23/11/2022

“Development of a CRISPR-Cas9 system to genetically modify epigenetic regulators KDM5A and KDM5B and study their effects on Breast Cancer Stem Cells”

Inter-institutional Interdepartmental Program of Postgraduate Studies
“Molecular and Cellular Biology and Biotechnology”

Biomedical Research Institute (BRI)/ FORTH

University of Ioannina

Master Student’s name: Eleftherios Sinanis, Biologist
Supervisor: Angeliki Magklara
University Campus of Ioannina, 45110 Ioannina, Greece,
(Angeliki Magklara’s lab, Biomedical Research Institute,
Foundation for Research and Technology –Hellas,
University of Ioannina

Ioannina 2022

Contents

Abstract

Cancer, in simple terms, is defined as “a group of diseases that are characterized by the uncontrolled growth and spread of abnormal cells” and has been torturing humanity from an early age. Breast cancer, in particular, is the most commonly diagnosed type of cancer in women (WHO 2020). Modern approaches have confirmed the presence of a particular subpopulation of cells in tumors, termed cancer stem cells (CSCs), possessing self-renewal and differentiation potential, as well as increased tumor-initiating ability and intrinsic resistance to conventional therapies compared to the bulk of tumor cells, and are thus considered major drivers of the disease. It is of extreme importance, therefore, the development of novel strategies targeting and eliminating this particular subpopulation of cells. Epigenetic mechanisms have been identified as a major contributor in regulating these cells’ most problematic features, like their stemness and therapy resistance properties. In the present thesis we targeted two epigenetic regulators, KDM5A and KDM5B, implicated in cancer and breast cancer biology. Our results revealed significant connections of these enzymes with the luminal breast cancer cell biology. Moreover, KDM5A was found to be an important regulator of CSC properties. These results reveal new research avenues for the investigation of the biological mechanisms that regulate the CSC phenotype and highlight novel potential therapeutic targets for the elimination of these cells.

Περίληψη

Ο καρκίνος, με απλούς όρους, ορίζεται ως «μια ομάδα ασθενειών που χαρακτηρίζονται από την ανεξέλεγκτη ανάπτυξη και διάχυση μη φυσιολογικών κυττάρων» και ιστορικά αποτελεί πρόβλημα για την ανθρωπότητα. Ο καρκίνος του μαστού συγκεκριμένα, αποτελεί τον πιο συχνό τύπο καρκίνου στις γυναίκες. Μοντέρνες προσεγγίσεις επιβεβαίωσαν την παρουσία ενός συγκεκριμένου υποπληθυσμού κυττάρων στους όγκους, τα επονομαζόμενα καρκινικά βλαστικά κύτταρα (ΚΒΚ), τα οποία διαθέτουν ογκογονικές ιδιότητες, την δυνατότητα αυτοανανέωσης και διαφοροποίησης, καθώς και έμφυτους μηχανισμούς ανοχής σε θεραπείες. Λόγω των ιδιοτήτων αυτών, θεωρούνται βασικοί παράγοντες προόδου της νόσου. Η ανάπτυξη νέων στρατηγικών θεραπείας για την αντιμετώπιση της συγκεκριμένης ομάδας κυττάρων αποτελεί μεγίστης σημασίας στόχο. Οι επιγενετικοί μηχανισμοί έχουν ταυτοποιηθεί σαν εξέχουσας σημασίας ρυθμιστές και διαμεσολαβητές των πιο προβληματικών ιδιοτήτων των ΚΒΚ, όπως είναι οι βλαστικές τους ιδιότητες και οι δυνατότητες ανοχής σε θεραπεία. Στην παρούσα εργασία μελετήσαμε δυο επιγενετικούς ρυθμιστές, τις KDM5A και KDM5B, ένζυμα ήδη συνδεδεμένα βιβλιογραφικά με τον καρκίνο του μαστού. Η μελέτη πραγματοποιήθηκε μέσω της γενετκής αποσιώπησης τους, με μεθόδους CRISPR και shRNA. Τα αποτελέσματα μας αποκαλύπτουν μια σημαντική σύνδεση αυτών των ενζύμων με τον luminal καρκίνο του μαστού και, συγκεκριμένα για την KDM5A, μια σημαντική συσχέτιση με τον καρκινικό βλαστικό φαινότυπο. Επιπλέον, Τα αποτελέσματα μας αποκαλύπτουν ευκαιρίες για περαιτέρω μελέτη, στην διαλεύκανση βιολογικών μηχανισμών ρύθμισης του ΚΒΚ φαινοτύπου, και, πιθανώς, νέους θεραπευτικούς στόχους.

1. Introduction

1.1 Breast Cancer

Cancer, in simple terms, is defined as “a group of diseases that are characterized by the uncontrolled growth and spread of abnormal cells” and has been torturing humanity from an early age. The name itself, «καρκίνος» comes from Hippocrates, who is widely regarded as the “father of medicine”. Breast cancer is the most frequently diagnosed cancer in women and ranks first among causes of cancer-related death in women globally (WHO, 2020). It is caused when the ductal or alveolar cells that form the mammary gland undergo malignant transformation. The mammary gland consists of two types of epithelial cells, the luminal cells, that cover the inner parts of the duct, and the cells that form the outer scaffold, basal and myoepithelial (Figure 1.I). These cells arise from multipotent stem cells (Figure 1.II), following distinct differentiation pathways. This innate heterogeneity caused by the developmental hierarchy makes the identification of the cell of origin quite important in disease, as any of these cells could transform and form a tumor (Figure 1.I and 1.II)[1]. Based on the expression of certain receptors, breast cancer (BC) could be categorized into 3 major groups: HR^+ BC expressing hormonal receptors ER (Estrogen Receptor) or PR (Progesterone Receptor), $HER2^+$, BC expressing human epidermal receptor 2 (HER2) and triple negative breast cancer (TNBC) (ER^-, PR^-, $HER2^-$). HR^+ breast cancer is the most prevalent type, accounting for 60-70% of breast cancer cases[2]. Both molecular and histological evidence indicate that each subtype can be further categorized into distinct subgroups as shown on Table 1.I.

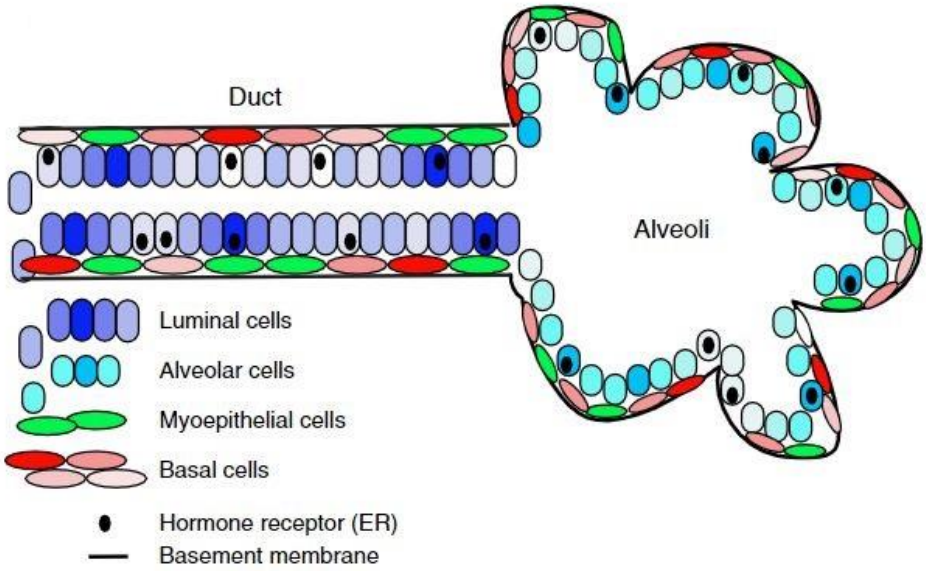


Figure 1.I: Mature Mammary Gland schematic: *Different cells, with different biological characteristics, constitute the mature mammary gland, any of which can be transformed to malignant cells, adding the first layer of heterogeneity within the grown tumor (modified from [3]).*

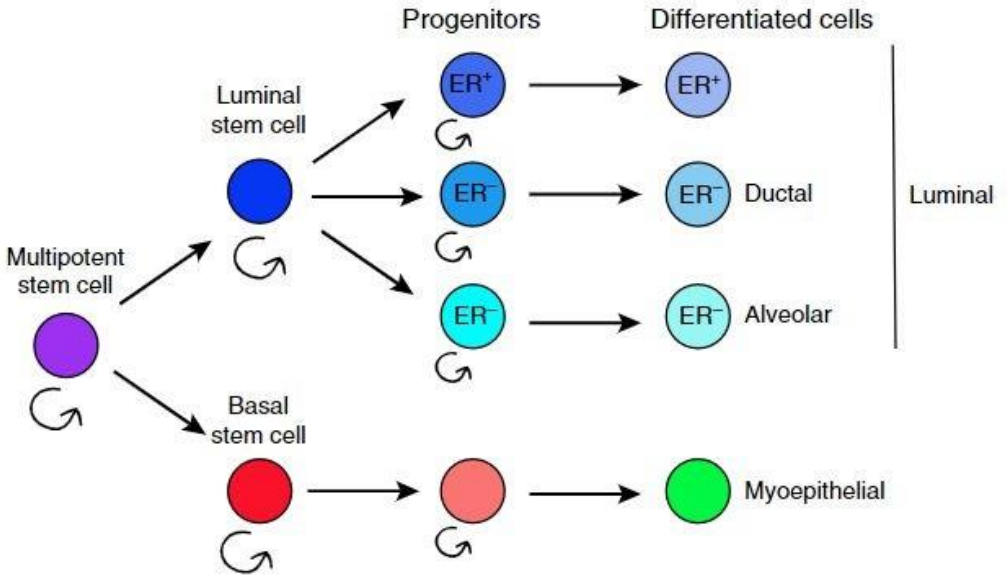


Figure 1.II: Mammary gland development: Simplistic model of *the mature mammary gland development, which consists of luminal and basal cells. These cells arise from distinct developmental programs from the same stem cell basis, forming a heterogeneous population in the mature organ (reproduced from* [1])

Table 1.I: Breast Cancer Subtypes

Brest cancer subtype	*Receptor Profile*	*Subtype Prevalence*	*Subcategories*
Hormone Positive	ER+ and/ or PR+	70%	Luminal A & B
HER2 Positive	HER2+	20%	-
Triple Negative	ER-, PR- AND HER2-	10%	Basal-like 1 (BL-1), basal-like 2(BL-2), Immunomodulatory (IM), mesenchymal (M), mesenchymal stem cell-like (MSL) and luminal androgen receptor (LAR)

1.2 Intratumor Heterogeneity

"All happy families are alike; each unhappy family is unhappy in its own way", wrote Leo Tolstoy in Anna Karenina. Just like the unhappy families of Tolstoy, tumors of the same type present distinct characteristics, both on the molecular and functional level, across patients (inter-tumor heterogeneity). Even within a specific tumor, cancerous cells tend to have genetic and non-genetic variations, a phenomenon termed intratumor heterogeneity (Figure 1. IV). Those variations in a heterogeneous tumor provide the background for the development and evolution of the disease, allowing for multiple subpopulations of cancer cells with complex phenotypes to emerge[4].

The origins of intratumor heterogeneity are a long-lasting point of research for scientists. The theory of clonal evolution of tumor cells suggests that the observed heterogeneity is the result of the accumulation of somatic mutations, so that the cells within a lesion have different mutational makeups. As a result, a subset of said mutations will offer a selective advantage to specific subpopulations, as they compete for resources[5]. On the other hand, the Cancer Stem Cell (CSC) model proposes that a hierarchy drives tumor organization, stemming from tumor stem cells and it is analogous to the one observed in tissues during development. Epigenetic mechanisms and changes in the tumor microenvironment (TME) can alter the functional characteristics of tumor cells reversibly, contributing to phenotypic plasticity at the level of individual cells [6,7]. Cancer Stem Cells (CSCs) and the different models of tumor heterogeneity will be discussed extensively in the next chapter of this thesis. It is important to note here that these models are not mutually exclusive, and it is possible for them to exist concurrently in a tumor. Tumors are typically heterogeneous both on the genetic and non-genetic level and these different types of intratumor heterogeneity might be related or independent (Figure1.IV).

Genetic heterogeneity refers to differences between cells on the DNA (genetic) level, specifically to stochastic mutations or chromosomic alterations that have occurred in different tumor cells independently. It has been recognized since the early days in biomedical research, around 1800s, when the pathologist Rudolf Virchow observed the morphological diversity of cells within tumors[8] . Since then, a lot of advancements have been made in the investigation of genetic heterogeneity through the development of immunohistochemistry and the discovery of histopathological markers, the karyotyping of genetic abnormalities and, most recently, the application of next generation sequencing[5,9,10]. Through pathological and cytogenetic analysis, intratumor heterogeneity was established as a hallmark of cancer and a major component of tumor evolutionary dynamics[9,10].

Non-genetic elements add another layer of heterogeneity in tumors. Unlike the genetic makeup, where mutations are non-reversible, the epigenomic and transcriptomic profiles of somatic cells are subject to changes, and heterogeneity therein can affect tumor cells morphology, hierarchy and evolutionary trajectory[6,11,12]. We can begin to fathom the importance and potential

of epigenetics in heterogeneity just by visualizing an adult human body. All the different organs, tissues and cells contain the same genetic code, propagated from the very first cell that was the steppingstone for the creation of the entire organism, the zygote. Alterations in the epigenome, usually a result of dysregulation of epigenetic enzymes (e.g., KDM5A and KDM5B), are predominant drivers of cell state dynamics[13,14]. Unlike genetic alterations, they affect multiple genomic loci simultaneously, rapidly changing regulatory programs and ultimately cell states. Intratumor epigenetic heterogeneity contributes, at least partly, to CSC phenotypes[15,16]. Dysregulation of epigenetic regulators allow tumor cells to hijack normal cellular processes and unlock embryogenesis-related programs during cancer progression[17]. Such events guide a cell state transition in tumor cells, conferring specific capabilities to a rare subset of cells, including asymmetric division and differentiation potential, mimicking properties of normal embryonic stem cells[18,19]. In some cases, resistance to anticancer drugs can also be conferred, even when tumor cells do not harbor pre-existing resistant mutations, presenting a major hurdle to developing efficient therapy strategies[20,21].

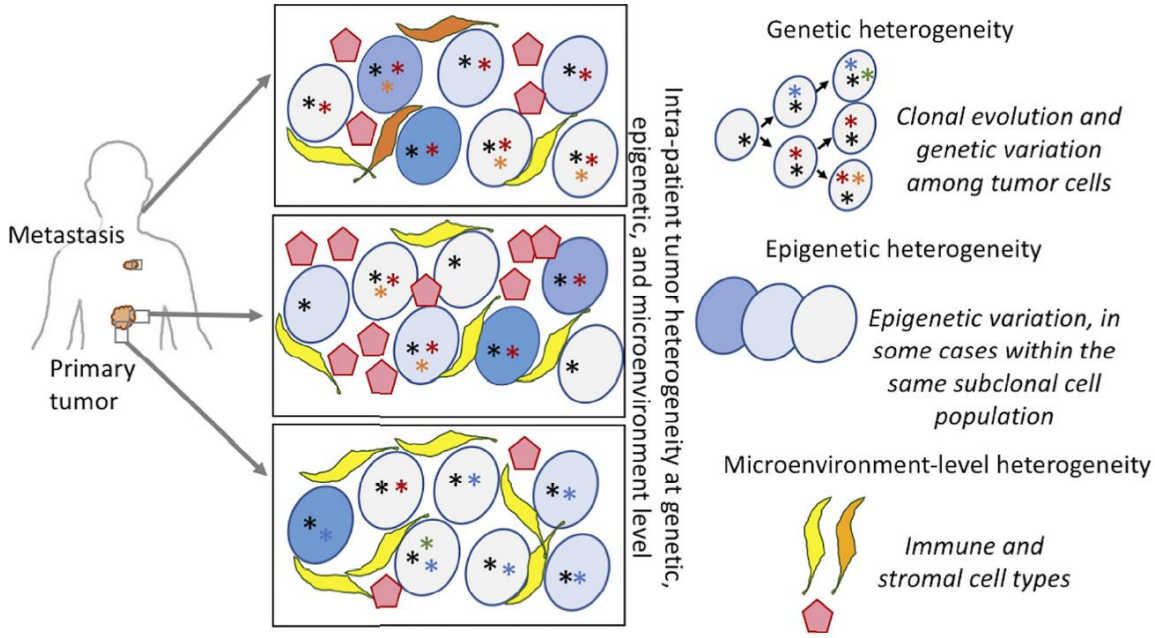


Figure 1.IV: Intratumor Heterogeneity: *Tumors consist of heterogeneous cell populations. The observed heterogeneity within a specific tumor can be attributed to genetic, epigenetic, or closely related micro-environmental factors (reproduced from[22]).*

1.3 Cancer Stem Cells

Cancer Stem Cells (CSCs) constitute a rare subtype of cells within a tumor with tumor-initiating ability, self-renewal and differentiation potential as well as intrinsic resistance to conventional therapies[18]. The Cancer Stem Cell (CSC) model has reshaped our view of cancer. Tumors are not just sprawling homogeneous masses of eternally proliferating cells. Analogous to the development and regeneration of healthy tissues, tumor growth is fueled by a small number of dedicated stem cells. Tumor development models mirror the scientific attempts to explain important observations, such as the intratumor heterogeneity, the almost inevitable recurrence after initially successful therapy as well as the phenomena of tumor dormancy and metastasis. CSCs have been identified in many common cancer types, including leukemia, prostate, brain and, most relevant for the present thesis, breast cancer[23] . As described in more detail in 1.3.3, $CD44^{high}CD24^{-/low}$ cells isolated from breast cancer tissues were observed having a substantially higher oncogenic potential in xenograft limiting dilution assays, compared to cells with alternate phenotypes, and therefore identified as breast cancer stem cells (BCSCs)[24].

1.3.1 The Classic CSC Model

The initial research on CSCs was majorly influenced by the concurrent breakthroughs in the field of hematopoietic stem cells[23] . The first proposed model was based on four premises. Firstly, the cellular heterogeneity observed in tumors is a direct result of its strict, unilateral hierarchical organization. Second, reminiscent to healthy tissue organization, cancer stem cells stand on the apex of the pyramid, giving rise to the entirety of the mature "structure" composed of non-CSCs and maintaining a pool of undifferentiated stem cells (self-renewal). There is limited plasticity, meaning that the CSC identity is hardwired, and non-CSC do not have the capacity to initiate tumors, a fact observed in xenograft assays[24]. Lastly, CSCs are resistant to common treatment modalities, thus they survive and are the source of tumor relapse after initial positive responses[18].

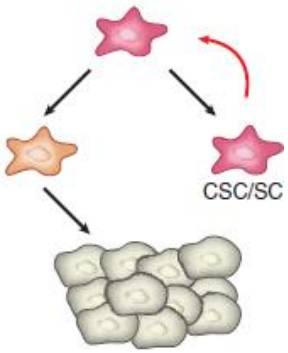


Figure 1.V: The standard CSC model: *This model suggests a unilateral relationship, where CSCs are the apex of the pyramid, maintaining the CSC pool by self-renewal mechanisms and differentiating to the rest of the bulk cells (reproduced from [18]).*

1.3.2 The Plastic CSC Model

The plastic CSC model is an evolution of the classic model. Novel studies revealed that CSCs and non-CSCs were capable of undergoing phenotypic transitions between them in response to stimuli. This notion was exemplified by an observation in breast cancer cell lines, where populations with distinct phenotypes were isolated by flow cytometry from a specific cell line and were allowed to expand. Cells displaying stem-, basal- and luminal-like phenotypes were shown capable of producing over time the observed heterogeneity identified in the initial cell line of origin[25]. Interestingly, in the same study, only stem-like cells generated xenografts in mice, functionally confirming the defining property of CSCs. The other populations were equally tumorigenic only after being triggered by external stimuli, hinting at a potential interconversion of phenotypes[25]. On such occasions, a strict model, like the classic CSC one, could not account for these observations. The plastic CSC model proposes such phenotypic changes of non-CSCs converting -either stochastically or after certain stimuli- to CSCs, highlighting at the same time the role of the tumor microenvironment (TME)[18].

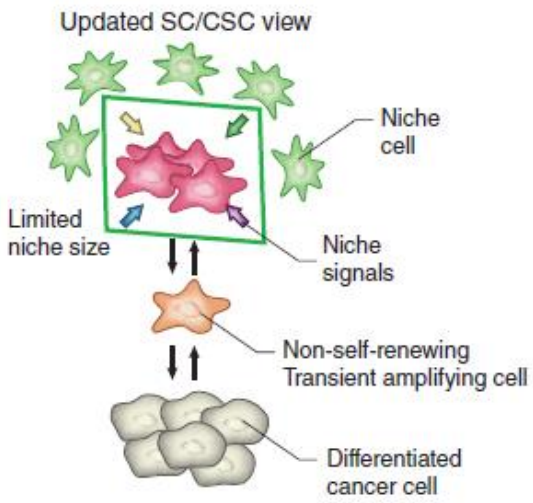


Figure 1.VI: The plastic CSC model: *This, updated version of the classic model, considers the bidirectional relationship between CSC – differentiated tumor cells, recognizing the ability of the latter to dedifferentiate. An important aspect of this model is the CSC niche and the signals generated from it (reproduced from[18]).*

1.3.3 Breast Cancer Stem Cells (bCSCs)

The development and establishment of specific biomarkers has made the identification and isolation of bCSC-enriched populations possible. In the first study that showed the existence of bCSCs in tumors, using the presence of cell surface markers ESA, CD44 and the absence of CD24, it was shown that as few as 200 cells of that phenotype ($ESA^+/CD44^+/CD24^-$), isolated from breast cancer tissues, were able to generate tumors in immunosuppressed mice, whereas 100-fold more cells of different phenotypes were non-tumorigenic[24]. Furthermore, the heterogeneity presented in the generated tumors was very similar to that observed in the original tumors[24]. Aldehyde Dehydrogenase (ALDH) expression may also be useful as an additional bCSC marker[26].

Functional assays are necessary to properly identify CSCs, as biomarker expression is mainly associated to cell – microenvironment interactions. The golden standard is the serial transplantation assay in limiting dilutions in vivo. Widely used in vitro assays include anchorage independent culture conditions, where bCSCs form floating colonies termed "mammospheres" and limiting dilution assays[27].

1.3.4 CSCs and Therapy Resistance

Tumor relapse is a common phenomenon in clinical practice. Current therapeutic strategies aim at eliminating highly proliferating bulk cells, which has limited benefits, since a small population of therapy-resistant cells are capable of repopulating the tumor (relapse)[28]. These residual cells are enriched in cancer stem cells, which, like normal stem cells, have intricate resistance to therapies via multiple independent mechanisms. In vitro studies in cell lines and primary cells have shown that breast cancer stem cells are relatively resistant to common therapeutics like radiation and cytotoxic chemotherapies[31]. $CD44^+CD24^{-/low}$ cells show a decrease in pro-oxidants, which could partly contribute to radiation resistance, or observed resistance could be mediated through the Wnt signaling pathway [32]. In an interesting neoadjuvant trial design, the frequency of the $CD44^+CD24^{-/low}$ cells was identified before and after cytotoxic therapy. After therapy, an increased percentage of $CD44^+CD24^{-/low}$ population was reported. Furthermore, the molecular profile of tumors obtained after chemotherapy closely resembled the gene expression profile of untreated $CD44^+CD24^{-/low}$ cells, revealing a positive selection and propagation of these clones after therapy[33].

In general, cell chemoresistance mechanisms include the upregulation of drug-efflux pumps, which actively transport drugs out of the cell (ABC transporters)[28] and alterations in DNA-damage response (DDR), exhibiting increased efficiency of the pathways governing it, by increasing the activation of the regulating factors (e.g., much higher phosphorylation levels of ATM, CHK1 and

CHK2 in the ATM-ATR pathway)[28]. Further, CSCs present an aberrant apoptotic pathway signaling, overexpressing pro-survival and anti-apoptotic molecules like the BCL-2 family, which plays an integral role in maintaining the balance between cell survival and apoptosis[28]. CSCs utilize important metabolic enzymes like Aldehyde Dehydrogenase (ALDH) which catalyzes aldehyde conversion to carboxylic acids[29], and confers resistance to certain chemotherapeutics[28]. Lastly, these cells exhibit a slower cell cycle, often adopting a quiescent state, passively driving drug resistance against common therapeutic strategies (Figure 1.VII) [30].

A number of such molecular mechanisms have been identified to contribute to the breast CSCs cytotoxic resistance. As previously mentioned, these cells express a variety of cellular transporters, including breast cancer resistance protein, a relative of the multidrug resistance protein[34]. Breast CSCs may also express ALDH and high levels of anti-apoptotic proteins like survivin and BCL-XL[35]. Lastly, these cells possess highly efficient DNA repair mechanisms and alterations in cell cycle kinetics, characteristics which may further render these cells resistant to cytotoxic agents[36].

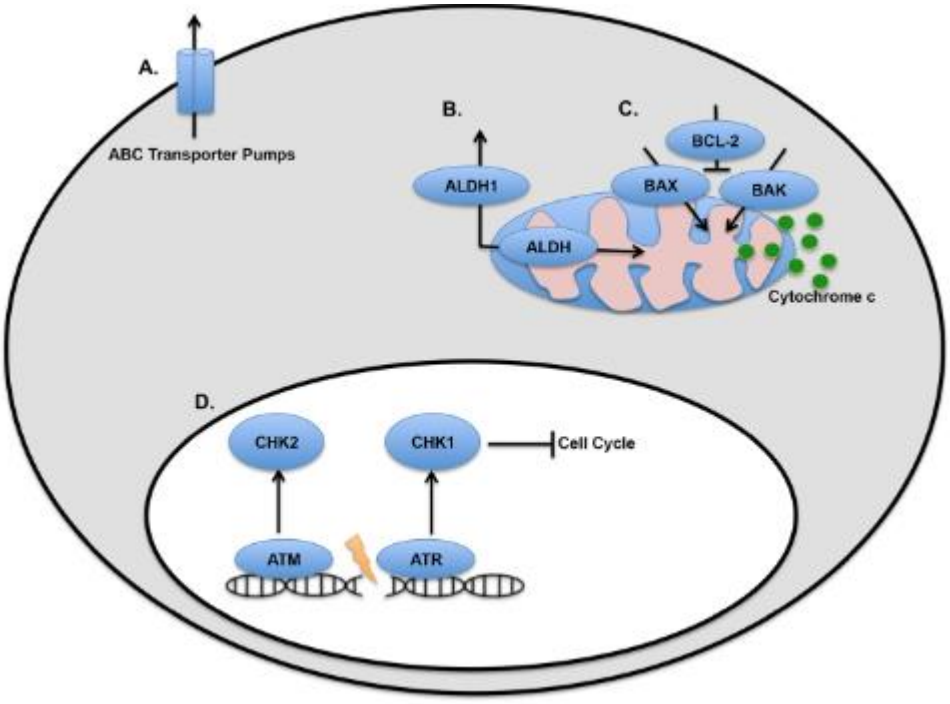


Figure 1.VII: Molecular Mechanisms of CSCs therapy resistance*: CSCs possess a plethora of mechanisms to resist therapy* ***A. Overexpression of*** *ABC transporters, pumping drugs out of the cell.* ***B.*** *Increased metabolic enzyme expression.* ***C.*** *Apoptotic signaling modifications, therefore resistance to apoptosis* ***D*** *Modified or more efficient DNA-repair mechanisms (reproduced from [28])*

1.3.5 Signaling Pathways in bCSCs

The aggressive nature of Cancer Stem Cells creates the need for development of agents to target this population. In order for this goal to bear fruit, we must understand the biology of CSCs and the mechanisms that regulate their various characteristics. One of the defining characteristics of these cells is their ability of self-renewal. Here, a number of developmental pathways that regulate this process in normal cells have been elucidated and appear to be exploited by CSCs. Interestingly, the malignant exploitation of these mechanisms occurs quite often as a downstream effect of epigenetic dysregulation, a phenomenon further explored in the next chapter of this thesis.

Stemness-associated pathways like Wnt, Notch and Hedgehog have been identified (Figure 1.VIII). Deregulation of these pathways in the mammary gland is able to produce breast cancers in mice[37–39]. The canonical Wnt/ β-catenin pathway regulates gene expression through the transcription factor β-catenin. In the absence of Wnt signaling, cytosolic β-catenin is degraded by a degradation complex (APC). Glycogen Synthase kinase 3 beta (GSK3β), a component of APC, phosphorylates β-catenin which leads to its ubiquitination and subsequent proteasomal degradation. When Wnt signaling molecule binds to Frizzled membrane receptors, the degradation complex is inactivated, allowing β-catenin's stabilization and nucleus translocation. In the nucleus, β-catenin induces gene expression for important genes, like CCND1 and MYC. Wnt pathway is of particular importance in normal tissue development and maintenance, regulating self-renewal and differentiation, characteristics that the CSCs hijack by exploiting this pathway[40,41].

