteration of epigenetic mechanisms regulates cancer stem cell emergence and tumor initiation.

Genetic alterations in chromatin lead to disruption of epigenetic regulation in either adult stem cells or committed cells and promote neoplastic transformation. The normal function of epigenetic mechanisms (in blue) and the consequences of epigenetic alterations induced by mutations (in red) are indicated[46].

Taken together, the past decade of discovery in cancer epigenetics has revealed a number of similarities between cancer genes and stem cell reprogramming genes, and the part played by chromatin structure in cellular plasticity and cancer development[44,46,47]. In fact, chromatin alterations cause oncogenic cellular reprogramming and induce cellular plasticity (**Figure 5**). The prominent role of epigenetic instability in the emergence of cancer stem cells and tumor evolution provides an opportunity to reverse drug resistance and deplete CSCs by inhibiting epigenetic mediators[44]. Thus, the use of epigenetic inhibitors as an adjuvant treatment might be more effective at earlier stages of cancer, not only because epigenetic alterations manifest as early events during carcinogenesis, but also because the burden of genomic alterations is lower. Moreover, with increased selectivity for cancer stem cells, these epigenetic inhibitors will have limited secondary effects on patients[47].

1.4. G9a-dependant H3K9 methylation signature in colorectal CSCs

The methylation of histone lysine residues, a highly dynamic mechanism, plays a central epigenetic role in the organization of chromatin domains and regulation of gene expression[50]. Histone 3 Lysine 9 methylation (H3K9), one of the best characterised histone methylation sites in eukaryotic chromatin, is a marker of condensed, inactive chromatin and correlates with specific gene silencing in tumors[51]. Mono- and di-methylation of H3K9 are catalyzed by the histone methyltransferase G9a (EHMT2)[52,53] **(Figure 6)**. In fact, G9a is expressed in stem and somatic cells and is known to cooperate with other transcription factors to regulate gene expression and is reportedly involved in a wide-range of functions influencing development, pluripotency and cellular differentiation[53,54].

A G9a-like protein (GLP) has also been identified, which directly interacts with G9a, forming a heterodimeric complex[55]. It has been shown that this structure is the predominant form, as well as the active state, of the methyltransferase *in vivo*. G9a/GLP catalyze mono- and dimethylation of histone 3 on Lysine 9 (H3K9me1/2) and can also methylate non-histone proteins. Interestingly, G9a/GLP was also suggested to take part in Polycomb group-mediated epigenetic silencing by catalyzing H3K27 mono-methylation, which would serve as a template for EZH2 recruitment[37, 56] (**Figure 6**). Recent studies have demonstrated that while G9a/GLP heterodimers seem to be required for H3K9 methylation, the activity of G9a is more important (over GLP) for the *in vivo* function of the complex[57].

It is clear that an appropriate level of G9a activity is required to maintain the normal phenotype in a cell. Interestingly, G9a depletion resulted in embryonic lethality with severe differentiation defects in embryonic stem cells, demonstrating that G9a is essential for the repression of developmental genes and is required during development[53]. Moreover, apoptotic cells were increased drastically in growth-arrested G9a knockout embryos, and this seemed to be the dominant cause of embryonic growth retardation[53]. Together, this suggests that G9a is necessary for important events during embryonic development or differentiation.

On the other hand, multiple studies have reported elevated G9a expression in various cancers and suggested functional linkages with malignant behaviors of cancer cells by silencing tumor suppressors and/or activating survival genes, leading to aberrant proliferation, metastasis, and chemoresistance [40,48,54,57]. A critical role for G9a in the regulation of cell stemness has been demonstrated. Specifically, G9a was established as a selective regulator of rapidly proliferating myeloid progenitors with no discernible function in hematopoietic stem cells (HSCs)[56]. In mouse models of acute myeloid leukemia (AML), loss of G9a significantly delays disease progression and reduces leukemia stem cell (LSC)

frequency[56]. These results highlight a potential clinical application for G9a inhibition as a means to counteract proliferation and self-renewal of AML cells.

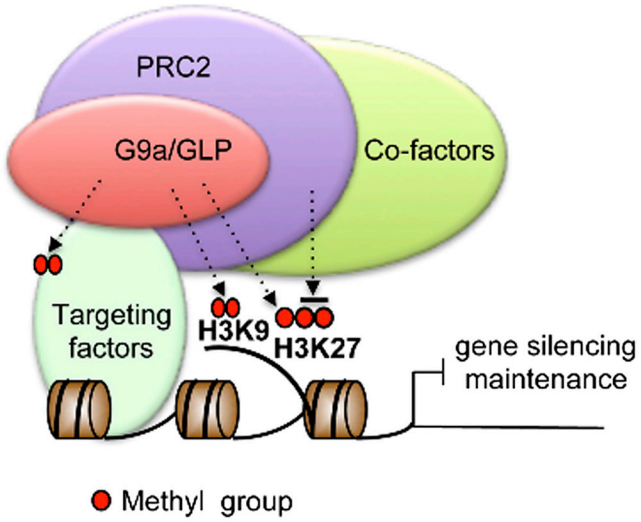


Figure 6. Histone methyltransferases G9a and GLP form heteromeric complexes.

G9a and G9a-like protein (GLP) form a catalytically active heterodimeric complex to catalyze H3K9 mono- and dimethylation, i.e., H3K9me1 and H3K9me2[37].

Previous studies have reported that G9a is overexpressed in a number of cancers, including hepatocellular carcinoma, aggressive lung cancer, brain cancer, and aggressive ovarian carcinoma and that overexpression is found to be associated with enhanced proliferation and metastasis[22,40,54,57,58]. Thus, high G9a expression could represent a marker for poor cancer prognosis. Specifically, elevated G9a levels are correlated with increased H3K9me2 deposition, leading to the silencing of important tumor suppressor genes[37, 40, 48,57]. This is supported by the finding that overexpression of G9a is often observed in aggressive and highly metastatic forms of cancer, indicating that G9a expression might be a key factor in the occurrence of metastasis[48, 57]. For instance, ovarian cancer xenografts have demonstrated a higher expression of G9a in metastatic lesions compared to the primary tumors[58]. However, knocking down G9a failed to reduce metastasis *in vivo* in lung cancer, suggesting that such an association between enhanced H3K9me2 deposition and metastasis is not necessarily applicable to all types of cancer[58]. The latest example highlights the fact that G9a inhibition leads to a cancer-specific response.

In colorectal cancer, it has been reported that G9a expression is required to preserve the *in vivo* tumorigenic capacity of the human cell line HT29[59]. However, little is known about the role of G9a in the regulation of gene networks and pathways that contribute to the maintenance of a neoplastic stem-like phenotype in colorectal cancer[48]. Considering the growing knowledge around G9a and its role in solid tumors[40,60,99], it is conceivable that targeting G9a and its associated epigenetic regulation could promote anticancer effects, such as the re-expression of tumor suppressor genes, inhibition of cell proliferation, and reduction of metastasis. Based on recent literature, these effects in human colorectal cancer could involve the suppression of CSC functions[57]. Surely, these observations suggest that pharmacological strategies designed to target G9a expression and/or activity may have clinical utility to improve survival in patients.

A number of small molecule inhibitors have been developed with the capacity to inhibit the catalytic activity of G9a and have been used in various *in vitro* and *in vivo* experiments[61-63]. One of the first molecules developed was BIX-01294 (diazepin-quinazolin-amine derivative), a competitive inhibitor of G9a and GLP[61]. BIX-01294 treatment was shown to reduce cell proliferation in leukemia cell lines, as well as in human germ cell tumors and squamous neck carcinoma[40]. However, BIX-01294 also showed intrinsic toxic effects. The molecule has since been optimized, leading to the synthesis of a second inhibitor, UNC0638, which exhibits high potency and specificity for G9a, combined with lower cell toxicity[40]. UNC0638 has been used in vitro to suppress cell proliferation in various cancer cell lines, including breast tumors, squamous head and neck carcinoma, hepatocellular carcinoma, acute myeloid leukemia, and cervical cancer[40].

Although UNC0638 displays improved chemical characteristics, it is affected by poor pharmacokinetics, which impedes its efficient use *in vivo*. Substantial progress has been made in developing small molecule inhibitors, which directly bind to G9a and impair its histone methyltransferase activity[62]. In 2013, Liu et al. reported the synthesis of a G9a and GLP inhibitor suitable for animal studies. The small molecule (UNC0642) demonstrates improved pharmacokinetics, while maintaining a high selectivity and low cell toxicity[62]. However, despite their specificity of interactions, systemic administration of these inhibitors is expected to exert toxicity by reducing global levels of H3K9me2 in normal tissues. Notably, liver-specific G9a knockout mice demonstrated significant deleterious effects in tissue maturation, lipid metabolism, and inflammatory response[64]. Also, conditional deletion mouse models showed that G9a is required for cardiomyocyte homeostasis in adult heart[65]. In another study, conditional knockout of G9a in the mouse hematopoietic system revealed failure in T helper cell differentiation, resulting in an impaired immunological response to common gastrointestinal parasite infections[66]. Moreover, Ugarte et al. showed that G9a inhibition using UNC-0638 delayed the normal

hematopoietic stem/progenitor cell differentiation capacity *in vitro*[67]. Altogether, these findings may explain the paucity of *in vivo* investigations and lack of clinical trials using these compounds. Thus, it is critical to work toward developing novel and more targeted approaches to mitigating G9a functions in cancer.

1.5. Phenotypic screening strategy to identify new colon CSC-bioactive compounds

As mentioned above, conventional CRC therapies including chemotherapy, surgery and radiation therapy have limited clinical efficacy[68]. Furthermore, many CRC patients struggle with devastating treatment failures and side effects that can destroy their overall health. Thus, the field of clinical oncology is in dire need of more efficient therapies to be developed.

Recent investigations have revealed that pro-oncogenic histone modification patterns are reversible and that such aberrations can be restored to nearly normal status through epigenetic therapies[36, 69, 70]. Thus, histone modification serves as a promising therapeutic target for treating various cancers and the inhibition of clinically relevant epigenetic processes is currently a subject of intensive investigations. Notably, histone deacetylation and methylation inhibitors are the most widely applied to colorectal cancer[36]. Therefore, we envision that future studies will lead to important advances in the design and development of specific epigenetic inhibitors with translational potential[63].

Since it is now clear that histone modifications are involved in the pathogenesis of CRC, it is important for us to apply this new understanding to develop novel therapeutic approaches for colorectal cancer[36]. Considering that the CSC epigenetic signature can be selectively targeted using small molecules[45], it is imperative to develop new epigenetic modulators that could be combined with existing chemotherapeutic regimens, leading to more effective treatment outcomes.

The main goal of our laboratory is to develop novel, translational, anticancer agents that target essential epigenetic features involved in CCSC development and/or maintenance while sparing normal progenitors. As presented and discussed in a later section, we used a phenotypic screen as a drug discovery approach that has the advantage to test the effects of small molecules on complete biological systems, such as mammalian cells, and lead to the identification of candidates with *bona fide*, pathologically-relevant molecular mechanisms of action[50,51]. The assay that was developed by our team aims to find compounds with the capacity to decrease the pluripotency state and H3K9me2 deposition in transformed human embryonic stem cells (t-hESC). Moreover, lead compounds identified from phenotypic screening are suggested to have higher chances of success in clinical trials[50,51]. Combined with the use of clinically approved or "repurposed" compounds, our methods should bolster the translational potential of discovery, since they are acknowledged to accelerate bench-to-bedside transition[52,53].

As a founding hypothesis, our lab proposes that pharmacological targeting of key elements of the CCSC epigenetic signature can lead to the suppression of essential cellular functions promoting development, maintenance, and renewal of CSC populations. Specifically, the development of new molecules that alter G9a functions could increase the sensitivity of CRC tumors when combined with standard chemotherapy. This strategy is expected to suppress the emergence and plasticity of a CSC phenotype, therefore, sensitizing resistant cells to cytoreductive therapies. This could be achievable through the re-expression of fundamental genes, such as important tumor suppressors.

2. Hypothesis

Here, I propose that the histone methyltransferase, G9a, represents a key enzyme contributing to the colorectal cancer stem cell (CCSC) epigenetic signature, that promotes self-renewal and tumorigenicity. Moreover, I explore the impact of pharmacological targeting of G9a expression and H3K9me2 deposition as a potent strategy to alter human CCSC functions, which could represent an appealing future strategy to treat human CRC.

3. Objectives

My work as a Master's student was focused on three main research objectives to elucidate the role of G9a in human CCSCs and to gain insights into better molecular strategies to adopt toward cancer-specific targeting of this HMTase:

3.1. To characterize the role of G9a in maintaining the molecular functions of cancer stem cells in colorectal cancer (CRC)

G9a activity seems to be essential for the maintenance of an embryonic-like transcriptional signature that promotes self-renewal, tumorigenicity and an undifferentiated state. Suppression of G9a activity reduces pluripotent identity and tumor growth capacity in transformed human embryonic stem cells (t-hESCs).

3.2. To study the impact of G9a inhibition on CCSC tumor-initiating capacities

G9a histone methyltransferase plays a key role in maintaining pro-oncogenic transcriptional programs responsible for compromising differentiation and promoting neoplastic stemness in CRC/intestinal

models. Additionally, pharmacological inhibition of its activity is expected to reduce the tumor-initiating capacity of primary CRC patient samples.

3.3. To identify new selective inhibitors contributing to the elimination of colorectal CSCs, by down-regulating G9a activity

Phenotypic screening will be used to identify novel anti-neoplastic agents. Specifically, experimental approaches were combined to design a multi-parametric phenotypic high-throughput screening pipeline, using pluripotency and a cancer-associated epigenome as readouts to facilitate the identification of new colorectal CSC-bioactive compounds.

4. Materials and methods

4.1. Tissue culture and reagents

Normal Human intestinal epithelial crypt cells (HIECs) were cultured in OptiMEM (Gibco) supplemented with 4% FBS (Wisent Premium), HEPES, GlutaMAX (Gibco), and 10ng/ml EGF as previously described [71]. Colorectal cancer cell lines HT29, HCT116, and SW480 were obtained and cultured according to recommendations from ATCC. 293FT cells for lentiviral particle production were purchased from ThermoFisher Scientific and cultured in DMEM supplemented with 10% FBS (Thermofisher), 1mM non-essential amino acids, 1mM L-Glutamine, and 1mM sodium pyruvate. Human embryonic stem cells (hESC: H9 line, WiCell) and transformed H9 ES cells (t-hESC: gift from Dr. Mickie Bhatia, McMaster University) were cultured on a Matrigel-coated culture plate in mTeSR media (Stemcell Technologies) according to previously established conditions[31]. All cell lines were authenticated and tested for mycoplasma. Moreover, cells used in this project are from the ATCC (American Type Culture Collection). They were used at low passages to avoid genetic drift and mycoplasma contamination.

Primary colorectal tumor tissues were obtained with patient consent, as approved by the University Health Network Research Ethics Board and from Celprogen Inc. (#36112-39P, Torrance, CA). Specimen origin, tumor stage and mutational status at diagnosis are presented in Fig.S4A. Tumor samples were processed as previously described [21]. Briefly, tissues were mechanically minced and incubated with Collagenase A (3 mg/mL) for 60 min at 37°C. Following dissociation, samples were filtered using a 70μm cell strainer. Red blood cells were removed using ammonium chloride solution (Stemcell Technologies). Isolated cells were maintained as spheres in ultra-low adhesion flasks with DMEM/F-12 (Gibco) supplemented with 1% pen-strep, L-glutamine (2 mM), nonessential amino acids (1X, Gibco), sodium pyruvate (1 mM), HEPES, heparin (4 μg/mL), B27 supplement (GIBCO), N2

supplement (GIBCO), lipids mixture (Sigma), EGF (20 ng/mL) and bFGF (10 ng/mL). Small molecule inhibitors BIX-01294, UNC0642, UNC1999, and GSK LSD1 were purchased from Tocris Bioscience and re-suspended in DMSO.

4.2. Lentiviral particle production and establishment of shRNA knockdowns

Lentiviral knockdown vectors pLKO.1-puro shG9a (TRCN0000115670 and TRCN0000115671, Sigma) and pLKO.1-puro Non-Mammalian shRNA control (shCTRL, Sigma #SHC002) were co-transfected with packaging vectors pMD2.G and psPAX2 into 293FT cells using Lipofectamine 2000 (ThermoFisher Scientific). Viral particles were harvested as previously described[45]. $5x10^5$ cells/well in 6-well plates were incubated with 0.5ml of viral suspension + 8ug/mL of polybrene (Millipore sigma) for 16 hours. Culture media containing 0.5 μg/ml of puromycin was used for 10 days of selection. Knockdown efficiency was determined by western blot.

4.3. Cell counts and drug dose-response analysis

Each cell model was plated at a density of $5x10^3$ cells/well in 96-well culture plates, 24h prior to drug treatments. Small molecule inhibitors BIX-01294, UNC0642, UNC1999, and GSK LSD1 were used at concentrations ranging from 0.02 to 20μM for 24 to 72 hours. Equivalent volumes of vehicle DMSO were used as control (≤0.1%). Cells were formalin-fixed, stained with Hoescht 33342, and plates were imaged with a Cellomics ArrayScan VTI High-Content imaging system (ThermoFisher Scientific). Images were analyzed HCS studio™ cell analysis software and half maximum effective concentration (EC50) values were calculated using GraphPad Prism.

