NORTHWESTERN UNIVERSITY

Interaction between DNA methylation and progesterone action during uterine leiomyoma stem cell

development

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULLFILMENT OF THE REQUIREMENT

For the degree

DOCTOR OF PHILOSOPHY

Driskill Graduate Program in Life Science

By

Shimeng Liu

EVANSTON, ILLINOIS

June 2020

**ABSTRACT**

**Interaction between DNA methylation and progesterone action during uterine leiomyoma stem cell development**

Shimeng Liu

Uterine leiomyoma (LM), the most common tumor of women, causes severe morbidity. LM cells can be separated into three molecularly and functionally distinct cell populations based on the expression pattern of CD34 and CD49b: stem (LSC, CD34+/CD49b+), intermediate (LIC, CD34+/CD49b-), and differentiated cells (LDC, CD34-/CD49b-). Progesterone via progesterone receptor (PR/PGR) expressed in differentiated LM cells stimulate paracrine signaling that targets PR-deficient stem cells to induce their proliferation, leading to tumor growth. Antiprogestins shrink LM but tumors regrow after treatment cessation possibly due to persisting stem cells. Thus, the molecular mechanism underlying LSC regulation is critical for understanding the LM tumorigenesis process and identifying potential novel therapeutic avenue for LM patients.

In this study, we analyzed the transcriptome landscape of LM cells at the three differentiation stages. The Receptor Activator of NF-κB Ligand (RANKL) was identified as one of the most significant differentially expressed genes among the three LM populations. As a novel progesterone/PR-responsive gene, we discovered that RANKL plays an important role in mediating the progesterone signals to the stem cell population via paracrine action through its receptor RANK. Upon progesterone stimulation, RANKL specifically upregulates LM stem cell proliferation through activation of Cyclin D1. Within the LM and normal myometrium (MM) system, RANKL expression was governed by the crosstalk among steroid hormone, genetic, and epigenetic signals. RANKL gene transcription was robustly induced by the progesterone agonist R5020, leading to a dramatically higher RANKL expression in LM compared to adjacent MM tissue. MethylCap-Seq revealed a differentially methylated region (DMR) adjacent to the distal PR-binding site (PRBS) 87kb upstream of the RANKL transcription start site. Hypermethylation of the DMR inhibited the recruitment of PR to the adjacent PRBS. Luciferase assays indicated that the DMR and distal PRBS

constitute a novel RANKL distal regulatory element that actively regulates RANKL expression. Furthermore, we discovered that PR physically interacts with MED12 protein, which is mutated in 70% LM tissue. The interaction between MED12 and PR, binding of PR and MED12 to PRBS, and RANKL gene expression are significantly higher in LM containing a distinct MED12 mutation (G44D) than in LM with wild-type MED12. These findings suggest that DNA methylation and MED12 mutation together constitute a complex regulatory network that affects progesterone/PR-mediated RANKL gene expression, with an important role in activating stem cell proliferation and fibroid tumor development.

Due to the importance of the crosstalk between DNA methylation and progesterone signaling on the *RANKL* gene regulation, we further investigated this interaction on a genome-wide scale by profiling DNA methylome in each LM population and progesterone receptor cistrome in LM tissues. Via vigorous bioinformatic analysis of PR ChIP-seq and DNA methylation data, we found that PR target genes were highly involved in the differentiation process and PR-deficient LM stem cells harbored a unique DNA methylation landscape with hypermethylation at the *PGR* gene locus and its genome-wide target regions, thereby suppressing the stem cell population's direct response to progesterone and maintaining its stemness. The DNA methylation inhibitor 5′-Aza upregulated the expression of PR, stimulated PR binding to its target genes, significantly depleted the LM stem cell population via accelerating its differentiation, and extensively reduced the tumor-initiating capacity of LM primary cells. Furthermore, PGR upregulated the expression of TET methylcytosine dioxygenases, suggesting a potential feedback loop between DNA methylation and hormone signaling to facilitate stem cell differentiation. Targeting stem cells using 5′-Aza also sensitized LM tissue to antiprogestins, which may lead to tumor eradication. This study provides new mechanistic insights into the stem cell differentiation process of hormone-sensitive tumors and may serve as an efficient method to identify potential treatment targeting the stem cell population.

**ACKNOWLEDGMENTS**

With the deepest gratitude, I would like to thank all the people who made this work possible and provided the intellectual and emotional support throughout my graduate training.