Notch is a transmembrane receptor involved in cell contact-dependent signaling[42]. Binding of ligands Jagged or Delta, triggers Notch receptor intracellular domain (NICD) cleavage by γ-secretase. NICD then translocates into the nucleus where it induces the transcription of genes like MYC and HES1[43]. Notch has important roles in normal development of various tissues and organs[44], while also regulating cell proliferation and differentiation across a wide range of cell types in different stages of cell lineage progression[43]. Moreover, Notch pathway regulates stem cell differentiation and self-renewal[45]. Studies have elucidated the role of Notch signaling in breast, colon, and esophageal CSCs[46–48].

Hedgehog (Hh) pathway functions guiding cell fate during organism development and maintaining adult tissue homeostasis[49,50]. It also regulates stem and progenitor cell proliferation and maintenance in several tissues[51]. When the ligand sonic hedgehog (Shh) interacts with the Patched receptor (PTCH1), Smoothened transmembrane protein (SMO) activates, releasing Gli transcription factors, a mechanism reminiscent of the Wnt pathway. Hh pathway has been linked to tumorigenesis in various tissues[52] and conferring stem properties (e.g., in Basal Cell Carcinoma[53]).

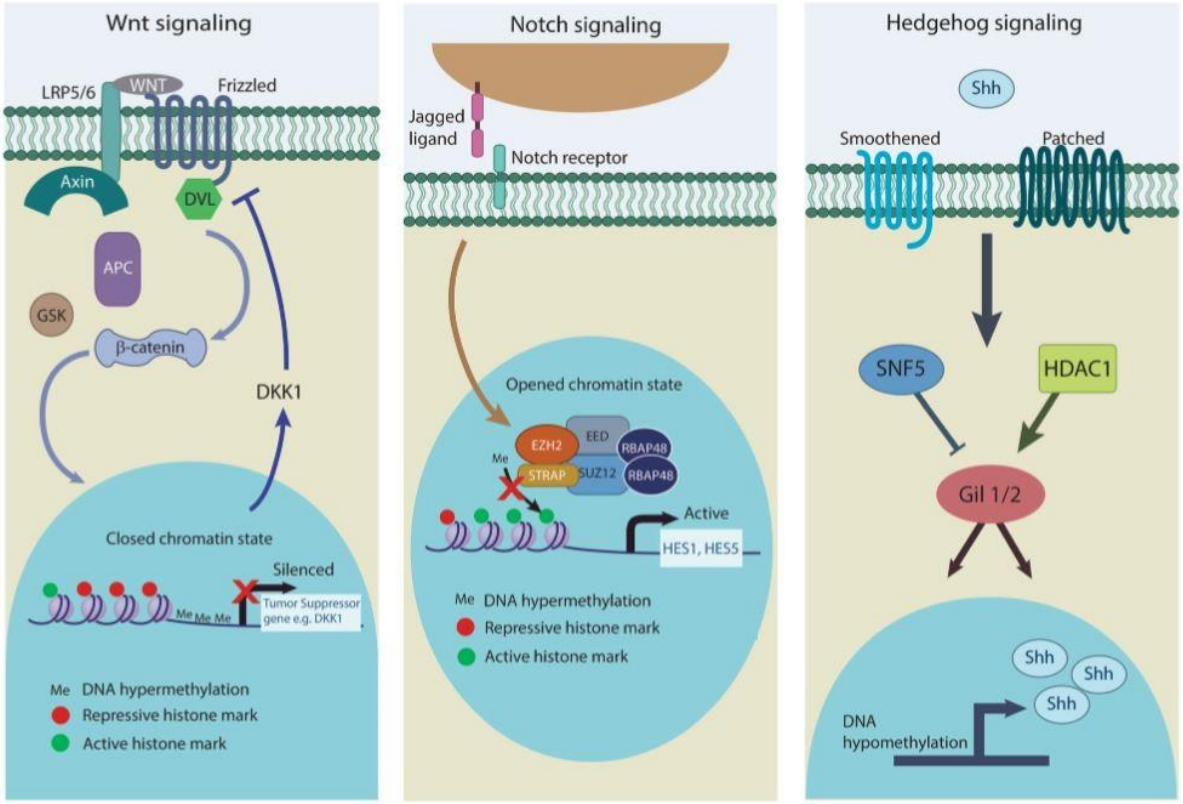


Figure 1.VIII: Regulation of key cancer stem cell signaling pathways by epigenetic mechanisms: A. *Wnt/β-catenin signaling can be enhanced by decreased expression of the DKK1 inhibitor through promoter hypermethylation and histone PTMs.* ***B.*** *Notch signaling target genes such as Hes1 and Hes5 can be activated by inhibition of inhibitory methylation marks at their promoter region by STRAP.* ***C.*** *Hedgehog signaling pathway can be activated in CSCs epigenetically by Shh promoter hypomethylation and increased HDAC1 expression. Epigenetic deregulation of CSC-related signaling pathways allows cancer cells to acquire self-renewal ability and drug resistance properties. (Reproduced from*[61]*)*

1.4 Epigenetics

Epigenetics include every heritable phenotypic modification occurring without interfering with the qualitative characteristics of the DNA. The Greek prefix epi- (έπι-, "over, outside or around") in epigenetics, implies features that are "on top of" or "in addition to" the traditional genetic base for inheritance. Such modifications may result in altering chromatin topology, architecture and, ultimately, gene expression[54]. DNA in the mammalian nucleus is packed densely around proteins, forming the structure called chromatin. The primary protein components of chromatin are histones. A histonic octamer consists of two sets of four histone cores, H2A, H2B, H3 and H4. A nucleosome is comprised of the DNA molecule wrapped around the histone octamers (Figure 1.VIII). Nucleosomes can then be compacted further when needed, in occasions like mitosis or during epigenetic control of gene regulation. Consequently, epigenetic factors may regulate the levels of density in different regions of chromatin, affecting the availability of DNA sequences to the transcriptional cell machinery. Tightly packed and inaccessible chromatin regions are termed heterochromatin, and genes found there are silenced. On the contrary, transcriptionally active DNA is found in loosely packed regions, which are called euchromatin. An important aspect of epigenetic control is its dynamic nature, allowing for a fast change in gene expression, by rapidly altering chromatin architecture, when environmental cues dictate it[55].

Epigenetic mechanisms, such as post-translational histone modifications (PTMs), DNA methylation, chromatin remodeling and changes in non-coding RNAs, together govern the epigenome landscape and dictate the outcome on cell phenotype, without changes in the DNA sequence (Figure 1.VIII). Such processes are important during normal mammalian development, ESC differentiation, and other important physiological processes[56]. Recent advances in epigenomic research have revealed that these mechanisms also play an important role in disease including cancer[57–59].

Post-translation modifications of the N-terminal histone tails are frequent and play significant roles in gene regulation and other DNA-templated processes[60]. Aberrations in this process can lead to deregulation of gene expression as seen in various human diseases and malignancies. The accumulation of such epigenetic abnormalities is considered to be an early event that predisposes tumor cells to acquire further mutations and genomic instability[57–59] . Furthermore, the epigenetic machinery is crucial for the maintenance of normal stem and progenitor cells[61] and the combination of mutations and epigenetic deregulation can lead to accumulation of cells with increased stemness properties and self-renewal ability, thus giving rise to CSCs[61].

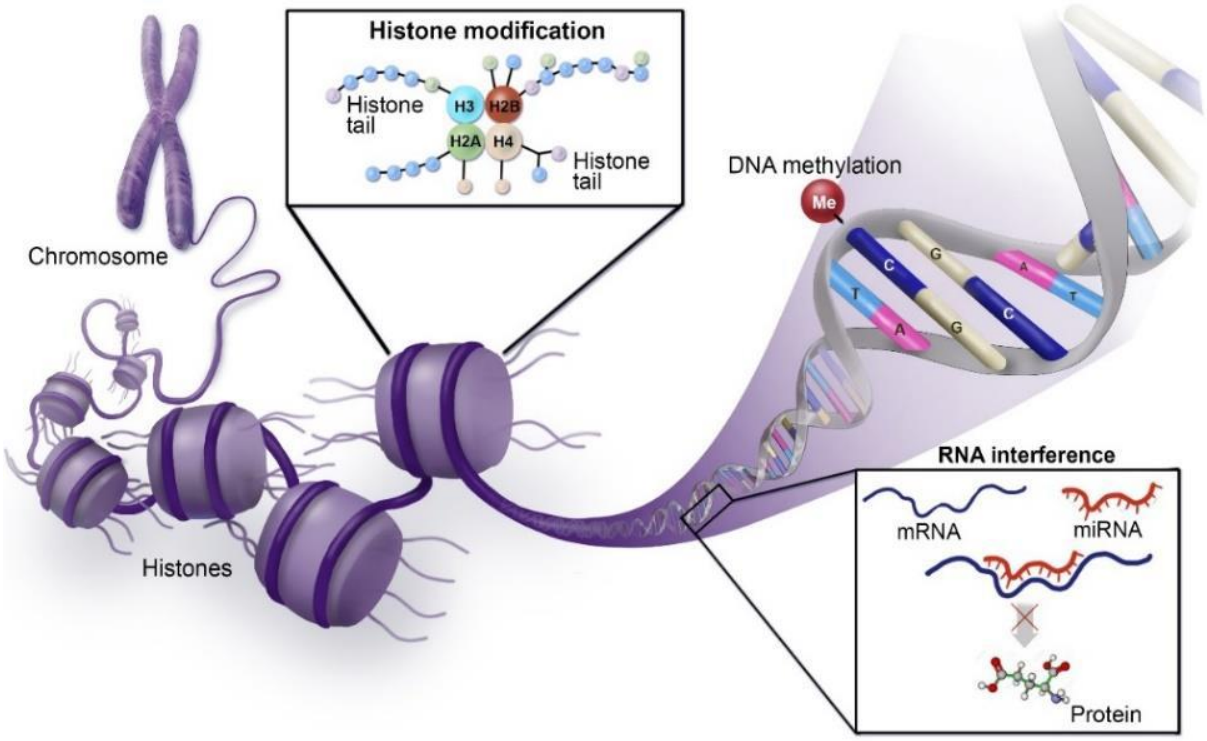


***Figure 1.VIII: DNA organization and Epigenetic Mechanisms**: Illustration of the nucleosome and its components and the main types of epigenetic mechanisms: DNA methylation, histone post-translational modifications and RNA interference (https://www.hematology.org/Research/Recommendations/Research-Agenda/3821.aspx)*

1.4.1 Epigenetic regulation of CSCs

Epigenetic regulation of the genome is one of the principal means by which the genetic code is altered to control cellular developmental hierarchies. Epigenetic mechanisms in their entirety govern the epigenome landscape that dictates the outcome of cell fate specifications. Such epigenetic changes are integral during normal organism development and ESCs differentiation.[56]

Epigenetic reprogramming is an indispensable requirement for the development of CSCs from their cell of origin, whichever that might be, an adult stem cell, a progenitor, or a differentiated somatic cell (Figure 1.IX). This reprogramming occurs in a stepwise fashion, where epigenetic factors are necessary to pave the way for stem-cell specific transcription factors to exude their function in their specific targets. This process is highly reminiscent to the somatic cell reprogramming (iPSCs), where an appropriate chromatin remodeling is necessary for the successful conversion of somatic cells into a stem state[62]. Epigenetic modulators act via two mechanisms during CSC development, they either facilitate chromatin availability to already

overexpressed transcription factors by exposing their target genomic cites, or they initiate the transcription-factor overexpression, playing the key functional role in driving the process[63].

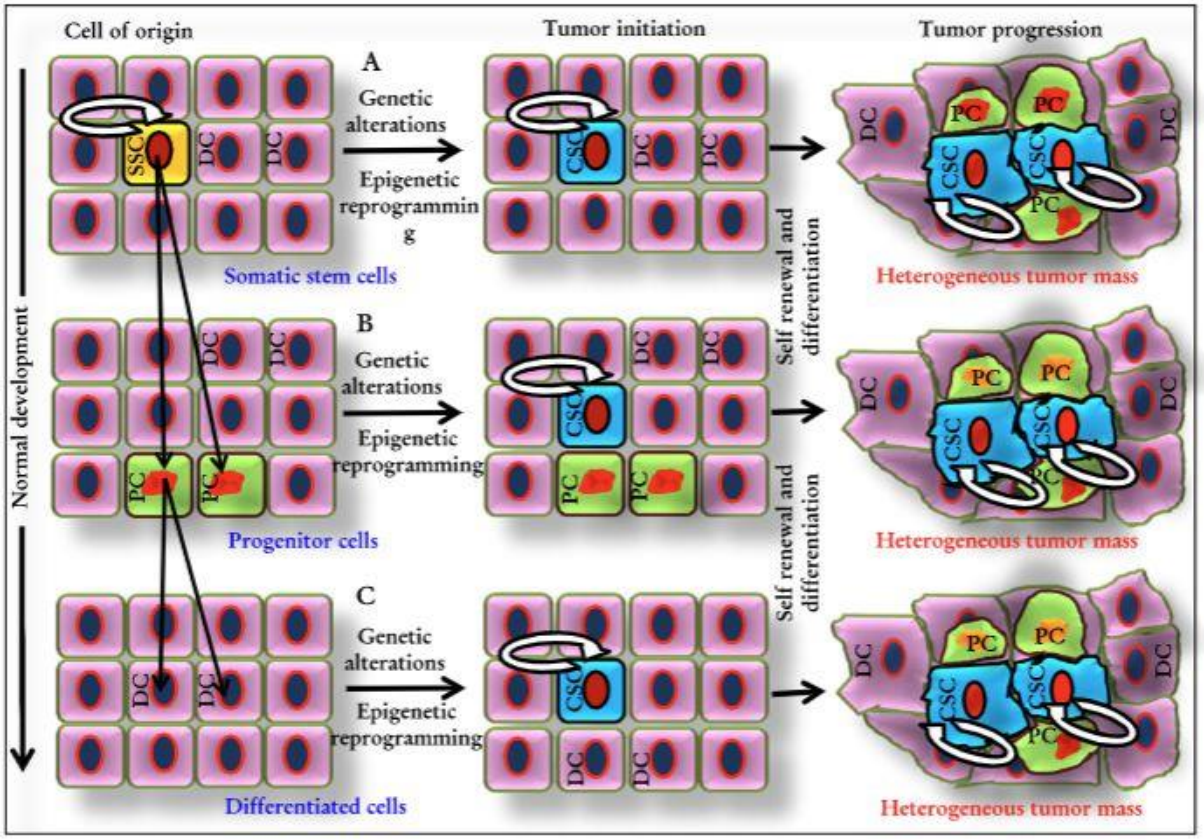


Figure 1.IX: Cancer Stem Cell origins: *Cancer Stem Cells may originate from* ***A. Normal Somatic Stem Cells (SSCs) B. Progenitor cells or C. Normal Differentiated cells*** *acquiring genetic alterations and epigenetic reprogramming that allows them to exploit developmental programs, giving rise to tumors (reproduced from*[63]*).*

Multiple studies have established the significance of a plethora of epigenetic aberrations in cancer. Chromatin remodeling is a process catalyzed by multi-enzyme complexes, known as chromatin remodeling complexes[64]. An aberrant function of the enzymatic parts of these complexes is often identified in CSCs. Brahma-related gene 1 (BRG1) is an integral subunit of SWI/SNF complex and was found to be indispensable for neuronal differentiation[65]. The SWI/SNF complex is the most frequently mutated chromatin remodeling complex in cancer, with almost 20% of tumors of all types harboring mutations in genes encoding for proteins of this complex[66]. BRG1 is a major driving force behind intestinal tumorigenesis, epigenetically regulating the Wnt pathway. Loss of this gene prevents intestinal adenoma formation and reduces the CSC population[65]. PRC-2, involved in trimethylation of H3K27, is required for embryonic development, ESC self-renewal and cancer. Pharmacological inhibition of PRC-2 leads to lesser tumorigenicity and reduced tumor progression in prostate CSCs[67]. Inhibitors of LSD1/KDM1 H3K4me2/3 demethylase were shown

to suppress the stem cell properties of bCSCs in vivo[68]. Overall, such observations prove the importance of chromatin alterations in setting the highly complex patterns of spatiotemporal gene expression during development and disease. Changes in the expression patterns of the proteins forming these complexes are capable of tumor initiation and CSC induction[63].

Another major epigenetic mechanism, and the most studied is DNA methylation. DNA methylation occurs at 5-position of cytosine residues of the CpG islands (CGIs) present in the mammalian DNA[69]. The methylation patterns of CGIs of normal somatic cells differ largely from those of ESCs, iPSCs and CSCs. While most of the CGIs remain unmethylated in normal somatic cells, the iPSCs and CSCs display hyper-methylation of CGIs in promoter regions of differentiation-specific and cancer-related genes[70]. *De novo* methylation has been shown to regulate stem cell characteristics[71]. A study displayed how tumors can be reprogrammed to lose their malignant behavior, by iPSC reprogramming in glioblastoma cells, where methylation induced silencing in cancer-promoting pathways[62]. An enrichment of DNMT3b, a principal enzyme catalyzing de novo DNA methylation, at the CpG-islands of hypermethylated genes suggests that this enzyme functions more importantly in furnishing the aberrant methylation pattern observed in CSCs, which enables them to retain their undifferentiated state[72]. Simultaneously, active DNA demethylation is considered a requirement for cells to regain stem-cell state during induction of pluripotency. This active demethylation is probably achieved through the conversion of 5-methyl cytosine (5-mC) to thymine by activation-induced deaminase (AID) and by ten-eleven translocation (TET) proteins which convert 5-mC to 5-hydroxymethylcytosine (5-hmC)[73,74]. Abnormal DNA methylation events occur early during carcinogenesis resulting in premalignant states and field cancerization[75,76].

Epithelial to mesenchymal transition (EMT) is a multi-step process in which cells adopt a mesenchymal phenotype and features like loss of cell polarity, increased cell motility and invasive properties[77,78]. Activation of this process can confer cells with CSC and tumor-initiating properties[79,80]. Several signaling pathways involved in normal embryonic development have been identified to regulate EMT, namely the Wnt, Hedgehog and Notch pathways (Analyzed in 1.3.5). Deregulation of these pathways, which often occurs as a result of epigenetic aberrations, can activate aberrant EMT induction, possibly resulting in tumor metastasis and contribute to poorer patient prognosis.

Several studies have shown that cells with both CSC and EMT-like phenotype are more resistant to chemotherapeutics, compared to other cancer cell populations[63,81]. Furthermore, in similar studies, it was shown that epigenetically driven CSC and EMT properties impact tumor cells' response to therapy[82]. Epigenetic deregulation has also been linked with other mechanisms of increased resistance found in CSCs, like the overexpression of drug-efflux pumps such as the ATP-binding cassette (ABC) family of transporters[83].

Collectively, these abnormal alterations in the epigenome, which lead to alterations in key proteins' expression and in key signaling pathways that contribute to CSC proliferation and maintenance, may serve as potential mechanisms for targeted therapy against CSCs.

1.5. Histone methylation and demethylation

For the purpose of this thesis, we will only focus on one of the primary mechanisms of epigenetic regulation, histone modifications and specifically histone methylation. Histone methylation occurs predominantly on Lysine (K) and Arginine (R) residues, and these methylation marks serve as docking sites for histone readers, enzymes that recognize and interact with this motif[84]. Histone lysine methylation can occur at three different levels: mono-, di- and tri-methylation. These modifications can impact gene expression positively or negatively based on the target of the modification. For instance, histone 3 lysine 4 (H3K4) methylation is associated with gene activation, while methylation on lysine 9 (H3K9) is linked to gene silencing[85] .

The methylation reaction is catalyzed by enzymes named lysine methyltransferases (HKMTs)[86]. Despite catalyzing a common chemical reaction, this family of HKMTs demonstrate large structural diversity of its active sites, allowing these enzymes to have high substrate specificity[87]. For example, DOT1L (KMT4) is a unique HKMT as it is currently the only known enzyme that methylates lysine 79 of histone H3 (H3K79)[88]. Similarly, methylation of H3K27 is only mediated by the catalytic subunit EZH2 (KMT6) of PRC2[87]. In contrast, some methylation marks can be catalyzed by several proteins, such as H3K9 methylation. These post-translational methylations of histones have important roles in regulation of gene expression, differentiation, DNA damage repair as well as in tumorigenesis[89,90]. Aberrant histone methylation can be due to gene mutations, overexpression or deregulated control of epigenetic modulatory enzymes involved.

H3K79 methylation by DOT1L is associated with transcriptional activation of genes under its regulation[91,92], and overexpression or aberrant DOT1L activity has been found in cancer, such as leukemia with mixed lineage leukemia (MLL) gene translocation. The MLL fusion protein can recruit DOT1L into a transcription complex, which subsequently methylates H3K79[93–96]. This leads to dysregulation and overexpression of many MLL-target genes, including Homeobox A9 (HoxA9) and Meis homeobox 1 (Meis1), which are key regulators of hematopoietic stem cell differentiation that contributes to leukemogenesis[86].

EZH2 is a member of PRC2, along with proteins embryonic ectoderm development protein (EED) and SUZ12, and is responsible for catalyzing H3K27 mono-, di- and tri-methylation[97–99]. Overexpression of EZH2 has been found in various cancers of the breast, lung, prostate, and hematological malignancies[100–104], and is associated with poor disease prognosis. Studies have also shown the role of EZH2 deregulation in tumor progression, metastasis[105,106] and maintenance of CSC self-renewal properties[107]. In glioblastoma multiforme (GBM), inhibition of EZH2 by S-adenosylhomocysteine hydrolase (SAH) inhibitor 3-deazaneplanocin A (DZNep) was able to reduce self-renewal and tumor-initiating capabilities of GBM CSCs in vivo via affecting transcriptional regulation of oncogene MYC[106].

H3K9 methyltransferases, such as euchromatic histone lysine methyltransferase 2 (G9a/EHMT2) and euchromatic histone lysine methyltransferase 1 (GLP/EHMT1), catalyze mono- and di-methylation of the lysine residue, while tri-methylation of H3K9 is mediated by Suppressor of variegation 3–9 homolog 1 (SUV39H1) and Suppressor of variegation 3–9 homolog 2 (SUV39H2) [108]. Upregulation of G9a activity has been linked to several types of cancer, including ovarian, lung, liver and bladder cancers[109–114].

The entirety of the epigenetic changes on histones are termed the "histonic code". All PTMs on histone tails are tightly regulated. Methylation levels of lysines on histones are also regulated by histone lysine demethylases (KDMs), acting as the counterweight on the methylation scale. This group of epigenetic erasers function in removing the methyl groups from lysine side chains on histones[115,116]. KDMs can be grouped into two families – the lysine-specific demethylase (LSD) family and Jumonji domain-containing (JmjC) family[117]. The LSD family are flavin adenine dinucleotide (FAD)-dependent amine oxidase that demethylates mono- and di-methyl lysine residues, while JmjC enzymes utilize 2-oxoglutarate and iron to oxidatively release methyl groups from all three methylation states at lysine residues[86,118].

Upregulated expression of LSD1 (KDM1A) has been found in various human cancers, including AML, ovarian, lung, bladder, and colorectal cancers[119,120]. LSD1 is important for normal hematopoiesis; loss of LSD1 has been found to inhibit differentiation and impair hematopoiesis[121]. This suggests a potential role of aberrant LSD1 activity in affecting stemness properties in tumor cells. Published work from our lab also indicates LSD1 as an integral modulator of bCSCs stemness properties and chemoresistance to doxorubicin[122]. Ubiquitously transcribed tetratricopeptide repeat X chromosome (UTX), also known as KDM6A, is responsible for demethylating H3K27[123–125], and loss of UTX activity has been found in multiple human malignancies, including multiple myeloma, esophageal squamous cell carcinoma and renal carcinoma[88]. However, no inhibitors of JmjC enzymes have advanced beyond biochemical studies[87]. KDM5A and KDM5B, the enzymes studied in the present work, also fall in this category, and are extensively discussed in the next pages of this thesis.

1.5.1 KDM5A and KDM5B

KDM5A (JARID1A/ RBP2) and KDM5B (JARID1B/ PLU1) are epigenetic enzymes. They belong to the KDM5 subfamily of histone demethylases, which consists of 4 highly related enzymes, KDM5A-D. KDM5s are part of the class of Jumonji-C containing lysine demethylases (JmjC-KDMs). These enzymes have been shown to remove di- and tri-methylation marks from H3K4, however *in vitro* data suggest that demethylation of H3K4me1 is also possible[126,127]. Catalytic activity is mediated by a composite JmjN/JmjC domain. A helical domain surrounding

aC5HC2 zinc finger motif is required for demethylation[128]. These regions make up a compact catalytic core[129] utilizing 2-oxoglutarate as a cofactor, as well as FE(II) and oxygen. KDM5s contain five conserved domains: ARID (DNA-binding domain), C5HC2 zinc finger, JmjC, JmjN and PHD (histone-binding domain). This subfamily is unique among the Jumonji containing demethylases due to the atypical ARID and PHD insertion into the Jumonnji domain, leading to the separation of said domain into two fragments (JmjC and JmjN) (Figure 1.IX)[130].

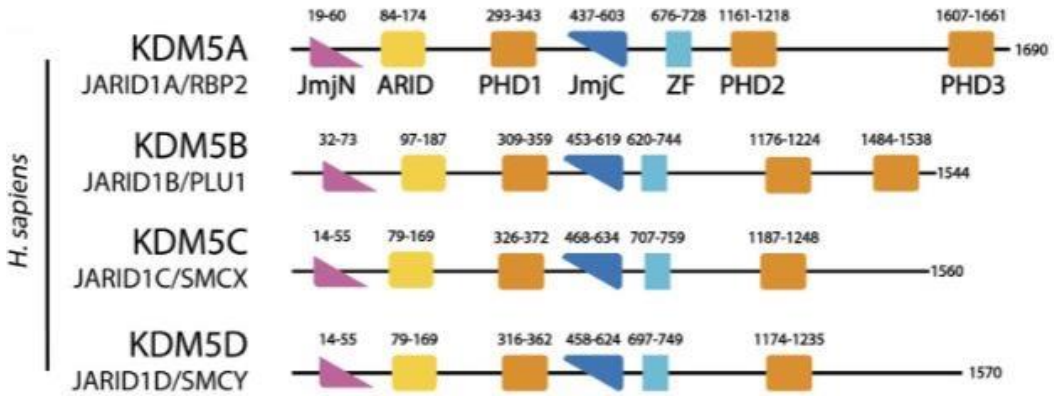


Figure 1.IX: KDM5A-D structure: *Amino-acid and region depiction of the enzymes. JmjN & JmJC: catalytic regions, ARID: DNA-binding domain, PHD1, PHD2 and PHD3: histone binding domains, ZF (Zinc Finger): DNA-interacting domain (reproduced from[130])*

Protein structure and function are interconnected. Understanding each domain and functional capacity of the KDM5 enzymes will shed light on the multifaceted roles they have in cells. Firstly, KDM5s serve as erasers of the histone code by removing methylation marks. Obviously, they have been studied mostly as transcriptional repressors, as H3K4me2/3 marks have been observed on transcriptionally active chromatin[131,132]. However, this may not be completely true, as changes in H3K4me3 levels not always correlate with transcriptional activation and it is not yet clear whether histone trimethylation is an indicator of active transcription or a result of it. Moreover, removal of H3K4me2/3 may have a positive effect on transcription, by suppressing cryptic transcription in the gene body or by activating specific gene elements; for example, the result of the KDM5 demethylase activity could be H3K4me1, which is a characteristic of enhancer regions[133].

The DNA binding ARID and the first PHD domain are partially dispensable for the catalytic activity of the enzyme[129] but likely play important roles in its allosteric regulation[134,135]. PHD1 plays an important role in substrate binding and activity regulation. This domain has a binding preference towards unmodified H3 peptides[136] and may also interact with methylated H3K9[134]. Interestingly, binding of H3 peptides that are unmethylated at K4 confer allosteric activation of KDM5A and B

demethylase activities[134,135]. PHD3 domain probably contributes to chromatin binding and substrate recognition of KDM5 enzymes. PHD2 is not yet biochemically or structurally characterized[130] (Figure 1.X).