4.4. Immunofluorescence

Formaldehyde-fixed paraffin-embedded human colon carcinoma tissue microarray sections, including normal colon tissues (US Biomax, #CO486) were rehydrated, quenched in 0.1M glycine buffer, and blocked with a 2.5% BSA-PBS solution prior to immunostaining, as previously described[72]. For staining of cells, culture wells were fixed with 2% formalin and incubated in Perm/Wash buffer (BD Biosciences) at 4^{o}C for 15-30 min prior to immunostaining[45]. Anti-H3K9me2 primary antibody **(Table-1)** was diluted at 1:500 in 1% BSA-PBS solution and incubated overnight at 4°C. Secondary antibody (anti-mouse Alexa Fluor 488) was used at 1:500 in a 1% BSA-PBS solution. For staining on tissue sections, slides were mounted using Vectashield mounting medium with DAPI (Vector labs). For cell staining, nuclei were stained with Hoechst 33342. Cells and tissue sections were imaged with a Cellomics ArrayScan VTI High-Content imaging system and fluorescence quantification was performed with HCS studio^{TM} cell analysis software for cell-based assays and Image J software (National Institutes of Health) for tissue sections.

4.5. Western blotting

Whole-cell extracts were prepared in Laemmli Sample Buffer (60mM Tris-HCL pH 6.8, 2% SDS, 10% glycerol, 5% β-mercaptoethanol, 0.01% bromophenol blue), sonicated and thermo-reduced/denatured (5 min, 95^{o}C) prior to electrophoresis on 12.5% or 15% polyacrylamide gels. Proteins were transferred onto nitrocellulose membranes and immuno-detected as previously described[45]. For dot blot experiments, whole-cell extracts were directly spotted onto membranes and air-dried for 30 min prior to blocking. Membranes were blocked in PBS containing 5% skim milk and 0.1% TWEEN 20. Primary and secondary antibodies used are described in **Table-1.** Full-range molecular mass marker (Full-Range

Rainbow Marker, VWR) was used as standard. Blot images were acquired using a ChemiDoc MP Imaging system. Quantitative optical densitometry analysis of bands was performed using Image J software (National Institutes of Health).

4.6. *In vivo* teratoma assays

Stable control and G9a-knockdown t-hESCs were treated with collagenase IV and harvested using TrypLE dissociation reagent (ThermoFisher) and injected intratesticularly into 12-week old male NOD/SCID mice. Specifically, $5x10^5$ t-hESC were injected into the left testicle of each mouse, as previously described[31]. Right testicles were injected with an equivalent volume of saline solution, as a "no-cell" injection control. At 14-day post-injection, mice were euthanized, and testicles size was measured for each animal using a caliper. Tumor burden was calculated as the size ratio of cell-injected (left) over control (right) testicles. Volume of each testicle was calculated as 1/2 x length x $width^2$.

4.7. Transcriptome profiling and quantitative PCR analysis

RNA was extracted using the total RNA purification kit by Norgen Biotech Corp, following the manufacturer's guidelines. Quantification of total RNA samples was performed with a Qubit HS RNA assay (ThermoFisher Scientific) and fragment size was evaluated with a Fragment Analyzer HS NGS assay (AATI). An RNA Quality Number (RQN) of 8.0 or higher was considered satisfactory for library construction. Library construction was performed with a Truseq RNA v2 (Illumina). Libraries were prepared with unique barcodes compatible with the Illumina NextSeq 500 platform. Quantification of the libraries was performed with a Qubit HS DNA assay and library fragment size was evaluated with a Fragment Analyzer HS NGS assay. Libraries were normalized to the same concentration, then samples

were pooled in equal amounts. Next-generation RNA sequencing was performed on an Illumina NextSeq 500 platform, according to 1 x 75bp cycles of single-end sequencing, yielding 25 million reads per sample. PhiX ssDNA was spiked in each sample and used as a technical control for clustering reactions. Upon alignment of sequencing data, control vs. treated, transcriptomes for each cell line were compared and significantly modulated genes ($p<0.05$) were identified using the Salmon transcript abundance method[73]. For quantitative PCR analysis, cDNA was synthesized from 1μg of total RNA using SuperScript VILO cDNA Synthesis Kit (Invitrogen). qPCR reactions were carried out using Power SYBR Green Master Mix (ThermoFisher Scientific) per manufacturer's recommendation. Amplification was performed using an ABI 7500 Real-Time PCR System. Primer sequences used in this study are presented in **Table-2.** All reactions were normalized to GAPDH as a reference gene and relative gene quantification was calibrated against vehicle/control-treated samples according to a method previously described[74].

4.8. Survival analysis

Hazard ratios relative to the expression of chromatin modifiers based on disease free survival were calculated using the platform GEPIA2[75]. Samples within the 85th percentile for expression of each chromatin regulators, determined by RNA-seq in TCGA cohorts were ranked as high-expressing tumors. Statistical significance cutoff was set at $p<0.05$.

4.9. Bioinformatics analysis and data repository

RNA-seq alignment and specialized bioinformatics analyses were performed with the support of the uOttawa Bioinformatics Core Facility. Data was analyzed using R. Tags Per Million values, genes and

patients of interest were downloaded from TCGA using TCGAbiolinks[76]. TPM values were multiplied by 1 million for ease of interpretation in spreadsheets. Clustering and heatmap visualization were performed using heatmap.2. Gene expression data for cell lines were presented as fold-change (Salmon transcript abundance method) or as log transformed row-mean normalized FPKM/RPKM. Data for HCT116 transcriptome profiling is from GSE61255 data set (control: GSM1500847, and BIX-01294: GSM1500848). Stem cell index calculation was performed as previously described[77]. GSEA analysis were performed as a pre-ranked analysis using GSEA software version 4.0.3 (www.broadinstitute.org/gsea) with 1000 permutations[78]. Normalized enrichment scores ± 1.5 and $p<0.05$ were considered significant. Information on gene signatures used in GSEAs is presented in **Table-3.** Gene ontology analysis was performed using the ICGC Data Portal (https://dcc.icgc.org/) where enrichment analysis was preformed against GO Molecular Function, GO Biological Process and GO Cellular Component with an FDR q-value threshold of 0.05 (**See Table-4**). ChIP-seq reads from GSE82131 data set were quantified in 1kb bins genome-wide for G9a and H3K9me2. Bins were filtered against the ENCODE blacklist and mitochondria as well as unmapped contig assemblies were excluded. Top 20K signal/bins for both G9a and H3K9me2 were overlapped, yielding 1,392 co-enriched regions. Nearest gene annotation was performed using the R ChIPseeker package. The accession number for the RNA-seq data newly reported in this paper is: GSE154057. Dendrograms of hierarchical gene set clustering were generated using Morpheus (https://software.broadinstitute.org/morpheus/).

4.10. Flow cytometry

Extracellular staining was performed in PBS supplemented with 3% FBS and 0.5mM EDTA (PEF) where 100K/ml cells of interest were incubated with antibodies for 1hour at 4°C, washed with 10 volumes of PEF and then stained with 7-amino actinomycin D (7-AAD, Immunotech) to eliminate dead

cells prior to analysis. Single-cell suspensions were stained using anti-CD133-PE and anti-CD44-APC **(see Table-1).** Flow cytometry analysis was performed on a BD LSR Fortessa 16-colour Analyzer and data analysis was conducted using FlowJo (Tree Star Inc)[45].

4.11. Serial organoid formation assay

Patient-derived spheres enriched for CCSCs were harvested and dissociated using TrypLE reagent (ThermoFisher) and passed through a 70-µm strainer to eliminate non-single-cell aggregates. Cell suspensions were mixed with Matrigel in sphere culture media (1:1 ratio) to get a 1-cell/µl density. Mixtures were immediately plated as 300µl domes in 6-well plates and incubated for 15 minutes at 37°C for Matrigel polymerization. 2.5mL of sphere media containing doses of UNC0642 (2.5 and 5µM) or vehicle control (DMSO) was added to each well. Drug treatment lasted for 7 days, followed by a 7-day drug-free incubation period. At day-14, plates were imaged using a Cellomics ArrayScan VTI platform to determine organoid counts and size using HCS Studio^{TM} software. For secondary passage experiments, control primary organoids and those remaining in treated wells were dissociated using a gentle dissociation reagent (STEMCELL Technologies) and re-plated as described above, but according to a 50-cell/µl density. Secondary organoids were grown for 14 days in sphere culture media with no further drug treatment. Then, secondary organoid counts were determined using a Cellomics ArrayScan VTI platform and HCS Studio^{TM} software.

4.12. Statistical analysis

Data is presented as mean ± SEM. P-values < 0.05 were considered significant. ‘‘n’’ denotes the number of times the data was replicated. Significant differences between groups were determined by 2-way ANOVA test and unpaired two-tailed Student t-test, using GraphPad Prism software

5. Results

5.1. Objective #1: To characterize the role of G9a in maintaining the molecular functions of cancer stem cells in colorectal cancer (CRC)

Rationale. Recent studies highlighted the possibility that G9a expression and/or HMTase activity could contribute to the elevated risk of cancer relapse by driving neoplastic stemness and establishing pluripotent-like transcriptional programs in human CRC[48,57]. In this objective, I investigated the impact of G9a inhibition in the regulation of oncogenic transcriptional networks in human pluripotent cells, as well as standard colorectal cancer cell lines. After establishing a link between G9a expression and pluripotent-like transcriptional signatures in primary colorectal tumors, I used transformed human embryonic stem cells (t-hESCs) to gain critical understandings on this relationship in a context of neoplastic stemness. Next, I explored the functional role of G9a HMTase activity in a colon CSC-like model (HT29) by down regulating its expression and HMTase activity.

I first started by figuring out which epigenetic pathways are associated with therapeutic failure and relapse in CRC. Relapse-free survival (RFS) analysis in various publicly-available cohorts determined the risk factors associated with 13 key chromatin regulators across different cancer types. I identified EHMT2/G9a (Hazard ratio (HR): 2.2, p=0.0036) and DNMT3A (HR: 1.9, p=0.047) overexpression to be significantly associated with shorter RFS in transcriptome profiling data from the TCGA colon adenoma patient cohort (COAD) (**Figure.7A** and **B**). In contrast, high expression of KDM4B - an H3K9-specific demethylase - was linked to enhanced RFS prognosis in colon tumors (HR: 0.27, p=0.028) (**Figure.7A, B**). The expression of EHMT1/GLP, which is a G9a partner potentiating its capacity to mono and di-methylate H3K9, was not a factor for RFS in CRC patients (**Figure. 7B**).

The next step was to assess the expression of the dimethylation marker, H3K9me2, in adenocarcinoma tissues compared to healthy colon tissues. We observed from COAD data that G9a and LSD1 transcript expression are significantly higher in CRC vs. normal colonic tissues (**Figure.7C**). We have not observed such a phenomenon for other H3K9me2 regulators (**Figure.7C**). Next, I used tissue sections from 39 primary CRC tumors and matched normal tissues to perform immunofluorescent staining of H3K9me2, which is mainly catalyzed by G9a HMTase activity. Global H3K9me2 levels were significantly increased in human colon adenocarcinomas vs. normal colonic mucosa (**Figure.7D**). This supports western blot results from three human CRC lines, where H3K9me2 was also higher compared to normal intestinal progenitor HIEC cells (**Figure.7E**). Quantitative PCR analysis of G9a transcript expression in CRC lines and normal HIEC cells also correlated with H3K9me2 deposition (**Figure.7E**). Interestingly, transcript expression of the lysine demethylase LSD1 followed the same pattern as G9a in CRC cell lines and HIEC (**Figure.7F**). Overall, these first observations suggest a higher expression of G9a and H3K9me2 deposition in CRC tumors compared to the normal tissues.

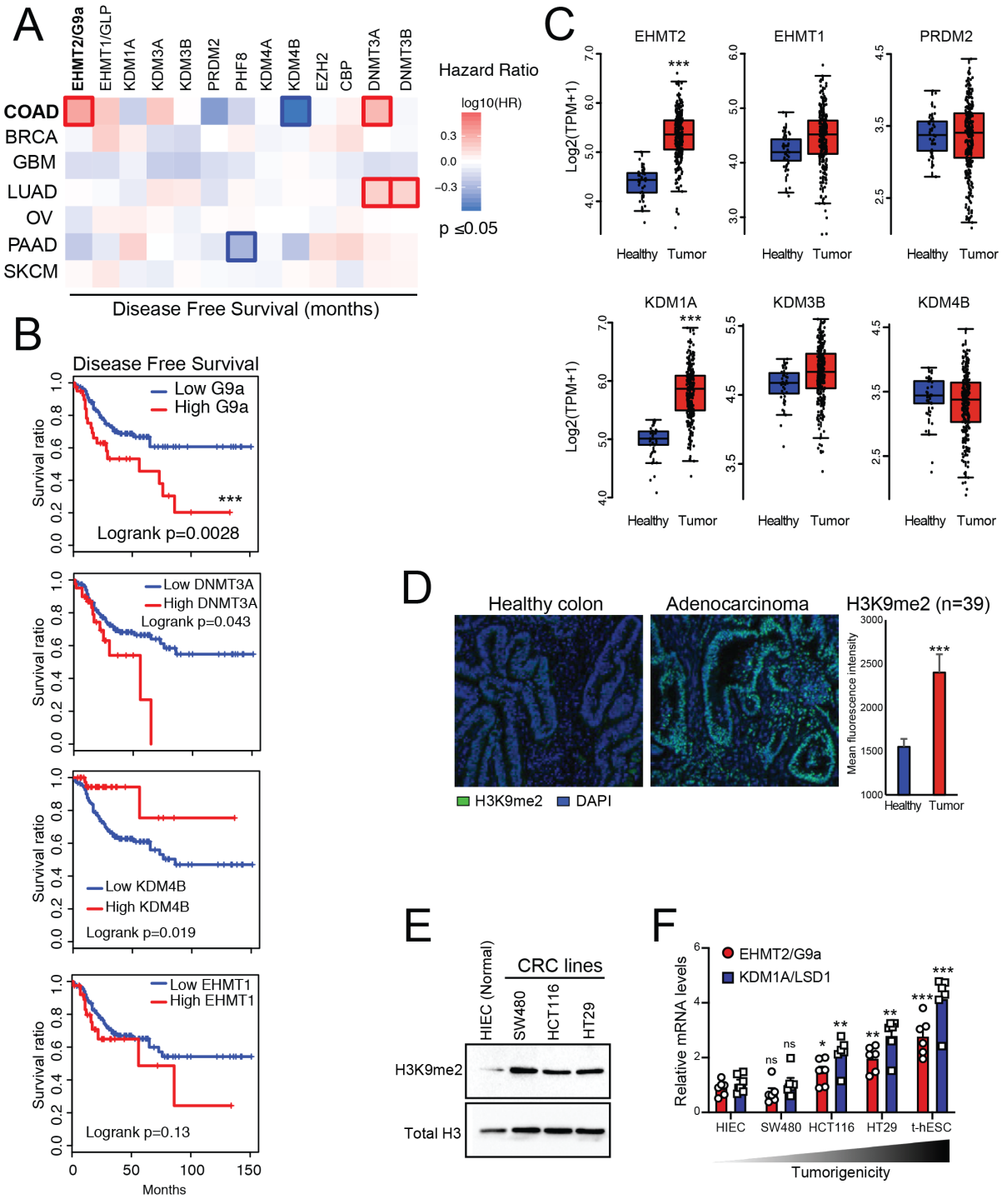

A
EHMT2/G9a
EHMT1/GLP
KDM1A
KDM3A
KDM3B
PRDM2
PHF8
KDM4A
KDM4B
EZH2
CBP
DNMT3A
DNMT3B
COAD
BRCA
GBM
LUAD
OV
PAAD
SKCM
Disease Free Survival (months)
Hazard Ratio
log10(HR)
0.3
0.0
−0.3
p ≤0.05
B
Disease Free Survival
Low G9a
High G9a
***
Logrank p=0.0028
Low DNMT3A
High DNMT3A
Logrank p=0.043
Low KDM4B
High KDM4B
Logrank p=0.019
Low EHMT1
High EHMT1
Logrank p=0.13
Survival ratio
Months
C
EHMT2
EHMT1
PRDM2
KDM1A
KDM3B
KDM4B
Log2(TPM+1)
Healthy
Tumor
D
Healthy colon
Adenocarcinoma
H3K9me2 (n=39)
Mean fluorescence intensity
H3K9me2
DAPI
E
HIEC (Normal)
CRC lines
SW480
HCT116
HT29
H3K9me2
Total H3
F
EHMT2/G9a
KDM1A/LSD1
Relative mRNA levels
HIEC
SW480
HCT116
HT29
t-hESC
Tumorigenicity

Figure 7. G9a expression is increased in human CRC vs. normal intestines and is associated with poor prognosis.

A) Survival map representing prognostic impact of genes encoding key chromatin regulators based on patient's disease/relapse free survival and expression levels across different cancer types: breast cancer (BRCA), glioblastoma (GBM), lung adenocarcinoma (LUAD), ovarian cancer (OV), pancreatic adenocarcinoma (PAAC) and skin cutaneous melanoma (SKCM). EHMT2/G9a expression levels display a negative impact on disease free survival in colon adenocarcinoma patients (TCGA COAD). The heat map shows the hazard ratios in log10 scale for tested genes. Red/blue frames indicate significant unfavorable/favorable risk in prognostic analyses (p-value ≤0.05).

B) Box plots of key H3K9me2 modulators profiled by RNA-seq in human colon adenocarcinoma (TCGA COAD, n=275 vs. normal, n=41). Values are expressed as TPM (***: p<0.0001, Log2FC cutoff >0.5).

C) Kaplan-Meier disease-free survival analysis of patients presenting high (top 15%, TPM+1) and low (bottom 85%, TPM+1) expression levels of G9a, DNMT3A, KDM4B, and GLP/EHMT1 in primary CRC tumors (TCGA COAD, n=275, Logrank p=0.0028: ***).

D) Immunofluorescent detection of H3K9me2 in human colon adenocarcinoma sections (n=39 patients) vs. normal colonic mucosa. Mean fluorescence intensity was quantified using ImageJ (p=0.0007, ***).