I would like to thank all the past and current members of the Bulun lab during my time here. In particular, I will be forever grateful to my mentors Serdar E. Bulun and Ping Yin, who have always believed in me, supported me, and guided me throughout these years. Thanks to Christia Sison, Hong Zhao, Sule Yildiz, Bahar Yilmaz, Tanvi Potluri, Ariel Dotts, Stacy A Kujawa, John S Coon V, Matthew Taylor, Amanda Kohlmeier, and Kaoru Miyazaki for their friendship, kindness, and willingness to provide scientific as well as personal advice. I am also grateful to the members of my thesis committee, Jindan Yu, Ali Shilatifard, and Christopher Payne for their helpful suggestions and collaborative efforts throughout my graduate training. I also wish to thank other members of the Division of Reproductive Science in Medicine: Julie Kim, Yong Wan, Daniela Matei, Debabrata Chakravarti, and Teresa Woodruff for their generous help and guidance.

I am fortunate to not only earn a Ph.D. while in graduate school but also to have gained so many lifelong friends along the way. Thanks to Yizhen Zhong, Yu Deng, Christia Sison, Sophy Gao, Yaqi Zhang, Guangyuan Zhao, Yinu Wang, Jia Xie, Susan Park, Ashley Haluck-Kangas, Mark Takesuye, Hyewon Kong, Galina Gritsina, Maxwell Edmonds, and Karl Guo for being a great source of support and fun. Graduate school becomes so much more fun with all these people around.

I am especially grateful to my parents, Hanxing Liu and Huiping Huang. They are the best parents that one can ever ask for. Thank you to them for their unconditional love and support, for leading me into the scientific world, and for allowing me to pursue my dream.

Most of all, I am grateful to my husband Haoming Liu. He supported me through every difficult challenge I have encountered during my graduate training. I could not have completed this work without his love and companionship.

**TABLE OF CONTENTS**

**LIST OF TABLES AND FIGURES**

**Figures**

**Tables**

**LIST OF COMMONLY USED ABBREVIATIONS**

| | |
|---|---|
| LM | Uterine leiomyoma |
| MM | Normal myometrium |
| LSC | Leiomyoma stem cells |
| LIC | Leiomyoma intermediate cells |
| LDC | Leiomyoma differentiated cells |
| PR | Progesterone receptor |
| *PGR* | Progesterone receptor gene |
| ESR1 | Estrogen receptor alpha |
| RANKL | The receptor activator of NF-κB ligand |
| TBP | Tata-binding protein |
| DNMT | DNA methyltransferases |
| TET | Ten-eleven translocation |
| MED12 | Mutation of mediator complex subunit 12 |
| GAPDH | Glyceraldehyde-3-phosphate dehydrogenase |
| mRNA | Messenger RNA |
| DMR | Differential methylated regions |
| DMG | Differential methylated genes |
| DEG | Differential expressed genes |
| PRBS | Progesterone receptor binding site |
| E2 | Estrogen |
| P4 | Progesterone |
| 5'-Aza | 5-aza-2'-deoxycytidine |
| HBSS | Hank's balanced salt solution |
| PBS | Phosphate buffered saline |
| DMEM | Dulbecco's Modified Eagle Medium |

| qPCR | Quantitative polymerase chain reaction |
|------|----------------------------------------|
| ChIP | Chromatin Immunoprecipitation |