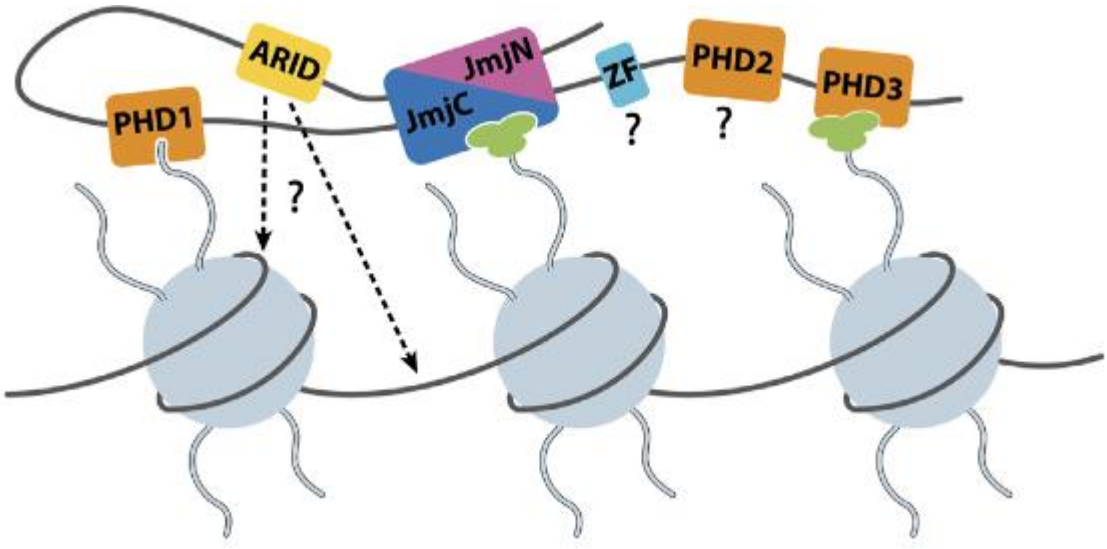


Figure 1.X: Graphic Illustration of KDM5A/B interaction with Histones and DNA*: PHD1 and PHD3 domains interact with histones and histone modifications, while the catalytic JmjC/ JmjN core forms. ARID probably interacts with DNA, while ZF and PHD2 domains' specific function is still being researched (reproduced from[130]).*

Besides their function as erasers, KDM5s may also act as readers of the histone code. As readers, KDM5s recognize histone PTMs and recruit proteins and multi-protein complexes to their genomic targets, while also mediating local allosteric activation or inhibition of the recruited enzymes[135].This adds another layer of complexity in the study of these enzymes that one should take into consideration when analyzing their effects in cell biology, as certain functions may be, at least partly, mediated independently of their catalytic activity[137].

1.5.2 KDM5A & KDM5B Functions: A Snippet

A growing body of literature describes the multifaceted roles of KDM5 demethylases in gene regulation, differentiation, developmental processes, and diseases, such as cancer.

KDM5A was initially identified as a Retinoblastoma (Rb) binding protein, hence named Rb Binding protein 2 (RBP2)[138]. KDM5A as a transcriptional regulator can act both as a repressor and as an activator. For example, it has been shown that it represses HOX genes[131], the Notch signaling pathway[139] and metabolic regulators, such as PGC1A[140]. It is also possible that some of these repressive effects may be demethylase-independent, as KDM5A has been shown to display transcriptional repression capabilities, even when a catalytically inactive enzyme is expressed[137]. On the other hand, KDM5A may mediate transcriptional activation, as was shown for cell cycle regulators[141] or cellular differentiation genes[142].

KDM5B (JARID1B) presents strong overall similarities with KDM5A, including an identical structure of protein domains (Figure 1.IX). Functionally, KDM5B has been shown to exert transcriptional repression of developmental transcription factors paired box 9 (PAX9) and brain-factor 1 (BF-1)[143], thus promoting differentiation and cell proliferation in ESCs [132]. It seems that KDM5B contributes to normal organism development by co-regulating important nuclear receptor signaling, like Estrogen (ERs), Progesterone (PRs) and Androgen (Ars)[134,144–146]; in addition, KDM5B null mice present developmental defects[147]. KDM5B positively regulates gene expression, by repressing intra-gene spurious transcription[148] and maintaining epigenome integrity of promoter and enhancer regions, contributing to the specific loci activity[149].

Another layer of complexity and potential of these enzymes is added when interacting proteins are cogitated. Epigenetic enzymes form a huge system of interacting complexes. Known interacting complexes include epigenetic enzymes like histone deacetylases (Sin3b, NuRD), H3K4me1 demethylase LSD1 as well as the Polycomb complex (Prc2) (Figure 1.XI A)[150,151]. They can also interact with transcription factors like C-Myc [152] and obviously, with the retinoblastoma protein (pRb)[138] (Figure 1.XI).

At this point, it is obvious that these enzymes are located at different sites within the genome and engage in various processes, and their functions are mediated through diverse molecular interactions with multiple phenotypic consequences microscopically within a cell or macroscopically in organisms. The complex but important role of these enzymes mandate the deciphering of the underlying mechanisms.

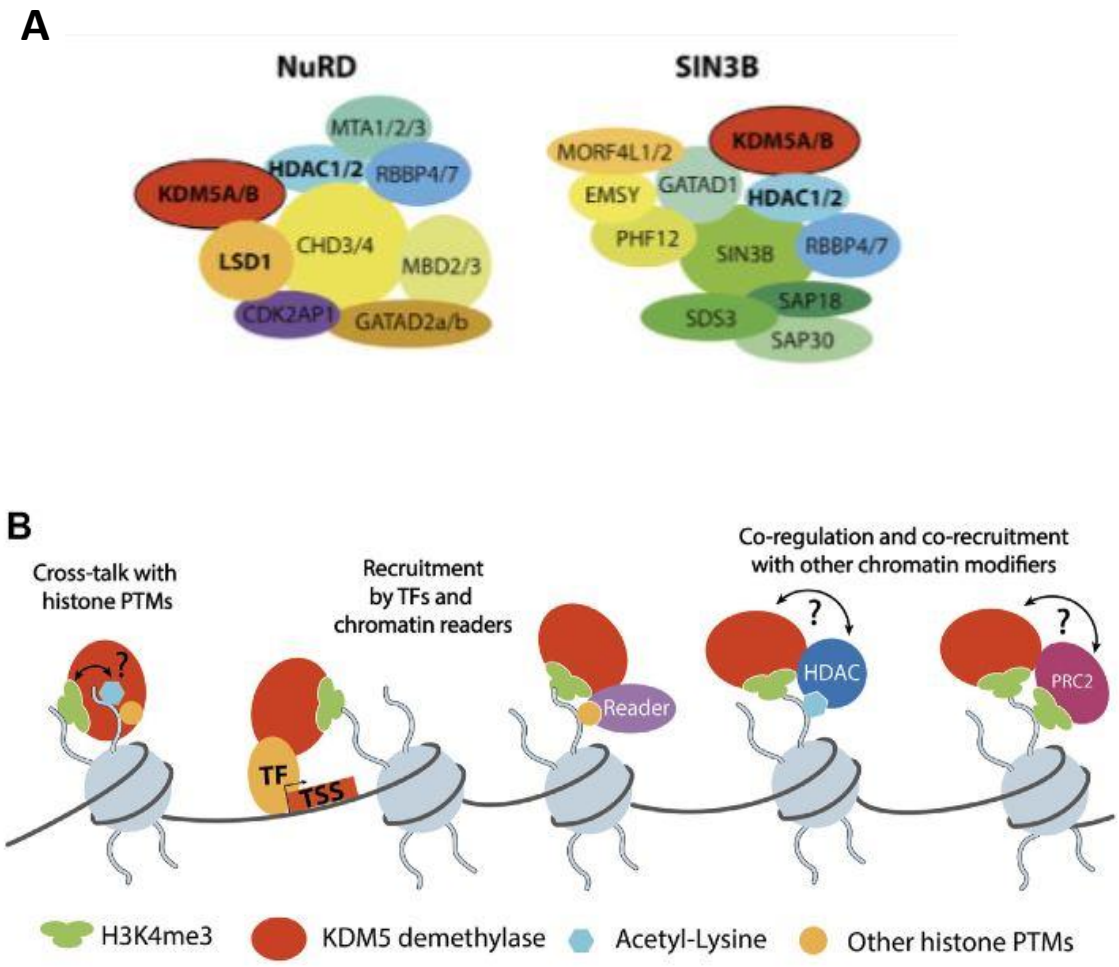


***Figure1.XI*: A.** *KDM5 demethylases were shown to physically and functionally interact with HDAC complexes. Interactions with the NuRD and SIN3B complexes have been shown for the mammalian KDM5A and B proteins. Note that, for reasons of clarity, the stoichiometry and detailed subunit composition of the complexes was neglected. For NuRD and SIN3B, composition and dynamics of subunits are subject to research and have not been definitely established. The placement of subunits and their proximity to each other and to the KDM5 proteins does not reflect experimentally verified proximity within the respective complexes.*

B. *Functional KDM5 interactions on chromatin. So far, only the binding of unmethylated H3K4 has been shown to regulate the demethylase activity of KDM5s. Given the potential interactions with HDACs, a direct or indirect responsiveness to other histone PTMs such as acetylated lysines, is conceivable. KDM5 proteins are recruited by transcription factors (TFs), reader domain proteins, or mediated by the association with other epigenetic regulators such as HDAC complexes or PRC2. The interaction and functional interplay of KDM5s with HDAC complexes and PRC2 suggests a potential mutual regulation of demethylase and other chromatin modifying activities. Such a direct interplay remains to be demonstrated experimentally.*

(Reproduced from[130]*).*

1.5.3 KDM5A & KDM5B in Cancer

Numerous observations strongly link KDM5 demethylases to cancer biology[20,130,153–155]. Mainly, KDM5 inhibition seems to affect tumor growth[156,157], while, in some instances, specific roles have been identified, where KDM5 demethylases control tumor phenotypic profile and therapeutic response[7].

It is currently known that the main driver behind cancer progression is genetic mutations. While several tumor phenotypes can be sufficiently explained through mutational mechanisms, there are certain characteristics that cannot, such as the rapid response and adaptability of cancer cells, particularly, to therapy resistance. Those specific characteristics are achieved through dysregulation of epigenetic mechanisms, such as the ones mediated by KDM5A and KDM5B. KDM5 demethylases are amplified or overexpressed in many types of cancer, including gastric[158], lung cancer[139] and breast[128,159].The result of the dysregulation in KDM5 activity is a cellular transcriptomic heterogeneity that results in variable phenotypic features in cancer cells, one of the key mechanisms underlying disease progression and therapeutic resistance[20]. This heterogeneity leads to the appearance of cellular populations with various oncogenic potential and/or resistance to therapies; subsets of tumor cells with such properties that are also bestowed with stemness traits constitute the cancer stem cells.

KDM5A has been studied mainly as a key determinant of therapeutic responsiveness. It is identified as a critical factor in drug-tolerant persister cells in non-small cell lung cancer (NSCLC)[21,160]. In breast cancer specifically, KDM5A is implicated in drug resistance of ER^+ breast cancer in tamoxifen treatment, regulating important cell signaling pathways[161]. Furthermore, KDM5A impacts cell proliferation, as it directly promotes G1-S cell cycle progression via increasing CyclinD1 levels, possibly through miRNA repression[162,163], and repression of cell cycle inhibitors gene expression, like p27[163]. It has also been shown to promote EMT, a prerequisite for metastasis and also a process that has been proven to generate cells with stem-like features[163,164]. Specifically, in TNBC, KDM5A promotes metastasis by expression of integrin β-1 (ITGB1), interestingly, in a demethylase-independent activity[165]. Finally, KDM5A seems to be implicated in DNA-damage (DDR) and replication stress response (RSR). In DDR, KDM5A proteins were shown to accumulate at γ-H2AX foci[166], demethylating H3K4me3, a process required for DSBs reparation[167]. KDM5A contributes to replication stress response by positively regulating Chk1 protein levels and developed tolerance to drugs like Hydroxyurea (HU) that induce replication stress requires KDM5A overexpression[168]. Taken together, these observations support an important role for KDM5A in aggressive breast cancers.

KDM5B has been widely studied in breast cancer, mainly as an oncogene[157,168,169]. It has been shown that KDM5B is important for cell proliferation, colony, and tumor formation, especially in ER^+ cell lines, where an estrogen-coupled activity is suspected, and KDM5B has been ultimately

described as a “luminal lineage driving oncogene”[145,159]. KDM5B has also been linked with therapy resistance and stem-like features in KDM5B overexpressing melanoma cells[170] and appears as a regulator of cancer stem cell properties in oral cancers[171].

In contrast, genome-wide gene expression analysis of MDA-MB-231 (triple-negative breast cancer) cells, showed that KDM5B knock-down increased expression of genes involved in cell proliferation, migration, and inflammatory response. Indeed, overexpression of KDM5B was shown to suppress migration and invasion of this cell line[134]. This corresponds with findings showing suppression of metastasis and angiogenesis by KDM5B in breast cancer cells[172]. Moreover, KDM5B facilitates genome stability, promoting DSB signaling and DNA repair, indicating it as an important genome caretaker[173].

It is evident that KDM5 enzymes have pivotal functions in different cellular processes in cancer. It is important, therefore, in order to understand the underlying mechanisms of their action, to elucidate the biological background and environment they act on.

1.6 CRISPR-Cas technology

The development of recombinant DNA technology marked the beginnings of a new era for biology. For the first time in biological research, scientists had the ability to manipulate DNA molecules, making the study of genes possible, as well as harnessing that knowledge for the development of novel medicine and biotechnology. Recent advances in genome engineering technologies are sparking another, similar revolution, in biological research. The CRISPR technology allows the direct and precise editing and modulation of DNA sequences in their endogenous context, enabling the elucidation of the functional organization of the genome at the systems level and the identification of causal genetic variations. Emmanuelle Charpentier and Jennifer A. Doudna were awarded the Nobel Prize in Chemistry 2020, for the development of the CRISPR-Cas system as a genome editing tool[174].

The work that eventually led to the discovery of the CRISPR-Cas9 system for genome editing began with the identification of repeated genome structures present in bacteria in 1987. The report noted an unusual, repeated structure in E. coli genome which included five highly homologous sequences of 29 base pairs[175]. Bioinformatics analyses revealed that these types of repeats were common in prokaryotes, and all consisted of the same features: a short, partially palindromic element organized in clusters and separated by unique intervening sequences of constant length, suggesting an ancestral origin and high biological relevance[175]. The term CRISPR was introduced, an abbreviation for clustered regularly interspaced short palindromic repeats[176].

The functional importance of the CRISPR loci remained elusive until 2005, when researchers noted that the unique CRISPR sequences were derived from transmissible genetic elements, such as bacteriophages and plasmids[177,178]. The hypothesis that CRISPR-Cas systems act as and adaptable defense system conferring resistance to invading foreign DNA was verified in 2007[179].

1.6.1 Molecular Understanding of the CRISPR mechanism

The structure of Cas9 consists of two distinct lobes in free form, the recognition (REC) lobe and the nuclease (NUC) lobe, with the HNH and RuvC nuclease domains contained in the latter. When Cas9 binds to sgRNA, it undergoes a structural rearrangement, with the REC lobe moving towards the HNH domain (Figure 1.XII).

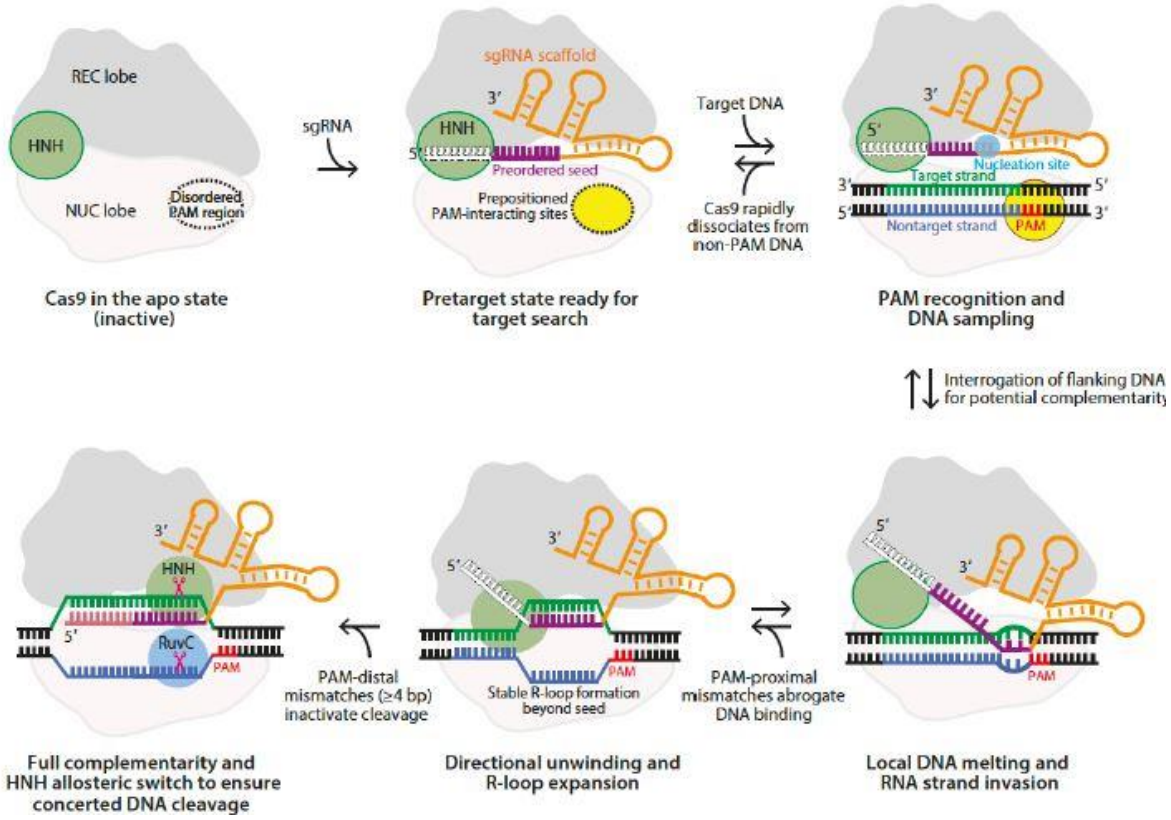


Figure 1.XII: A schematic representation of the mechanism by which CRISPR-Cas9 recognizes and targets DNA for cleavage*. Binding of the sgRNA molecule leads to a large conformational change in Cas9. In the activated conformation, the PAM-interacting site (dotted circle), scans the genome for PAM sequences, and the seed sequence of sgRNA is positioned to interact with adjacent DNA for complementarity to sgRNA. When a PAM sequence is recognized, Cas protein contributes to local DNA melting and RNA strand invasion. There is a stepwise elongation of the R-loop formation and a conformational change in the HNH domain, positioning the DNA strands in the Cas nuclease active sites, to ensure concerted DNA cleavage.*

For target recognition, the 20-nt spacer (sgRNA) sequence must form complementary base pairs with the protospacer (target) sequence. Cas9 forms a complex with sgRNA and is positioned to engage with the target sequence in DNA, via a 10-nt seed sequence in the 3' end[180,181]. PAM (Protospacer Adjacent Motifs) are short sequence motifs adjacent to the target sequence, necessary for target recognition and cleavage[182–184]. The Cas9 protein first searches for the PAM sequence, and once found, interrogates the proximal DNA for complementarity to the sgRNA. Once a stable RNA-DNA duplex is formed, Cas9 is activated for DNA cleavage. Each of the two nuclease domains cleaves one strand of the target DNA at a specific site 3 bp from the 5'-NGG-3' PAM sequence, forming blunt ends.

1.6.2 Application of the CRISPR-Cas9 technology

In their work, Charpentier and Doudna generated a simple two-component system that could be programmed for highly specific cleavage of target DNA and thereby started a revolution in genome editing. CRISPR-Cas9 has found many unique applications for genome engineering[185]. In the present thesis, CRISPR-Cas9 was used to generate KDM5A and KDM5B knock-down cell lines.

To achieve this goal, CRISPR-Cas9 was used to introduce double-stranded breaks (DSBs) in the genomic loci of these proteins. Double-stranded breaks lead to either non-homologous end joining (NHEJ) repair or homology-directed repair (HDR). The HDR pathway uses a homologous DNA sequence as a template to repair the break (Figure 1.XIII). By introducing modified genetic sequences as templates for the HDR, it is possible to introduce specific changes in the genomic loci. In our case, a template was not provided and CRISPR was introduced in non-proliferating cells, to push NHEJ repair. In this type of DNA repair, the ends are directly ligated back together, and the process usually results in a small insertion or deletion of DNA at the break, frequently causing frame shifts in coding sequences and loss of protein expression via introduction of premature stop codons (Figure 1.XIII).

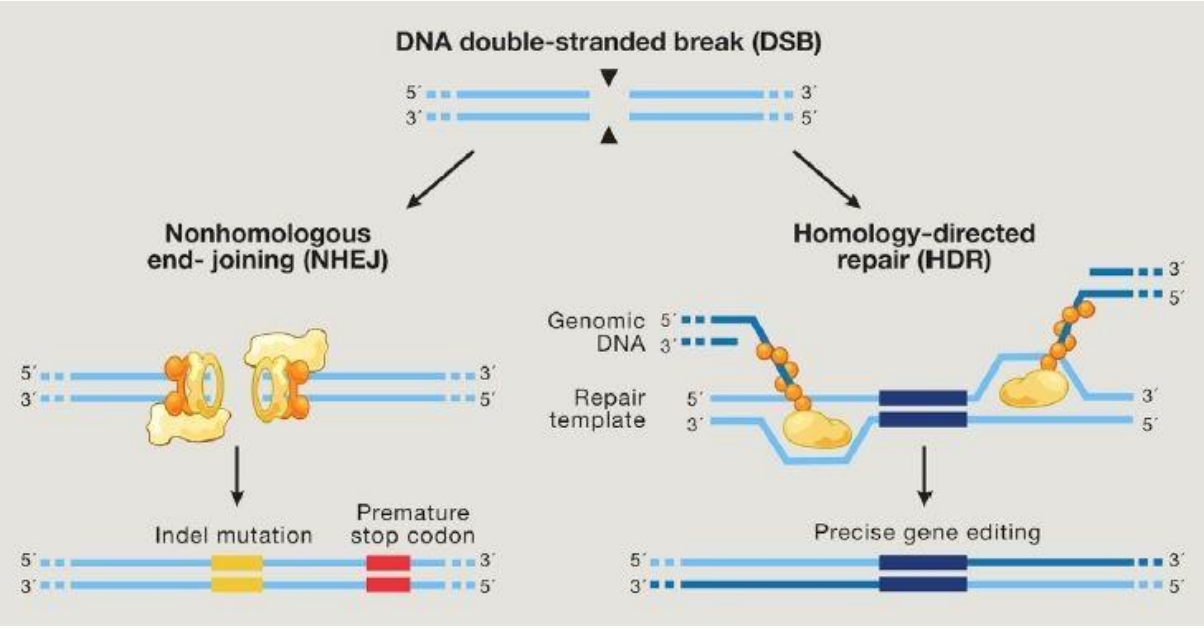


Figure 1.XIII: DNA double-strand breaks (DSBs) *are typically repaired by nonhomologous end-joining (NHEJ) or homology-directed repair (HDR). In the error prone NHEJ pathway, indels are introduced during the process of end resection and misaligned repair due to microhomology, often leading to frameshift mutations and the creation of premature stop codons. Alternatively, DSB ends are repaired by recruiting factors that direct genomic recombination with homology arms on an exogenous repair template. Bypassing the matching sister chromatid facilitates the introduction of precise gene modifications (reproduced from[185]).*

Aim of the Study

Breast cancer is, unfortunately, a common disease in our days, affecting women globally. Therapeutic strategies currently implemented fail to completely eradicate the disease, aiming at prolonging the patient's life expectancy. All tumors are highly heterogeneous, which is the principal factor behind discrepant clinical outcomes. Cancer heterogeneity can be explained by two models, the clonal evolution, and the Cancer Stem Cell model. Cancer Stem Cells (CSCs) constitute a small, rare sub-population of cells within tumors, with stemness properties and tumorigenic potential. Thus, elucidating their distinct biological make-ups is of particular importance, in our attempts to understand and cure cancer.

Epigenetic regulation has been identified as a major aspect in CSC biology, governing many of their properties. Histone Lysine Demethylases A and B (KDM5A and KDM5B) are histone demethylases associated with gene regulation. Their association with important developmental processes and cancer has already been established. KDM5A and KDM5B are overexpressed in Breast Cancer and have been shown to contribute to the CSC phenotype in other types of cancer, regulating important features of these cells, like their stemness properties and resistance to therapeutics. Current bibliography has contributed majorly to cancer research, however the predominant use of chemical inhibitors to study the role of these enzymes could be problematic, due to the high structural similarities members of this enzymatic family share, and their contribution in the specific context of luminal HR^+ breast cancer cells and breast cancer stem cells appears less investigated.

Based on the information above, the aim of this study was to:

- Develop a CRISPR-Cas9 system to specifically target and genetically modify KDM5A and KDM5B.
- Utilizing CRISPR-Cas9, produce luminal breast cancer cell lines with reduced KDM5A and KDM5B expression (knock-down/ knock-out).
- Investigate the effects of KDM5A and KDM5B absence in luminal breast cancer cells.
- Investigate the effects of KDM5A and KDM5B absence in luminal breast cancer stem cells (bCSCs).

2. Materials and Methods

1. MCF-7 cell culture

MCF-7 (ATCC – HTB-22) cells were used as the system of study of the present thesis. These cells were isolated for the first time in 1973, from a 69-year-old woman with metastatic breast cancer. MCF-7 is one of the most frequently used breast cancer cell lines. It belongs to the Luminal A subtype, expressing Estrogen (ER+) and Progesterone Receptors (PR+). Cells were cultured in Dulbecco's Modified Eagle's Medium-high glucose (DMEM) media (Sigma-Aldrich, D6429), enriched with 10% Fetal Bovine Serum (FBS) (Gibco FBS, Qualified, Standard, REF 10270-106) and 1% penicillin/ streptomycin antibiotics (Biosera, XC-A4122/100). Cells were incubated at 37°C, 5% CO2. Passaging was performed with Trypsin EDTA (1x) (Biosera, LM-T1705/100) after wash steps with Dulbecco's Phosphate Buffered Saline (PBS) (Sigma-Aldrich D8537).

2. Drugs

Doxorubicin

Adriamycin (Pfizer) is a chemotherapeutic drug. Its active substance, doxorubicin, is a cytotoxic antibiotic (carbocyclin) isolated from *Streptomyces* peucetius var. caesius. It acts as an antimitotic – cell-toxic drug, interfering within topoisomerase II – DNA complexes during the DNA replication phase, inducing replication stress. It was used at concentrations ranging from 10 nM – 5 μM.

Paclitaxel

Taxol (Paclitaxel) is a chemotherapeutic drug. Taxol targets tubulin, blocking mitotic spindle formation, microtubule polymerization and, ultimately, mitosis progression. This effect leads the cells to apoptotic cell death or reversion to the G0-phase of the cell cycle. Paclitaxel was used at concentrations ranging from 1 nM – 100 nM.

3. Mammosphere Formation Assay

Mammosphere formation is an in vitro assay to gauge breast cancer cell stemness. Cells are cultured in media containing EGF, FGF and B27 and in the absence of FBS, in low attachment plates. Cancer Stem Cells (CSCs) survive under these conditions and form spheroid structures in the culture, the mammospheres, while the rest of the cells die.

Spheres culture media: Dulbecco's Modified Eagle's Medium/F12 (DMEM/F12) (Biosera, LM-D1220/500), enriched with 1% glutamine , 2% B27 (Gibco B27 Supplement, REF 12587-010) and EGF (Immunotools Cat N. 11343406, FGF (20 ng/ml) (Immunotools Cat. N. 11343625) growth factors.

1st generation mammospheres

For this purpose, passage 5, 70-80% confluency MCF-7 cells were used. Cells were detached using Trypsin and collected in a falcon tube. Cells were counted using Neubauer plates. A suitable number of cells was then centrifuged at 1500 rpm for 5 min. The supernatant was discarded, cells were resuspended in mammosphere medium, transferred in bacterial (low attachment) plates, and incubated at 37°C, 5% CO_2, for 7 days. Starting at day 2, the number of mammospheres was counted daily. Spheres in six different, random focal points for each condition were measured and the mean number of spheres was calculated. In this way, Mammosphere Formation Efficiency (MFE) was calculated based on the formula:

M.F.E. = (number of mammospheres per well / number of cells seeded per well) x 100

At day 7, the spheres were collected and used in subsequent experiments.