E) Western analysis of H3K9me2 levels in whole-cell lysates from human normal intestine epithelial crypt cells (HIEC) vs. CRC cell lines SW480, HCT116, and HT29. Total histone H3 was used as loading control.

F) Quantitative PCR analysis of G9a and LSD1 expression in normal HIEC, SW480, HCT116, HT29 CRC lines, as t-hESCs. GAPDH was used as a reference gene (n=6, ***: p≤0.001; **: p≤0.007; *: p=0.027).

Considering the importance of G9a for oncogenic self-renewal[56] and developmental cell fate regulation[80], I scrutinized the potential relationship of G9a with a pluripotent-like transcriptional

signature and tumor stemness in human CRC[30,77]. I performed a clustering analysis of chromatin modifier expression across colon adenocarcinoma (COAD) patients and observed that EHMT2/G9a and KDM1A/LSD1 mRNAs are particularly abundant in a specific subset of CRC patients (**Figure. 8A**). With the help of the uOttawa Bioinformatics core facility, we applied a one-class logistic regression machine-learning approach, to determine the degree of transcriptome similarity between individual CRC tumors samples and human pluripotent stem cells. This computational approach developed by Malta et al.[77] enabled us to assign a transcriptional stem cell index (mRNAsi) to 121 COAD tumors. Thus, CRC samples were categorized into "High mRNAsi" (≥0.6) and "Low mRNAsi" (≤0.3) groups. Interestingly, 14 out of 16 samples falling into the High mRNAsi group (≥0.6, n=16) were found within the cluster of high G9a/LSD1-expressing patient samples (**Figure.8A, B**). This coincides with elevated transcript levels of these chromatin modifiers in the highly tumorigenic CRC lines HCT116 and HT29, both recognized to contain large subpopulations of CSC-like cells, compared to normal HIEC and less-tumorigenic SW480 cells (**Figure.7F**). We also observed that EZH2-overexpressing tumors were also correlated to high mRNAsi (**Figure.8B**). This supports previous findings identifying EZH2 as another chromatin regulator maintaining colorectal cancer stem cell (CCSC) functions[49]. When using pharmacological inhibitors of G9a and LSD1 on CRC cells, we observed lower half maximal effective concentration (EC_{50}) values for BIX-01294 and UNC0642 (G9a inh: 1.65μM and 2.87μM respectively) compared to LSD1 inhibition (GSK-LSD1: 22.1μM) (**Figure. 8C**). G9a inhibition displayed comparable effects on cell growth vs. EZH2 inhibition (UNC1999, EC_{50}: 1.59μM) used as a positive control in this experiment (**Figure. 8C)**. Thus, our data suggest that G9a expression could contribute to the elevated risk of relapse by driving tumor stemness and establishing pluripotent-like transcriptional programs in human CRC.

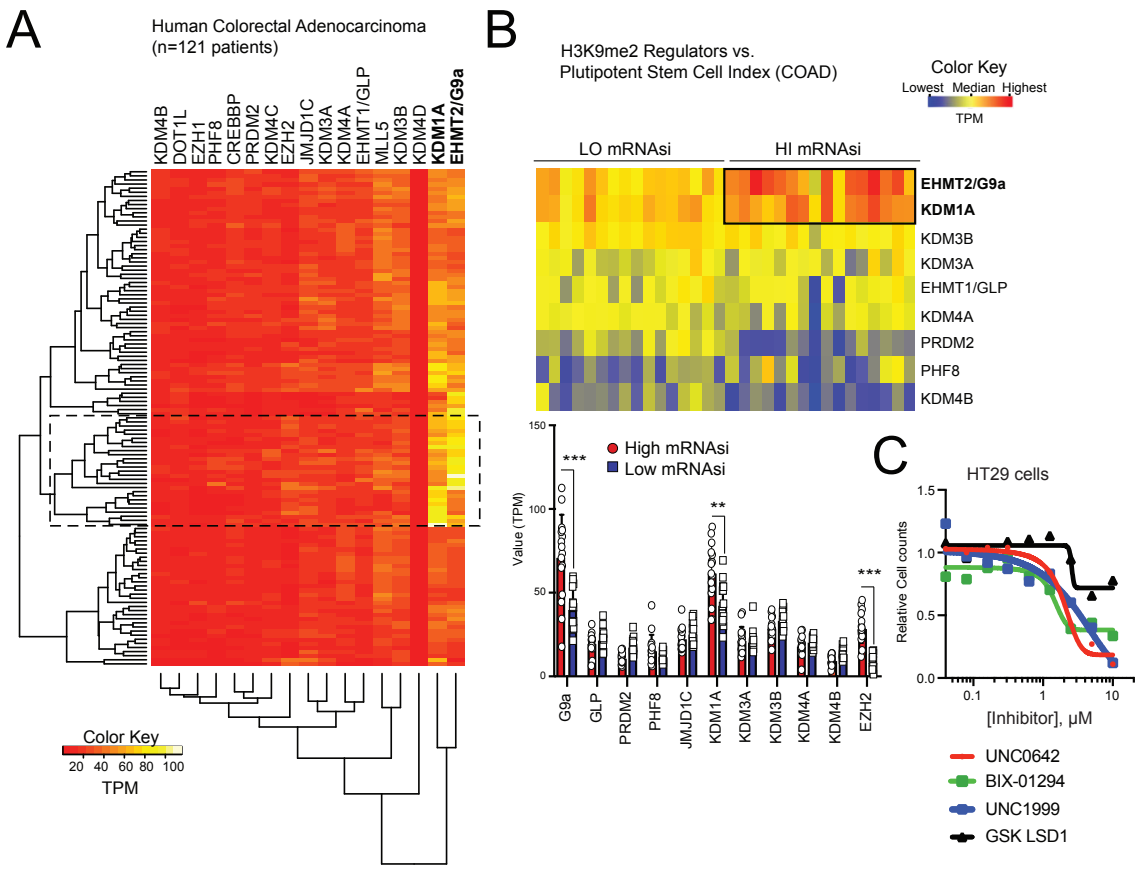


Figure 8. G9a expression is associated with a pluripotent-like transcriptional program in CRC.

A) RNA-seq profiling of key chromatin regulators in human colon adenocarcinoma (TCGA COAD, n=121). Values are expressed as TPM.

B) Expression of main H3K9me2 regulators in COAD samples displaying "high" (>0.6, n=16) and "low" (<0.3, n=16) mRNA expression-based pluripotent stem cell indices (mRNAsi). TPM values were compared between HI and LO mRNAsi groups for each H3K9me2 regulator (***: p≤0.0001; **: p=0.0009).

C) Dose-response experiment assessing the impact of G9a (UNC0642), EZH2 (UNC1999), and LSD1 (GSK-LSD1) inhibition on HT29 cell growth (n=4).

As a next step, I explored the potential role of G9a in the context of neoplastic stemness using a transformed variant of H9 human ES cells (t-hESC), as a well-documented surrogate model for CSC in adherent culture conditions[31,45,81]. t-hESCs faithfully recapitulate the main functional, transcriptional, and epigenetic characteristics of CSCs in culture and *in vivo*[29,31,45,81].

Specific shRNAs were used to knockdown G9a expression in t-hESC. Cells were infected with lentiviral particles containing constructs expressing small hairpin RNA targeting G9a mRNA (shG9a), as well as scramble protein (shCTRL). Knockdown of G9a yielded significant reductions of H3K9me2 and global DNA methylation (5-methylcytosine: 5meC) in t-hESCs (**Figure.9A, B**). Notably, G9a knockdown reduced the growth rate of t-hESCs (**Figure.9C**) and induced morphological aspects of non-transformed, parental H9 cells (**Figure.9D**).

To evaluate the impact of G9a knockdown on t-hESC tumorigenicity, *in vivo* teratoma assays were performed and the size of testicular tumors were monitored 14 days post-inoculation. We observed significantly reduced tumor growth in G9a knockdown injected animals vs. control group (scrambled shRNA) (**Figure.9E**). This suggests that supressing G9a expression, and consequently its chromatin modifying activity, in a CSC context represents an attractive strategy to block pro-oncogenic functions.

Similar to the G9a knockdown in the transformed pluripotent model, I wanted to determine the effect of G9a inhibition in cancer stem cell models. I treated t-hESC with two pharmacological inhibitors of G9a HMTase activity (BIX 01294 and UNC 0642). By western blot analysis, I observed that G9a inhibition significantly reduced the levels of H3K9me2 and global DNA methylation in t-hESCs (**Figure.10A**). Interestingly, G9a inhibition exhibited no effects on H3K27 mono-methylation and EZH2-catalyzed H3K27me3 (**Figure.10B**).

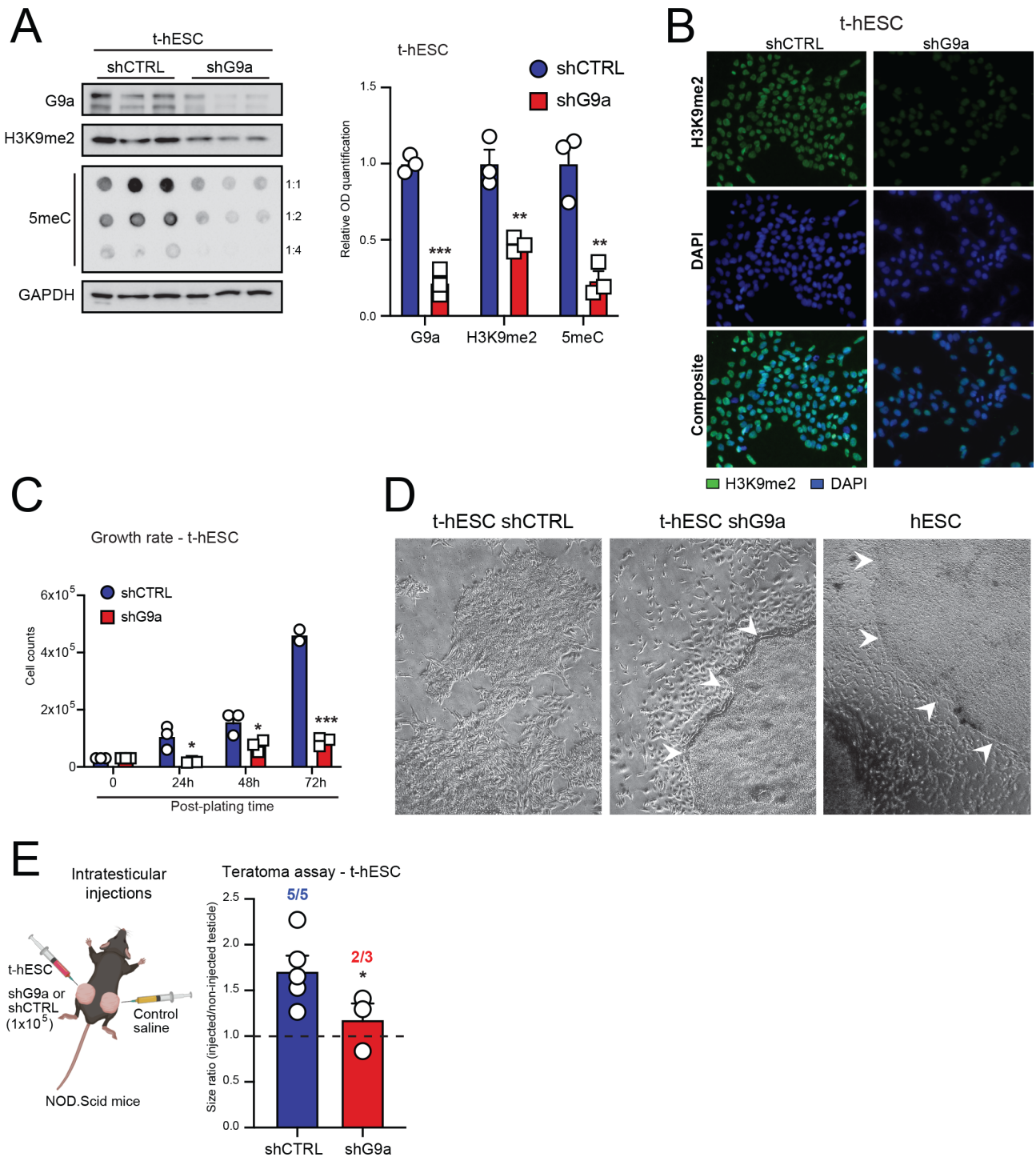

A
t-hESC
shCTRL
shG9a
G9a
H3K9me2
5meC
1:1
1:2
1:4
GAPDH
t-hESC
shCTRL
shG9a
Relative OD quantification
1.5
1.0
0.5
0.0
***
**
**
G9a
H3K9me2
5meC
B
t-hESC
shCTRL
shG9a
H3K9me2
DAPI
Composite
H3K9me2
DAPI
C
Growth rate - t-hESC
shCTRL
shG9a
Cell counts
6x10^5
4x10^5
2x10^5
0
*
*
***
0
24h
48h
72h
Post-plating time
D
t-hESC shCTRL
t-hESC shG9a
hESC
E
Intratesticular injections
t-hESC
shG9a or shCTRL (1x10^5)
Control saline
NOD.Scid mice
Teratoma assay - t-hESC
Size ratio (injected/non-injected testicle)
2.5
2.0
1.5
1.0
0.5
0.0
5/5
2/3
*
shCTRL
shG9a

Figure 9. Suppression of G9a expression reduces transformed human embryonic stem cell (t-hESC) growth and tumorigenesis.

A) Western blot analysis of G9a, H3K9me2, and dot blot analysis of 5-methyl cytosine (5meC: methylated DNA) levels in G9a knockdown t-hESCs vs. scramble control shRNA (shCTRL). GAPDH was used as loading control. Relative OD signal quantification vs. GAPDH intensity is presented (n=3, ***: p=0.0001; **: p≤0.0057).

B) Immunofluorescence detection of the H3K9me2 mark in non-silencing shCTRL and shG9a t-hESCs. Nuclei were counterstained with DAPI (20X magnification).

C) Growth rate assessment of shG9a vs. shCTRL t-hESC cultures. Cell counts were acquired at 24h (*: p=0.021), 48h (*: p=0.034), and 72h (***: p=0.0003) post-seeding (30K cells/well, n=3).

D) Phase-contrast micrographs of shCTRL and shG9a t-hESCs compared to normal human ES cells (hESC) (10X magnification). Arrowheads are marking the presence of sharp colony edges.

E) Testicular teratoma assay using shCTRL (n=5) and shG9a (n=3) t-hESCs over 14 days. $1x10^5$ cells (left testicle) or saline (right testicle) were injected in each mouse. Tumor formation frequencies and tumor/control size ratios were presented for both groups (*: p=0.043).

Both histone marks were previously suggested to depend on G9a activity for Polycomb group-mediated silencing of developmental regulator genes in mouse embryonic stem cells[37]. Next, I performed transcriptome-profiling experiments on control and BIX-01294 treated (1μM, 48h) t-hESC to determine the impact of G9a inhibition on gene networks associated with neoplastic stemness. Differentially expressed genes in BIX-01294 treated group vs. DMSO control and showing a p-value lower than 0.05 were considered for subsequent analysis (**Figure.10C**). Transcriptomic data was used to assess changes in t-hESC mRNAsi (or stem cell index) upon G9a inhibition. I observed a downshift in t-hESCs mRNAsi score in response to BIX treatments (**Figure.10D**). Gene set enrichment analysis (GSEA), performed by Christopher Bergin, on genes significantly modulated by BIX in t-hESCs revealed negative correlation patterns with an ES cell transcriptional signature and targets upregulated by the proto-oncogene c-Myc (**Figure.10E**)[30]. In addition, GSEA suggests the restoration of p53 target expression upon G9a inhibition in t-hESCs (**Figure.10E**)[82]. Accordingly, the inhibition of H3K9me2 deposition in t-hESCs was marked by a significant upregulation of genes restricting proliferation, such as p21 (CDKN1A), and the repression of genes associated with cell cycle progression (*e.g.* CDC45) (**Figure.10F**). Consistent with its impact on mRNAsi, G9a inhibition also induced the expression of genes regulating cell fate and differentiation (**Figure.10F**). Despite the impact of G9a inhibition on t-hESC transcriptional networks, the expression of core pluripotency factors was not altered upon BIX-01294 treatments (**Figure.10G**). Thus, our observations regarding the role of G9a in the context of neoplastic pluripotency are consistent with risks associated with its overexpression in CRCs and support the relevance of such an epigenetic mechanism as a potential therapeutic axis.