# CHAPTER 1    Introduction

## 1.1    Introduction of uterine leiomyoma

Uterine leiomyomas (LM), also known as fibroids,  are benign smooth muscle tumors originating from the myometrium of the uterus [1]. LM is the most common tumor in women, affecting approximately 80% of women by the age of 50 years [2, 3]. LM can cause significant morbidity, including cause excessive uterine bleeding, anemia, labor obstruction, pelvic discomfort, and urinary incontinence, and may delay the diagnosis of ovarian cancer [3]. African American women develop larger LM at a higher rate and at earlier ages than Caucasian women and develop more severe symptoms [4, 5] (Figure 1-1). To remove or destroy LM, approximately 250,000 hysterectomies or myomectomies, and many less invasive but morbid procedures are performed in the US annually, costing up to $34.4 billion [6]. No new pharmaceutical has been approved in the US for LM treatment since the 1990s, and that treatment, a GnRH agonist, has limited efficacy and significant side effects. There is an urgent need for new medical options to treat LM [7].
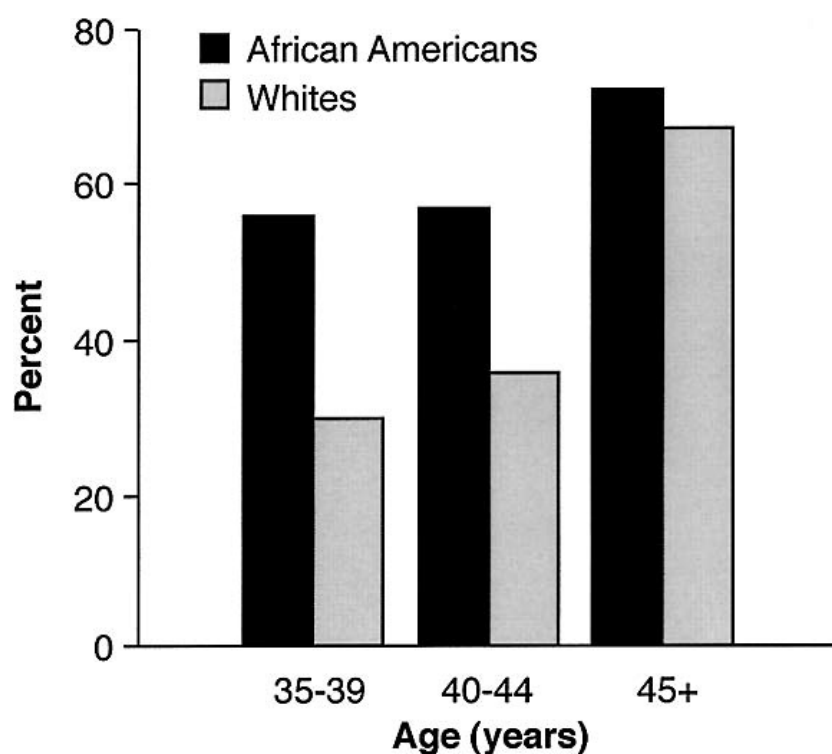


**Figure 1-1 Age difference of LM onset between Africa American and Caucasian women**

The proportion with newly detected LM among women with no previous diagnosis of leiomyoma, by age group. The *black bar* represents black women; the *gray bar* represents white women. (Adapted from ref. [4])

### 1.1.1 Uterine leiomyoma subtypes
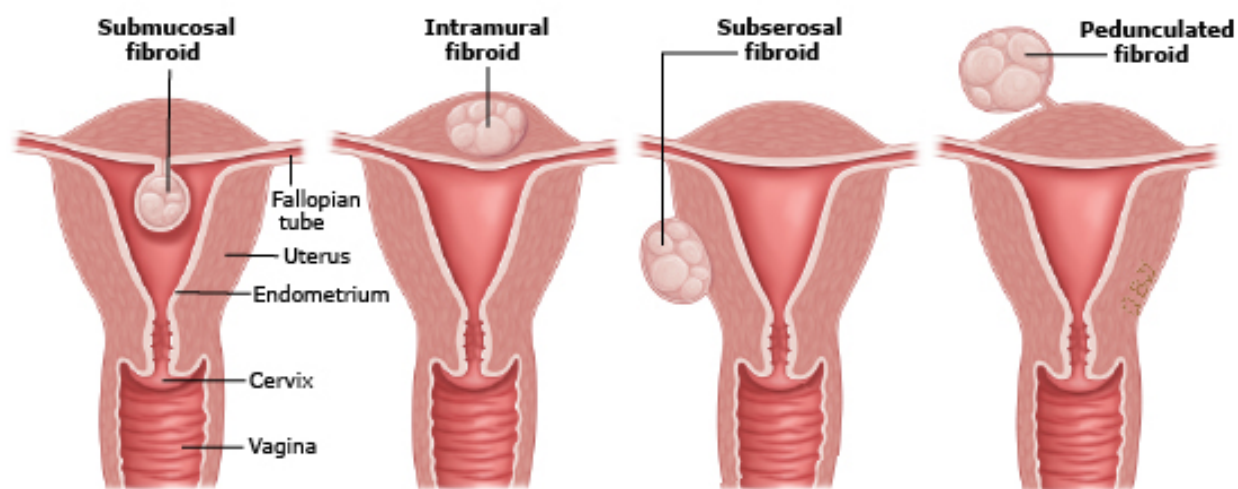


**Figure 1-2 Uterine leiomyoma locations.**
Intramural myomas, originated within the uterine wall, are the most common LM, which account for 80% of all cases. Submucosal myomas are less common; they derive from myometrial cells right under the endometrium, which is the lining of the uterine cavity. Subserosal myomas originate from the myometrium at the serosal surface of the uterus. They may have a broad or pedunculated base and maybe intraligamentary (Adapted from ref. [8]).