4. Cell Population Growth

The Incucyte Zoom System (Essen BioScience, Hertfordshire, United Kingdom) was used for the proliferation assays. Incucyte is an automated cell observation machine, which allows the study of cells in real time and for long periods. It also records phenotypic differences between conditions. Cell proliferation is calculated based on cell confluency changes due time. Actively replicating cells, result in an increased density on the culture plate. Here, MCF-7 cells were cultured in triplicates, in 96-well plates. Photographs were taken in various timepoints to calculate cell density and establish cell growth, as the difference in cell confluency (%) every 24 hours.

5. Chemoresistance Assay

Similarly, Incucyte was used to establish IC_{50} values for doxorubicin. Following the same principal as in cell population growth assay, 1500 cells were incubated in duplicates and cell confluency (%) differences were established in 2 days. IC_{50} was automatically calculated using Graphpad Prism software from each sample's Normalized Cell Growth% - log($C_{doxorubicin}$) curve.

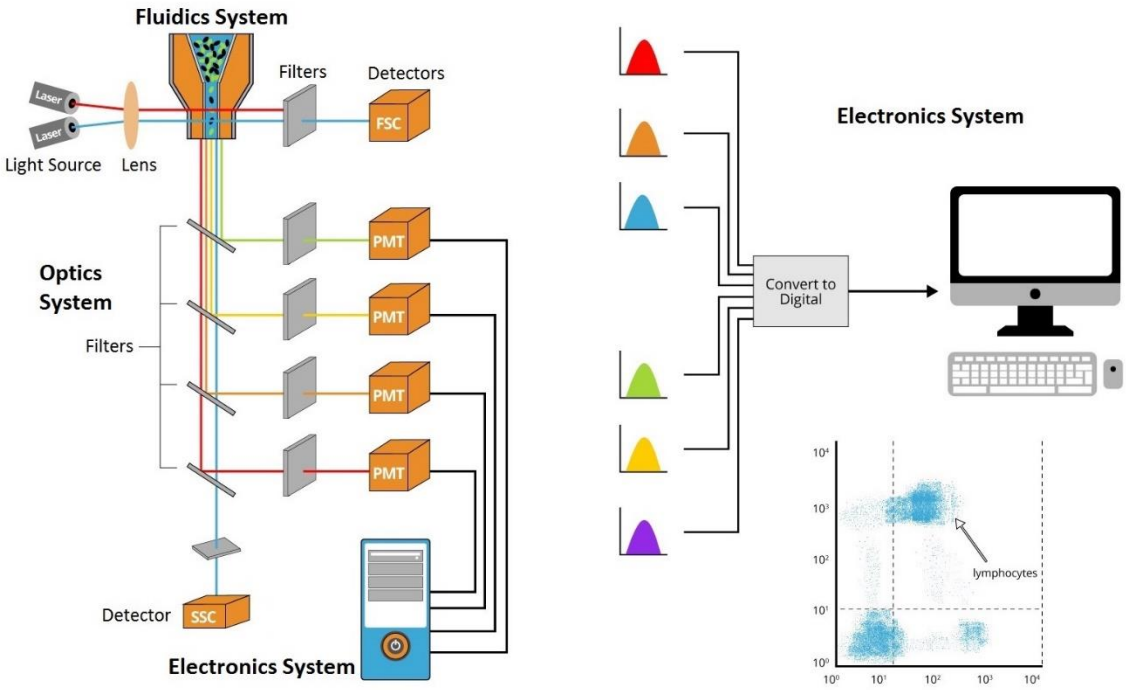
6. Cell Death Assay

Cell death was quantified utilizing the fluorescence detection feature of Incucyte and propidium iodide (PI) staining in the culture. PI is a nuclear dye that penetrates only dead cells and is excluded by living cells. 10.000 cells were seeded in duplicates, in 96-well plates. Every 24 hours, the medium was replaced with fresh medium containing 2 μg/ml

of PI. Photos of the well were taken 2 hours after medium change, and the percentage of fluorescent cells was calculated in each well. Dead cells (%) ratio was established as the ratio (% fluorescent cell confluency / % total cell confluency).

7. Fluorescence activated cell sorting (FACS)

Flow cytometry is a complex, multiparameter and automated method of measuring multiple physicochemical and/or phenotypic characteristics of cells or cellular organelles (nuclei, mitochondria, lysosomes). The characteristics are determined directly and distinctly for each of the cells of the tested sample. Its main advantage is the ability to simultaneously analyze more than one parameter. It is a quantitative method, characterized by high analytical capacity, accuracy and reliability compared to conventional techniques, such as microscopy. Flow cytometry provides the ability to determine cell size and granulation, assays for cell populations or subpopulations expressing characteristic membrane proteins, and so on. The analysis is performed using composite flow cytometer devices. A flow cytometer consists of three basic systems: a) the hydrodynamic flow system b) the optical system and c) the electronic data analysis system as it is shown in Figure 2.5. The hydrodynamic system is a hydraulic system that directs the cells or organelles of the suspended specimens to flow behind each other (flowing filamentous flux) in front of a focused laser beam absorbing them from the suspension. This is accomplished by inserting the sample into the center of a channel, surrounded by flowing inert liquid (Sheath fluid) along the channel. The channel is placed into the flow chamber where hydrodynamic focus is achieved in order for the cell to come into contact with the laser beam. The optical system collects the light signals emitted by the cells from the incident laser beam on them. The analysis is done on the basis of the collected radiation, in particular the scattered light and the fluorescence emitted. When the cell becomes a receiver of the vertical beam, as compared to the laser beam flow direction, a portion of the incident is absorbed by the cell while the remainder is scattered. Radiation scattered in the laser beam direction is called front scattering (FSC) and gives size indication. Radiation scattered in a direction perpendicular to the laser beam axis is called lateral scattering (SSC), indicating granulation. The fluorescence emitted is collected, analyzed, and measured by appropriate dichroic mirror and filter systems. Given the operating principles of a flow cytometer the necessary conditions for sample preparation and measurement are the presence of the test cell population in suspension and the labeling of cells with antibody-coupled fluorescent dyes.



Figurec.I: Schematic representation of the Flow Cytometer. *On the left side of the figure are showed the central parameters of the flow cytometer will on the right side is observed the image of the analysis with the electronic system (https://www.bosterbio.com/protocol-and-troubleshooting/flow-cytometry-principle).*

In our experiments, Flow cytometry was used to measure the percentage of the CSCs sub-population under different conditions using fluorescent-conjugated antibodies against the surface markers CD44 and CD24.

MCF-7 mammospheres were centrifuged at 800 rpm, 3min and collected in a falcon tube. Subsequently, spheres were dissociated into single cells by mechanical force. Cells were measured using a Neubauer plate, and an appropriate number of cells was centrifuged and resuspended in PBS with FBS (2%). The next step was the addition of the antibodies, starting with the antiCD44-PE conjugated (BD Pharmingen, Cat 555479). The cells were incubated with the antibody for 20 minutes at 4° C, under rotation. Then, antiCD24- FITC conjugated antibody (BD Pharmingen, Cat 555427) was added, and the sample was incubated under the same conditions. In order to set the appropriate parameters-conditions for the FACS analysis we used specific fluorescent-conjugated IgG control antibodies (PE and FITC isotype controls). The staining with these antibodies was performed by their simultaneous addition to the samples. The cells were then incubated

for 20 minutes at 4° C under rotation. In all cases, the staining of the cells was followed by two washes with PBS-2% FBS. Finally, the cells were centrifuged (1500 RPM, 5 minutes, 4o C) and then resuspended in 200 μl PBS-2% FBS.

Cell Cycle Analysis

Cells ($5x10^5$) were collected and washed with 8 mL of PBS. Then, they were centrifuged at 1000 rpm for 10 minutes and the supernatant was aspirated. Ice cold 80% Ethanol was added dropwise into the cell pellet while vortexing and cells were incubated at -20°C for at least 2 hours. After incubation, cells were washed twice with PBS and staining buffer. Following, cells were centrifuged (1500 rpm, 10 min) and stained using PI/ RNAase staining (BD Pharmigen, Cat: 550825) (15 min, RT). Tubes were protected from light, at 4oC and analyzed using FACS within an hour after staining.

Apoptotic Population Analysis

FACS was also utilized for apoptotic population analysis. Cells were collected and washed with PBS twice. Then, they are centrifuged (1500 RPM, 5 min) and resuspended 50 ul binding buffer. 4ul of Annexin V – FITC (BD Pharmigen, Cat:556419) conjugate are added and incubated for 15 min, 20°C, in darkness. After incubation, 250 ul binding buffer are added and the samples are analyzed in FACS.

8. RNA isolation

Total RNA isolation from the cells was done using Trizol (Life technologies) or Nucleozol (Macherey-Nagel) reagents. Following the protocol proposed by the company we accomplished to isolate total RNA from different samples. RNA quality and concentration were measured using Nanodrop (NanoDrop™ 2000/2000c Spectrophotometers) or a plain photometer. To estimate the amount of RNA in each sample using photometry, we measured the absorbance at 260 nm. The concentration of RNA in the sample was calculated by the formula

$$\text{RNA } (\mu g / mL) = \text{OD260nm} \times (\text{dilution factor}) \times 40\ \mu g / mL$$

9. Reverse Transcription (cDNA)

Reverse Transcription (RT) is called the synthesis of a complementary DNA strand (cDNA) having as template an RNA molecule. This reaction is catalyzed by the enzyme reverse transcriptase, which is naturally found in RNA-viruses (retroviruses) such as HIV. In the present thesis, for cDNA preparation we used the PrimeScript reverse transcriptase from TAKARA, following the manufacture's protocol.

10. Polymerase Chain Reaction (PCR)

PCR is a laboratory-controlled, in-vitro polymerization reaction, which mimics to some extent, the natural process of DNA replication. In particular, it is catalyzed by a DNA-dependent polymerase which, on the basis of a double-stranded, locally truncated DNA template and in the presence of a pair of suitable primers, the four dNTPs and Mg^{++} syntheses in vitro a huge number of DNA replicons (millions to hundreds of millions), with respect to that part of the original template, that is attributed to the primers. In the present study, we examined the mRNA level of many different genes using the KAPA Taq PCR kit (KAPA BIOSYSTEMS) and we followed the protocol that was proposed by the company.

11. Real-time PCR (RT-PCR) – Quantitative PCR (q-PCR)

In order to detect the expression level of several genes but also to compare them we chose to perform quantitative PCR (q-PCR) where the, DNA amplification is monitored at each cycle of PCR. Specifically, in this type of PCR, a fluorescent reporter is used in the reaction and when the DNA is in the log linear phase of amplification, the amount of fluorescence increases above the background. The point at which the fluorescence becomes measurable is called the threshold cycle (CT) or crossing point. In order to apply that technique in our study we used the KAPA SYBR® FAST qPCR Master Mix (2X) Kit and followed the protocol proposed by the company. As template, we used cDNA prepared from total RNA isolated from the cancer cells samples obtained. For analysis and quantitation of the data we used the $2^{-\Delta CT}$ method. GAPDH mRNA levels were used as a control and for normalization.

12. Protein Isolation

Whole cell extracts were used for protein level examination. The culture plate was washed with ice-cold PBS, before the cells were scraped in ice-cold lysis buffer and transferred in an Eppendorf tube. Samples were then constantly agitated for 30 minutes at 4^{o}C, before centrifuged at 16.000 x g for 20 minutes. The supernatant was again transferred in a clean Eppendorf and measured for protein concentration. It was then stored at -80^{o}C.

Cell Lysis buffer

- 150 mM NaCl
- 1 mM EDTA

- 50 mM Tris-HCl pH 7.4
- 0.5% sodium deoxycholate
- 1% NP-40 (Igepal CA-630) or Triton-X
- 0.1% SDS
- Protease inhibitors (c0mpleteTablets, mini, EDTA-free, EASYpack, Roche)

BCA protein Assay

Quantification with the BCA Protein Assay (ThermoScientific) was performed following the protocol proposed by the company. Total protein concentration was calculayed based on the absorbance od the samples at 562 nm.

13. Western Blot

Western Blotting is an analytical method with high sensitivity that is used to detect and identify proteins, providing on the same time information on their molecular size. Western Blotting takes advantage of the antigen-antibody recognition specificity and combines the distinctive power of electrophoresis, antibody specificity and sensitivity of enzyme assays. The proteins of the sample are electrophoretically separated under denaturing conditions and then transferred to a nitrocellulose membrane by application of electricity. The detection of the proteins occurs after incubation with specific antibodies and the reaction with a particular chromogen or fluorescent substrate. In the present thesis, the experimental protocol we followed for the western blot analysis will be described below. For SDS-Page electrophoresis, at least 25 μg of total protein were used for each sample, followed by their transfer on a Nitrocellulose membrane (pore size 0.45μm, Porablot NCP) that lasted for 1,5 hours at 4^{o} C (250 mA). After the transfer of the proteins, the membrane was blocked for 1 hour in 5% milk (in TBST) at room temperature on a shaker. The next step was the incubation with the 1^{st} antibodies the duration of which was for a night at 4^{o} C or 90 min at RT. After the incubation with the 1st antibody, the membrane was washed for three times with TBST (10 minutes, RT, on a shaker). Incubation with the 2nd antibody (HRP-conjugated) followed (1 hour, RT, on a shaker). Finally, the membrane was washed for three more times with TBST on the same conditions and then was incubated with the ECL reagent followed by the appearance of the results on.

14. CRISPR – Cas9 Protocol

sgRNAs were designed using the Broad Institute tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). This software provides a file (sgRNA – designs) that includes all the candidate sequences and their characteristics. The user picks the most suitable sgRNA after considering:

- On – target efficacy score: the degree of interaction with the targeting sequence
- Target Cut %: the percentage % of the point of the gene the CRISPR system will act on, creating the cut.
- Off-Target rank: Rank of the options based on unwanted off-target interactions in the genome
- After considering the above, the user proceeds by performing a BLAST search on the candidate sequences to identify the point of recognition (for the wanted gene and its isoforms) and possible off-target interactions.

In this study we picked the following sgRNA targeting KDM5A.

KDM5A specific sequences:	
sgRNAKDM5A_F	5' CACCGGTGTCCTAAATGTGTCGCCG3'
sgRNAKDM5A_R	3'CCACAGGATTTACACAGCGGCCAAA5'
KDM5B specific sequences:	
sgRNAKDM5B_1F	5'CACCGGCAGTGGGCTCACATATCAG3'
sgRNAKDM5B_1R	3'CCGTCACCCGAGTGTATAGTCCAAA5'

sgRNA KDM5A Characteristics: On-target efficacy: 65,07%. Target cut: 20.2%. No off-target sequences were identified. The chosen oligos were annealed and cloned into the lentiCRISPRv2 (Figure 2.I) (Addgene) vector.

sgRNA KDM5B Characteristics: On-target efficacy: 69.94%. Target cut: 11%. No off-target sequences were identified. The chosen oligos were annealed and cloned into the lentiCRISPRv2 (Figure 2.I) (Addgene) vector.

Cloning

Digestion and dephosphorylation of the vector, was performed utilizing the BsmBI restriction enzyme 30 minutes, 37^{o}C. Components included 5ug of the lentiCRISPRv2 vector, 3ul BsmBI (ThermoScientific, REF: ER0451), 3ul FastAP (Fermentas), 1x FastDigest Buffer, 100 mM DTT and H_2O It was then isolated through gel electrophoresis and a gel clean-up kit. The sgRNA oligos were annealed and phosphorylated. Components of the reaction : 1ul Oligo 1 (100 uM), 1ul Oligo 2 (100 uM), 1x T4 Ligation Buffer, T4 PNK enzyme and H_20. The last step, includes the ligation of the annealed oligos to the digested vector,

adding the BsmbI digested vector (50 ng), 1 ul diluted oligo duplex (1:200 into sterile water), 1x Quick Ligase Buffer (NEB), Quick Ligase (NEB M2200S) and H_20 resulting in the vector-insert plasmid.

Bacterial Transformation

The vector-insert plasmid was transformed into RecA- DH10B bacteria. The bacteria were thawed on ice. Plasmid DNA was added to the bacterial mix while on ice for 20 minutes. Tubes were then transferred to a heating block at 42°C for 1 minute, LB growth medium was added and then the cells were spread onto selection plates (agar + ampicillin). Single, pure colonies were selected, propagated, and tested for plasmid production. Finally, the selected plasmid was sequenced to verify plasmid-insert integrity.

Lentivirus production

Lentiviruses were utilized to stably transduce MCF-7 cells with the CRISPR system. The production of lentiviral particles containing our plasmid was conducted in HEK293 cells. A 3rd generation Lentivirus system was used. The plasmids psPAX (6μg), PMD2.G (2μg) (Addgene) and the CRISPR:sgRNA plasmid (10μg) were chemically transfected ($CaCl_2$ transfection) to the cells, in medium without serum (changed from full medium 2h before transfection). Cell medium was changed to full after overnight incubation. Cells were then left for 48h, when the virus-containing cell medium was collected in Eppendorf tubes and stored at -80°C until use.

Cell Transduction

MCF-7 cells were seeded in 6-well plates. Different virus to medium dilutions were used in each well, ranging from 1:2 – 1:5 ratios. Selection with puromycin began 48 hours after virus incubation. Puromycin selection was performed with progressively higher antibiotic concentrations, ranging from 0.5 ug/ml – 3 ug/ml, for at least 7 days.

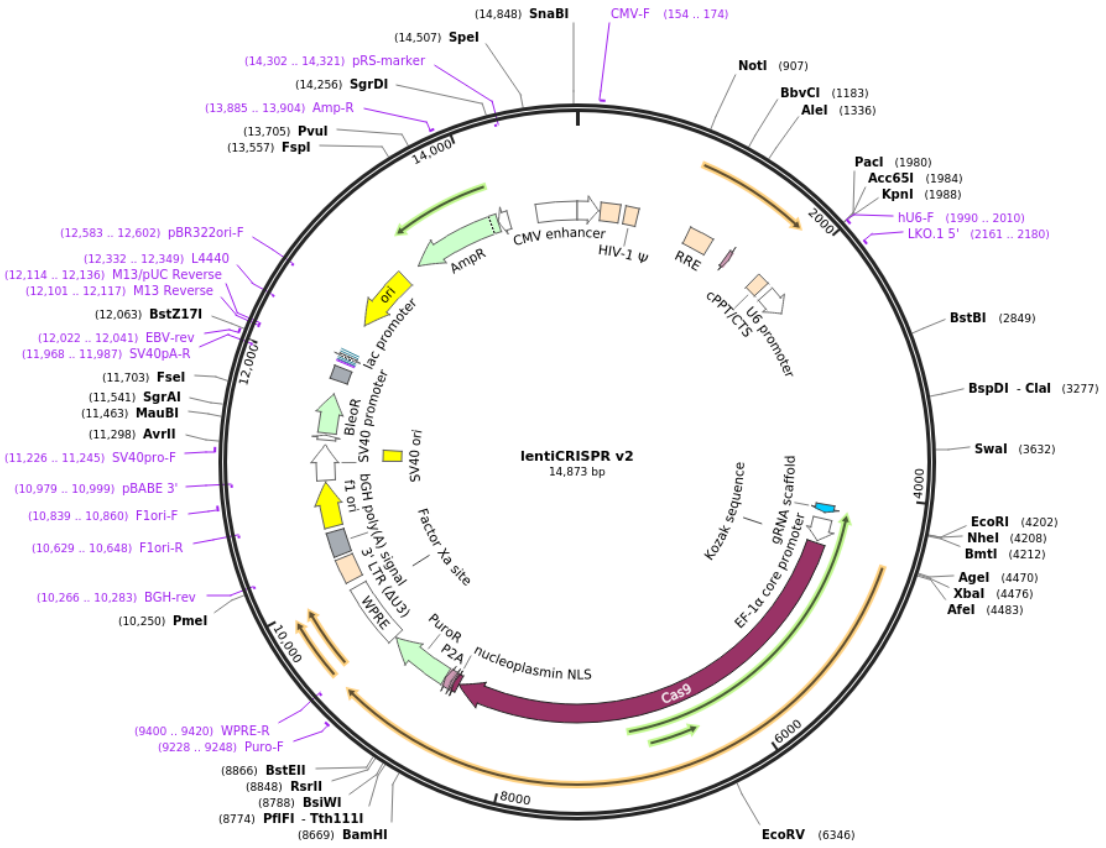


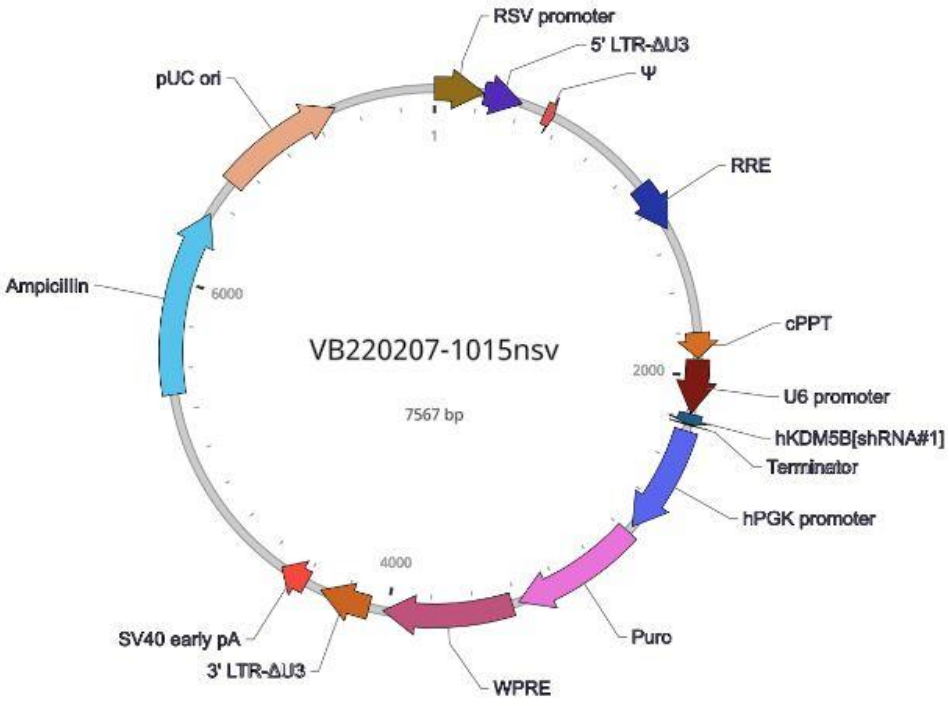
Figure 2.I: LentiCRISPRv2 vector: *Map of the vector displaying the functional regions and regions recognized by popular restriction enzymes.*

15. Extreme Limiting Dilution Assay (ELDA)

In vitro ELDA is a functional assay where cells in progressively higher dilutions are tested for their stemness and oncogenic potency. Cells are grown in non- attachment conditions and spheres formed are considered tumor "substitutes". MCF-7 mammospheres were generated as described above. A 96-well low-attachment plate (ThermoScientific 174925, 96U Bottom Plate) was used to seed 1000, 500, 100 and 50 cells in 12 replicates for each cell line. The number of wells containing mammospheres was counted after 7 days. The results were analyzed using the Institute of Medical Research Walter+Eliza Hall software[186].

16. shRNA

shRNA (short hairpin RNA) is a commonly used strategy for gene silencing. A plasmid containing the genetic information encoding a specific artificial RNA molecule with tight hairpin turn is transfected into the desired cell population. Upon expression, the RNA molecule acts as a mediator of RNAi on complementary mRNA sequences, leading to the silencing of said mRNA. Specific shKDM5B (GGACAACAGAACCTCATATTT) (Figure 2.III) and shScramble (CCTAAGGTTAAGTCGCCCTCG) (Figure 2.IV) were purchased from VectorBuilder. The vectors came in the form of E. coli strains which were cultured, and the vector was isolated using Plasmid Kit (Qiagen) following the manufacturer's protocol.



Lentiviral particles containing each vector were produced and used to transduce MCF-7 cells, following the previously mentioned protocol (Materials and Method 2.14).

Figure 2.III: sh_KDM5B vector: *Map of the vector displaying the functional regions (Vector Builder).*

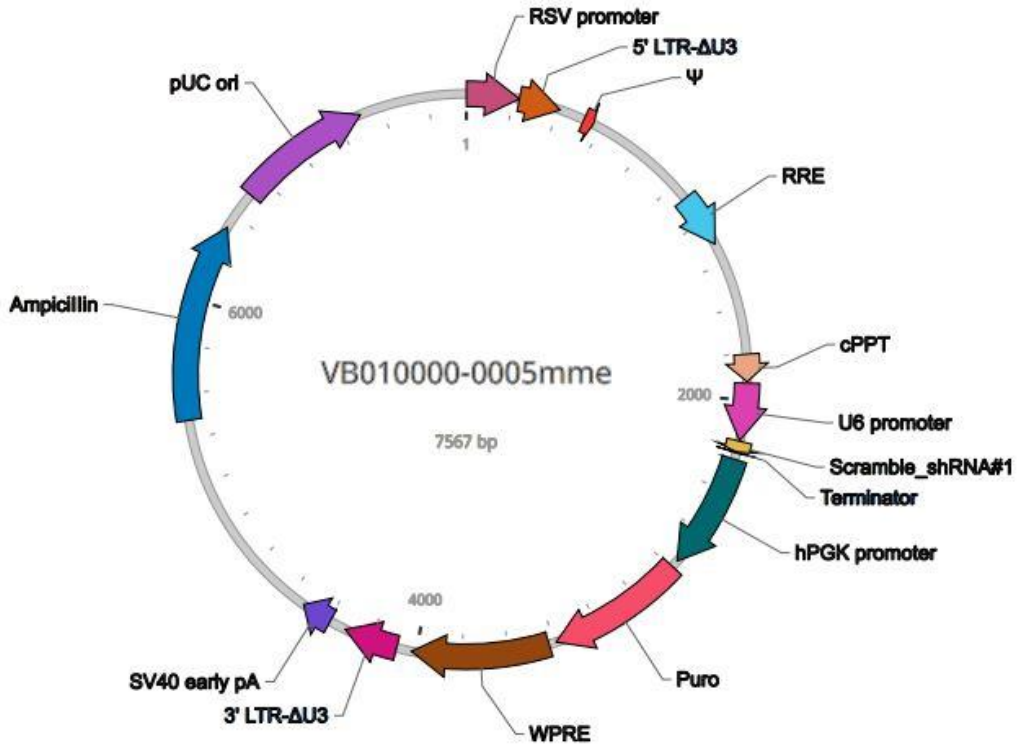


Figure 2.IV: sh_Scramble vector: *Map of the vector displaying the functional regions (Vector Builder).*

17. Bacterial Culture

Bacterial culture was used for the production of vectors used in this work. Bacterial strains were grown in 250 mL cone flasks in 100 mL LB media broth in the presence of ampicillin, 37^{o}C under constant agitation. Vectors were isolated using Qiagen plasmid purification kits.

18. Reagents

The different reagents used in the present work are listed below. For the cell culture experiments the reagents used are shown in **Table 2.1.**

Table 2.1: Cell culture medium and reagents

Cell culture		
Medium	Dulbecco's Modified Eagle's Medium-DMEM	Sigma-Aldrich/ D5697
	Dulbecco's Modified Eagle's Medium Ham's F12-DMEM F/12	biosera/LM-D112
Supplements-Growth Factors	B27 Supplement	Gibco/17504044
	Recombinant Human Fibroblast Growth Factorbasic (rh FGF-b / FGF-2)	Immunotools 11343625
	Recombinant Human Epidermal Growth Factor (rh EGF)	Immunotools 11343406
	Insulin	Sigma-Aldrich/ 9011-M
	Fetal bovine serum /FBS	Gibco 16140
Drugs	Doxorubicin	Adriblastina (Doxorucin Hydrochloride) 10mg/5ml VIAL Pfizer
	Paclitaxel (Taxol)	PATAXEL VIAL 30MG X5ML BIANEΞ A.E.
Others	Ultra-low attachment plates	Corning CLS3473
	Dulbecco's Phosphate Buffered Saline-PBS	Sigma-Aldrich/D8537

The reagents as well as the kits we used for the molecular experiments are listed in **Table 2.2.**

Table 2.2: Reagents and Kits used for Molecular experiments

Molecular Reagents		
RNA isolation	Trizol	TRIzol™ Reagent ThermoFisher Scientific 15596026
	Nucelozol	NucleoZOL reagent MACHEREY-NAGEL 740404.200

cDNA preparation	PrimeScript RTase	PrimeScript Rtase TAKARA 2680A
PCR	KAPA Taq	KAPA Taq PCR Kit, 500 U KAPA BIOSYSTEMS KK1016
RT-PCR	KAPA Sybergreen	KAPA SYBR® FAST Qpcr Master Mix (2X) Kit KAPA BIOSYSTEMS KK4604

In **Table 2.3** are listed all the antibodies used for FACS as well as western Blot analysis

Table 2.3 Antibodies used for FACS and Western Blot

Antibodies		
FACS	anti-CD44	PE Mouse Anti-Human CD44 Clone 515 (RUO) BD Pharmingen 550989
	anti-CD24	FITC Mouse Anti-Human CD24 Clone ML5 (RUO) BD Pharmingen 560992
Western Blot	anti-KDM5A	Anti-KDM5A, Abcam, ab70892
	anti-KDM5B	Anti-KDM5B, Abcam, ab181089
	anti-actin	Anti-Actin a.a. 50-70, clone C4 Millipore U.S.A. MAB1501
	anti-rabbit-HRP	Anti-rabbit IgG, HRP-linked Antibody Cell Signaling 7074
	anti-mouse-HRP	Anti-mouse IgG, HRP-linked Antibody Cell Signaling 7076

In the present study we performed RT-PCR. In the **Table 2.4** below are listed the sequences of the primers used in each case.