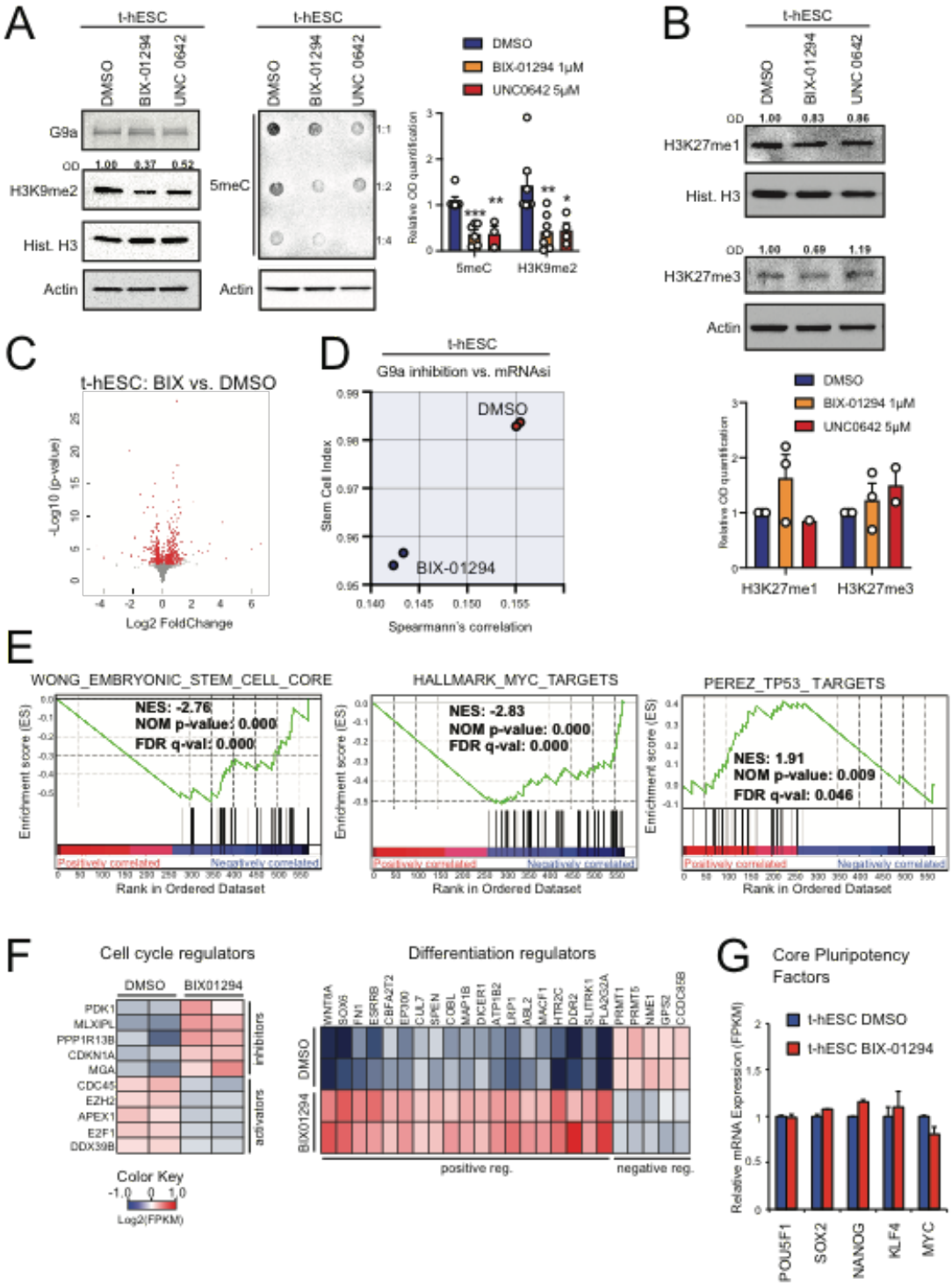

A
t-hESC
DMSO
BIX-01294
UNC 0642
G9a
OD 1.00 0.37 0.52
H3K9me2
Hist. H3
Actin
5meC
1:1
1:2
1:4
DMSO
BIX-01294 1µM
UNC0642 5µM
Relative OD quantification
5meC
H3K9me2
B
OD 1.00 0.83 0.86
H3K27me1
OD 1.00 0.69 1.19
H3K27me3
C
t-hESC: BIX vs. DMSO
-Log10 (p-value)
Log2 FoldChange
D
G9a inhibition vs. mRNAsi
Stem Cell Index
Spearmann's correlation
E
WONG_EMBRYONIC_STEM_CELL_CORE
NES: -2.76
NOM p-value: 0.000
FDR q-val: 0.000
HALLMARK_MYC_TARGETS
NES: -2.83
NOM p-value: 0.000
FDR q-val: 0.000
PEREZ_TP53_TARGETS
NES: 1.91
NOM p-value: 0.009
FDR q-val: 0.046
Enrichment score (ES)
Positively correlated
Negatively correlated
Rank in Ordered Dataset
F
Cell cycle regulators
DMSO
BIX01294
PDK1
MLXIPL
PPP1R13B
CDKN1A
MGA
CDC45
EZH2
APEX1
E2F1
DDX39B
inhibitors
activators
Color Key
Log2(FPKM)
Differentiation regulators
positive reg.
negative reg.
G
Core Pluripotency Factors
t-hESC DMSO
t-hESC BIX-01294
Relative mRNA Expression (FPKM)
POU5F1
SOX2
NANOG
KLF4
MYC

Figure 10. Inhibition of G9a activity alters the pluripotent-like gene signature and induces differentiation programs in t-hESCs.

A) Western blot analysis of H3K9me2 and dot blot analysis of 5meC levels upon pharmacological inhibition of G9a HMTase activity (BIX-01294: 1μM, UNC0642: 5μM, 48h) in t-hESCs (H3K9me2: n≥4, **: p=0.0067; *: p=0.031; 5meC: n≥3, ***: p=0.00057; **: p=0.0083).

B) Western blot analysis of H3K27me1 and H3K27me3 levels in BIX-01294 (1μM, 48h) and UNC0642 (5μM, 48h) t-hESCs vs. DMSO controls. Total Histone H3 and GAPDH were used as loading controls. Relative OD signal quantification vs. GAPDH intensity is presented in the bar graph.

C) Volcano plot displaying differentially expressed genes in BIX-01294 (1μM, 48h) treated vs. DMSO. Statistically significant genes ($p<0.05$) with <5% FDR value (highlighted in red) were used in subsequent experiments.

D) mRNAsi from RNA-seq transcriptome profiling was determined for BIX-01294 (1μM) and DMSO treated t-hESCs.

E) GSEA revealed negative correlations between significant transcriptional modulations induced by BIX-01294 in t-hESCs (vs. DMSO, $p<0.05$) and genes upregulated in embryonic stem cells, genes induced by MYC, as well as p53 targets.

F) Heat maps representing expression of significantly modulated genes ($p<0.05$) associated with cell cycle (Negative Reg.: GO:0045786, Positive Reg.: GO:0045787) and differentiation regulation (Positive Reg./Stem Cell Diff.: GO:0045597, GO:0048863, Negative Reg.: GO:0045596) in BIX-01294-treated t-hESCs vs. DMSO controls.

G) Relative mRNA expression of core pluripotency factors in t-hESCs treated with BIX-01294 (1μM, 48h) or vehicle control (DMSO) determined by RNA-seq profiling.

As previously shown by qPCR analysis (**Figure.7F**), G9a mRNA expression is significantly higher in colon CSC-like HT29 cells compared to normal progenitor HIECs (**Figure.11A**, $p<0.0001$). Similar to our approach in transformed pluripotent cells, we investigated the impact of G9a knockdown and HMTase activity inhibition on the regulation of oncogenic functions in CRC. We used the same shRNA hairpins as in t-hESCs to knockdown G9a expression in HT29 cells, which demonstrated robust reductions of G9a protein expression (**Figure.11B**). I also observed decreased levels of H3K9me2 deposition in G9a knockdown HT29 cells (**Figure.11B**). I observed that G9a knockdown induces a substantial growth reduction compared to control shRNA, when cells were counted 72 hours after initial plating (**Figure. 11C**).

Considering the important difference in G9a expression observed between cancer and normal models, I compared the impact of G9a pharmacological inhibition on growth of CSC-like (t-hESCs and HT29 cells) vs. related normal models (H9 hESCs and intestinal progenitor HIEC cells). As observed in t-hESCs, I confirmed that both BIX-01294 and UNC0642 substantially reduce H3K9me2 deposition in HT29 cells (**Figure.11D**). In all cases, pharmacological inhibition of G9a led to selective toxicity toward neoplastic tissues, when comparing the EC_{50} calculated in normal cells over the EC_{50} calculated in transformed models (**Figure.11E**). Next, I tested whether increasing H3K9me2 deposition by inhibiting KDM4B activity could also have an impact on CRC and HIEC cell growth. Using the small molecule NCGC00244536 to inhibit KDM4B demethylase activity, I observed no effects on H3K9me2 levels in HT29 cells (5μM, 48h), potentially due to a plateau in such histone mark deposition (**Figure.11F**). No significant effects on the growth rate of HT29s were observed at doses of NCGC below 10μM (**Figure.11G**). However, when normal HIEC cells were treated with the KDM4B inhibitor, I observed an increase in H3K9me2 deposition and a ~40% decrease in cell counts (**Figure.11F, G**). To determine whether an increase in H3K9me2 deposition can modify the normal intestinal progenitor

long-term growth rate versus controls (H3K9me2^{Low}), I maintained surviving NCGC-treated HIEC cells (H3K9me2^{High}) in culture to allow them to recover from the treatment (24 hours drug-free recovery). I observed that H3K9me2^{High} HIECs retained substantially slower growth rates compared to H3K9me2^{Low} cells at 72h post-plating (**Figure.11H**). This suggests that increasing H3K9me2 deposition in normal intestinal cells is not sufficient to confer a long-term growth rate advantage.

Considering the important shift in transcriptional programs upon G9a inhibition in t-hESCs, we tested the impact of either knocking down G9a or inhibiting its HMTase activity, on the differentiation state of HT29 cells. In standard culture conditions, HT29 cells maintain CSC-like phenotype. However, both G9a knockdown and HMTase inhibition induced intestinal differentiation in this cancer cell line (**Figure.11I**). This is based on increased expression of differentiation markers such as KRT20, DPP4, KLF4, PTK6, and OSGIN1 (**Figure.11I**). Importantly, no significant expression changes were observed for those differentiation markers in normal HIEC cells treated with G9a inhibitors BIX-01294 and UNC0642 (**Figure.11I**). Altogether, our observations in CRC/intestinal models confirm the key role of G9a in maintaining the pro-oncogenic transcriptional programs responsible for compromising differentiation and promoting neoplastic stemness in colorectal tumors.

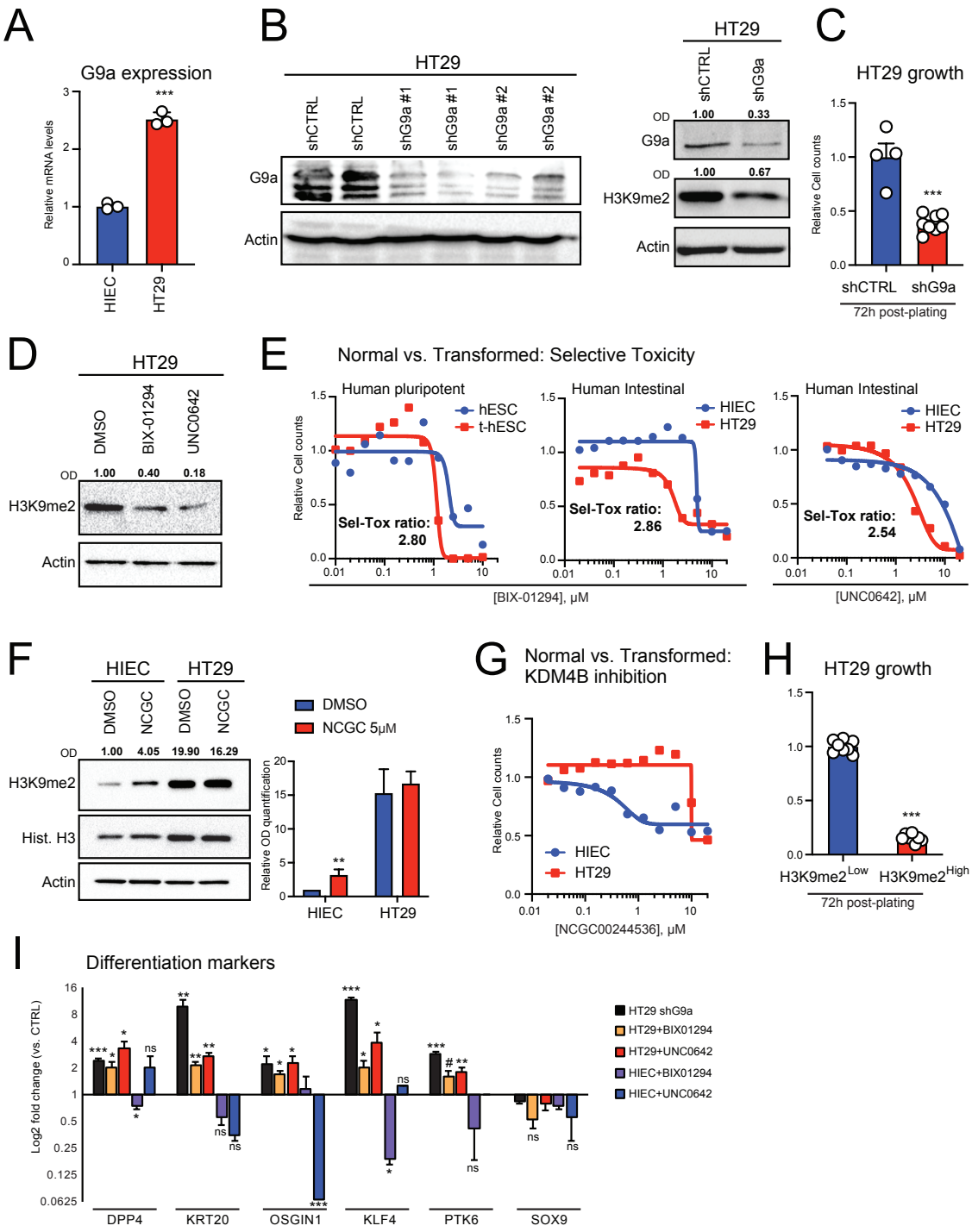


Figure 11. G9a inhibition selectively induces an intestinal differentiation program in human CRC cells.

A) qPCR assessment of G9a expression in HT29 vs. HIEC. GAPDH was used as housekeeping gene (n=3, ***: $p<0.0001$).

B) Validation of G9a knockdown efficiency in HT29 cells using shRNA clones #TRCN0000115670 (shG9a #1) and # TRCN0000115671 (shG9a #2) vs. scramble control shRNA #SHC002 (shCTRL) (left panel). G9a knockdown reduces H3K9me2 deposition in HT29 cells (right panel).

C) Growth rate assessment of G9a knockdown (shG9a, n=8) vs. non-silencing control (shCTRL, n=4) HT29 cells at 72h post-seeding (***: $p\leq0.0001$).

D) Western analysis of H3K9me2 levels in shCTRL vs. shG9a and BIX-01294 (1μM, 48h), UNC0642 (2.5μM, 48h) and DMSO-treated HT29 cells. Actin was used as loading control. Relative OD signal quantification for H3K9me2 vs. Actin intensity is presented.

E) Dose-response curves assessing selective toxicity (Sel-Tox) of G9a inhibition (BIX-01294) in t-hESCs vs. normal human ES cells (hESC), as well as cancer stem-like HT29 vs. normal HIEC cells (BIX-01294 and UNC0642). Sel-Tox ratios were determined as follow: EC50Normal/EC50Cancer (n=3). The data were best-fitted in a nonlinear curve fitting model.

F) Western blot analysis of H3K9me2 levels upon pharmacological inhibition of KDM4B demethylase activity using NCGC00244536 (5μM, 48h) in HIEC and HT29 cells. Histone H3 and Actin were used as loading controls. Relative OD signal quantification vs. Actin intensity is presented in the bar graph (n=4, **: $p=0.0016$).

G) Dose-response curves assessing toxicity of KDM4B inhibition (NCGC00244536) in CRC stem-like HT29 vs. normal HIEC cells.

H) Growth rate assessment of G9a knockdown (shG9a, n=8) vs. non-silencing control (shCTRL, n=4) HT29 cells, and H3K9me2High (n=8) vs. H3K9me2Low (n=8) HIEC cells at 72h post-seeding (***: $p<0.0001$).

I) Quantitative PCR assessment of intestinal differentiation markers expression in G9a knockdown, BIX-01294 (1μM, 48h) and UNC0642-treated (2.5μM, 48h) HT29 cells, as well as BIX-01294 (1μM, 48h) and UNC0642-treated (2.5μM, 48h) HIEC cells (n=3, ***: $p\leq0.0008$; **: $p\leq0.008$; *: $p\leq0.047$; #: $p=0.052$). Results are expressed as mRNA fold-change vs. respective controls (shCTRL and DMSO). GAPDH was used as a reference gene.

5.2. Objective #2: To study the impact of G9a inhibition on CCSC tumor-initiating capacities

Rationale. In this objective, I worked in collaboration with a Ph.D. student in the lab (Christopher Bergin), using primary CRC tissue samples to determine the potential for G9a inhibition to affect the tumor-initiating capacity of colon cancer stem cells. Then, I used an integrative genomic approach to study key functions altered by G9a inhibition in human colorectal cancer stem cells, and potentially impacting tumor-initiating and self-renewal capacities.