Although all LMs are thought to be derived from myometrium cells, LM can take place in multiple regions inside a uterus. The clinical effects of LM also depend on both the size and the location of the LM. Based on the International Federation of Gynecology and Obstetrics classification system, LM can be specified into three main categories based on their locations: intramural myomas, submucosal myomas, and subserosal myomas (Figure 1-2) [8, 9]. Intramural myomas, originated within the uterine wall, are the most common LM, which account for 80% of all cases. This type of LM can enlarge sufficiently inside the

myometrial layer and may distort the uterine cavity or serosal surface. In some cases, intramural myomas can be transmural and extend from the serosal to the mucosal surface [8]. Submucosal myomas are less common; they derive from myometrial cells right under the endometrium, which is the lining of the uterine cavity. Due to the location, this type of neoplasms can protrude into the uterine cavity, and the extent of this protrusion is important for predicting outcomes of myomectomy surgeries [8]. This type of LM is more likely to disturb pregnancy. Subserosal myomas originate from the myometrium at the serosal surface of the uterus. They may have a broad or pedunculated base and maybe intraligamentary.

### 1.1.2   Genetic features of uterine leiomyomas

Somatic mutations and chromosomal aberrations are associated with LM development [3]. One of the first major chromosomal aberrations found associated with the pathogenesis of LM is around the *HMGA2* gene. The *HMGA2* gene encodes the High-Mobility Group AT-hook 2 protein. As its name indicates, HMGA2 binds to the minor groove of AT-rich DNA fragments with three DNA-binding domains known as the AT-hooks in order to induce conformational changes of chromatin structure and regulate target gene expression [10]. Around 7.5% of LMs were found containing chromosomal rearrangement involving chromosome 12q14-15, which locates at the 5' of the *HMGA2* gene; this alternation gives rise to a very significant increase of HMGA2 level in LMs [3, 11].  Overexpression of the *HMGA2* gene itself has been proved to be sufficient to induce leiomyoma-like tissue formation in normal myometrium [12]. Several studies showed that overexpressed HMAG2 contributes to multiple aspects of tumor development in LM. Elevated HMGA2 level can suppress senescence through downregulating p14(Arf) in LM tumors [13]. Interestingly, this may also contribute to the disturbed Let7- HMGA2 - p14(Arf) pathway in LM, which is important in regulating the self-renewal of the fibroid stem cell population [14]. Furthermore, it was reported that the aberrant expression of HMGA2 in LM was associated with loss of TSC2 tumor suppressor gene function [15].
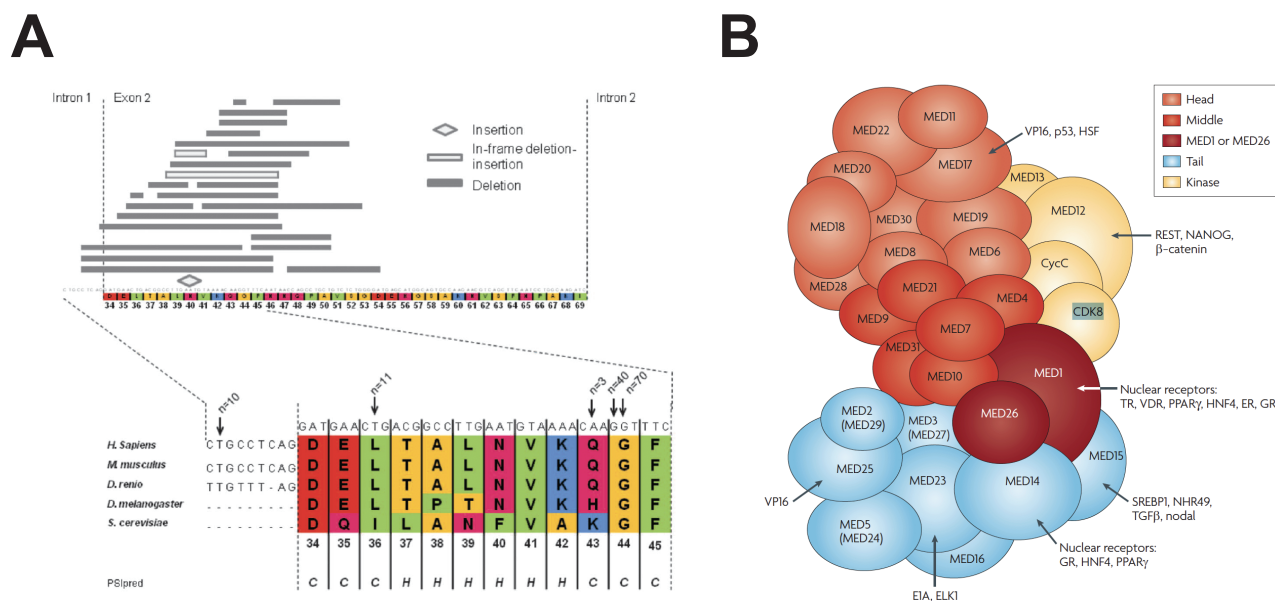
**Figure 1-3 MED12 mutation in LM.**