Table 2.4: RT-PCR primer sequences

RT-PCR primers		
GENE	Sequence 5'-3'	
KDM5A	Forward	TGTGGTCGGGGAAACAATGA
	Reverse	GTTTGCTACATTCCTCGGCG
KDM5B	Forward	ACCAAGATGGGGTTTGCTCC
	Reverse	GCAAACACCTTAGGCTGTCTCC

19. Statistical analysis

The experimental data that were obtained in the present thesis, were analyzed using 1 tailed paired TTEST.

3. Results

3.1 Development of the CRISPR-Cas9 system for KDM5A targeting

To accurately assess the role of the epigenetic enzyme KDM5A in cancer stem cells, KDM5A knock-out (k/o) cell lines were generated by genetically modifying the KDM5A locus, utilizing the CRISPR-Cas9 technology. Specific sgRNAs were designed using Broad Institute's tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). Factors like on-target efficacy, target cut% and off-target recognition were taken under consideration to select the optimal ones. Finally, one was used in subsequent experiments.

The sgRNA oligos were cloned into the appropriate plasmid backbone; specifically, the lentiCRISPRv2 plasmid vector (Addgene) was used, which expresses the components of the CRISPR-Cas9 system and was transformed into competent ($RecA^-$) bacteria. After bacterial culture, the lentiCRISPRv2:sgKDM5A plasmid was isolated and sequenced to verify plasmid-insert integrity. Lentiviruses containing the CRISPR-Cas9:sgKDM5A plasmid were produced and propagated, using HEK293T cells. Then, they were used to transduce MCF-7 luminal breast cancer cells, integrating the genetic elements to the cell genome, so that the cells would stably express the CRISPR components, Cas9 and the sgRNA.

3.2 Establishment and characterization of stable KDM5A knock-out cell lines using the CRISPR system

3.2.1 Generation of MCF-7 KDM5A knock-out cell lines

MCF-7 cells were transduced by lentiviral particles containing the CRISPR:sgKDM5A vector and selected using puromycin; only successfully transduced cells expressed the antibiotic resistance gene and survived under these conditions. Single resistant cells were transferred into 96-well plates, where they were incubated and left to grow and form colonies. Growing colonies were transferred successively into bigger plates till they were moved to a 35cm-plate, where they were evaluated for KDM5A protein expression.

Utilizing this strategy, two KDM5A knock-out cell lines were eventually generated, named E8 and F5 (Figure 3.I). In culture, E8 and F5 presented distinct features compared to parental MCF-7, including a slower growth rate, an increase in cell death (both analyzed and discussed later) and reduced plate attachment. Alterations in cell morphology were also observed. KDM5A k/o cells appeared smaller and with well-defined boundaries compared to MCF-7 parental cells (Figure 3.I).

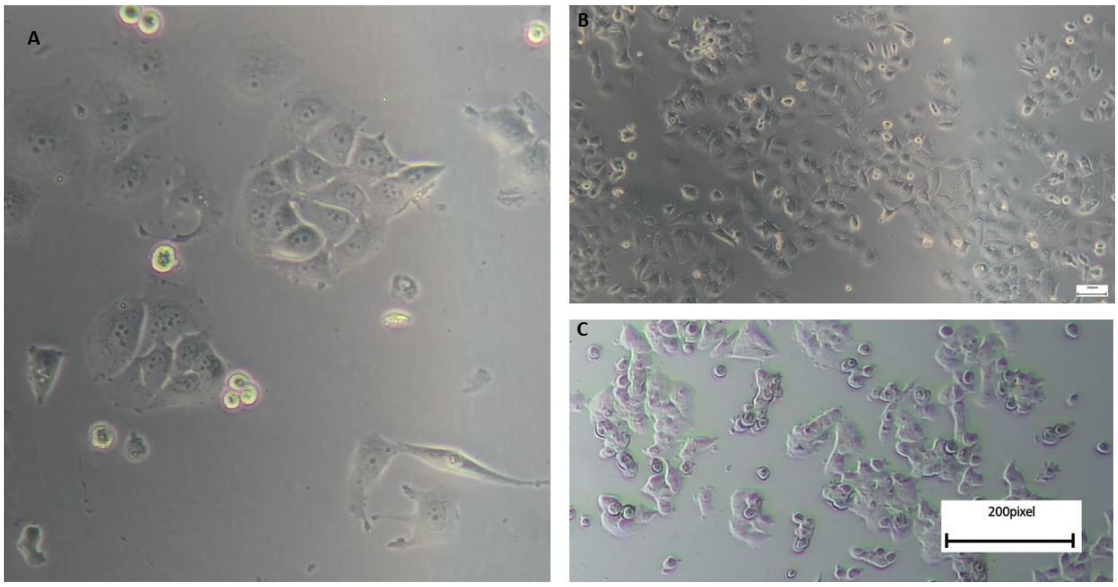


Figure 3.I: *Parental and KDM5A knock-out cell lines growing in culture.* ***A.*** *Parental MCF-7 cells* ***B.*** *F5 clone of KDM5Ako cells* ***C.*** *E8 clone of KDM5Ako cells. Photos under microscope (magnification* ***A*** *8x* ***B&C*** *4x)*

3.2.2 Knock-out validation: mRNA & Protein levels

To validate the generated knock-out cell lines, the absence of KDM5A was verified both on the transcriptional, as well as on the translational level. Whole-cell RNA was extracted from cells using established protocols and the levels of KDM5A expression were quantified by using RT-qPCR. KDM5A mRNA levels were significantly reduced in both cell lines, displaying a 96% knock-down in the E8 clone and a 98% in the F5 one, a direct indication of reduced protein expression (Figure 3.II.A). Protein expression levels were evaluated by Western Blot. Both approaches uniformly indicated a successful KDM5A knock-out. Specifically, KDM5A appeared completely abolished in E8 cells, while the F5 k/o clone presented a 95% reduction in KDM5A protein levels, compared to control MCF-7 cells (Figure 3.II B, C).

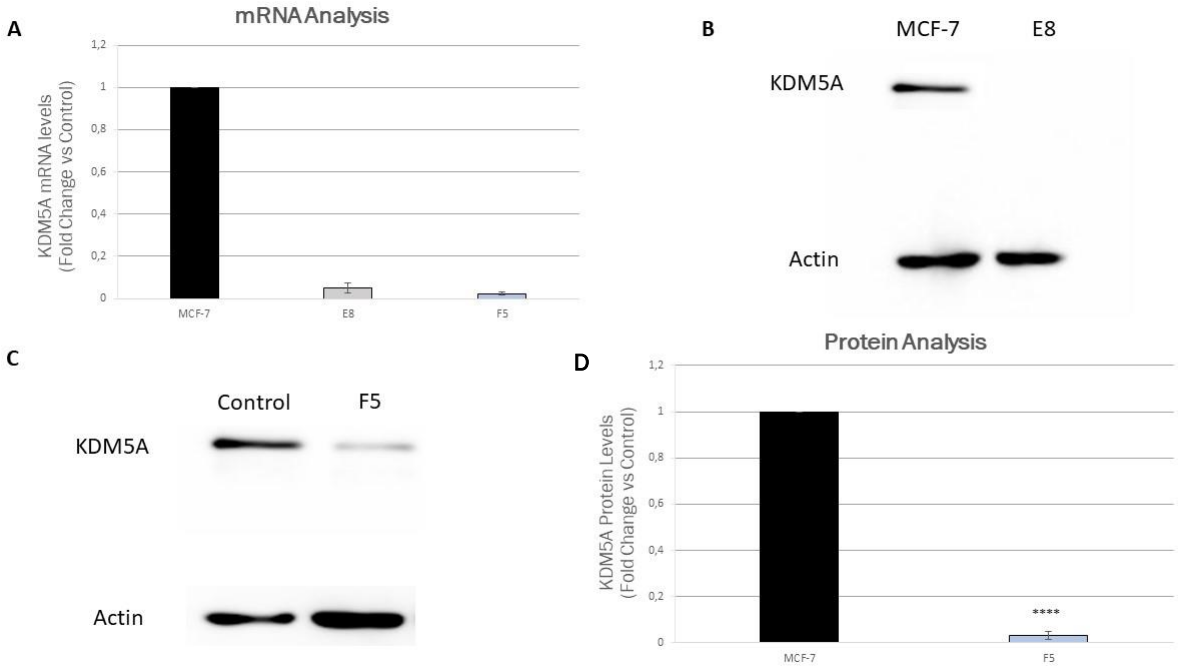


Figure 3.II: KDM5A expression levels in parental and KDM5A knock-out MCF-7 cell lines: ***A****. KDM5A mRNA levels in the two k/o clones compared to control MCF-7.* ***B. Western blot for*** *KDM5A protein levels in parental MCF-7 and the E8 k/o clone.* ***C. Western blot for*** *KDM5A protein levels in parental MCF-7 and the F5 clone.* ***D****. Quantification of the WB presented in (C). The error bars represent the SEM from 3 independent experiments. (****: denotes a p- value < 0.001).*

3.3 Investigation of the Role of KDM5A in MCF-7 cells grown as monocultures or as CSC-enriched mammospheres

3.3.1 KDM5A knock-out leads to a reduction in the MCF-7 cell population growth

As previously noted, the KDM5A k/o E8 and F5 cell lines displayed differences in their growth rates in culture, in comparison to parental MCF-7 cells. We investigated this phenomenon and tried to understand the causal effect behind it. Cell population growth as a term includes the entire process of population growth, the net result of the dynamics between cell proliferation and cell death.

Initially, we performed a cell growth assay to demonstrate experimentally the differences that were observed in culture by visual inspection. Incucyte allows live monitoring of cells and quantifies the percentage of the area these cells cover on the plate (% cell confluency). A specific number of cells, the same for each cell line, was seeded in triplicate, in 96-well plates, and population growth was measured using Incucyte in the course of 5 days. As seen in Figure 3.III, KDM5A k/o cells did present a significantly slower cell population growth rate compared to parental MCF-7 cells (p-value <0.001)

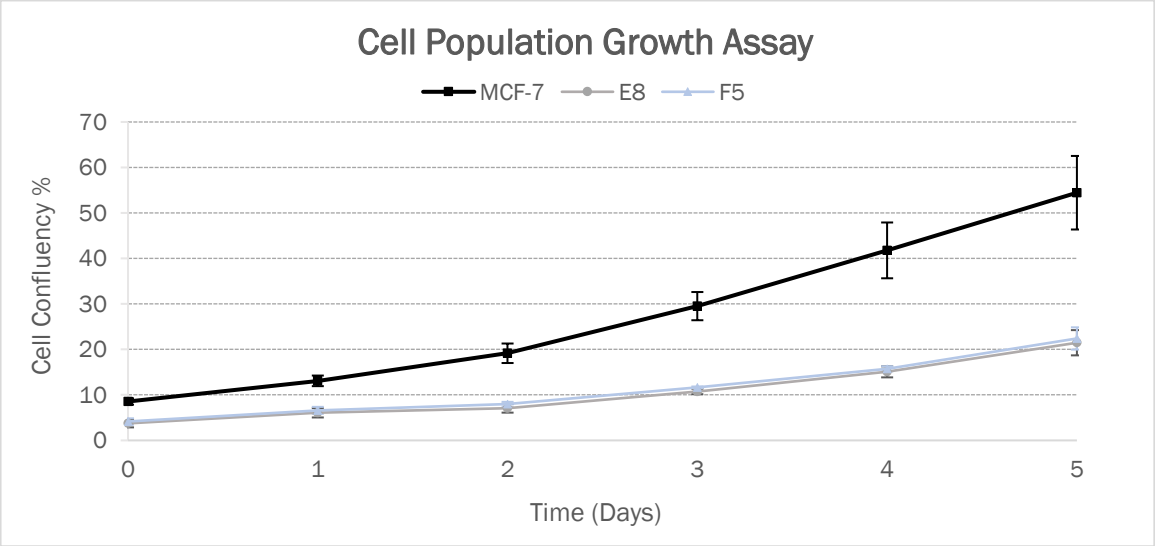


Figure 3.III: Cell Population Growth Assay for parental MCF-7 vs KDM5Ako cells: *The cell population growth of the E8 and F5 cell lines was measured during the course of 5 days by Incucyte. The error bars represent the SEM from 3 independent experiments.*

We then sought to discover the driving force of this delayed cell growth. KDM5A has been shown to promote cell cycle directly by promoting pro-proliferative signaling molecules, as discussed in section 1.5.3 of the present thesis. So, after confirming our observations, we proceeded by performing FACS to analyze the cell cycle dynamics of each cell line. During this process untreated cells were stained with PI, a popular red- fluorescent nuclear dye. PI interacts with the DNA by intercalating between the bases, which is why it is commonly used in quantitative DNA assays. Actively proliferating cells will have a higher population percentage in S-phase onwards, thus with duplicated DNA, a qualitative trait that we can detect with this method. As it is shown in Figure 3.IV, about 40% of parental MCF-7 cells were detected with duplicated DNA and were concluded to be actively proliferating. Percentage of proliferating cells drops at 32.6% in E8 (p-value <0.05) and 24% in F5 (p-value <0.001) KDM5A k/o cell lines. Thus, the cell cycle analysis revealed a small, but significant, stall in G1 phase in the E8 and F5 KDM5A k/o cell lines, compared to the parental cells.

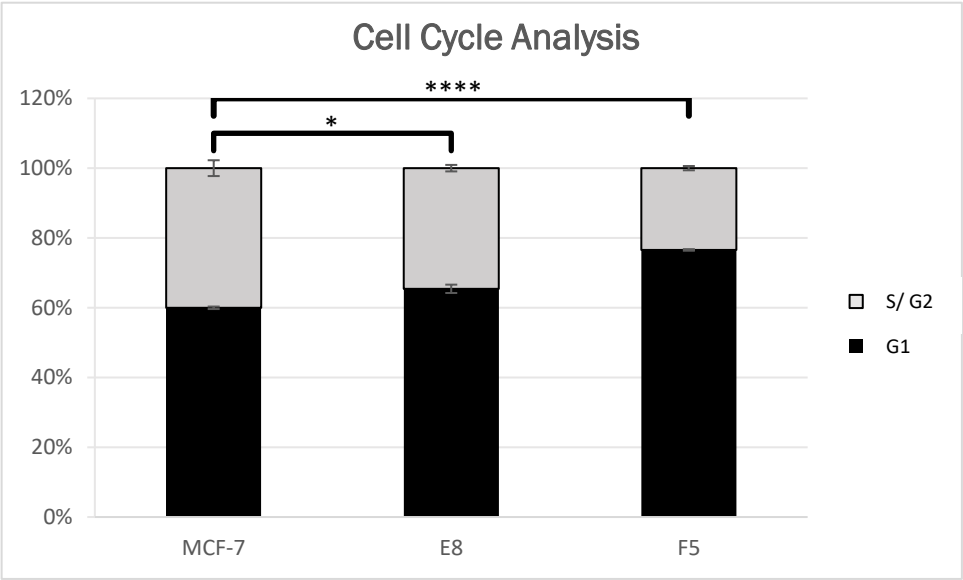


Figure 3.IV: Cell Cycle Analysis: *KDM5Ak/o cell cycle dynamics were analyzed with PI nuclear dye and compared to MCF-7 parental cells by FACS. The graph depicts the percentage of cells in each specific cell cycle phase based on the amount of DNA identified by PI fluorescence. G1 phase (black – non-duplicated DNA) and S/ G2 phase (gray – duplicated DNA). E8 and F5 cell lines display an increased percentage of cells in G1 phase compared to MCF-7 cells. The error bars represent the SEM from 3 independent experiments (*: denotes a p- value < 0,05, ****: denotes a p- value < 0,001)*

The small G1 stall did not appear to be the sole explanation for the differences in cell growth. Given that an increase in cell death was visually observed in cell culture, we proceeded to examine other possible reasons behind the delay in cell population growth. Apoptosis is the most ordinary form of cell death. Annexin V – FITC conjugates bind to phosphatidylserine phospholipids in the outer cellular membrane, a principal apoptotic signaling molecule. Utilizing FACS, the fluorescent-conjugated molecule can be detected, and the fluorescent-high population reveals the apoptotic cells in the sample. FACS analysis for parental MCF-7 and the KDM5A k/o cell lines did not reveal a significant increase in the apoptotic cell population, neither in E8 nor in F5 cells (Figure 3.V).

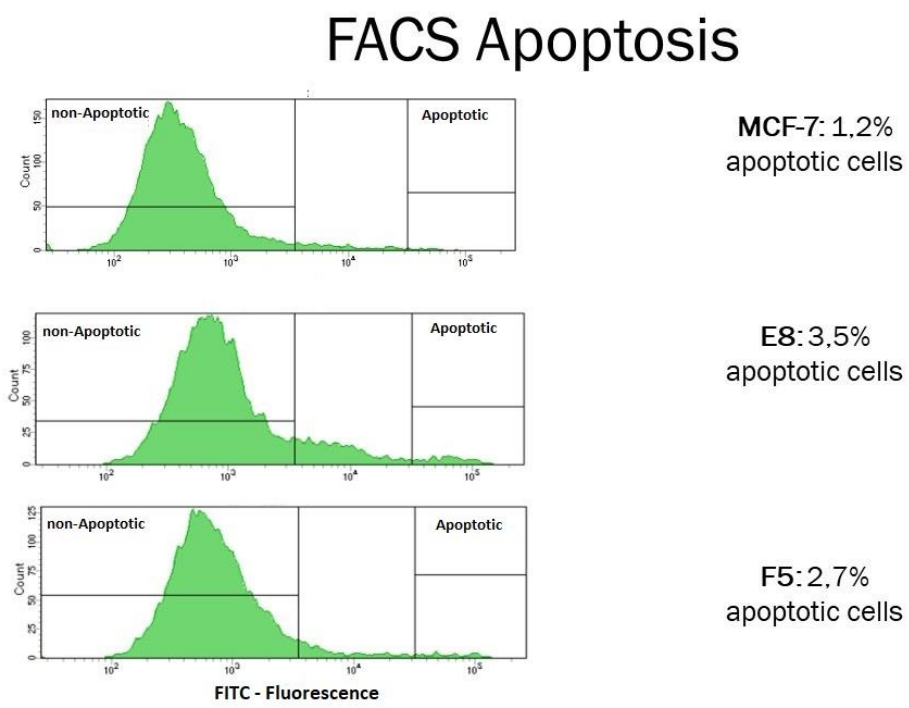


Figure 3.V: Apoptotic Cell Population Analysis: *Apoptotic population in each cell line was revealed using Annexin V conjugated with FITC and analyzed by FACS. An apex in the diagram, in the apoptotic cell area is not formed in any of the samples and a significant difference in the apoptotic populations between the cell lines was not identified.*

The apoptosis assay indicated that this form of cell death was not a principal factor in E8 and F5 cells delayed growth. As Figure 3.V depicts, an apoptotic population apex is not formed in the E8 and F5 cell lines and the small increase in the percentage of these cells was concluded to be insignificant. However, there are multiple other forms of cell death which could affect these cells. To evaluate this hypothesis, total cell death was quantified. As previously mentioned, PI is a fluorescent nuclear dye. PI is excluded by intact cell membranes but penetrates dead cells. Utilizing this feature, the percentage of dead cells in each sample can be quantified by measuring the percentage of fluorescent cells in the plate. This quantification was done by Incucyte, after seeding cells in triplicate in 96-well plates. Figure 3.VI depicts the fold-change differences in the percentage

of total dead cell ratio (percentage of fluorescent-positive cells relative to percentage of total cell confluency) in each cell line at a specific timepoint. E8 (KDM5A k/o) cells displayed the highest degree (7.5-fold on average) (p-value <0,05) of cell death in culture. F5 cells that express a small amount of KDM5A depicted a smaller increase (2.3-fold on average) in cell death percentages (p-value <0,005), confirming visual observations during culture procedures.

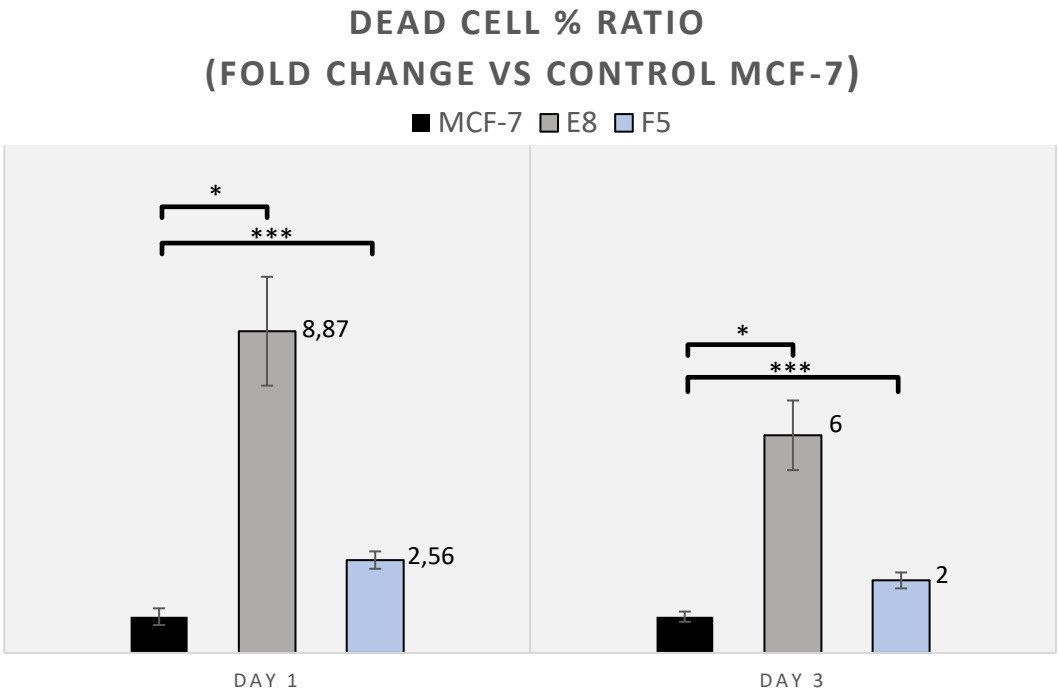


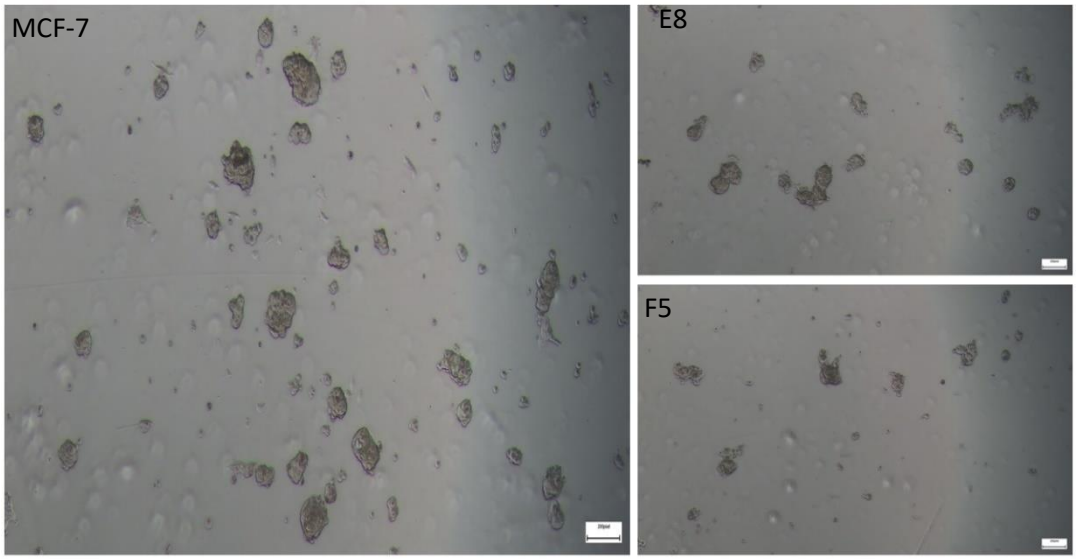
Figure 3.VI: Cell Death Assay: Quantification of the dead cells population in parental and KDM5A knock-out MCF-7 cell lines: *Cells were treated with PI, a dye permeable only to dead cells. The percentage of necrotic cells was quantified as a percent of the entire population (dead cell % ratio). The graph depicts the fold change in necrotic ratios for each cell line relative to control MCF-7 cells (set to 1). The error bars represent the SEM from 3 independent experiments (*:denotes a p- value < 0.05, ***: denotes a p- value < 0.005)*

In conclusion, the small stalling in G1 phase and the significantly higher death rate of the KDM5A k/o cell lines, revealed by these assays, could explain the differences noticed in the overall cell growth between the cell lines. Our results indicate a potential correlation between KDM5A expression levels and MCF-7 cell viability and cell cycle promotion.

3.3.2 KDM5A Knock-out depletes the Cancer Stem Cell Population

CSCs constitute a small subpopulation of cells within the tumor bulk that possess self-renewal capabilities and a higher oncogenic potential, in comparison to the rest of the tumor cells. To study this subpopulation of cells in vitro, it is essential to apply CSC-enrichment methodologies. In the present thesis, the cells were enriched in CSCs by low-attachment 3-D cell culture. MCF-7 cells growing under low-attachment conditions form 3-D spheroid structures termed mammospheres, and these structures are a direct correlation to CSCs' presence, as non-CSCs cannot grow under such conditions, due to the anoikis phenomenon. The E8 and F5 cell lines formed mammospheres with altered morphology (less well-defined, round structure) and reduced efficiency (p-value <0.005) compared to the parental MCF-7 cells (Figure 3.VII).

A



B

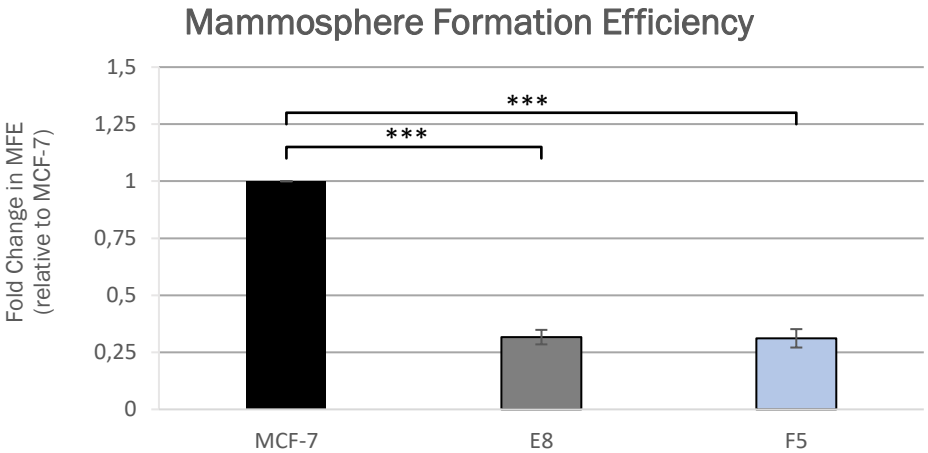


Figure 3.VII: Mammosphere Formation Efficiency: A. *Mammosphere culture for each cell line. Photos under microscope (magnification 4x)* ***B.*** *Quantified Mammosphere Formation Efficiency fold change differences between MCF-7 (set to 1) and KDM5Ak/o cell lines. The error bars represent the SEM from 3 independent experiments (p- value < 0.005)*

Mammosphere Formation Efficiency (MFE) provides a hint to the size of the CSC population within a cell line. However, mammospheres need to be analyzed further utilizing FACS to exude an informed conclusion on the potential stem cell effects. The cell surface molecules CD44 and CD24 have been thoroughly studied in the cancer stem cell context as specific biomarkers of the subpopulation. It is well established that breast CSCs present the $CD44^{high}/CD24^{low/-}$ phenotype[24]. Using specific fluorescent-conjugated antibodies, the presence of these markers was detected using FACS. FACS revealed a significant reduction in the $CD44^{high}/CD24^{low/-}$ cells in the KDM5A knock-out lines, by 98% in E8 and 95% in F5 cell lines, compared to parental MCF-7 cells (p-value <0.005) (Figure 3.VIII).