To further explore the functional role of G9a HMTase activity in CCSC populations from patient-derived tumor samples, we used a serial organoid formation assay and tested the potential for G9a inhibition to affect the tumor-initiating capacity of primary CCSCs. The course of this organoid formation assay is illustrated in **Figure.12A**. One of the key characteristics of CSCs resides in their capacity to initiate tumor formation[14,20] and Patient-Derived Organoids (PDO) enable the study of human solid tumors initiated from a single stem cell in a 3D system[83]. PDOs recapitulate functional and morphological characteristics of their primary tissue of origin, and can be used to predict drug responses in a pre-clinical or clinical *in vivo* setting[83]. PDO formation assays using primary human colon cancer specimens was shown to be a powerful method to test the impact of drugs on tumor-initiation activity[49,84,85]. Briefly, surgically-resected pieces of colon tumors from consenting patients were minced and enzymatically dissociated in Collagenase A. Once put in ultra-low adhesion culture conditions in a serum-free media supplemented with bFGF, EGF, and a cocktail of soluble factors (N2 and B27 supplements), CSCs were enriched from bulk cell suspensions via formation of spheroids[21]. I confirmed CSC enrichment from spheroid cultures based on flow cytometry profiling of key surface markers ($CD133^+/CD44^+$) characteristic of CCSCs (**Figure.12B**)[16,86,87]. Functionally, CD44 is a cell surface protein that plays an important role in extracellular matrix functions. On the other hand, CD133 is a

transmembrane glycoprotein organizing the cell membrane. Both surface markers are being a characteristic for cancer stem cells[16]. I could observe that CCSC populations were enriched through the formation of spheroids in suspension, as marked by the large proportion of $CD133^{+}/CD44^{+}$ cells, comparable to the CSC-like HT29 cell line (**Figure.12B**)[16,45]. In contrast, HIEC cells, which display no organoid-initiating capacity present a low frequency of CD133/CD44 double-positive cells (**Figure.12B**). Patient specimens used in the context of my work were obtained **from Dr. Catherine O'Brien** (University of Toronto) and via collaboration with **Dr. Rebecca Auer** (Ottawa Hospital Research Institute (OHRI)). The clinical information about the three patient samples used in our study is available in **Table-5.** To establish tumor organoid cultures, CSC-enriched spheroids were dissociated and passed through a 70-μm strainer to eliminate non-single-cell aggregates. Patient-specific cell suspensions were mixed with Matrigel in spheroid culture media. Cell-Matrigel mixtures were immediately plated as 300 μl domes in 6-well plates. By opposition to spheroids, 3D organoids were described as mini tumors, maintaining primary patient tumor heterogeneity[49]. Accordingly, we observed higher transcript expression of the CCSC marker LGR5 in patient-derived spheroids vs. bulk 3D organoids resulting from single-cell seeding. Specifically, this protein is a member of the G-protein coupled receptor family that acts as a biomarker of adult and cancer stem cells in colon tissue. (**Figure.12C**)[16,88]. Moreover, we observed reduced levels of the differentiation marker DPP4, along with a clear enrichment of G9a expression in patient-derived spheroids vs. bulk organoids (**Figure.12C**). Two different working doses of the G9a inhibitor UNC0642 (2.5 – 5μM vs. DMSO control) were added to each organoid-containing dome and incubated for 7 days[84], followed by a 7-day drug-free incubation (**Figure.12A**). At day-14, organoids were imaged and PDO frequency was determined for each well. G9a inhibition in a primary series of patient-derived 3D organoids resulted in a significant decrease of frequency compared to vehicle-treated groups (**Figure.12D**). Residual primary organoids were

dissociated and re-seeded in a secondary series. Importantly, no further drug treatments were performed on this secondary series of PDOs, enabling *bona fide* assessment of persisting tumor-initiating cell populations in samples previously treated with a G9a inhibitor (vs. DMSO control) (**Figure.12A**). Thus, we observed that UNC0642-treated primary organoids had lower tumor-initiating capacity when plated in a secondary assay (**Figure.12E**). Overall, these experiments confirm the potential for G9a inhibition to restrict tumor-initiating functions in colorectal cancer, which represents a key hallmark of CSCs.

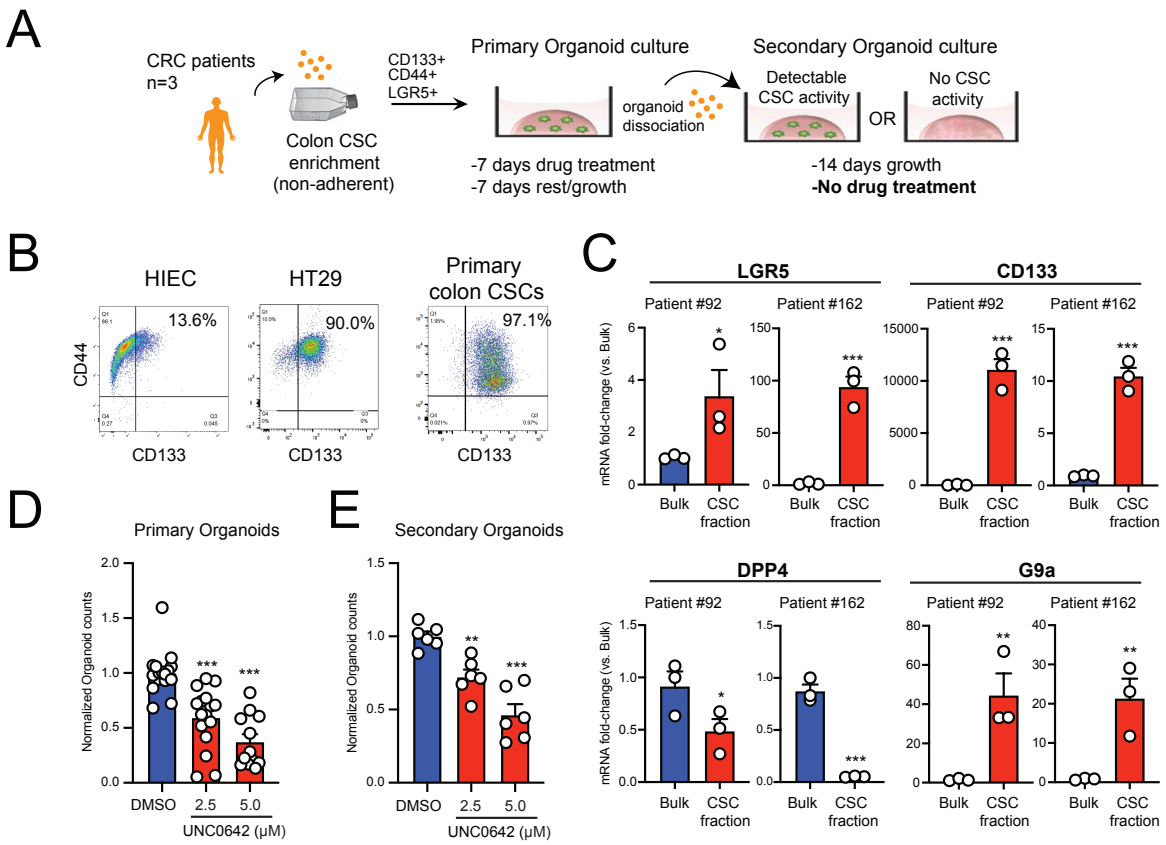


Figure 12. Pharmacological inhibition of G9a HMTase activity is reducing tumor-initiating capacity in primary CRC patient samples

A) Schematic representation of serial organoid formation assay using primary human CRC tissues. Cancer stem cell (CSC) fraction is enriched in non-adherent sphere cultures.

B) Flow cytometry profiling of patient-derived spheroid cultures for CCSC markers CD133 and CD44. Normal progenitor cells (HIEC) and CCSC-like HT29 were presented as controls.

C) Quantitative PCR assessment of CCSC markers LGR5, CD44, and CD113, as well as G9a expression in bulk primary organoids vs. CSC-enriched fractions (spheroid cultures) (2 patients tested, n=3 per patient, ***: $p \leq 0.0004$; **: $p \leq 0.0093$; *: $p \leq 0.041$).

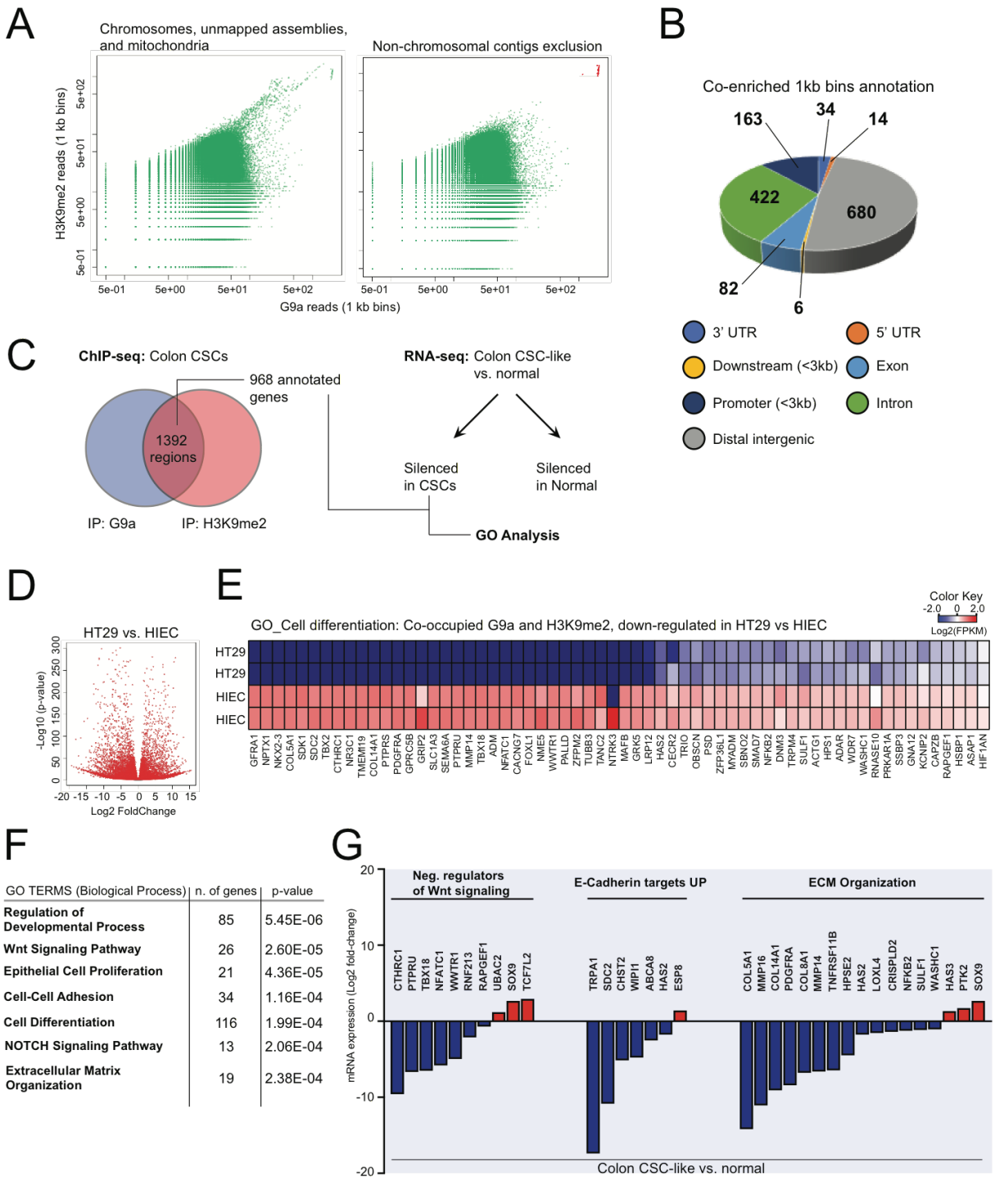
D) Primary organoid formation frequency observed upon UNC0642 treatments (2.5μM and 5μM, 7 days). Organoid counts were normalized vs. DMSO controls (3 patients, $n \geq 12$, ***: $p < 0.0001$).

E) Organoid formation frequencies observed in secondary plating assays, for DMSO, UNC0642 2.5μM and UNC0642 5μM groups. Organoid counts were normalized vs. DMSO controls (3 patients, n=6, ***: $p < 0.0001$; **: $p = 0.005$).

Mechanistically, I searched for potential pathways involved in the observed G9a-mediated maintenance of CSCs in CRC patient samples. Analysis of chromatin immunoprecipitation (ChIP-seq) data performed on primary CCSC fragmented chromatin, and using anti-G9a and anti-H3K9me2 antibodies revealed that G9a and H3K9me2 are significantly co-enriched at 1392 genomic regions (1kb) in patient-derived CCSCs (**Figure.13A, C**)[57]. A majority of the 968 annotated G9a/H3K9me2 co-enriched sites were mapped in introns (422) and distal intergenic regions (680) (**Figure.13B**). This is consistent with methylated DNA immunoprecipitation (MeDIP) assays performed by Joshua Haebe, a master student in our lab, who demonstrated that DNA methylation was heavily impacted following G9a inhibition at long interspersed element-1 (LINE1) genomic elements. These retrotransposable elements are observed at a high frequency within intronic and intergenic space[89].

Next, I performed an integrative analysis of transcriptionally silenced genes in CSC-like HT29 cells vs. normal HIECs (RNA-seq) that are simultaneously enriched/co-occupied with G9a and H3K9me2 in patient-derived CCSCs (ChIP-seq) (**Figure.13C**). A total of 6963 genes presented lower expression levels in HT29 cells vs. normal HIECs, with a p-value below 0.05 (**Figure.13D**). By performing our integrative outline, I identified a list of 183 annotated entities presenting lower expression in CRC (vs. normal intestine progenitors) and co-occupied by G9a and H3K9me2 in primary CCSCs (**Figure.13C**). Among these genes, several were associated with the cell differentiation process, as presented in **Figure.13E**. We perform a gene ontology analysis on the list of 183 genes and found that genes putatively silenced by G9a in CCSCs were also associated with processes such as the regulation of development (GO:0050793), negative regulation of canonical Wnt signaling (GO:0030178), and extracellular matrix organization (GO:0031012) (**Figure.13F, G**). We also observed that several genes involved in the repression of epithelium-to-mesenchyme transition (E-cadherin targets up)[90] were silenced in CRC (vs. normal) and co-enriched with G9a and H3K9me2 in patient-derived CCSCs

(**Figure.13G**). Taken together, our data on primary CRC patient samples confirm that G9a is playing a critical role as an epigenetic factor maintaining CSC activity in human CRCs. With the help of François Desrochers, a graduate student in our lab, I assembled a recapitulative schema of G9a functions through its associated chromatin mark H3K9me2 in human CCSC biology (**Figure.14**).



GO TERMS (Biological Process)	n. of genes	p-value
Regulation of Developmental Process	85	5.45E-06
Wnt Signaling Pathway	26	2.60E-05
Epithelial Cell Proliferation	21	4.36E-05
Cell-Cell Adhesion	34	1.16E-04
Cell Differentiation	116	1.99E-04
NOTCH Signaling Pathway	13	2.06E-04
Extracellular Matrix Organization	19	2.38E-04

Figure 13. G9a-dependent H3K9me2 deposition is regulating key molecular pathways in CCSCs

A) Read counts for 1Kb regions detected by G9a and H3K9me2 ChIP-seq in human primary CCSCs. 1Kb regions were filtered against the ENCODE blacklist. Left panel represents chromosome, unmapped assemblies, and mitochondria regions. Right panel represents only chromosome-mapped 1Kb regions, which were used for subsequent analyses.

B) Pie chart summarizing genomic distribution of retained 1392 co-occupied (G9a/H3K9me2) 1Kb regions in primary CCSCs.

C) Integrative analysis using ChIP-seq and RNA-seq transcriptome profiling to identify G9a and H3K9me2 co-enriched annotated genes in primary human CCSCs (GSE82131) that are silenced in HT29 (CCSC-like) vs. normal progenitors (HIEC).

D) Volcano plot displaying differentially expressed genes between HT29 and HIEC cells. Statistically significant genes ($p<0.05$) with <5% FDR value is highlighted in red and were used in related analyses.

E) Heat map representing RNA-seq profiling of genes involved in cell differentiation (GO_Cell_Differentiation: GO:0030154), co-enriched with G9a and H3K9me2 in human CCSCs (GSE82131) and silenced in HT29 vs. HIEC cells ($p<0.05$). Relative mRNA levels of G9a in HT29 vs. HIEC is also presented (n=3, ***: $p<0.0001$).

F) Gene ontology analysis of G9a/H3K9me2 co-enriched genes and silenced in CCSC (vs. normal) reveals the implication of key pathways linked to tumorigenicity.

G) Diagram illustrating relative mRNA expression of G9a/H3K9me2 co-enriched genes in human CCSCs (vs. HIEC, $p<0.05$) across key functional categories (GO:0030178, M18757, GO:0031012).

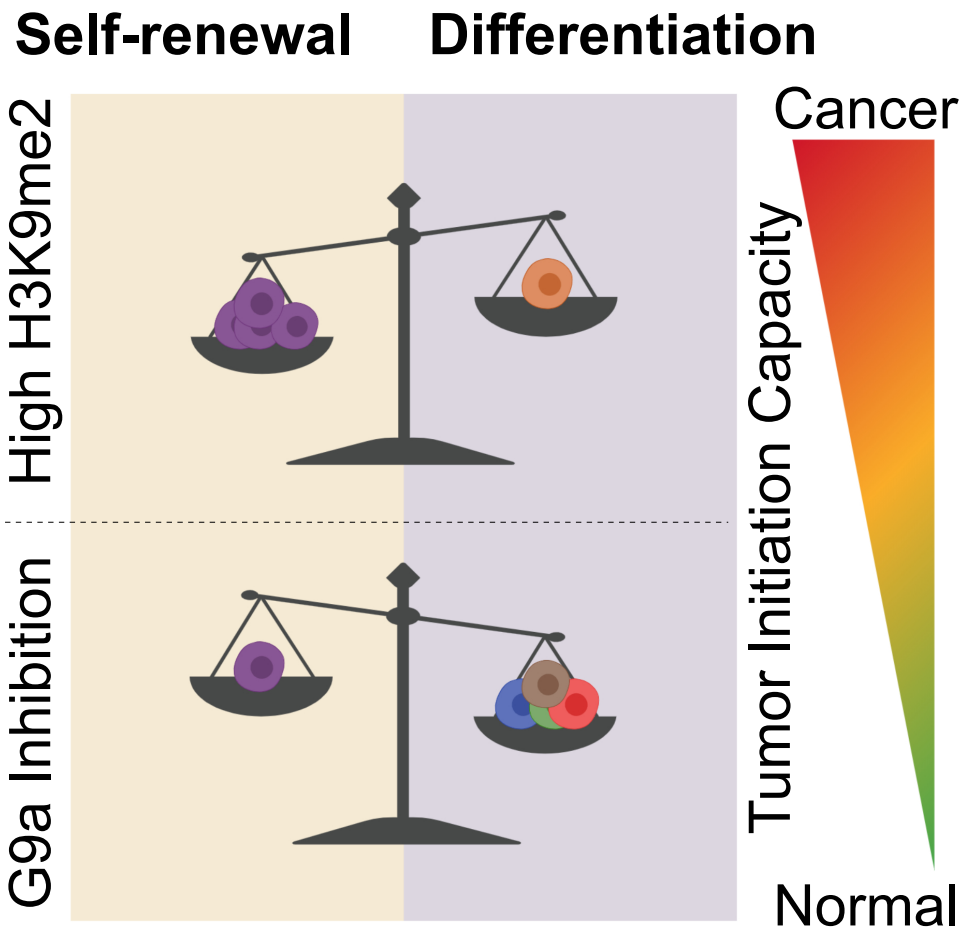


Figure 14. Recapitulative schema of G9a functions through its associated chromatin mark H3K9me2 in human CCSC biology. Elevated G9a expression, as observed in CRC promotes self-renewal at the expense of differentiation in CSC populations. This is accompanied by enhanced tumor-initiation capacity. G9a inhibition appears as an interesting strategy to restore such biological functions to normal-like levels.

5.3. Objective #3: To identify new selective inhibitors contributing to the elimination of colorectal CSCs by down-regulating G9a activity

Rationale. The need for next-generation chemotherapeutics targeting the small reservoirs of CSCs is still largely unmet, since the drugs currently used in the clinic to treat colon cancer tend to be ineffective against such a portion of tumor heterogeneity. Although we confirmed that G9a inhibition represents an interesting strategy to curb CSC activity in colorectal cancer, we also observed that pharmacological inhibitors of G9a are not making their way to clinical trials[61-63]. We sought to address this issue by using a stem cell-based, high-throughput phenotypic screening assay to identify new translational compounds to block G9a functions in CCSCs. The ultimate goal of this objective is to bring forward new drug candidates that could be used in the future as cornerstone components of chemotherapeutic regimens to target CSC populations and improve patient survival.

As outlined in previous sections, the histone methyltransferase G9a represents an attractive therapeutic target to impact CCSC functions. Considering the absence G9a inhibitors in current and previous clinical trials, our lab initiated a chemical screening campaign to identify clinically approved drugs with the capacity to downregulate G9a activity in human CSCs.

In this phenotypic screening project, we used a chemical library of 1847 clinically approved molecules (MedChem Express, HY-L022M) to screen for specific drugs causing growth inhibition (↓cell counts), loss of pluripotency (↓OCT-4 expression) and altered H3K9me2 deposition upon drug treatment vs. control t-hESCs (**Figure.15A**). One of the main advantages for choosing a library of clinically approved (or repurposed) drugs refers to their previously validated safety parameters for use in human patients, which expedites further development toward clinical applications[91,92]. On the other hand, phenotypic screening has the advantage to test the effect of small molecules on complete biological systems, such as

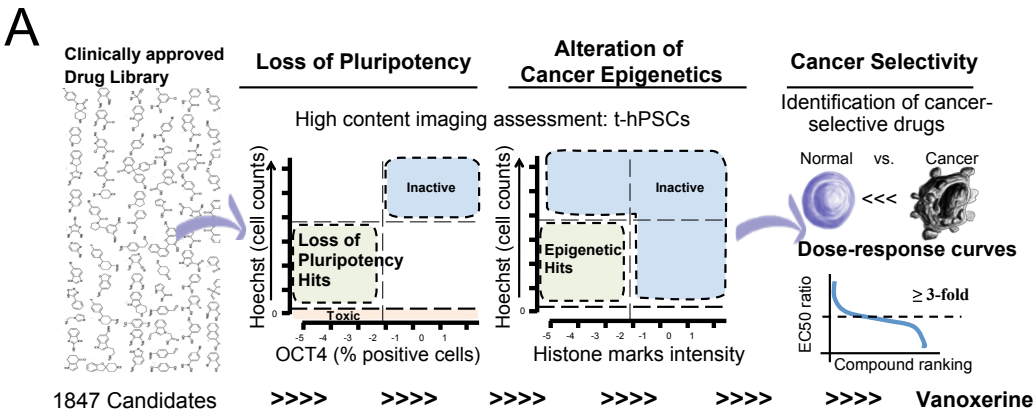
mammalian cells, and lead to the identification of candidates with *bona fide* pathologically-relevant molecular mechanisms of action[93,94]. Thus, lead compounds identified from phenotypic screening are suggested to have higher chances of success in clinical trials[93,94].

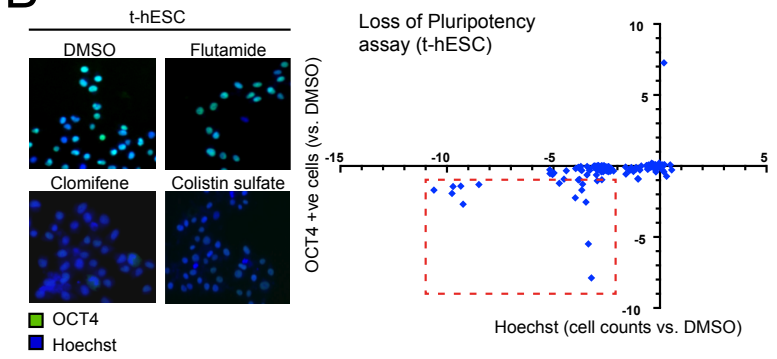
First, I used high-content imaging to measure cell counts (Hoechst nuclear staining) and the percentage of OCT4-positive cells (immunofluorescence) following 48h drug treatments (5μM vs. controls) in a 96-well format. Entities falling outside of -1.0 log2 fold-change from control DMSO for OCT-4 positivity, and -2.0 log2 fold-changes for cell counts (marked by a red dashed rectangle) were scored as CSC-bioactive hits (**Figure.15B**). Next, I selected 32 compounds inducing loss of pluripotency in the first screening step and used them individually to treat t-hESCs (5μM, 48h) in 96-well plates. Treated cells were immunostained to detect H3K9me2 fluctuations in response to treatments. DMSO was used as negative control, while BIX-01294 and CWP232228 were used as positive controls, decreasing H3K9me2 deposition and cell counts respectively (**Figure.15C, D**). High-content imaging analysis revealed that certain loss of pluripotency inducers could decrease H3K9me2 along with decreased cell counts (**Figure.15C, D**). Drugs outperforming positive controls for H3K9me2 and growth inhibition were considered as hits (**Figure.15D**, red circles). This is the case for Vanoxerine (VXN), which successfully passed the first two screening assays (**Figure.15E**). VXN was subjected to a selective toxicity test, measuring its impact on growth inhibition in 3 colon cancer cell lines relative to normal intestinal epithelium progenitors (**Figure.15A**). Dose-response experiments demonstrated that VXN has a lower EC_{50} in all 3 tested CRC cell lines (HT29=1.15μM, SW480=1.51, HCT116=1.18) compared to its calculated EC_{50} in HIEC cells (27.05μM) (**Figure.15F**). Considering that all four cell models exhibit similar doubling time, ranging between 55 and 65 hours, we could determine that VXN displays selective toxicity against colorectal cancer cells vs. normal intestinal progenitor cells.

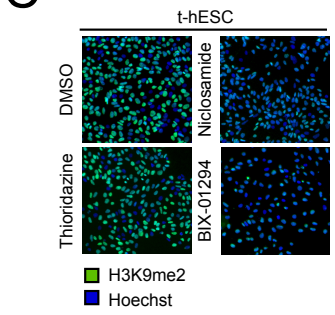
With the help of Joshua Haebe, a master student in the lab, I sought to confirm the capacity of VXN to inhibit H3K9me2 deposition. Western blot experiments performed in a dose-response format showed that VXN reduces H3K9me2 levels in HT29 cells after 48 hours of treatment (**Figure.15G**). As an explanation for this phenomenon, I observed that VXN is causing a decrease of G9a expression in HT29 cells, while no impact was seen in HIEC cells (**Figure.15H**). In fact, G9a protein expression is almost undetectable in normal HIEC cells (**Figure.15H**), which could represent a key criterion for neoplastic selective toxicity (**Figure.15F**).

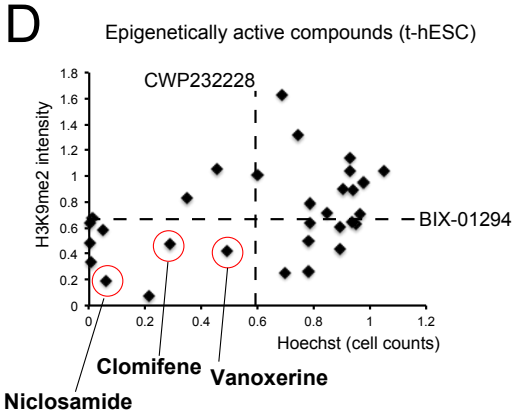
Additional work from other lab members using VXN revealed that this repurposed compound acts via the dopamine transporter SLC6A3 to block the tumor-initiating capacity of human CCSCs. This was shown using the previously described serial 3D organoid formation assay, involving primary tumor samples, as well as through *in vivo* serial tumor transplantation experiments in a syngeneic mouse model.

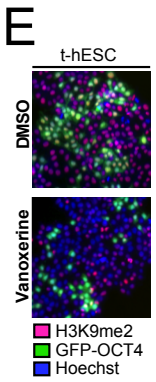
Altogether, my work led to compelling data confirming that one of our repurposed drug candidates (VXN) is altering the function of the histone methyltransferase, G9a, specifically in colon cancer stem-like cell models. Importantly, we expect this compound to have a higher chance of success in future clinical trials for G9a inhibition in cancer, since this repurposed drug was already deemed safe for use in human patients[95].

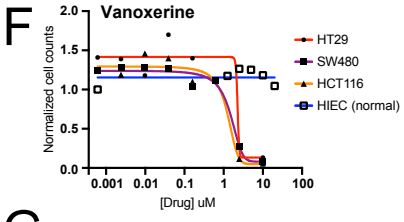

A
Clinically approved Drug Library
Loss of Pluripotency
Alteration of Cancer Epigenetics
Cancer Selectivity
High content imaging assessment: t-hPSCs
Identification of cancer-selective drugs
Hoechst (cell counts)
Inactive
Loss of Pluripotency Hits
Toxic
OCT4 (% positive cells)
Epigenetic Hits
Histone marks intensity
Normal vs. Cancer
Dose-response curves
EC50 ratio
≥ 3-fold
Compound ranking
1847 Candidates
>>>>
Vanoxerine

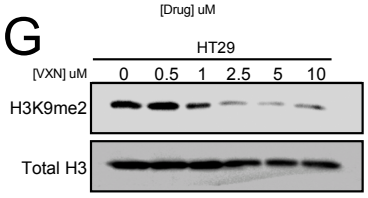

B
t-hESC
DMSO
Flutamide
Clomifene
Colistin sulfate
OCT4
Hoechst
Loss of Pluripotency assay (t-hESC)
OCT4 +ve cells (vs. DMSO)
Hoechst (cell counts vs. DMSO)


C
t-hESC
DMSO
Niclosamide
Thioridazine
BIX-01294
H3K9me2
Hoechst


D
Epigenetically active compounds (t-hESC)
CWP232228
BIX-01294
H3K9me2 intensity
Hoechst (cell counts)
Clomifene
Vanoxerine
Niclosamide


E
t-hESC
DMSO
Vanoxerine
H3K9me2
GFP-OCT4
Hoechst


F
Vanoxerine
Normalized cell counts
HT29
SW480
HCT116
HIEC (normal)
[Drug] uM


G
HT29
[VXN] uM
0 0.5 1 2.5 5 10
H3K9me2
Total H3

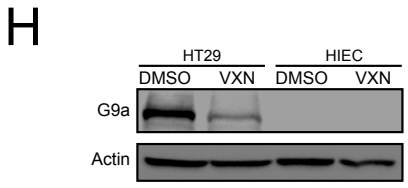

H
HT29
HIEC
DMSO VXN DMSO VXN
G9a
Actin

Figure 15. Phenotypic screening strategy to identify new colon CSC-bioactive compounds.

A) Phenotypic screening strategy to identify repurposed epigenetically active drugs, exhibiting a selective activity in human colon cancer stem cells vs. normal intestinal progenitors. Transformed human embryonic stem cells (t-hESC) were used to screen a chemical library of 1847 clinically approved molecules.

B) Representative images of OCT4 staining in t-hESCs upon drug treatments (5μM, 48h) and DMSO control (left panel). Quantification of OCT4-positive cells vs. cell counts was scored for each compound relative to DMSO control (right panel). Red dashed box indicates drugs inducing loss of pluripotency (↓OCT4+ve) and decreased cell counts (log2 fold-changes).

C) Representative images of H3K9me2 staining in t-hESCs following treatments (5μM, 48h) with drugs inducing loss of pluripotency.

D) Quantification of H3K9me2 signal in t-hESC cells vs. cell counts was scored for each loss of a pluripotency inducer relative to DMSO control. CWP232228 (100nM) and BIX-01294 (1μM) were used as positive controls for cell counts and H3K9me2 inhibition, respectively (dashed lines). Three main hit compounds (Vanoxerine (VXN), Niclosamide, Clomifene) were identified in red circles.

E) Fluorescence images of H3K9me2 immunostaining (red) and OCT4-GFP reporter in DMSO vs. VXN-treated (5μM, 48h) t-hESC.

F) Demonstration of selective toxicity for VXN in human CRC cell lines HT29 (n=4), SW480 (n=3), HCT116 (n=3) vs. normal intestinal progenitors HIEC (n=6). Data were best-fitted using nonlinear curve fitting model.

G) Western blot analysis confirming that VXN is inhibiting H3K9me2 deposition in a dose-dependent manner in HT29 cells (vs. total histone H3).

H) Western blot analysis showing that VXN is decreasing G9a expression in HT29 CRC line. G9a protein is undetectable in HIECs.

6. Discussion

6.1. Summary

Herein, I established a fundamental role for G9a histone methyltransferase as a regulator of pluripotency networks in colon CSCs. Indeed, G9a expression levels are significantly associated with poor clinical outcomes and correlate with a pluripotent transcriptional signature, which was recently demonstrated as a robust indicator of tumor aggressiveness and neoplastic stem cell phenotype[77]. Moreover, our results established that G9a regulates primary human colorectal cancer cells, including functionally defined CSCs. The identification of a molecular regulator of CSCs enabled us to successfully test the relevance of the CSC model in CRC; if CSCs are clinically relevant, then therapeutically targeting them should result in effective and durable therapies.

As for previous studies using the same class of small molecule inhibitors (dual G9a/GLP inhibitors)[12, 56, 57], we generalized our findings to G9a functions, since the knockdown experiments recapitulated chemical treatments. Moreover, genetic knockdown of G9a delayed tumor recurrence and pharmacologic inhibition of G9a prevented the growth of recurrent tumors. Taken together, we found that G9a-dependent epigenetic reprogramming promotes cancer recurrence.

Further, our survival analysis identified G9a overexpression as a risk factor for relapse in CRC, supporting H3K9me2 enrichment as a biomarker of the CRC epigenetic signature[30, 96]. We showed that knocking down G9a resulted in genome-wide methylation and expression of multiple genes, which shows that it is a key epigenetic regulator of oncogene expression in CSCs derived from patient

samples. Further, pharmacological inhibition of G9a stimulates differentiation and restricts self-renewal and tumorigenicity in CCSCs.

6.2. To characterize the role of G9a in maintaining the molecular functions of cancer stem cells in colorectal cancer (CRC)

Human pluripotent cells were shown useful for the study of cancer stemness and were used to identify and characterize thioridazine and CWP232228 as pharmacological agents suppressing CSC activity in human leukemia[45,81]. Additionally, the embryonic stem-like transcriptional signature in somatic tissues is recognized to elicit oncogenic self-renewal[30]. I decided to study in depth the function of G9a HMTase in this cell model since its activity was extensively linked to early embryonic stages and the regulation of pluripotent gene networks[53]. Both, knockdown and chemical inhibition of G9a induced a decrease of H3K9me2 and DNA methylation levels in t-hESCs. Moreover, our survival analysis identified G9a overexpression as a risk factor for relapse in CRC, supporting H3K9me2 enrichment as a biomarker of the CRC epigenetic signature.

Herein, we established a fundamental role for G9a/H3K9me2 as a regulator of pluripotency networks in CCSCs. Despite, G9a HMTase activity being extensively linked to early embryonic stages and the regulation of pluripotent gene networks[37,48,53], its inhibition in t-hESCs had no effects on pluripotency factor expression, including c-Myc. Still, we observed a robust inhibition of c-Myc targets when H3K9me2-deposition was inhibited, which strengthen the concept that G9a plays a key role in the establishment of an oncogenic pluripotent gene signature in CSCs.

Our results suggest that epigenetic remodelling is a major feature of tumor recurrence, and this remodeling is associated with an acquired dependence on G9a HMTase. Interestingly, G9a activity and

its relationship with pluripotency networks in CRC show similarities with another recent study relating Enhancer of zeste homolog 2 (EZH2) activity to hedgehog signaling, core pluripotency factors, and the maintenance of tumor-initiating functions in CCSCs[49]. Further, interplay between G9a and EZH2-dependent chromatin patterning was documented, where G9a is involved in H3K27 mono-methylation, serving as a precursor to EZH2 recruitment and transcriptional silencing[37,97, 98]. Altogether, this supports the existence of an epigenetic regulatory hub involving G9a, which could be targeted to eliminate CSC populations, either alone or in combination with chemotherapy.