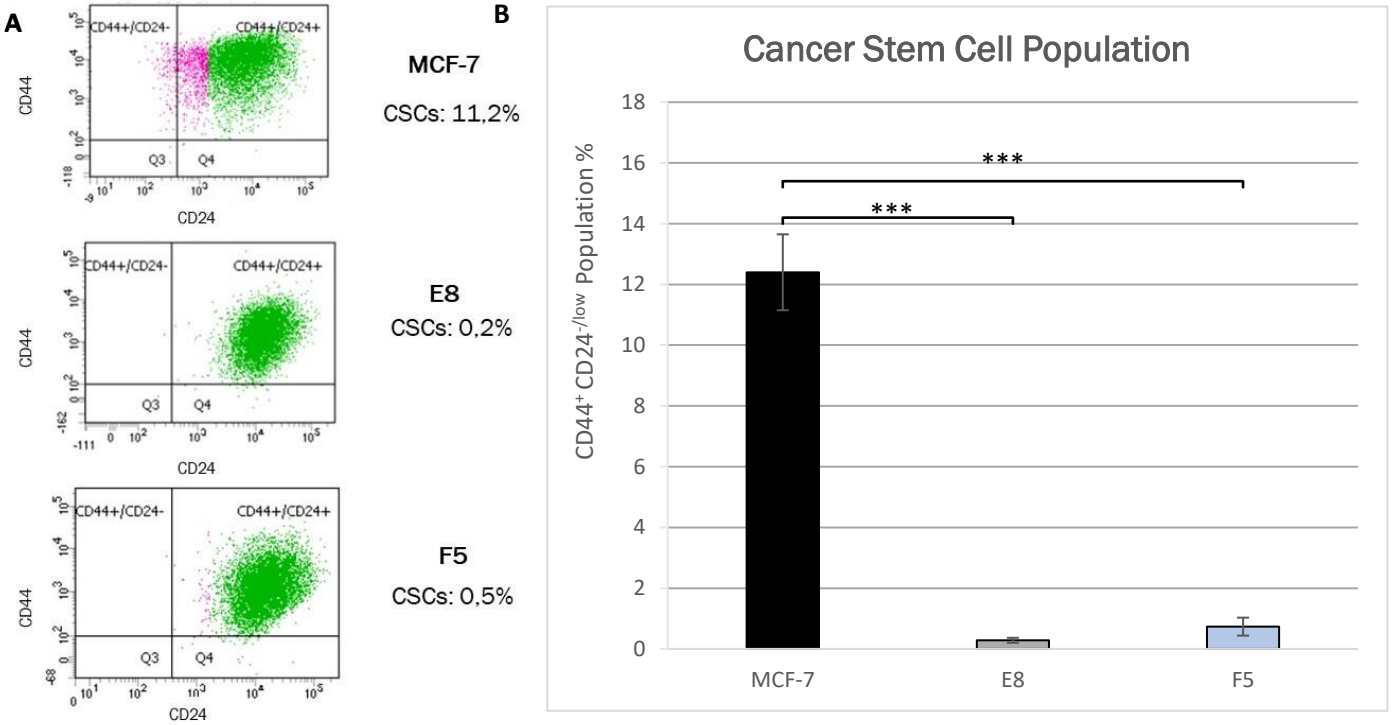
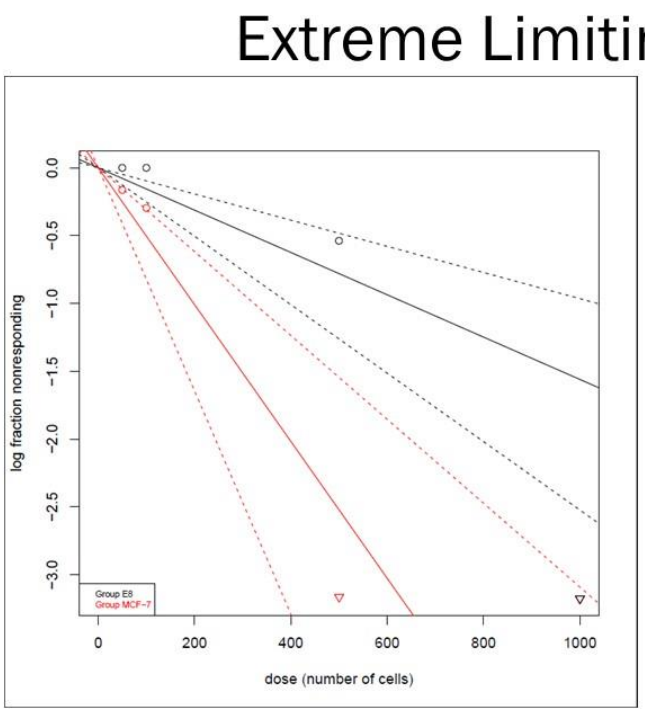


Figure 3.VIII: CSCs population quantification in parental and KDM5A knock-out cell lines: A. *FACS figures displaying the difference in* $CD44^{high}/CD24^{low/-}$ *population* ***B.*** *Graphical display of FACS results. The error bars represent the SEM from 3 independent experiments (***: denotes a p- value < 0,005).*

Cancer Stem Cells constitute the apex of tumor development due to their higher oncogenic potential, being capable of repopulating tumors in very small numbers compared to the differentiated tumor cells. As such, functional evidence is required to display their tumor-initiating capacity, complementing the results indicated by phenotypic assays. Extreme limiting dilution assays have been used extensively as mechanistic evidence of CSC presence, both in vivo and in vitro. Here, we adopted an in vitro ELDA (Extreme Limiting Dilution Assay) approach[186] to further

support our findings. In this experiment, successive dilutions of E8 and parental MCF-7 cells were seeded in low-attachment 96-well plates in mammosphere-special medium, and the non-mammosphere-forming cell fraction was quantified (non-responding) for each cell line. The results are displayed in Figure 3.IX, where the non-responding cells log fraction is plotted against the total number of cells seeded in each condition. A steeper line reveals a decrease in the number of wells, where mammospheres were not formed, hence a higher mammosphere formation potency (Figure 3.IX). As few as a hundred MCF-7 cells were capable of forming mammospheres, while at least five hundred KDM5A k/o cells were needed to replicate this result. As depicted in Figure 3.IX, E8 cells were significantly less potent in producing mammospheres, with an estimated 1/641 cells being a stem cell. In comparison, CSCs were three times more frequent in MCF-7 parental cells.



Overall test for differences in stem cell frequencies between any of the groups		
Chisq	DF	P.value
10.4	1	0.00123

Confidence intervals for 1/(stem cell frequency)			
Group	Lower	Estimate	Upper
E8	1037	641	396
MCF-7	323	198	122

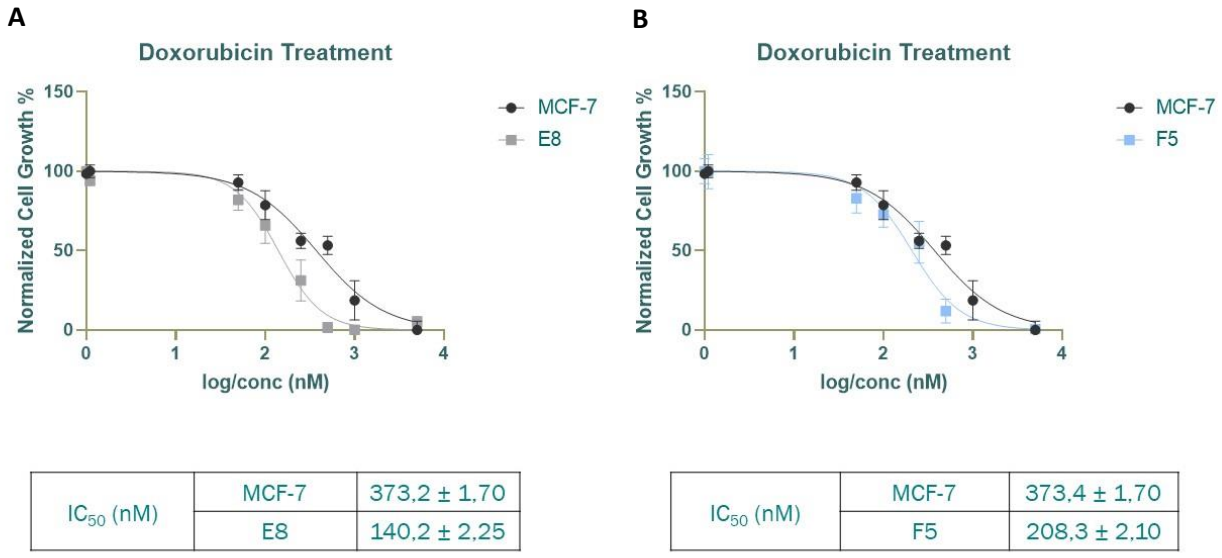
Figure 3.IX: Extreme Limiting Dilution Assay in parental MCF-7 and in the E8 KDM5A knock-out cell lines: A. *Graph depicting the log fraction of non-responding cells relative to the total cell number in each condition.* ***B.*** *Statistical analysis and quantification of CSC estimate for each group. CSCs are three times more frequent in parental MCF-7 cells.*

All the above data strongly suggest that there is a significant association between KDM5A expression and the CSC population in MCF-7 cells.

3.3.3 KDM5A Knock-out cells are more sensitive to chemotherapeutics

A huge hurdle in restricting and -ultimately- curing cancer is the various mechanisms of resistance the tumors develop. Therapy resistance is also a hallmark and an innate characteristic of CSCs, which are often considered the drivers of resistance and tumor recurrence after initially successful therapy (section 1.3.4).

After showing the reduction in the CSC- population in the KDM5A knock-out lines, we sought to discover whether these cell lines would present a lower chemoresistance dynamic relative to the parental MCF-7 cells. To this end, the cells were seeded in 96-well plates in duplicate in the presence of a commonly used chemotherapeutic, doxorubicin. The IC_{50} for each cell line was calculated, measuring the difference in % cell confluency in the course of 2 days using the Incucyte system, as it was described before (Figure 3.X). The E8 and F5 cells were significantly more sensitive to doxorubicin treatment compared to parental MCF-7 cells (p-value <0.05), 1.8 to 2.66-fold as revealed by the IC_{50} values.



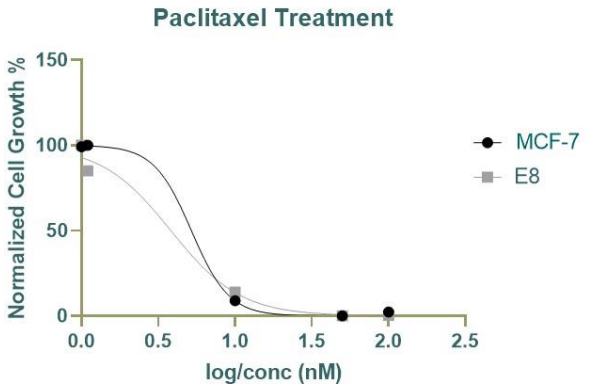
IC_{50} (nM)	MCF-7	373,2 ± 1,70
	E8	140,2 ± 2,25

IC_{50} (nM)	MCF-7	373,4 ± 1,70
	F5	208,3 ± 2,10

Figure 3.X: Cell growth assay in parental and KDM5A knock-out cell lines after doxorubicin treatment: *Cell growth relative to various concentrations of doxorubicin measured by Incucyte.* ***A****. MCF-7 parental vs KDM5 Ak/o line E8.* ***B.*** *MCF-7 parental vs F5 KDM5A k/o line. E8 & F5 cell lines were more sensitive to Doxorubicin treatment. The IC_{50} values were calculated using GraphPad. The error bars represent the SEM from 3 independent experiments (p- value <0.05).*

The increased sensitivity to doxorubicin coupled with the reduction in the CSC population could pinpoint KDM5A as a potential CSC-specific target for combination therapy with this drug in luminal breast tumors.

Chemosensitivity was also tested for another commonly used chemotherapeutic drug, paclitaxel. The E8 clone that presented a total KDM5A depletion was used (Figure 3.XI). The experiment was conducted once, and the results revealed a slight decrease (37%) in the IC_{50} values, compared to parental MCF-7 cells, which was deemed insignificant and was not pursued further.



IC50 (nM)	MCF-7	5,2
	E8	3,8

Figure 3.XI: Cell growth assay in parental and KDM5A knock-out cell lines after paclitaxel treatment: *Cell growth relative to various concentrations of paclitaxel measured by Incucyte. MCF-7 parental vs KDM5 Ak/o line E8.*

The fact that the KDM5A k/o clones are more sensitive to doxorubicin than paclitaxel is probably attributed to the different mechanisms of action of each drug. KDM5A may be implicated in Replication Stress Response (RSS), contributing on cell response to stressors occurring during DNA replication (S-phase). However, no such implication has been observed in the M-phase checkpoint, where paclitaxel is known to act.

3.4 Establishment and characterization of stable KDM5B knock-down cell lines

3.4.1 Development of the CRISPR-Cas9 system for KDM5B targeting

Our approach in studying KDM5B mirrored the one for KDM5A, attempting to deplete KDM5B protein levels utilizing the CRISPR-Cas9 technology. To this end, specific sgRNAs were designed using Broad Institute's tool (https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design). Factors like on-target efficacy, target cut% and off-target recognition were taken under consideration to select the optimal ones. Finally, one was used in subsequent experiments.

The sgRNA oligos were cloned into the appropriate plasmid backbone; specifically, the lentiCRISPRv2 plasmid vector (Addgene) was used, which expresses the components of the CRISPR-Cas9 system and was transformed into competent ($RecA^-$) bacteria. After bacterial culture, the lentiCRISPRv2:sgKDM5B plasmid was isolated and sequenced to verify plasmid-insert integrity. Lentiviruses containing the CRISPR-Cas9:sgKDM5B plasmid were produced and propagated, using HEK293T cells. Then, they were used to transduce MCF-7 luminal breast cancer cells, integrating the genetic elements to the cell genome, so that the cells would stably express the CRISPR components, Cas9 and the sgRNA. Unfortunately, the same result was not replicated in this situation, as none of the 96 stable cell clones produced presented KDM5B knock-out or sufficient knock-down, tested both on the mRNA level using q-PCR and on the protein level by western blot.

3.4.2 MCF-7 KDM5B Knock-down cell lines

To study the role of KDM5B in MCF-7 CSCs, and since the CRISPR-Cas9 was not successful, we altered our approach by adopting the shRNA gene silencing method aiming to generate a stable knock-down of the desired protein. shKDM5B plasmids (VectorBuilder) were propagated and isolated from already transformed E.coli bacteria. Using a simultaneous (3-plasmid) transfection of the shRNA vectors and the lentiviral genome vectors, our construct of interest was packaged in lentiviruses and used to permanently transduce MCF-7 cells, expressing the corresponding shRNA, thus achieving a stable knock-down. This process was replicated thrice, generating three cell lines, named sh_Scr, where MCF-7 cells were transduced with a scrambled shRNA sequence, targeting no known gene, and lines sh_KDM5B1 and sh_KDM5B2 (Figure 3.XII).

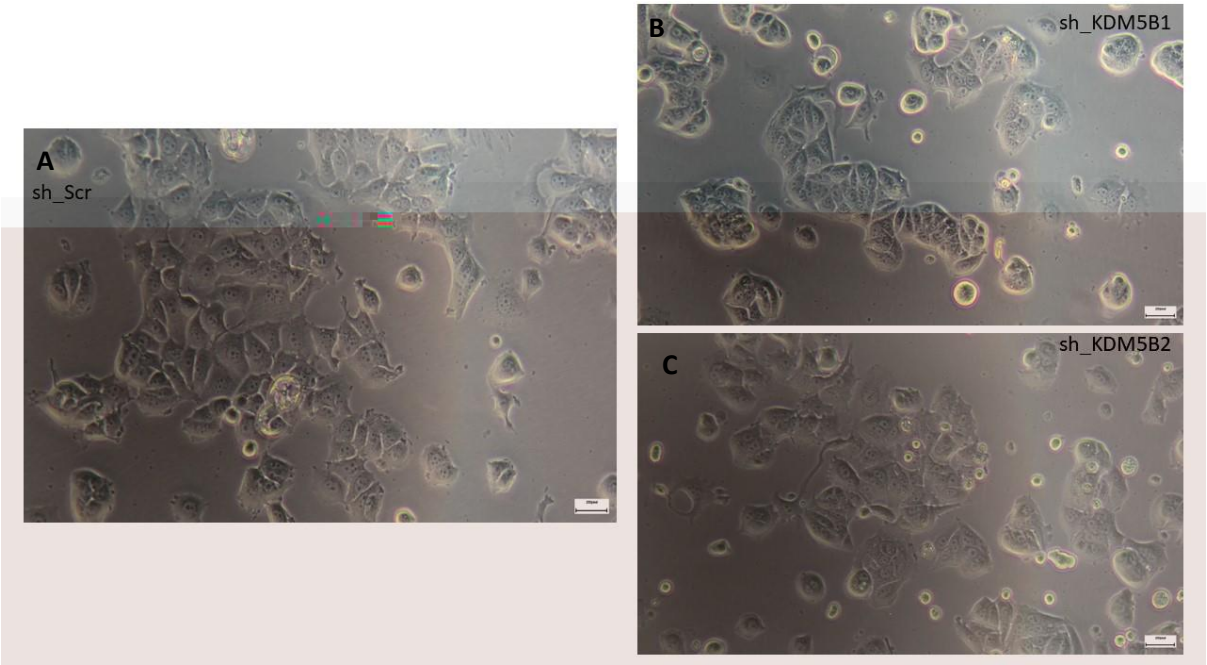


Figure 3.XII: *sh_Scramble and sh_KDM5B MCF-7 transduced cell lines growing in culture.* ***A.*** *sh_scr cells* ***B.*** *sh_KDM5B1 cells* ***C****. Sh_KDM5B2 cells. Photos under microscope (magnification 8x)*

KDM5B gene expression was analyzed on both the transcriptional and the translational level. mRNA analysis revealed a successful knock-down compared to scrambled shRNA, around 50% in both cell lines (Figure3.XIII A). Western blotting also revealed a significant reduction in KDM5B protein levels, almost at a 95% silencing for sh_KDM5B1 cell line and 70% in sh_KDM5B2 cells (Figure3.XIII B).

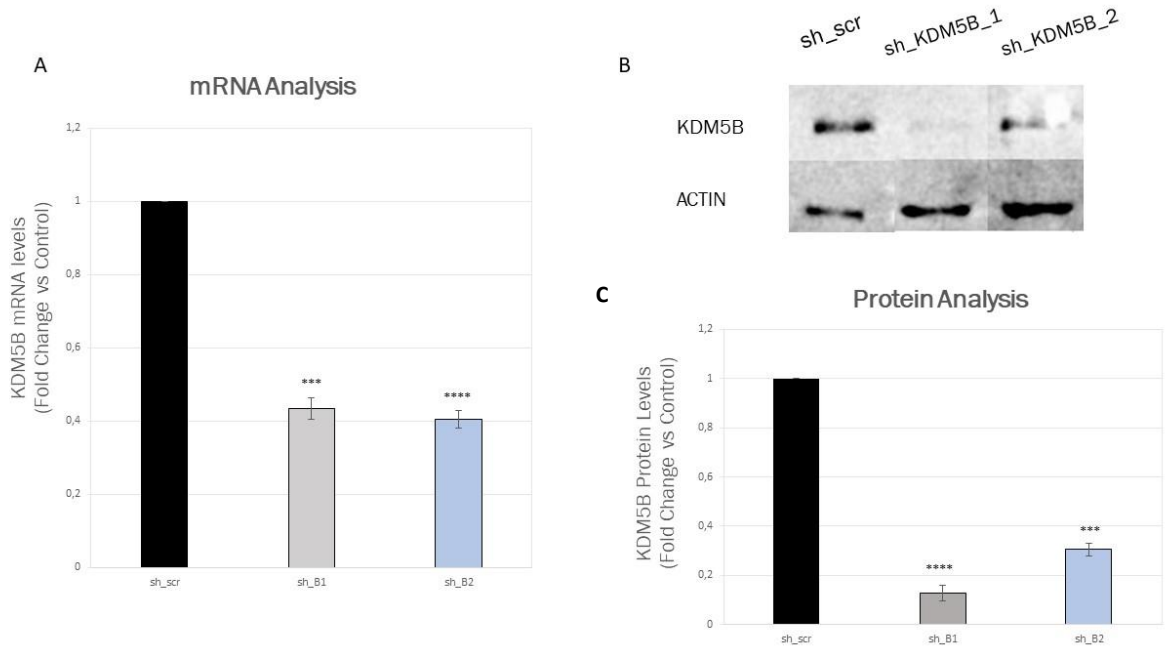


Figure 3.XIII: KDM5B expression levels in sh_scramble and KDM5B knock-down MCF-7 cell lines: ***A****. Fold change as indicated by qPCR detection of KDM5B mRNA levels on the two k/d clones relative to control sh_scramble (set to 1).* ***B. Western blot for*** *KDM5B protein levels in sh_scramble and sh_KDM5B transduced clones.* ***C.*** *Graphical representation of the fold change in protein expression between sh_scramble (set to 1) and shKDM5B cell lines. The error bars represent the SEM from 3 independent experiments. (***: denotes a p-value <0.005; ****: denotes a p- value < 0.001).*

The above results serve as a validation of our approach, also displaying a gene silencing potency for shRNA approaching that of CRISPR-Cas9, as sh_B1 protein levels resemble the results of the F5 cell line in KDM5A (Figure 3.II).

3.5 Investigation of the Role of KDM5B in MCF-7 cells grown as monocultures or as CSC-enriched mammospheres

3.5.1 shKDM5B knock-down leads to a reduction in cell population growth in MCF-7 cells

KDM5B has been studied in various biological environments, where it presents distinct functions; it is often being studied in tandem with KDM5A, especially in cancer research[20,130,154,168]. To provide an initial analysis of its roles in our system, KDM5B was approached the same way KDM5A was. Cell population growth analysis with Incucyte, revealed slower population growth dynamics for sh_B1 and sh_B2 KDM5B k/d, compared to sh_scr cell lines (p-value <0.001) (Figure 3.XIV).

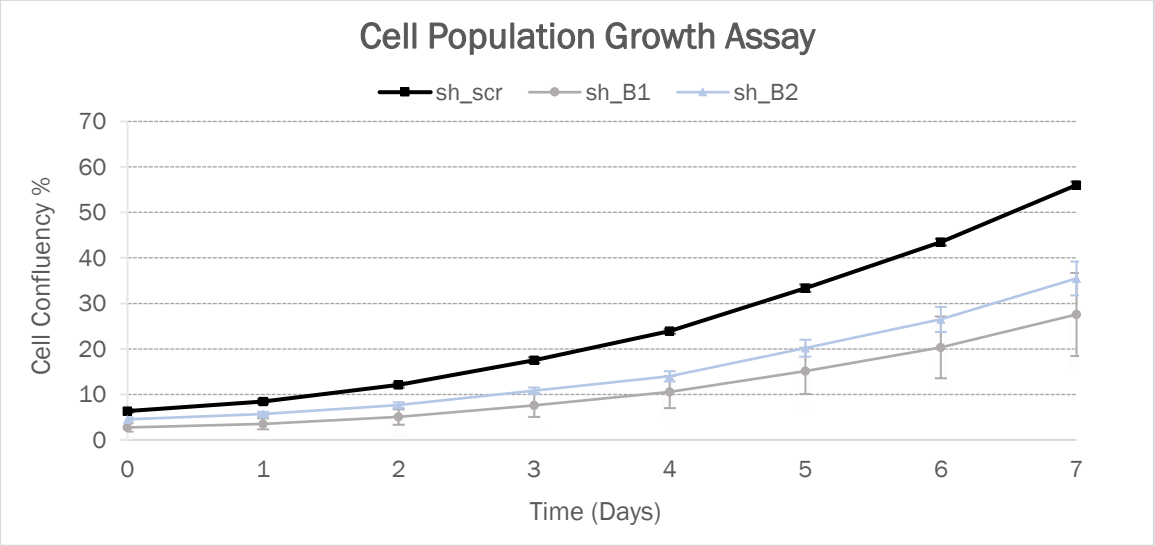


Figure 3.XIV: Cell Population Growth Assay in sh_Scramble and KDM5B knock-down cell lines*: Graph depicts the cell population growth in cell confluency% in the course of 7 days. Measurements were made by Incucyte. KDM5Bk/d lines present a slower cell growth. The error bars represent the SEM from 3 independent experiments. (p- value < 0.001)*

The sh_KDM5B1 line, which presents the higher degree of knock-down, also displays a slightly higher effect in growth dynamics. This phenomenon was also observed in the apoptotic analysis (Figure 3.XV). Using Annexin V - FITC staining and FACS we identified the apoptotic cells in each cell line. sh_B1 cells display a significantly higher apoptotic population (2,33-fold) compared to both sh_scr (p-value <0.05) and sh_B2 cells, which do not present a significant difference in apoptotic cells relative to control.

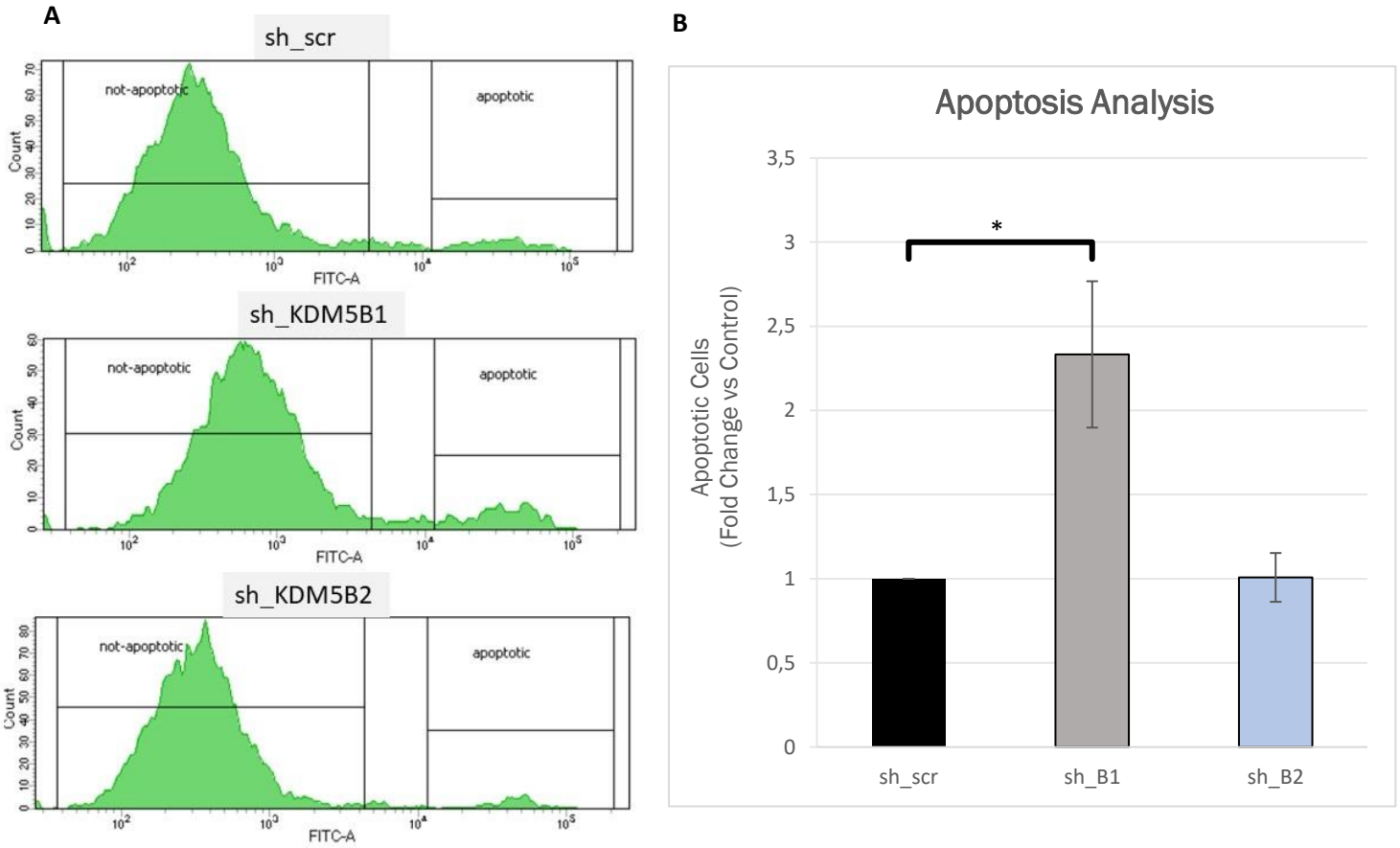


Figure 3.XV: Apoptotic Cell Population Analysis in sh_scramble and KDM5B knock-down cell lines: A. FACS figures *sh_scr, sh_KDM5B1 and sh_KDM5B2 stained with Annexin V to reveal the apoptotic population. An apoptotic population apex formation tendency is noticed in sh_B1 cells.* ***B.*** *Graph displaying the fold change in apoptotic populations. Sh_KDM5B1 cell line presents a significant increase in the apoptotic population compared to sh_scr (set to 1). A slight but non-significant increase was found in sh_KDM5B2 cells. The error bars represent the SEM from 3 independent experiments. (*: denotes a p- value < 0.05).*

The above data may illustrate a role for KDM5B in cell viability or apoptosis evasion in MCF-7 cancer cells displayed in extremely low levels of protein presence.

3.5.2 KDM5B knock-down effect on the MCF-7 Cancer Stem Cell (CSC) Population

As discussed in section 1.5.3, KDM5B has been identified both as a positive and a negative regulator of cancer and cancer stem cells, depending on the biological context. $CD44^{high}/CD24^{low/-}$ population analysis was performed using FACS. Our results on the cancer stem population of MCF-7 cells after KDM5B silencing, revealed mixed results, displaying no effect on sh_B1 and a tendency of increase on sh_B2 cell lines. The sh_KDM5B2 line displayed a significant, 50% increase on the $CD44^{high}/CD24^{low/-}$ cells, while the sh_B1 CSC population difference relative to control was not significant (Figure 3.XVI). MFE was not calculated in these experiments, as problematic low-attachment-plate coating led to cells attaching on the plate, presenting inflated cell numbers and preventing the establishment of clear images.

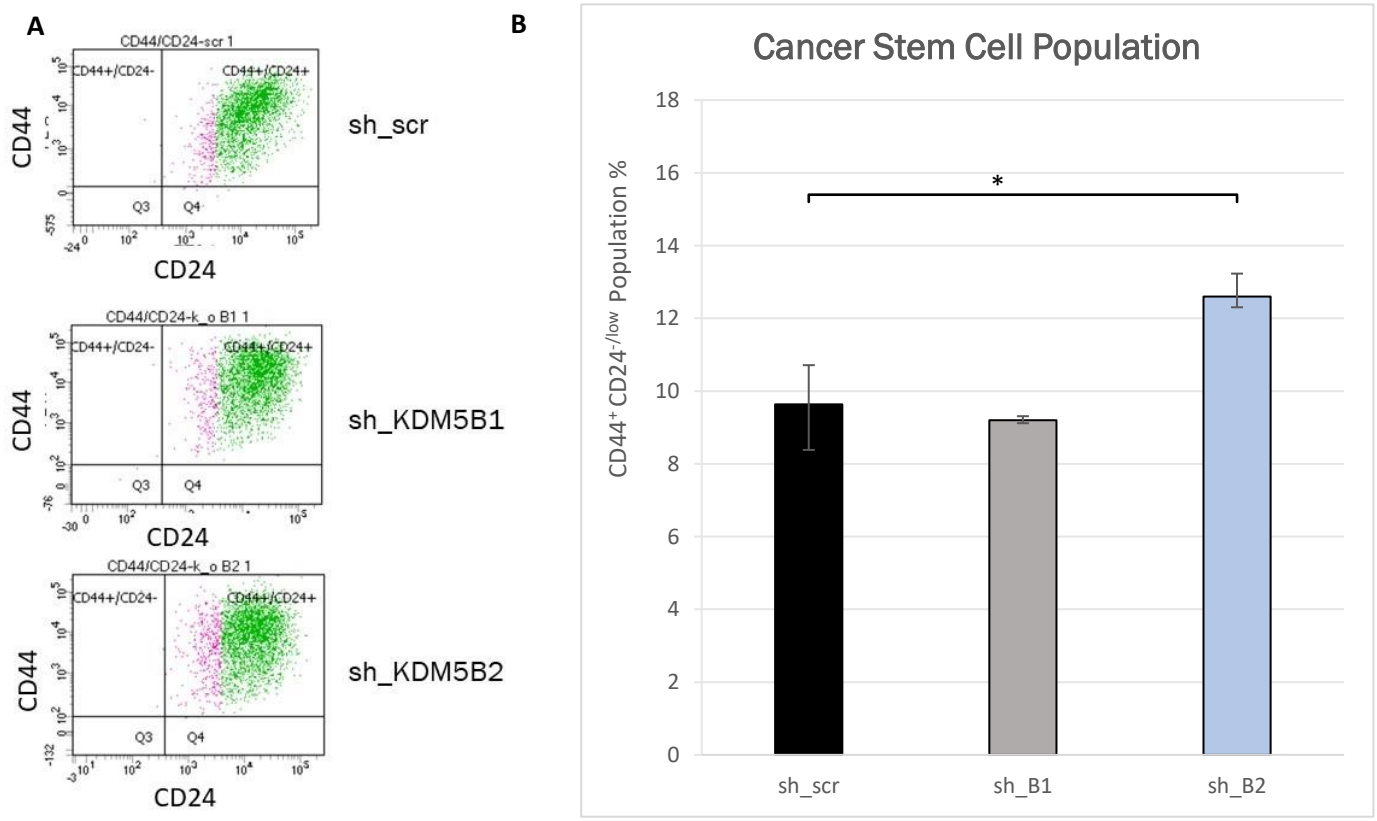


Figure 3.XVI: CSC population analysis in sh_scramble and KDM5B knock-down cell lines: A. *FACS figures displaying the difference in* $CD44^{high}/CD24^{low/-}$ *population* ***B.*** *Graphical display of FACS results. A significant increase in the CSC population is shown for sh_KDM5B2 cells, but not for sh_KDM5B1 cells. The error bars represent the SEM from 3 independent experiments. (*: denotes a p value < 0,005).*

The above data suggest that KDM5B is not implicated in the regulation of the CSC-subpopulation in MCF-7 cells and no other experiments were pursued in this direction.

3.5.3 KDM5B knock-down cells are more resistant to chemotherapeutics

Next, we sought to discover whether sh_KDM5B1 and sh_KDM5B2 cell lines would present differences regarding their sensitivity to therapy, compared to sh_scr cells. Cell growth assays after doxorubicin treatment illustrated an increased chemoresistance to the drug as shown by the IC_{50} values. Cell population growth analysis with Incucyte revealed that sh_scr cells had an IC_{50} similar to MCF-7 parental cells (Figure 3.X), while KDM5B knock-down cells were significantly more resistant to the drug, about 2-fold in sh_B2 and 3-fold in sh_B1, compared to sh_scr cell lines (p-value <0.05) (Figure 3.XVII).

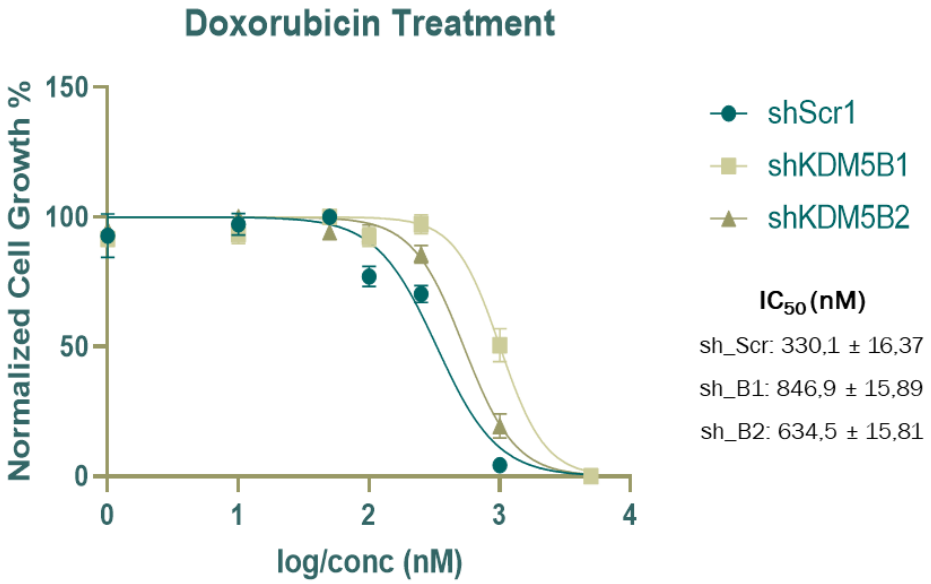


Figure 3.XVII: Cell growth assay in sh_scramble and KDM5B knock-down cell lines after doxorubicin treatment: *Cell growth was estimated after 2 days in the presence of various concentrations of Doxorubicin, for sh_scr and KDM5Bk/d cell lines, sh_B1 and sh_B2. Sh_B1 and sh_B2 cell lines are more resistant to Doxorubicin treatment as showcased by the IC_{50} values. The error bars represent the SEM from 3 independent experiments (p value < 0.05).*

A clear correlation can be drawn between KDM5B absence and increased drug resistance. The IC_{50} values closely relate to the levels of KDM5B protein present in each cell line, possibly suggesting a causal effect between the two phenomena.

4. Discussion

Cancer has been a major cause of concern in our time, for physicians and researchers alike. What was considered as a mass of uncontrollably replicating cells, tumors have now been proven to be much more sophisticated and structured than initially thought (Introduction 1.1). The basis of tumor structure is the Cancer Stem Cells (CSCs), that, similar to normal stem cells, possess self-renewal capacity, replenishing the CSC pool, and differentiation capabilities, giving rise to the sum of the heterogeneous population of cells found within a tumor. The necessity for novel treatment strategies that target this specific subpopulation of cells is already adamant, as well as the scientific pursuit of uncovering the biological mysteries governing their unique abilities (Introduction 1.2 and 1.3). Breast cancer in particular, is the leading cause of cancer-related death in women, and the most prevalent type of cancer in the world as of 2020 (WHO). Moreover, about 70% of breast cancer cases fall under the luminal phenotype, which was studied in the present thesis (Introduction 1.1).

A predominant factor of CSC biology is epigenetics. It has been well established that epigenetic dysregulation, most commonly caused by aberrant epigenetic enzymes expression, is a major contributor to CSCs' most problematic features, their stem capabilities and resistance to therapy (Introduction 1.4 and 1.5). In light of these facts, and previous work in the lab ("Isolation and characterization of Breast Cancer Stem Cells. The role of histone demethylase LSD1 in the biology of Breast Cancer Stem Cells", J. Verigos, 2019; "The Role of KDM5 Demethylases in Breast Cancer Stem Cells", I. Zikopoulou, 2020), two epigenetic enzymes already implicated in cancer biology were chosen, KDM5A and KDM5B. These enzymes are responsible for the methylation status of H3K4 within the genome, a locus tightly paired with gene activity and their overexpression in breast cancer is already shown. Researchers have revealed some of their multifaceted roles in disease and many cancer types. For KDM5A, existing data was more focused on its impact to therapy resistance and novel inhibitor development, while it was lacking in the area of luminal breast cancer, specifically in the context of CSCs. KDM5B is highly related to KDM5A and has been connected to CSC biology in other types of cancer and subtypes of breast cancer, although a clear role is still to be identified. Moreover, it appears to play a unique role in ER^+ breast cancer cells, rendering it an appealing target of study in our chosen system (Introduction 1.5).

4.1 KDM5A

With our focus to provide more insight into the aforementioned areas, we used MCF-7 (ER^+, luminal breast cancer) cells and developed a CRISPR-Cas9 system to construct two stable KDM5A knock-out cell lines, E8 & F5 (Results 3.1 and 3.2). The impact on these cells was immediately noticed during culture procedures, as the delay in their cell growth was obvious. This observation

was properly characterized, by performing a cell growth assay on KDM5A knock-out and MCF-7 wild type cells, where the slow-down in cell growth was established (Results 3.3.1, Figure 3.III). Cell growth is the net result of two opposing dynamics, cell proliferation and cell death. KDM5A has been shown to contribute directly to cell proliferation, by upregulating CyclinD1 protein levels and inhibiting cell cycle inhibitors' expression, thus promoting the cell cycle (Introduction 1.5). Our initial thought was that absence of KDM5A could cause a stall in the G1 phase. This theory was tested by a cell cycle analysis. Cell cycle analysis in the E8 and F5 cell lines did reveal a stall in these cells, but solely it did not satisfy the difference observed in cell growth (Figure 3.IV). More revealing was the second aspect impacting cell growth, cell death. A noticeable increase in dead cells (not attributed to apoptosis), compared to MCF-7 control cells, was observed in E8 and F5, up to 8-fold higher for E8 KDM5A knock-out cell line and approximately 2-fold in F5 KDM5A knock-down cells (Figure 3.V). The huge contribution of KDM5A to DNA repair mechanisms has already been proven (Introduction 1.5), so a less efficient response to stressors due to KDM5A absence, could be an explanation to the increase in cell death. In fact, the huge difference in the total increase in dead cells for E8, could explain the unexpectedly smaller stall (compared to F5 cells) observed in cell cycle analysis, as in that type of analysis dead cells are excluded. In other words, if in the total absence of KDM5A, DNA-repair mechanisms were more severely impacted, the fraction of living cells in the G1 phase would be reduced, while simultaneously, that of the dead cells would increase, possibly explaining the results of these assays.

After establishing cell growth dynamics, our work focused on the main goal of this thesis, the exploration of the role of KDM5A in breast cancer stem cells. We already know that CSCs hijack mechanisms utilized by normal stem cells and both enzymes studied are shown to contribute to such mechanisms. KDM5A specifically controls developmental pathways (Notch, Wnt, Hedgehog) and the connection with a major CSC hallmark, chemoresistance, has already been established in other cancer types and cell lines (Introduction 1.5). Our experiments targeting this population revealed a significant impact of KDM5A absence. To begin with, we utilized an already developed system of study enriched in CSCs, mammospheres. During this procedure, cells are grown in non-anchorage conditions, non-permitting for the growth of bulk tumor cells. Mammosphere Formation Efficiency (MFE) tests showed a noticeable decrease in the ability of E8 and F5 cells to form spheres, a direct correlation to the size of the CSC subpopulation in these lines (Figure 3.VII). Further phenotypic assessment, using the predominantly used $CD44^{high}/CD24^{low/-}$ markers (Introduction 1.3), not only verified, but revealed a significantly higher decrease of these cells than expected from MFE (Figure 3.VIII), approximately 96-98%. Lastly, an in vitro functional Extreme Limiting Dilution Assay (ELDA) was performed. Mammosphere formation ability for MCF-7 and E8 cells was tested for 50, 100, 500 and 1000 cells. The results further verified our observations, as KDM5A knock-out line, E8, needed at least 500 cells to form spheres, while as much as 100 MCF-7 cells were sufficient (Figure 3.IX). The quantification revealed a significant difference in the two

populations (Figure 3.IX – Chisq) with an estimated stem cell every 198 for MCF-7 cells and one in 641 cells for E8 (Figure 3.IX – estimated stem cell frequency). Based on these observations we conclude that KDM5A is an integral factor in the CSC biology of MCF-7 cells, which, to the best of our knowledge, is a novel establishment.

Perhaps the most problematic part of CSCs is their ability to escape or resist different forms of therapy. Therapy resistance is a hallmark of CSCs (Introduction 1.3) and KDM5A has been shown to contribute to it (Introduction 1.5). Thus, it was of importance to study whether the absence of KDM5A and the reduction in the CSC population would render the KDM5A knock-out cell lines more sensitive to a commonly used chemotherapeutic, doxorubicin. To this end, cells were treated with doxorubicin for 2 days and the IC_{50} for each cell line was calculated (Figure 3.X). Based on the IC_{50}, F5 cells appear to be 45% more sensitive to treatment and E8 cells 63%. Doxorubicin acts by inhibiting topoisomerase II, an enzyme integral in DNA replication during S-phase, (Materials & Methods 2.2), thus inducing replication stress. The role of KDM5A in Replication Stress Response (RSS) has been established (Introduction 1.5). It is possible that the difference in IC_{50} for doxorubicin between the cell lines is attributed to problematic RSS, decreased CSC population fraction possessing other resistant mechanisms to this drug, or a combination between the two, with KDM5A being part of a "resistant mechanism" to cytotoxic drugs for CSCs. In every case, KDM5A appears important in the survival of MCF-7 cells treated with doxorubicin and is painted as a probable candidate to target in combinational treatments.

A similar tendency, but a highly blunter effect (37% reduction in IC_{50}) was revealed when E8 KDM5A k/o cell line was tested using the chemotherapeutic agent paclitaxel. Paclitaxel is a cytoskeletal drug, known to target microtubule polymerization, thus inhibiting cells from proceeding past the mitotic (M) checkpoint and leading them to death by apoptosis (Materials 2.2). The mechanism of action is of great importance here, as this drug targets a different cell cycle checkpoint than the one KDM5A has been shown to contribute to (Introduction 1.5). This observation leads to the conclusion that, the increased sensitivity displayed in doxorubicin treatment is majorly KDM5A-dependent and only in part attributed to other therapy resistance mechanisms, as the effect was not replicated in paclitaxel treatment, which targets KDM5A independent pathways. The slight decrease displayed in the IC_{50} value of paclitaxel could be explained by the decrease in the population of CSCs observed in KDM5A k/o cell lines possessing other resistance mechanisms, or it could be a side-effect of the generally higher cell death ratio observed in E8 cells and cannot be distinguished in the assay performed here.

To summarize, our work established KDM5A as an integral factor in the biology of MCF-7 cancer cells. Cells without KDM5A present reduced proliferation dynamics and tumor-initiating potential in vitro, and they appear to be more sensitive to the chemotherapeutic agent doxorubicin, compared to parental MCF-7 cells. The cancer stem cell population is especially affected, as its fraction diminishes, when KDM5A is absent, and its stemness potential is impaired, as manifested

by the mammosphere formation assay and the ELDA. RNA-seq experiments are currently underway to elucidate the gene networks regulated by KDM5A in MCF-7 cells.

4.2 KDM5B

Existing bibliography concerning KDM5B is focused mainly on the enzyme's role as an oncogene in various types of cancer, the development of novel inhibitors targeting it, and the stem-like abilities it confers to cancer cells. However, there is an ever-growing number of articles contradicting the above statements, showcasing its tumor-suppressive roles. It is evident that KDM5B serves different purposes in different biological contexts (Introduction 1.5). Specifically, in ER^+ breast cancer, previous works established KDM5B as an important oncogene and a principal facilitator of the ERα–mediated expression program (Introduction 1.5). Here, a gene silencing shRNA approach was adopted to study the role of KDM5B in MCF-7 cells (Materials & Methods 2.13). Two stable knock-down cells lines were created, sh_KDM5B1 (B1) and sh_KDM5B2 (B2). Gene expression was verified on both the mRNA and protein level, where the B1 line showed an 80% and the B2 a 70% knock-down. Catchpole S. et al, with their work in MCF-7 cells, identified KDM5B as an important gene in these cells, co-regulating ERα expression program and stimulating ERα-mediated cell growth. Our results on cell growth verified these findings, as sh_KDM5B cell lines displayed a blunter cell growth curve compared to sh_scramble transduced cells (Figure 3.XIII), and expanded, by revealing a significant increase in the apoptotic population in the B1 cell line, that also presents the higher levels of KDM5B knock-down (Figure 3.XIV). This finding could illustrate KDM5B as an important factor in cell viability, or as contributor to the apoptotic balance scale of MCF-7 cells, an effect highlighted only when extremely low amounts of the enzyme are present. Furthermore, as cell death is an important factor in cell population kinetics, the difference in the growth curves between B1 and B2 cell lines could be explained by the difference they present in the apoptotic cell population.

While KDM5B's contribution to organism development has been shown and its connection with stem properties is known in cancer cells (e.g., melanoma cells), similar results concerning ER^+ luminal breast cancer cells are lacking. Mirroring our work with KDM5A, cells were grown in low-attachment conditions to form mammospheres. Phenotypic analysis on these spheres revealed a tendency of increased CSC population in sh_KDM5B2 cells, which was statistically significant. The same was not true for sh_KDM5B1 cells (Figure 3.XIV). In the present thesis, stable cell lines were created utilizing lentiviral transduction of MCF-7 cells. While viral vectors are highly convenient to stably transduce cells and express the construct of interest, the genomic loci in which the viral genome will integrate in the cell cannot be predetermined. It is a possibility, that the differences observed in phenotypic CSC analysis between the two clones are a side-effect of the point of integration of the virus in one of the cell lines. A genome sequencing approach would be required

to establish the points of viral genome insertion and test this theory; however, such analysis was not conducted in the framework of this thesis.

Further characterization of these cell lines included their response to chemotherapeutics, where doxorubicin was utilized. Interestingly, both sh_KDM5B cell lines displayed an increased resistance to the drug, a 2-fold increase in B2 and an a 2.5-fold in B1 cells (Figure 3.XV), following the levels of KDM5B expression. A probable scenario is that the increased IC_{50} appears due to the slower cell cycle the sh_KDM5B cells present, although a plethora of other drug resistance mechanisms could contribute.

Our results verified already known functions of KDM5B in MCF-7 cells, namely a decreased proliferation potency, while also revealing an increase in apoptotic cells when KDM5B is knocked-down. Slower cell cycle is a characteristic cancer stem cells often adopt, but we failed to identify such a relation here, as B1 cells (presenting the higher decrease in KDM5B levels) phenotypic analysis did not reveal a correlation. A significant increase was found in the ability of sh_KDM5B1 and sh_KDM5B2 cells to survive doxorubicin treatment, which could be, in part, a downstream effect of their reduced cell growth rate, as slower proliferating cells are less influenced by cytotoxic drugs. The aforementioned results could not establish a clear role for KDM5B in MCF-7 cancer stem cells. The ambiguity of our results could be a side-effect of the viral means used to transduce the cells.

4.3 Closing Thoughts and Future Perspectives

The importance of epigenetics in disease and, specifically, cancer is a constantly growing and contemporary topic of research. Simultaneously, major advancements have been made in identifying Cancer Stem Cells (CSCs) and understanding their biology. In this diploma thesis we set to contribute towards these ends, by studying the effects of epigenetic factors in breast cancer cells and breast CSCs. KDM5A and KDM5B are epigenetic enzymes that have been implicated in various aspects of cancer and are known to regulate processes related to stemness and stem characteristics in healthy and malignant cells. Furthermore, specific roles of these enzymes have been identified in breast cancer. On that aspect, the luminal breast cancer subtype, although the most prevalent, was the least explored and the effects observed in bibliography were not directly linked to CSCs or analyzed under a CSC scope.

CRISPR-Cas9 was adopted as the preferred means of protein deletion as it allows for precise genome editing and generation of stably knock-out cells. While the use of specific enzyme inhibitors is common and such experiments have already been conducted in our lab (“The role of KDM5 Demethylases in Breast Cancer Stem Cells”, I. Zikopoulou, 2020), the high relation found within the KDM5 enzymatic family, sets the specific inhibition of a single member a tenuous task, leaving questions as to whether the effects observed are attributed to a single enzyme. Moreover,

these enzymes play pivotal roles in cell biology regardless of their enzymatic activity, an aspect that cannot be explored by simple inhibition. A CRISPR-Cas9 system targeting KDM5A and KDM5B was successfully created and the generation of knock-out cell lines was attempted. MCF-7 KDM5A k/o cell clones were generated, but the process was not as successful in our attempt to produce KDM5B knock-outs. Thus, shRNA technology was used to imitate CRISPR-generated KDM5B absence and KDM5B knock-down cell lines were created.

In the present work, we managed to identify some already known roles of KDM5A and KDM5B in cancer cells and expanded that knowledge in the field of luminal breast cancer. The results on cell growth and cell cycle verified our system of study. Further, a novel reveal was the effect of KDM5A absence in the viability of MCF-7 cells in an apoptosis-independent manner, as was shown via a cell death and a cell apoptosis assay. Here, we established an integral connection between KDM5A presence and luminal bCSCs, with KDM5A regulating their phenotype and functional features like resistance to doxorubicin and tumorigenic potency which, to the best of our knowledge, has not been reported again. Our approach verified the necessity of KDM5A in luminal bCSCs. The results of the present thesis present many opportunities for further research, diving deeper into the causal effects of our observations and attempting to uncover specific biological mechanisms governing them. KDM5A overexpression experiments in MCF-7 cells, would complement existing results as to the potency of KDM5A in inducing the CSC phenotype and features. A phenotypic rescue approach could even be implemented, if the E8 KDM5A k/o cells were to be used for the overexpression. Moreover, as KDM5A acts principally as a transcriptional regulator, it would be really interesting to establish differences in the transcriptome via RNA sequencing between KDM5A k/o and parental MCF-7 cells, in an effort to pinpoint important aspects governing these features. Through this analysis and further research based on its results, the mechanism or the axis through which KDM5A exudes these effects may be elucidated. Acquired knowledge in these areas could annotate potential molecules for CSC-targeting therapy or KDM5A combinational therapy, in an effort to eradicate this population of cells. Most importantly, such a finding could potentially be applied in as much as 70% of breast cancer patients.

Our results regarding KDM5B were more ambiguous. While known effects of KDM5B inhibition were verified in our knock-down system, a correlation between KDM5B and CSCs was not established. KDM5B inhibition has been shown to decrease proliferation in cancer cells, however the aspect of increased apoptotic population percentages as a contributor to reduced cell kinetics that was established here, has not been revealed again in our bibliographical research. Another novelty in our findings was the relation between KDM5B knock-down and increased resistance to doxorubicin, as, commonly, high levels of KDM5B are linked to therapy resistance. KDM5B has been studied extensively in the CSC context in other types of cancer and has been shown to regulate stem pathways. Surprisingly, our results failed to establish a correlation between

KDM5B absence and the CSC phenotype. The $CD44^{high}/CD24^{-/low}$ phenotypic analysis results in the clone presenting lower KDM5B levels showed no difference relative to control, while an increase of the population was found on the clone with the lower KDM5B knock-down. Since these results are conflicting and are not following KDM5B expression, we are led to conclude that they are either KDM5B independent or there is a problem with our system of study. Analyzing the bibliography, our approach, and the protocols followed to generate the knock-down cell lines, the most probable scenario is that the lentiviruses used, caused unforeseen problems in one of the cell lines by integrating in an important genomic locus. For the validation of the integrity of the present results a genome sequencing of the shKDM5B stable cell lines is mandatory. Further experiments could be completed, after new attempts to generate KDM5B knock-out cell lines using CRISPR, which was our initial goal. Based on the observations on the aforementioned topics, an approach similar to KDM5A could be followed. Namely, KDM5B overexpression in MCF-7 parental or KDM5B k/o cells to establish the reversal of the cell phenotype, and RNA sequencing, in an attempt to unravel the potential biological mechanisms governing the observed results.

5. Bibliography

1. *Cristea, S., Polyak, K. Dissecting the mammary gland one cell at a time. Nat Commun **9**, 2473 (2018).*

2. Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi-Kalbolandi S, Safari E, Farahmand L*. Breast cancer: Biology, biomarkers, and treatments. Int Immunopharmacol.* 2020 Jul;84:106535.

3. Koren, S. & Bentires-Alj, M. Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. *Mol Cell* **60**, 537–546 (2015).

4. Koren, S. & Bentires-Alj, M. Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. *Molecular Cell* vol. 60 537–546 (2015).

5. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613–628 (2017).

6. Brock, A., Chang, H. & Huang, S. Non-genetic heterogeneity — a mutation-independent driving force for the somatic evolution of tumours. *Nat Rev Genet* **10**, 336–342 (2009).

7. Hinohara, K. *et al.* KDM5 Histone Demethylase Activity Links Cellular Transcriptomic Heterogeneity to Therapeutic Resistance. *Cancer Cell* **34**, 939-953.e9 (2018).

8. David, H. Rudolf Virchow and Modern Aspects of Tumor Pathology. *Pathol Res Pract* **183**, 356–364 (1988).

9. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer* **12**, 323–334 (2012).