As for previous studies using the same class of small molecule inhibitors (dual G9a/GLP inhibitors)[12,22,56,57], we generalized our findings to G9a functions, since 1) knockdown experiments recapitulated chemical treatments and 2) no impact on relapse free survival (RFS) was observed for GLP overexpression in CRC patients. Moreover, in a preliminary experiment, I suppressed the expression of GLP in CSC-like HT29 cells using a chemical inhibitor (MS012). As a result, blocking this histone methyltransferase reduced HT29 cell counts. Consequently, we cannot exclude the possibility that GLP modulation alone would show effects on CCSC biology.

Although previous reports highlighted G9a-mediated H3K9me2 deposition as a driver of CSC biology in other cancers[22,56], a better understanding of its function regulating CCSC epigenetic signature was necessary.

6.3. Impact of G9a inhibition on CCSC tumor-initiating capacity

Our survival analysis identified G9a overexpression as a risk factor for relapse in CRC, supporting H3K9me2 enrichment as a biomarker of the CRC epigenetic signature[96,99]. DNMT3A overexpression also correlated with a higher risk for relapse in colon cancer[100]. In fact, G9a is a known stabilizing

factor, promoting the recruitment of DNMT3A to imprinted loci in ES cells, independent of its HMTase activity[55]. Further, G9a HMTase activity was linked to cancer stemness in lung carcinoma by maintaining oncogenic DNA methylation patterns via interaction with DNMT1, but not DNMT3A[22]. Both knockdown and chemical inhibition of G9a induced a decrease of DNA methylation levels in t-hESCs. Hence, other investigations would be necessary to determine the nature of the interactions taking place between G9a and DNA methylation machinery in CSCs.

We also identified Long interspersed element -1 (LINE1) as a target of genomic hypomethylation induced by G9a inhibition in CSCs. Transcriptional activation of LINE1 retrotransposable elements, and anti-tumorigenic effects observed following G9a inhibition are consistent with a previous study showing that the reactivation of endogenous retro-elements can induce viral mimicry and block tumor-initiating functions in CRC patient samples[101].

Although previous reports highlighted G9a-mediated H3K9me2 deposition as a driver of CSC biology in other cancers[22,56], a better understanding of its function regulating CCSC epigenetic signature was necessary. Specifically, the findings on primary CRC patient samples showing that G9a chemical inhibition with UNC0642 induced a decrease of mini tumors or organoid size and a lower tumor-initiating capacity in primary organoids when plated in a secondary assay established a fundamental role for G9a/H3K9me2 as a regulator of pluripotency networks in CCSCs.

We are aware that Organoid Formation Frequencies observed in Patient Derived Organoids (PDO) assays can be due to simple cytotoxic effects on bulk cancer cells during the treatment window. Secondary organoid plating using residual drug-treated primary PDOs addresses this caveat by measuring the persistence of tumor-initiating cell populations over time. Considering that PDOs can accurately predict drug response *in vivo*, the completion of this objective will set the foundations for

future experiments in human-to-mouse serial transplantation assays, testing the potential of our repurposed drugs for clinical application.

Furthermore, G9a knockdown in a colon cancer stem cell-like model (HT29) significantly decreased cell growth and the level of H3K9me2, underscoring the relevance of G9a as key enzyme regulator for cancer epigenetic regulation and presenting G9a as an attractive therapeutic target. We also observed that the knockdown of this histone methyltransferase induces the expression of differentiation markers in shG9a-treated HT29s, namely markers such as KRT20, DPP4, KLF4, PTK6, and OSGIN1. These observations in CRC/intestinal models confirm a key role for G9a in maintaining pro-oncogenic transcriptional programs responsible for compromising differentiation and promoting neoplastic stemness in colorectal tumors.

Our integrative multi-omics approach identified canonical Wnt/β-catenin signaling, extracellular matrix organization, and epithelium-to-mesenchyme transition as key networks regulated by G9a HMTase activity. These networks represent hallmarks of neoplastic self-renewal and de-differentiation (**Figure.14)**[45,54,57,102,103]. Collectively, our data supports recent findings connecting G9a expression to cell proliferation and canonical Wnt activity in CRC[12,48,57], as well as previous demonstration that G9a expression enhances *in vivo* tumorigenicity and resistance to ionizing radiation in HT29 cells[104].

Beyond direct impacts on CSC biology, our results exposed the cancer-selective nature of G9a inhibition, when comparing growth and differentiation responses in neoplastic versus normal tissues. Specifically, inhibition of G9a with a chemical drug (BIX-0194) induced selective toxicity in colon cancer stem cell models (HT29 and t-hESC) compared to normal lines (HIEC and hESC), showing the selective effect of G9a inhibition. This result illustrates the specific targeting of G9a affecting the cancer stem cells while sparing the normal progenitors. This aligns with a rising trend suggesting G9a as an attractive therapeutic target, which would exhibit limited toxicity to healthy tissues. Furthermore, our

study also demonstrates that chemical induction of H3K9me2 deposition in normal intestinal progenitor cells (HIEC) is not sufficient to solely drive the acquisition of oncogenic characteristics such as enhanced growth and tumor-initiation capacity.

Collectively, our data indicate that colorectal cancer cells are highly dependent on G9a, which functions as a pleiotropic regulator that maintains the viability and proliferative capacity of colorectal cancer cells in general but also governs CSC self-renewal. The effectiveness of transient inhibition suggests that CRCs may be reliant on G9a to sustain growth and clonal maintenance.

6.4. Identification of new selective inhibitors contributing to the elimination of colorectal CSCs by down regulating G9a activity

The development of therapies targeting specific epigenetic changes holds promise as a new category of treatments for CRC alone or in combination with other treatments to overcome drug resistance[3,45,59,70.] Precisely, targeting CSC stemness determinants including self-renewal and *in vivo* tumorigenicity has been proposed as a therapeutic goal and our study provides strong evidence for this concept. A number of molecules and gene signatures that underlie stemness have been identified and future studies should be undertaken to identify and target other key pathways driving self-renewal in the CSC subset. For that, our study provides a template for preclinical evaluation of other CSC targets that focuses on primary samples and clonal *in vitro* CSC assays.

The approach proposed herein relies on targeting context-specific molecular pathways to which CCSCs are dependent and representing points of vulnerability for CSCs. Specifically, as it has been demonstrated earlier, the epigenetic regulator G9a plays a central part in colorectal CSC self-renewal and presents a direct avenue to target the stem-like functions of CSCs with small-molecule inhibition[22, 54, 56, 57]. This approach will facilitate their development toward clinical applications.

Critically, we characterized the impact of 3 clinically approved drugs with anti-CSC properties in CRC models; Vanoxerine, Niclosamide and Clomifene. Our 3 repurposed candidates are clinically approved and, consequently, deemed safe for uses in human patients. Vanoxerine is a high-affinity dopamine reuptake inhibitor that was used in the treatment of cocaine addiction. Niclosamide is used as an oral anthelminthic to treat parasitic infections. Lastly, Clomifene is used to induce ovulation in women who have infertility problems. Afterwards, transcriptome profiling analysis revealed that one repurposed drug candidate; Vanoxerine, impacts the expression of the histone methyltransferase G9a and this result was confirmed in a western blot experiment where this drug decreases the levels of G9a and H3K9me2 selectively in CSCs models (HT29) while sparing the normal progenitors (HIEC). Beyond direct impacts on CSC biology, our results exposed the cancer-selective nature of G9a inhibition, when comparing growth and differentiation responses in neoplastic versus normal tissues. This aligns with a rising trend suggesting G9a as an attractive therapeutic target, which would exhibit limited toxicity to healthy tissues.

Overall, inhibition of biomedically important epigenetic processes is currently a subject of intensive investigations[12,41,47], therefore, we envision that future studies will lead to important advances in design and development of specific inhibitors of the therapeutic potential. Towards this aim, our work highlights that targeting the histone lysine methyltransferases is an alternative strategy to commonly used approaches that target histone substrates.

However, we noted a paucity of *in vivo* investigations and lack of clinical trials using these inhibitors, potentially due to poor pharmacokinetics parameters. Despite such *in vivo* hurdles, additional work on alternative ways to regulate G9a HMTase expression or activity in cancer is still critical to develop novel, effective approaches targeting G9a for future therapies against CSCs.

As a next critical step for this project, pre-clinical *in vivo* testing of novel colon CSC-targeting agents will be performed in serial xenotransplantation models. We strongly believe that our work will shed light on other types of human malignancies and provide momentum for significant advances in other areas of cancer research.

7. Conclusions and future directions

In this project, I investigated the role of G9a in regulating stemness and *in vivo* tumorigenicity and its role on genome-wide epigenetic reprogramming of colorectal cancer. Firstly, I showed that G9a/H3K9me2 is a regulator of pluripotency networks in CCSCs where G9a expression levels are significantly associated with poor clinical outcomes. Then, using patient-derived CSC, I determined that G9a is involved in sphere forming and growth capacities of CSCs and its inhibition results in decreased stemness and tumorigenicity *in vitro*. Further, I showed that knocking down G9a resulted in genome-wide methylation and expression of multiple genes, which shows that it is a key epigenetic regulator of oncogene expression in colon CSC.

Altogether, these critical findings support G9a as a potential biomarker and key regulator of pro-oncogenic pathways maintaining CSC activity in colon tumors. Little is known about the functional role of G9a in CCSCs and its associated H3K9me2 as a potential target for cancer therapy[12]. Although I used specific G9a chemical inhibitors in this study, current G9a inhibitors are not suitable for use in humans due to poor bioavailability and modest potency[48,61,63]. The development of G9a inhibitors that are more potent and have favourable pharmacokinetics may lead to an effective therapy for this cancer. Overall, by targeting G9a alone or in combination with other therapies may be a novel approach to treating CRC patients[12,22].

8. References:

1 Yarbro, J. W. Oncogenes and cancer suppressor genes. *Semin Oncol Nurs* **8**, 30-39, doi:10.1016/0749-2081(92)90006-o (1992).

2 Brenner, H., Kloor, M. & Pox, C. P. Colorectal cancer. *Lancet* **383**, 1490-1502, doi:10.1016/S0140-6736(13)61649-9 (2014).

3 Jung, G., Hernández-Illán, E., Moreira, L., Balaguer, F. & Goel, A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. *Nat Rev Gastroenterol Hepatol* **17**, 111-130, doi:10.1038/s41575-019-0230-y (2020).

4 Smith, G. *et al.* Mutations in APC, Kirsten-ras, and p53--alternative genetic pathways to colorectal cancer. *Proc Natl Acad Sci U S A* **99**, 9433-9438, doi:10.1073/pnas.122612899 (2002).

5 de Sousa e Melo, F. *et al.* A distinct role for Lgr5. *Nature* **543**, 676-680, doi:10.1038/nature21713 (2017).

6 Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759-767, doi:10.1016/0092-8674(90)90186-i (1990).

7 Verma, M. The Role of Epigenomics in the Study of Cancer Biomarkers and in the Development of Diagnostic Tools. *Adv Exp Med Biol* **867**, 59-80, doi:10.1007/978-94-017-7215-0_5 (2015).

8 Baylin, S. B. DNA methylation and gene silencing in cancer. *Nat Clin Pract Oncol* **2 Suppl 1**, S4-11, doi:10.1038/ncponc0354 (2005).

9 Vaiopoulos, A. G., Athanasoula, K. C. h. & Papavassiliou, A. G. Epigenetic modifications in colorectal cancer: molecular insights and therapeutic challenges. *Biochim Biophys Acta* **1842**, 971-980, doi:10.1016/j.bbadis.2014.02.006 (2014).

10 Phan, T. G. & Croucher, P. I. The dormant cancer cell life cycle. *Nat Rev Cancer* **20**, 398-411, doi:10.1038/s41568-020-0263-0 (2020).

11 Varghese, A. Chemotherapy for Stage II Colon Cancer. *Clin Colon Rectal Surg* 28, 256-261, doi:10.1055/s-0035-1564430 (2015).

12 Wang, H. *et al.* Overexpression of PSAT1 regulated by G9A sustains cell proliferation in colorectal cancer. *Signal Transduct Target Ther* **5**, 47, doi:10.1038/s41392-020-0147-5 (2020).

13 Meacham, C. E. & Morrison, S. J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328-337, doi:10.1038/nature12624 (2013).

14 Visvader, J. E. & Lindeman, G. J. Cancer stem cells: current status and evolving complexities. *Cell Stem Cell* **10**, 717-728, doi:10.1016/j.stem.2012.05.007 (2012).

15 Yu, Z., Pestell, T. G., Lisanti, M. P. & Pestell, R. G. Cancer stem cells. *Int J Biochem Cell Biol* **44**, 2144-2151, doi:10.1016/j.biocel.2012.08.022 (2012).

16 O'Brien, C. A., Pollett, A., Gallinger, S. & Dick, J. E. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106-110, doi:nature05372 [pii]10.1038/nature05372 (2007).

17 Yeung, T. M., Gandhi, S. C., Wilding, J. L., Muschel, R. & Bodmer, W. F. Cancer stem cells from colorectal cancer-derived cell lines. *Proc Natl Acad Sci U S A* **107**, 3722-3727, doi:10.1073/pnas.0915135107 (2010).

18 Roberts, C. M., Cardenas, C. & Tedja, R. The Role of Intra-Tumoral Heterogeneity and Its Clinical Relevance in Epithelial Ovarian Cancer Recurrence and Metastasis. *Cancers (Basel)* **11**, doi:10.3390/cancers11081083 (2019).

19 Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* **3**, 730-737 (1997).

20 Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275-291, doi:10.1016/j.stem.2014.02.006 (2014).

21 Kreso, A. & O'Brien, C. A. Colon cancer stem cells. *Curr Protoc Stem Cell Biol* **Chapter 3**, Unit 3.1, doi:10.1002/9780470151808.sc0301s7 (2008).
22 Pangeni, R. P. *et al.* G9a regulates tumorigenicity and stemness through genome-wide DNA methylation reprogramming in non-small cell lung cancer. *Clin Epigenetics* **12**, 88, doi:10.1186/s13148-020-00879-5 (2020).
23 Dawood, S., Austin, L. & Cristofanilli, M. Cancer stem cells: implications for cancer therapy. *Oncology (Williston Park)* **28**, 1101-1107, 1110 (2014).
24 Snippert, H. J. *et al.* Prominin-1/CD133 marks stem cells and early progenitors in mouse small intestine. *Gastroenterology* **136**, 2187-2194 e2181, doi:10.1053/j.gastro.2009.03.002 (2009).
25 Munro, M. J., Wickremesekera, S. K., Peng, L., Tan, S. T. & Itinteang, T. Cancer stem cells in colorectal cancer: a review. *J Clin Pathol*, doi:10.1136/jclinpath-2017-204739 (2017).
26 Kreso, A. *et al.* Self-renewal as a therapeutic target in human colorectal cancer. *Nat Med* **20**, 29-36, doi:10.1038/nm.3418 (2014).
27 Eppert, K. *et al.* Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med* **17**, 1086-1093, doi:10.1038/nm.2415 (2011).
28 Ng, S. W. *et al.* A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433-437, doi:10.1038/nature20598 (2016).
29 Werbowetski-Ogilvie, T. E., Morrison, L. C., Fiebig-Comyn, A. & Bhatia, M. In vivo generation of neural tumors from neoplastic pluripotent stem cells models early human pediatric brain tumor formation. *Stem Cells* **30**, 392-404, doi:10.1002/stem.1017 (2012).
30 Wong, D. J. *et al.* Module map of stem cell genes guides creation of epithelial cancer stem cells. *Cell Stem Cell* **2**, 333-344, doi:10.1016/j.stem.2008.02.009 (2008).
31 Werbowetski-Ogilvie, T. E. *et al.* Characterization of human embryonic stem cells with features of neoplastic progression. *Nat Biotechnol* **27**, 91-97, doi:10.1038/nbt.1516 (2009).
32 Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**, 499-507, doi:10.1038/ng.127 (2008).
33 Goel, A. & Boland, C. R. Epigenetics of colorectal cancer. *Gastroenterology* **143**, 1442-1460.e1441, doi:10.1053/j.gastro.2012.09.032 (2012).
34 Toh, T. B., Lim, J. J. & Chow, E. K. Epigenetics in cancer stem cells. *Mol Cancer* **16**, 29, doi:10.1186/s12943-017-0596-9 (2017).
35 Sawan, C. & Herceg, Z. Histone modifications and cancer. *Adv Genet* **70**, 57-85, doi:10.1016/B978-0-12-380866-0.60003-4 (2010).
36 Qin, J., Wen, B., Liang, Y., Yu, W. & Li, H. Histone Modifications and their Role in Colorectal Cancer (Review). *Pathol Oncol Res*, doi:10.1007/s12253-019-00663-8 (2019).
37 Mozzetta, C. *et al.* The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol Cell* **53**, 277-289, doi:10.1016/j.molcel.2013.12.005 (2014).
38 Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074-1080, doi:10.1126/science.1063127 (2001).
39 Reddy, M. A., Park, J. T. & Natarajan, R. Epigenetic modifications and diabetic nephropathy. *Kidney Res Clin Pract* **31**, 139-150, doi:10.1016/j.krcp.2012.07.004 (2012).
40 Casciello, F., Windloch, K., Gannon, F. & Lee, J. S. Functional Role of G9a Histone Methyltransferase in Cancer. *Front Immunol* **6**, 487, doi:10.3389/fimmu.2015.00487 (2015).
41 Shi, Y. Histone lysine demethylases: emerging roles in development, physiology and disease. *Nat Rev Genet* **8**, 829-833, doi:10.1038/nrg2218 (2007).
42 Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425-432, doi:10.1038/nature05918 (2007).