10. Heppner, G. H. & Miller, B. E. Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer and Metastasis Reviews* **2**, 5–23 (1983).

11. Nam C, Ziman B, Sheth M, Zhao H, Lin DC. Genomic and Epigenomic Characterization of Tumor Organoid Models. Cancers (Basel). 2022 Aug 24;14(17):4090.

12. Hinohara, K. & Polyak, K. Intratumoral Heterogeneity: More Than Just Mutations. *Trends Cell Biol* **29**, 569–579 (2019).

13. Feinberg, A. P., Koldobskiy, M. A. & Göndör, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet* **17**, 284–299 (2016).

14. Jones, P. A., Issa, J.-P. J. & Baylin, S. Targeting the cancer epigenome for therapy. *Nat Rev Genet* **17**, 630–641 (2016).

15. Aguirre-Ghiso, J. A. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer* **7**, 834–846 (2007).

16. Maley, C. C. *et al.* Classifying the evolutionary and ecological features of neoplasms. *Nat Rev Cancer* **17**, 605–619 (2017).

17. Ratajczak, M. Z., Bujko, K., Mack, A., Kucia, M. & Ratajczak, J. Cancer from the perspective of stem cells and misappropriated tissue regeneration mechanisms. *Leukemia* **32**, 2519–2526 (2018).

18. Batlle, E. & Clevers, H. Cancer stem cells revisited. *Nat Med* **23**, 1124–1134 (2017).

19. Carvalho J. Cell Reversal From a Differentiated to a Stem-Like State at Cancer Initiation. Front Oncol. 2020 Apr 15;10:541.

20. Hinohara, K. *et al.* KDM5 Histone Demethylase Activity Links Cellular Transcriptomic Heterogeneity to Therapeutic Resistance. *Cancer Cell* **34**, 939-953.e9 (2018).

21. Sharma, S. v. *et al.* A Chromatin-Mediated Reversible Drug-Tolerant State in Cancer Cell Subpopulations. *Cell* **141**, 69–80 (2010).

22. Biswas, A. & De, S. Drivers of dynamic intratumor heterogeneity and phenotypic plasticity. *Am J Physiol Cell Physiol* **320**, C750–C760 (2021).

23. Dalerba, P., Cho, R. W. & Clarke, M. F. Cancer stem cells: Models and concepts. *Annu Rev Med* **58**, 267–284 (2007).

24. Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A. 2003 Apr 1;100(7):3983-8.

25. Gupta, P. B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011).

26. Charafe-Jauffret, E. *et al.* Breast cancer cell lines contain functional cancer stem sells with metastatic capacity and a distinct molecular signature. *Cancer Res* **69**, 1302–1313 (2009).

27. Akbarzadeh, M. *et al.* Current approaches in identification and isolation of cancer stem cells. *J Cell Physiol* **234**, 14759–14772 (2019).

28. Abdullah, L. & Chow, E. Chemoresistance in cancer stem cells. *Clin Transl Med* **2**, 1–9 (2013).

29. Ikawa, M., Impraim, C. C., Wang, G. & Yoshida, A. Isolation and characterization of aldehyde dehydrogenase isozymes from usual and atypical human livers. *Journal of Biological Chemistry* **258**, 6282–6287 (1983).

30. Aguirre-Ghiso, J. A. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer* **7**, 834–846 (2007).

31. Dave, B. & Chang, J. Treatment resistance in stem cells and breast cancer. *Journal of Mammary Gland Biology and Neoplasia* vol. 14 79–82

32. Woodward, W. A. *et al. WNT/-catenin mediates radiation resistance of mouse mammary Proc Natl Acad Sci U S A. 2007 Jan 9;104(2):618-23.*

33. Creighton, C. J. *et al. Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features*. Proc Natl Acad Sci U S A. 2009 Aug 18;106(33):13820-5.

34. Engelmann, K., Shen, H. & Finn, O. J. MCF7 side population cells with characteristics of cancer stem/progenitor cells express the tumor antigen MUC1. *Cancer Res* **68**, 2419–2426 (2008).

35. Wang, S. & El-Deiry, W. S. TRAIL and apoptosis induction by TNF-family death receptors. *Oncogene* vol. 22 8628–8633, (2003).

36. Frosina, G. DNA repair and resistance of gliomas to chemotherapy and radiotherapy. *Molecular Cancer Research* vol. 7 989–999

37. Huelsken J, Vogel R, Brinkmann V, Erdmann B, Birchmeier C, Birchmeier W. Requirement for beta-catenin in anterior-posterior axis formation in mice. J Cell Biol. 2000 Feb 7;148(3):567-78.

38. Soriano, J. v., Uyttendaele, H., Kitajewski, J. & Montesano, R. Expression of an activated Notch4(int-3) oncoprotein disrupts morphogenesis and induces an invasive phenotype in mammary epithelial cells in vitro. *Int J Cancer* **86**, 652–659 (2000).

39. Vorechovský I, Benediktsson KP, Toftgård R. The patched/hedgehog/smoothened signalling pathway in human breast cancer: no evidence for H133Y SHH, PTCH and SMO mutations. Eur J Cancer. 1999 May;35(5):711-3.

40. Hoffmeyer, K. *et al.* Wnt/β-Catenin Signaling Regulates Telomerase in Stem Cells and Cancer Cells. *Science (1979)* **336**, 1549–1554 (2012).

41. Myant, K. B. *et al.* ROS production and NF-κB activation triggered by RAC1 facilitate WNT-driven intestinal stem cell proliferation and colorectal cancer initiation. *Cell Stem Cell* **12**, 761–773 (2013).

42. Andersson, E. R. & Lendahl, U. Therapeutic modulation of Notch signalling — are we there yet? *Nat Rev Drug Discov* **13**, 357–378 (2014).

43. Andersson, E. R., Sandberg, R. & Lendahl, U. Notch signaling: Simplicity in design, versatility in function. *Development* vol. 138 3593–3612 (2011).

44. Liu, J., Sato, C., Cerletti, M. & Wagers, A. Chapter Twelve - Notch Signaling in the Regulation of Stem Cell Self-Renewal and Differentiation. in *Current Topics in Developmental Biology* vol. 92 367–409 (2010).

45. Androutsellis-Theotokis, A. *et al.* Notch signalling regulates stem cell numbers in vitro and in vivo. *Nature* **442**, 823–826 (2006).

46. D'Angelo, R. C. *et al.* Notch reporter activity in breast cancer cell lines identifies a subset of cells with stem cell activity. *Mol Cancer Ther* **14**, 779–787 (2015).

47. Jin, L., Vu, T., Yuan, G. & Datta, P. K. STRAP promotes stemness of human colorectal cancer via epigenetic regulation of the NOTCH pathways. *Cancer Res* **77**, 5464–5478 (2017).

48. Wang, Z. *et al.* Notch signaling drives stemness and tumorigenicity of esophageal adenocarcinoma. *Cancer Res* **74**, 6364–6374 (2014).

49. Ingham, P. W. & McMahon, A. P. Hedgehog signaling in animal development: Paradigms and principles. *Genes and Development* vol. 15 3059–3087 (2001).

50. Ma, J. *et al.* Mammalian target of rapamycin regulates murine and human cell differentiation through STAT3/p63/Jagged/Notch cascade. *Journal of Clinical Investigation* **120**, 103–114 (2010).

51. Beachy, P. A., Karhadkar, S. S. & Berman, D. M. Tissue repair and stem cell renewal in carcinogenesis. *Nature* **432**, 324–331 (2004).

52. Teglund, S. & Toftgård, R. Hedgehog beyond medulloblastoma and basal cell carcinoma. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **1805**, 181–208 (2010).

53. Sellheyer, K. Basal cell carcinoma: cell of origin, cancer stem cell hypothesis and stem cell markers. *British Journal of Dermatology* **164**, 696–711 (2011).

54. Sell, S. On the stem cell origin of cancer. *American Journal of Pathology* **176**, 2584–2594 (2010).

55. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet* **20**, 207–220 (2019).

56. Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432 (2007).

57. Jones, P. A. & Laird, P. W. Cancer-epigenetics comes of age. *Nat Genet* **21**, 163–167 (1999).

58. Feinberg, A. P., Ohlsson, R. & Henikoff, S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* **7**, 21–33 (2006).

59. Jones, P. A. & Baylin, S. B. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* **3**, 415–428 (2002).

60. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* vol. 128 693–705 (2007).

61. Toh, T. B., Lim, J. J. & Chow, E. K. H. Epigenetics in cancer stem cells. *Mol Cancer* **16**, 1–20 (2017).

62. Papp, B. & Plath, K. Reprogramming to pluripotency: Stepwise resetting of the epigenetic landscape. *Cell Research* vol. 21 486–501 (2011).

63. Shukla, S. & Meeran, S. M. Epigenetics of cancer stem cells: Pathways and therapeutics. *Biochimica et Biophysica Acta - General Subjects* vol. 1840 3494–3502 (2014).

64. Muñoz, P., Iliou, M. S. & Esteller, M. Epigenetic alterations involved in cancer stem cell reprogramming. *Molecular Oncology* vol. 6 620–636 (2012).

65. Holik, A. Z. *et al.* Brg1 Loss Attenuates Aberrant Wnt-Signalling and Prevents Wnt-Dependent Tumourigenesis in the Murine Small Intestine. *PLoS Genet* **10**, (2014).

66. Lee, R. S. & Roberts, C. W. M. Linking the SWI/SNF complex to prostate cancer. *Nat Genet* **45**, 1268–1269 (2013).

67. Crea, F. *et al.* Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. *Mol Cancer* **10**, (2011).

68. Wang, J. *et al.* Novel histone demethylase LSD1 inhibitors selectively target cancer cells with pluripotent stem cell properties. *Cancer Res* **71**, 7238–7249 (2011).

69. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999 Oct 29;99(3):247-57.

70. Doi, A. *et al.* Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* **41**, 1350–1353 (2009).

71. van Vlerken, L. E., Hurt, E. M. & Hollingsworth, R. E. The role of epigenetic regulation in stem cell and cancer biology. *J Mol Med* **90**, 791–801 (2012).

72. Jin, B. *et al.* Linking DNA Methyltransferases to Epigenetic Marks and Nucleosome Structure Genome-wide in Human Tumor Cells. *Cell Rep* **2**, 1411–1424 (2012).

73. Bhutani, N. *et al.* Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042–1047 (2010).

74. Wang, P. *et al.* 'TET-on' pluripotency. *Cell Res* **23**, 863–865 (2013).

75. Esteller, M. Epigenetic lesions causing genetic lesions in human cancer: promoter hypermethylation of DNA repair genes. *Eur J Cancer* **36**, 2294–2300 (2000).

76. Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome-biological and translational implications. *Nature Reviews Cancer* vol. 11 726–734 (2011).

77. Heerboth, S. *et al.* EMT and tumor metastasis. *Clin Transl Med* **4**, (2015).

78. Thiery, J. P., Acloque, H., Huang, R. Y. J. & Nieto, M. A. Epithelial-Mesenchymal Transitions in Development and Disease. *Cell* vol. 139 871–890 (2009).

79. Morel, A. P. *et al.* Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS One* **3**, (2008).

80. Mani, S. A. *et al.* The Epithelial-Mesenchymal Transition Generates Cells with Properties of Stem Cells. *Cell* **133**, 704–715 (2008).

81. Shah, A. N. *et al.* Development and characterization of gemcitabine-resistant pancreatic tumor cells. *Ann Surg Oncol* **14**, 3629–3637 (2007).

82. Arumugam, T. *et al.* Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. *Cancer Res* **69**, 5820–5828 (2009).

83. Wu, C.-P., Calcagno, A. M. & Ambudkar, S. v. *Reversal of ABC drug transporter-mediated multidrug resistance in cancer cells: Evaluation of current strategies Curr Mol Pharmacol. 2008 Jun;1(2):93-105*.

84. Stallcup, M. R. *Role of protein methylation in chromatin remodeling and transcriptional regulation*. *Oncogene* **20**, 3014–3020 (2001).

85. Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science (1979)* **293**, 1074–1080 (2001).

86. Song Y, Wu F, Wu J. Targeting histone methylation for cancer therapy: enzymes, inhibitors, biological activity and perspectives. J Hematol Oncol. 2016 Jun 17;9(1):49

87. Finley, A. & Copeland, R. A. Small molecule control of chromatin remodeling. *Chemistry and Biology* vol. 21 1196–1210 (2014).

88. Popovic, R. & Licht, J. D. Emerging epigenetic targets and therapies in cancer medicine. *Cancer Discovery* vol. 2 405–413 (2012).

89. Copeland, R. A., Solomon, M. E. & Richon, V. M. Protein methyltransferases as a target class for drug discovery. *Nat Rev Drug Discov* **8**, 724–732 (2009).

90. Bhaumik, S. R., Smith, E. & Shilatifard, A. Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* **14**, 1008–1016 (2007).

91. Steger, D. J. *et al.* DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. *Mol Cell Biol* **28**, 2825–2839 (2008).

92. Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* **40**, 897–903 (2008).

93. Mueller, D. *et al.* A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. (2007) Blood. 2007 Dec 15;110(13):4445-54.

94. Krivtsov, A. v. *et al.* H3K79 Methylation Profiles Define Murine and Human MLL-AF4 Leukemias. *Cancer Cell* **14**, 355–368 (2008).

95. Bitoun, E., Oliver, P. L. & Davies, K. E. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum Mol Genet* **16**, 92–106 (2007).

96. Zhang, W., Xia, X., Reisenauer, M. R., Hemenway, C. S. & Kone, B. C. Dot1a-AF9 complex mediates histone H3 Lys-79 hypermethylation and repression of ENaCα in an aldosterone-sensitive manner. *Journal of Biological Chemistry* **281**, 18059–18068 (2006).

97. Cao, R. *et al.* Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science (1979)* **298**, 1039–1043 (2002).

98. Müller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell. 2002 Oct 18;111(2):197-208.

99. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the enhancer of zeste protein. *Genes Dev* **16**, 2893–2905 (2002).

100. Wang, X., Zhao, H., Lv, L. *et al.* Prognostic Significance of EZH2 Expression in Non-Small Cell Lung Cancer: A Meta-analysis. *Sci Rep* **6**, 19239 (2016).

101. Kleer, C. G. *et al. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells*. Proc Natl Acad Sci U S A. 2003 Sep 30;100(20):11606-11.

102. Simon, J. A. & Lange, C. A. Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **647**, 21–29 (2008).

103. Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).

104. Li Z, Wang Y, Qiu J, Li Q, Yuan C, Zhang W, Wang D, Ye J, Jiang H, Yang J, Cheng J. The polycomb group protein EZH2 is a novel therapeutic target in tongue cancer. Oncotarget. 2013 Dec;4(12):2532-49.

105. Shi, J. *et al.* The Polycomb complex PRC2 supports aberrant self-renewal in a mouse model of MLL-AF9;NrasG12D acute myeloid leukemia. *Oncogene* **32**, 930–938 (2013).

106. Suvà, M. L. *et al.* EZH2 is essential for glioblastoma cancer stem cell maintenance. *Cancer Res* **69**, 9211–9218 (2009).

107. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H. & Helin, K. Genome-wide mapping of polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* **20**, 1123–1136 (2006).

108. 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 (2002).

109. Kondo Y, Shen L, Ahmed S, Boumber Y, Sekido Y, Haddad BR, Issa JP. Downregulation of histone H3 lysine 9 methyltransferase G9a induces centrosome disruption and chromosome instability in cancer cells. PLoS One. 2008 Apr 30;3(4):e2037.

110. Kondo, Y. *et al.* Alterations of DNA methylation and histone modifications contribute to gene silencing in hepatocellular carcinomas. *Hepatology Research* **37**, 974–983 (2007).

111. Chen, M. W. *et al.* H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. *Cancer Res* **70**, 7830–7840 (2010).

112. 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-IN10 (2011).

113. Hua, K. T. *et al.* The H3K9 methyltransferase G9a is a marker of aggressive ovarian cancer that promotes peritoneal metastasis. *Mol Cancer* **13**, (2014).

114. Bai, K. *et al.* Association of histone methyltransferase G9a and overall survival after liver resection of patients with hepatocellular carcinoma with a median observation of 40 months. *Medicine (United States)* **95**, (2016).

115. Højfeldt, J. W., Agger, K. & Helin, K. Histone lysine demethylases as targets for anticancer therapy. *Nat Rev Drug Discov* **12**, 917–930 (2013).

116. Dimitrova, E., Turberfield, A. H. & Klose, R. J. Histone demethylases in chromatin biology and beyond. *EMBO Rep* **16**, 1620–1639 (2015).

117. Arrowsmith, C. H., Bountra, C., Fish, P. v., Lee, K. & Schapira, M. Epigenetic protein families: A new frontier for drug discovery. *Nature Reviews Drug Discovery* vol. 11 384–400 (2012).

118. Klose, R. J., Kallin, E. M. & Zhang, Y. JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* **7**, 715–727 (2006).

119. Harris, W. J. *et al.* The Histone Demethylase KDM1A Sustains the Oncogenic Potential of MLL-AF9 Leukemia Stem Cells. *Cancer Cell* **21**, 473–487 (2012).

120. Konovalov S, Garcia-Bassets I. Analysis of the levels of lysine-specific demethylase 1 (LSD1) mRNA in human ovarian tumors and the effects of chemical LSD1 inhibitors in ovarian cancer cell lines. J Ovarian Res. 2013 Oct 29;6(1):75.

121. Sprüssel, A. *et al.* Lysine-specific demethylase 1 restricts hematopoietic progenitor proliferation and is essential for terminal differentiation. *Leukemia* **26**, 2039–2051 (2012).

122. Verigos, J. *et al.* The histone demethylase LSD1/KDM1A mediates chemoresistance in breast cancer via regulation of a stem cell program. *Cancers (Basel)* **11**, (2019).

123. Lan, F. *et al.* A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689–694 (2007).

124. Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731–734 (2007).

125. Lee, M. G. *et al.* Demethylation of H3K27 Regulates Polycomb Recruitment and H2A Ubiquitination. *Science (1979)* **318**, 447–450 (2007).

126. Metzger, E. *et al.* Phosphorylation of histone H3T6 by PKCB i controls demethylation at histone H3K4. *Nature* **464**, 792–796 (2010).

127. Kristensen, L. H. *et al.* Studies of H3K4me3 demethylation by KDM5B/Jarid1B/PLU1 reveals strong substrate recognition in vitro and identifies 2,4-pyridine-dicarboxylic acid as an in vitro and in cell inhibitor. *FEBS Journal* **279**, 1905–1914 (2012).

128. Yamane, K. *et al.* PLU-1 Is an H3K4 Demethylase Involved in Transcriptional Repression and Breast Cancer Cell Proliferation. *Mol Cell* **25**, 801–812 (2007).

129. Johansson, C. *et al.* Structural analysis of human KDM5B guides histone demethylase inhibitor development. *Nat Chem Biol* **12**, 539–545 (2016).

130. Pavlenko E, Ruengeler T, Engel P, Poepsel S. Functions and Interactions of Mammalian KDM5 Demethylases. Front Genet. 2022 Jul 11;13:906662.

131. Christensen, J. *et al.* RBP2 Belongs to a Family of Demethylases, Specific for Tri-and Dimethylated Lysine 4 on Histone 3. *Cell* **128**, 1063–1076 (2007).

132. Dey, B. K. *et al.* The Histone Demethylase KDM5b/JARID1b Plays a Role in Cell Fate Decisions by Blocking Terminal Differentiation. *Mol Cell Biol* **28**, 5312–5327 (2008).

133. Heintzman, N. D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* **39**, 311–318 (2007).

134. Klein, B. J. *et al.* The histone-H3K4-specific demethylase KDM5B Binds to its substrate and product through distinct PHD fingers. *Cell Rep* **6**, 325–335 (2014).

135. Torres, I. O. & Fujimori, D. G. Functional coupling between writers, erasers and readers of histone and DNA methylation. *Current Opinion in Structural Biology* vol. 35 68–75 (2015).

136. Zhang, Y. *et al.* The PHD1 finger of KDM5B recognizes unmodified H3K4 during the demethylation of histone H3K4me2/3 by KDM5B. *Protein Cell* **5**, 837–850 (2014).

137. Klose, R. J. *et al.* The Retinoblastoma Binding Protein RBP2 Is an H3K4 Demethylase. *Cell* **128**, 889–900 (2007).

138. Defeo-Jones, D. *et al.* Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product. *Nature* **352**, 251–254 (1991).

139. Oser, M. G. *et al.* The KDM5A/RBP2 histone demethylase represses NOTCH signaling to sustain neuroendocrine differentiation and promote small cell lung cancer tumorigenesis. (2019) Genes Dev. 2019 Dec 1;33(23-24):1718-1738.

140. Váraljai, R. *et al.* Increased mitochondrial function downstream from KDM5a histone demethylase rescues differentiation in pRB-deficient cells. *Genes Dev* **29**, 1817–1834 (2015).

141. Brier, A. S. B. *et al.* The KDM5 family is required for activation of pro-proliferative cell cycle genes during adipocyte differentiation. *Nucleic Acids Res* **45**, 1743–1759 (2017).

142. Benevolenskaya, E. v., Murray, H. L., Branton, P., Young, R. A. & Kaelin, W. G. Binding of pRB to the PHD protein RBP2 promotes cellular differentiation. *Mol Cell* **18**, 623–635 (2005).

143. Zhao, L. H. & Liu, H. G. Immunohistochemical detection and clinicopathological significance of JARID1B/ KDM5B and p16 expression in invasive ductal carcinoma of the breast. *Genetics and Molecular Research* **14**, 5417–5426 (2015).

144. Krishnakumar, R. & Kraus, W. L. PARP-1 Regulates Chromatin Structure and Transcription through a KDM5B-Dependent Pathway. *Mol Cell* **39**, 736–749 (2010).

145. Catchpole, S. *et al.* PLU-1/JARID1B/KDM5B is required for embryonic survival and contributes to cell proliferation in the mammary gland and in ER+ breast cancer cells. *Int J Oncol* **38**, 1267–1277 (2011).

146. Vicent, G. P. *et al.* Unliganded progesterone receptormediated targeting of an RNA-containing repressive complex silences a subset of hormone-inducible genes. *Genes Dev* **27**, 1179–1197 (2013).

147. Albert, M. *et al.* The Histone Demethylase Jarid1b Ensures Faithful Mouse Development by Protecting Developmental Genes from Aberrant H3K4me3. *PLoS Genet* **9**, (2013).

148. Xie, L. *et al.* KDM5B regulates embryonic stem cell self-renewal and represses cryptic intragenic transcription. *EMBO Journal* **30**, 1473–1484 (2011).

149. Kidder, B. L., Hu, G. & Zhao, K. *KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation*. *Genome Biol* **15**, R32 (2014).

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

151. Allen, H. F., Wade, P. A. & Kutateladze, T. G. The NuRD architecture. *Cellular and Molecular Life Sciences* vol. 70 3513–3524 (2013).

152. Secombe, J., Li, L., Carlos, L. & Eisenman, R. N. The Trithorax group protein Lid is a trimethyl histone H3K4 demethylase required for dMyc-induced cell growth. *Genes Dev* **21**, 537–551 (2007).

153. Yang, G. J. *et al.* The emerging role of KDM5A in human cancer. *J Hematol Oncol* **14**, 1–18 (2021).

154. Plch, J., Hrabeta, J. & Eckschlager, T. KDM5 demethylases and their role in cancer cell chemoresistance. *Int J Cancer* **144**, 221–231 (2019).

155. Kirtana R, Manna S, Patra SK. Molecular mechanisms of KDM5A in cellular functions: Facets during development and disease. Exp Cell Res. 2020 Nov 15;396(2):112314

156. Paroni, G. *et al.* HER2-positive breast-cancer cell lines are sensitive to KDM5 inhibition: definition of a gene-expression model for the selection of sensitive cases. *Oncogene* **38**, 2675–2689 (2019).

157. Akinyi, F. gland The role of the histone demethylase KDM5B, in the normal and malignant mammary gland, e-thesis, King's College London, 2018.

158. Zeng, J. *et al.* The Histone Demethylase RBP2 Is Overexpressed in Gastric Cancer and Its Inhibition Triggers Senescence of Cancer Cells. *Gastroenterology* **138**, 981–992 (2010).

159. Yamamoto, S. *et al.* JARID1B is a luminal lineage-driving oncogene in breast cancer. *Cancer Cell* **25**, 762–777 (2014).

160. Vinogradova, M. *et al.* An inhibitor of KDM5 demethylases reduces survival of drug-tolerant cancer cells. *Nat Chem Biol* **12**, 531–538 (2016).

161. Choi HJ, Joo HS, Won HY, Min KW, Kim HY, Son T, Oh YH, Lee JY, Kong G. Role of RBP2-Induced ER and IGF1R-ErbB Signaling in Tamoxifen Resistance in Breast Cancer. J Natl Cancer Inst. 2018 Apr 1;110(4).

162. Mitra, D., Das, P. M., Huynh, F. C. & Jones, F. E. Jumonji/ARID1 B (JARID1B) protein promotes breast tumor cell cycle progression through epigenetic repression of MicroRNA let-7e. *Journal of Biological Chemistry* **286**, 40531–40535 (2011).

163. Teng, Y. C. *et al.* Histone demethylase RBP2 promotes lung tumorigenesis and cancer metastasis. *Cancer Res* **73**, 4711–4721 (2013).

164. Zhou D, Kannappan V, Chen X, Li J, Leng X, Zhang J, Xuan S. RBP2 induces stem-like cancer cells by promoting EMT and is a prognostic marker for renal cell carcinoma. Exp Mol Med. 2016 Jun 10;48(6):e238.

165. Gale M, Sayegh J, Cao J, Norcia M, Gareiss P, Hoyer D, Merkel JS, Yan Q. Screen-identified selective inhibitor of lysine demethylase 5A blocks cancer cell growth and drug resistance. Oncotarget. 2016 Jun 28;7(26):39931-39944.

166. Seiler, D. M. *et al.* Double-strand break-induced transcriptional silencing is associated with loss of tri-methylation at H3K4. *Chromosome Research* **19**, 883–899 (2011).

167. Gong, F. & Miller, K. M. Double duty: ZMYND8 in the DNA damage response and cancer. *Cell Cycle* **17**, 414–420 (2018).

168. Gaillard S, Charasson V, Ribeyre C, Salifou K, Pillaire MJ, Hoffmann JS, Constantinou A, Trouche D, Vandromme M. KDM5A and KDM5B histone-demethylases contribute to HU-induced replication stress response and tolerance. Biol Open. 2021 May 15;10(5):bio057729.

169. Xue, S. *et al.* Histone lysine demethylase KDM5B maintains chronic myeloid leukemia via multiple epigenetic actions. *Exp Hematol* **82**, 53–65 (2020).

170. Roesch, A. *et al.* A Temporarily Distinct Subpopulation of Slow-Cycling Melanoma Cells Is Required for Continuous Tumor Growth. *Cell* **141**, 583–594 (2010).

171. Facompre, N. D. *et al.* JARID1B enables transit between distinct states of the stem-like cell population in oral cancers. *Cancer Res* **76**, 5538–5549 (2016).

172. Li, Q. *et al.* Binding of the JmjC demethylase JARID1B to LSD1/NuRD suppresses angiogenesis and metastasis in breast cancer cells by repressing chemokine CCL14. *Cancer Res* **71**, 6899–6908 (2011).

173. Li, X. *et al.* Histone demethylase KDM5B is a key regulator of genome stability. *Proc Natl Acad Sci U S A* **111**, 7096–7101 (2014).

174. Royal Swedish Academy of Sciences, T. *Nobel Prize*. (2020).

175. Mojica, F. J. M., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Molecular Microbiology* vol. 36 244–246 (2000).

176. Jansen, R., van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565–1575 (2002).

177. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol* **60**, 174–182 (2005).

178. Bolotin, A., Quinquis, B., Sorokin, A. & Dusko Ehrlich, S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (N Y)* **151**, 2551–2561 (2005).

179. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (1979)* **315**, 1709–1712 (2007).

180. Jinek, M. *et al.* Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science (1979)* **343**, 1247997 (2014).

181. Jiang, F., Zhou, K., Ma, L., Gressel, S. & Doudna, J. A. A Cas9–guide RNA complex preorganized for target DNA recognition. *Science (1979)* **348**, 1477–1481 (2015).

182. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J Bacteriol* **190**, 1390–1400 (2008).

183. Wang, J. *et al.* Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. *Cell* **163**, 840–853 (2015).

184. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).

185. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* vol. 157 1262–1278 (2014).

186. Hu, Y. & Smyth, G. K. ELDA: Extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods* **347**, 70–78 (2009).