43 Hong, S. N. Genetic and epigenetic alterations of colorectal cancer. *Intest Res* **16**, 327-337, doi:10.5217/ir.2018.16.3.327 (2018).
44 Feinberg, A. P., Koldobskiy, M. A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet* **17**, 284-299, doi:10.1038/nrg.2016.13 (2016).
45 Benoit, Y. D. *et al.* Sam68 Allows Selective Targeting of Human Cancer Stem Cells. *Cell Chem Biol* **24**, 833-844.e839, doi:10.1016/j.chembiol.2017.05.026 (2017).
46 Suvà, M. L., Riggi, N. & Bernstein, B. E. Epigenetic reprogramming in cancer. *Science* **339**, 1567-1570, doi:10.1126/science.1230184 (2013).
47 Wainwright, E. N. & Scaffidi, P. Epigenetics and Cancer Stem Cells: Unleashing, Hijacking, and Restricting Cellular Plasticity. *Trends Cancer* **3**, 372-386, doi:10.1016/j.trecan.2017.04.004 (2017).
48 Bergin, C. J. & Benoit, Y. D. G9a Is SETting the Stage for Colorectal Oncogenesis. *Genes (Basel)* **11**, doi:10.3390/genes11060616 (2020).
49 Lima-Fernandes, E. *et al.* Targeting bivalency de-represses Indian Hedgehog and inhibits self-renewal of colorectal cancer-initiating cells. *Nat Commun* **10**, 1436, doi:10.1038/s41467-019-09309-4 (2019).
50 Alam, H., Gu, B. & Lee, M. G. Histone methylation modifiers in cellular signaling pathways. *Cell Mol Life Sci* **72**, 4577-4592, doi:10.1007/s00018-015-2023-y (2015).
51 Chaturvedi, C. P. *et al.* Maintenance of gene silencing by the coordinate action of the H3K9 methyltransferase G9a/KMT1C and the H3K4 demethylase Jarid1a/KDM5A. *Proc Natl Acad Sci U S A* **109**, 18845-18850, doi:10.1073/pnas.1213951109 (2012).
52 Kramer, J. M. Regulation of cell differentiation and function by the euchromatin histone methyltranserfases G9a and GLP. *Biochem Cell Biol* **94**, 26-32, doi:10.1139/bcb-2015-0017 (2016).
53 Tachibana, M. *et al.* G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16**, 1779-1791, doi:10.1101/gad.989402 (2002).
54 Liu, S. *et al.* G9a is essential for EMT-mediated metastasis and maintenance of cancer stem cell-like characters in head and neck squamous cell carcinoma. *Oncotarget* **6**, 6887-6901, doi:10.18632/oncotarget.3159 (2015).
55 Zhang, T. *et al.* G9a/GLP Complex Maintains Imprinted DNA Methylation in Embryonic Stem Cells. *Cell Rep* **15**, 77-85, doi:10.1016/j.celrep.2016.03.007 (2016).
56 Lehnertz, B. *et al.* The methyltransferase G9a regulates HoxA9-dependent transcription in AML. *Genes Dev* **28**, 317-327, doi:10.1101/gad.236794.113 (2014).
57 Kato, S. *et al.* Gain-of-function genetic alterations of G9a drive oncogenesis. *Cancer Discov*, doi:10.1158/2159-8290.CD-19-0532 (2020).
58 Watson, Z. L. *et al.* Histone methyltransferases EHMT1 and EHMT2 (GLP/G9A) maintain PARP inhibitor resistance in high-grade serous ovarian carcinoma. *Clin Epigenetics* **11**, 165, doi:10.1186/s13148-019-0758-2 (2019).
59 Luo, C. W. *et al.* G9a governs colon cancer stem cell phenotype and chemoradioresistance through PP2A-RPA axis-mediated DNA damage response. *Radiother Oncol* **124**, 395-402, doi:10.1016/j.radonc.2017.03.002 (2017).
60 Cho, H. S. *et al.* Enhanced expression of EHMT2 is involved in the proliferation of cancer cells through negative regulation of SIAH1. *Neoplasia* 13, 676-684, doi:10.1593/neo.11512 (2011).
61 Kubicek, S. *et al.* Reversal of H3K9me2 by a small-molecule inhibitor for the G9a histone methyltransferase. *Mol Cell* **25**, 473-481, doi:10.1016/j.molcel.2007.01.017 (2007).
62 Liu, F. *et al.* Discovery of an in vivo chemical probe of the lysine methyltransferases G9a and GLP. *J Med Chem* **56**, 8931-8942, doi:10.1021/jm401480r (2013).
63 Lenstra, D. C., Al Temimi, A. H. K. & Mecinović, J. Inhibition of histone lysine methyltransferases G9a and GLP by ejection of structural Zn(II). *Bioorg Med Chem Lett* **28**, 1234-1238, doi:10.1016/j.bmcl.2018.02.043 (2018).

64 Lu, H., Lei, X. & Zhang, Q. Liver-specific knockout of histone methyltransferase G9a impairs liver maturation and dysregulates inflammatory, cytoprotective, and drug-processing genes. *Xenobiotica*, 1-13, doi:10.1080/00498254.2018.1490044 (2018).

65 Papait, R. *et al.* Histone Methyltransferase G9a Is Required for Cardiomyocyte Homeostasis and Hypertrophy. *Circulation* **136**, 1233-1246, doi:10.1161/CIRCULATIONAHA.117.028561 (2017).

66 Lehnertz, B. *et al.* Activating and inhibitory functions for the histone lysine methyltransferase G9a in T helper cell differentiation and function. *J Exp Med* **207**, 915-922, doi:10.1084/jem.20100363 (2010).

67 Ugarte, F. *et al.* Progressive Chromatin Condensation and H3K9 Methylation Regulate the Differentiation of Embryonic and Hematopoietic Stem Cells. *Stem Cell Reports* **5**, 728-740, doi:10.1016/j.stemcr.2015.09.009 (2015).

68 Dylla, S. J. *et al.* Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy. *PLoS One* **3**, e2428, doi:10.1371/journal.pone.0002428 (2008).

69 Roberti, A., Valdes, A. F., Torrecillas, R., Fraga, M. F. & Fernandez, A. F. Epigenetics in cancer therapy and nanomedicine. *Clin Epigenetics* **11**, 81, doi:10.1186/s13148-019-0675-4 (2019).

70 Richart, L. & Margueron, R. Drugging histone methyltransferases in cancer. *Curr Opin Chem Biol* **56**, 51-62, doi:10.1016/j.cbpa.2019.11.009 (2020).

71 Perreault, N. & Beaulieu, J. F. Use of the dissociating enzyme thermolysin to generate viable human normal intestinal epithelial cell cultures. *Exp Cell Res* **224**, 354-364, doi:10.1006/excr.1996.0145 (1996).

72 Benoit, Y. D. *et al.* Cooperation between HNF-1alpha, Cdx2, and GATA-4 in initiating an enterocytic differentiation program in a normal human intestinal epithelial progenitor cell line. *Am J Physiol Gastrointest Liver Physiol* **298**, G504-517, doi:10.1152/ajpgi.00265.2009 (2010).

73 Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419, doi:10.1038/nmeth.4197 (2017).

74 Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**, e45 (2001).

75 Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* **47**, W556-W560, doi:10.1093/nar/gkz430 (2019).

76 Colaprico, A. *et al.* TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* **44**, e71, doi:10.1093/nar/gkv1507 (2016).

77 Malta, T. M. *et al.* Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* **173**, 338-354.e315, doi:10.1016/j.cell.2018.03.034 (2018).

78 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

79 Raskov, H., Søby, J. H., Troelsen, J., Bojesen, R. D. & Gögenur, I. Driver Gene Mutations and Epigenetics in Colorectal Cancer. *Ann Surg* **271**, 75-85, doi:10.1097/SLA.0000000000003393 (2020).

80 Feldman, N. *et al.* G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8**, 188-194, doi:ncb1353 [pii]10.1038/ncb1353 (2006).

81 Sachlos, E. *et al.* Identification of drugs including a dopamine receptor antagonist that selectively target cancer stem cells. *Cell* **149**, 1284-1297, doi:10.1016/j.cell.2012.03.049 (2012).

82 Perez, C. A., Ott, J., Mays, D. J. & Pietenpol, J. A. p63 consensus DNA-binding site: identification, analysis and application into a p63MH algorithm. *Oncogene* **26**, 7363-7370, doi:10.1038/sj.onc.1210561 (2007).

83 Tuveson, D. & Clevers, H. Cancer modeling meets human organoid technology. *Science* **364**, 952-955, doi:10.1126/science.aaw6985 (2019).

84 van de Wetering, M. *et al.* Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* **161**, 933-945, doi:10.1016/j.cell.2015.03.053 (2015).

85 Crespo, M. *et al.* Colonic organoids derived from human induced pluripotent stem cells for modeling colorectal cancer and drug testing. *Nat Med* **23**, 878-884, doi:10.1038/nm.4355 (2017).

86 Sikandar, S. S. *et al.* NOTCH signaling is required for formation and self-renewal of tumor-initiating cells and for repression of secretory cell differentiation in colon cancer. *Cancer Res* **70**, 1469-1478, doi:0008-5472.CAN-09-2557 [pii]10.1158/0008-472.CAN-09-2557 (2010).

87 Benoit, Y. D. *et al.* Pharmacological inhibition of polycomb repressive complex-2 activity induces apoptosis in human colon cancer stem cells. *Exp Cell Res* **319**, 1463-1470, doi:10.1016/j.yexcr.2013.04.006 (2013).

88 de Sousa e Melo, F. *et al.* A distinct role for Lgr5(+) stem cells in primary and metastatic colon cancer. *Nature* **543**, 676-680, doi:10.1038/nature21713 (2017).

89 Ardeljan, D., Taylor, M. S., Ting, D. T. & Burns, K. H. The Human Long Interspersed Element-1 Retrotransposon: An Emerging Biomarker of Neoplasia. *Clin Chem* **63**, 816-822, doi:10.1373/clinchem.2016.257444 (2017).

90 Onder, T. T. *et al.* Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* **68**, 3645-3654, doi:10.1158/0008-5472.CAN-07-2938 (2008).

91 Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nat Rev Drug Discov* **10**, 507-519, doi:10.1038/nrd3480 (2011).

92 Chen, H., Wu, J., Gao, Y. & Zhou, J. Scaffold Repurposing of Old Drugs Towards New Cancer Drug Discovery. *Curr Top Med Chem* **16**, 2107-2114 (2016).

93 Swinney, D. C. Phenotypic vs. target-based drug discovery for first-in-class medicines. *Clin Pharmacol Ther* **93**, 299-301, doi:10.1038/clpt.2012.236 (2013).

94 Kotz, J. 3 (Nature Publishing Group, Science-Business eXchange, 2012).

95 Dittrich, H. C. *et al.* COR-ART: A multicenter, randomized, double-blind, placebo-controlled dose-ranging study to evaluate single oral doses of vanoxerine for conversion of recent-onset atrial fibrillation or flutter to normal sinus rhythm. *Heart Rhythm* **12**, 1105-1112, doi:10.1016/j.hrthm.2015.02.014 (2015).

96 Nakazawa, T. *et al.* Global histone modification of histone H3 in colorectal cancer and its precursor lesions. *Hum Pathol* **43**, 834-842, doi:10.1016/j.humpath.2011.07.009 (2012).

97 Curry, E. *et al.* Dual EZH2 and EHMT2 histone methyltransferase inhibition increases biological efficacy in breast cancer cells. *Clin Epigenetics* **7**, 84, doi:10.1186/s13148-015-0118-9 (2015).

98 Wang, Z. & Patel, D. J. Small molecule epigenetic inhibitors targeted to histone lysine methyltransferases and demethylases. *Q Rev Biophys* **46**, 349-373, doi:10.1017/S0033583513000085 (2013).

99 Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40**, 897-903, doi:10.1038/ng.154 (2008).

100 Subramaniam, D., Thombre, R., Dhar, A. & Anant, S. DNA methyltransferases: a novel target for prevention and therapy. *Front Oncol* **4**, 80, doi:10.3389/fonc.2014.00080 (2014).

101 Roulois, D. *et al.* DNA-Demethylating Agents Target Colorectal Cancer Cells by Inducing Viral Mimicry by Endogenous Transcripts. *Cell* **162**, 961-973, doi:10.1016/j.cell.2015.07.056 (2015).

102 Benoit, Y. D., Guezguez, B., Boyd, A. L. & Bhatia, M. Molecular pathways: epigenetic modulation of Wnt-glycogen synthase kinase-3 signaling to target human cancer stem cells. *Clin Cancer Res* **20**, 5372-5378, doi:10.1158/1078-0432.CCR-13-2491 (2014).

103 Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871-890, doi:10.1016/j.cell.2009.11.007 (2009).

104 Luo, C.-W. *et al.* G9a governs colon cancer stem cell phenotype and chemoradioresistance through PP2A-RPA axis-mediated DNA damage response. *Radiotherapy and Oncology* **124**, 395-402, doi:https://doi.org/10.1016/j.radonc.2017.03.002 (2017).

9. Appendix

Supplemental Table- 1: List of antibodies used in this study

Antibodies	**Application**	**Source**	**Identifier**
Mouse Anti-Actin Monoclonal Antibody	Wb	Millipore	MAB1501
Mouse Anti-5-methylcytosine, clone 33D3	Db	Millipore	MABE146
Mouse Monoclonal Anti-GAPDH	Wb	Abcam	Ab-2107448
Rabbit Anti-G9a	Wb	Abcam	Ab-40542
Mouse Anti-H3K9me2	Wb,IF	Abcam	Ab-1220
Rabbit Anti-H3	Wb	Abcam	Ab-176842
Anti-Rabbit IgG (H+L), HRP Conjugate	Wb	Biorad	170-6515
Anti-Mouse secondary antibody	Wb	Promega	W402B
Goat anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	IF	Invitrogen	A-11029
Rabbit Anti-H3K27me1	Wb	Invitrogen	39377
Rabbit Anti-H3K27me3	Wb	Invitrogen	PA5-31817
OCT4	IF	Cell signaling	C30A3
Hoecht 33342	IF	Invitrogen	H3570
Anti-Human CD44 Mouse Monoclonal Antibody, APC conjugated, Clone G44-26	Flow Cytometry	Invitrogen	12-0441-82
Anti-Human CD133 Mouse Monoclonal Antibody, PE conjugated, Clone W6B3C1	Flow Cytometry	Invitrogen	17-1338-41

Supplemental Table -2: List of qPCR primers used in this study

Oligonucleotides

G9a	Forward	AGTGATGATGTCCACTCACTGGGA
	Reverse	AGAGACTGAAGTCATCACCCACCA
LSD1	Forwards	GCCTGAAGAACCATCGGGGCA
	Reverse	AGTCGGCTCTGGAAAGCTGCG
Dpp4	Forward	CCTTCTACTCTGATGAGTCACTGC
	Reverse	GTGCCACTAAGCAGTTCCATCTTC
KRT20	Forward	ACTAACGGAGCTGAGACGCA
	Reverse	GTAACGGGCCTTGGTCTCCT
OSGN1	Forward	AAACATGAAGTCGGTCCTCAC
	Reverse	GAAGACCTCTTCGCTTCTTCTG
KLF4	Forward	GCCACCCACACTTGTGATTA
	Reverse	GTGGTAAGGTTTCTCACCTGT
PTK6	Forward	AGGCCATTACTCCACCAAATC
	Reverse	GTTGGACATGCCTGGGTA
SOX9	Forward	GTA CCG CACTTGCACAAC
	Reverse	TCTCGCTCTCGTTCAGAAGTC
LGR5	Forward	TGCTCTTCACCAACTGCATC
	Reverse	CTCAGGCTCACCAGATCCTC
CD133	Forward	TTTGGTGCAAATGTGGAAAA
	Reverse	TTGAAGCTGTTCTGCAGGTG
GAPDH	Forward	GAAATCCCATCACCAATCTTCCAGG
	Reverse	GCAATTGAGCCCCAGCCTTCTC

Supplemental Table-3: Information on gene signatures used in GSEAs

Signature ID	MSigDB Systematic #	Exact source
WONG_EMBRYONIC_STEM_CELL_CORE	M7079	PMID 18397753
HALLMARK_MYC_TARGETS_V1	M5926	N/A
GRADE_COLON_AND_RECTAL_CANCER_UP	M16740	PMID 17210682
PEREZ_TP53_TARGETS	M4391	PMID 17563751
ONDER_CDH1_TARGETS_1_UP	M18757	PMID 18483246
GO_NEGATIVE_REGULATION_OF_WNT_SIGNALING_PATHWAY	M16074	GO:0030178

Supplemental Table- 4: GO terms highlighted in gene ontology analysis

GO Term	GO ID	P-value	Adjusted P-value	Ontology
Regulation of developmental process	GO:0050793	5.45E-06	1.12E-02	Biological Process
Wnt signaling pathway	GO:0016055	2.60E-05	1.72E-02	Biological Process
cell-cell signaling by wnt	GO:0198738	2.79E-05	1.72E-02	Biological Process
Epithelial cell proliferation	GO:0050673	4.36E-05	2.00E-02	Biological Process
cell population proliferation	GO:0008283	4.16E+01	2.31E-02	Biological Process
Cell-cell adhesion	GO:0098609	1.16E-04	2.21E-02	Biological Process
Cell differentiation	GO:0030154	1.99E-04	2.64E-02	Biological Process
regulation of cell differentiation	GO:0045595	3.23E-04	3.37E-02	Biological Process
Notch signaling pathway	GO:0007219	2.06E-04	2.64E-02	Biological Process
negative regulation of Notch signaling pathway	GO:0045746	1.36E-04	2.21E-02	Biological Process
Extracellular matrix organization	GO:0030198	2.38E-04	2.67E-02	Biological Process

Supplemental Table- 5: Clinical information for CRC patients involved in this study

Patient ID	Origin	Tumor Stage	Mutation status	Age	Gender
92	Primary colon	IV	APC, P53, BRAF	N/A	N/A
162	Primary colon	III	APC, P53, KRAS, BRAF	63	F
146	Primary colon	III	P53, KRAS, BRAF	65	M