**A**) Alignment of the mutation hot spots near the beginning of MED12 exon 2. Arrows indicate observed missense and intronic mutations, and numbers above the arrows indicate how many times each mutation was detected in the 225 tumors included in this study. (Adapted from ref. [16]). **B**) Components of the MED12 complex and their subcategories (Adapted from ref. [17])

Besides chromosomal aberrations, somatic point mutations are also commonly found in LM tumors. The *MED12* (the Mediator Complex Subunit 12) gene, located on the chromosome Xq13.1 is the most highly mutated gene in LM. Around 70% of all LMs contain mutations within the MED12 gene and it has been reported that those mutations are mutually exclusive from the previously described HMGA2 rearrangement [16, 18]. The Mediator complex is a 26-subunit transcriptional regulator that serves as a bridge between regulatory sequences and RNA polymerase; as a component in the Mediator complex, MED12 is heavily involved in transcription activities [17]. Structurally, the components of the Mediator complex components are categorized into 4 subunits: head, middle, tail, and kinase [19]. The enzyme module contains MED12, MED13, Cyclin C, and CDK8 (Figure 1-3 B). Previous researches have demonstrated that this subunit has roles in transcriptional repression, but it can also serve as a positive regulator [20, 21]. The Mäkinen group reported that most of the MED12 mutations in LM take place in exon 2, specifically at codon 43 and 44

(Figure 1-3 A). MED12 G44D mutation is the most common mutation identified in LM; it can cause disrupted CDK activity in the Mediator complex [22]. Turunen et al demonstrated that, under normal conditions, MED12 indirectly activates CDK8 enzymatic activity by binding to a conserved surface groove on Cyclin C; G44D mutation disrupted the MED12-Cyclin C interaction resulting a loss of function of the CDK activity. The mechanism underlying MED12 mutation-initiated LM tumorigenesis requires further investigation.

### 1.1.3   Leiomyoma treatment options

There is no single best approach to treat LM. Treatments strategies toward LM depends on the severity of the symptoms, the size and location of LM, age of the patient, and the patient's desire for future fertility. As previously mentioned, 80% of women have at least one fibroid by the age of 50, but only 15% have symptoms [3]. When a patient with LM has little to no discomfort, watchful waiting is the best option [7].  LM growth is highly dependent on steroid hormones, especially progesterone. Thus, GnRH agonists (e.g. Lupron, Eligard, etc.) and progesterone antagonists (e.g. Mifepristone, Asoprisnil, etc.) are commonly used to manage LM in patients.  GnRH agonists block the production of estrogen and progesterone in patients and effectively shrink the existing tumors. Previous studies have shown that GnRH can increase apoptosis and decrease angiogenesis in LM lesions [23].  Although initially effective, the effect of both GnRH agonists and progesterone antagonists are temporary and tumor recurrences after treatment cessation have been consistently reported [24]. Also, long-term use of GnRH agonists can also cause severe side effects, including bone loss.  Currently, there is no available long-term drug for LM management approved by the Food and Drug Administration. Due to this reason, the gold standard of LM treatment is still surgical intervention. Hysterectomy is the definitive surgical operation. However, for patients who wish to preserve future fertility, myomectomy is also commonly preferred. Due to the heavy surgical burden and huge social economic-impact caused by LM, there is an urgent need to identify novel non-surgical options that can eradicate the existing tumors and prevent the initiation of new ones.

## 1.2 Progesterone and progesterone receptors

### 1.2.1 Progesterone production and its function

Progesterone is one of the most critical regulators during female reproductive system development, with diverse tissue-specific effects in humans. Progesterone is a member of the progestogen steroid hormone group. Like all other steroid hormones, progesterone is also generated from cholesterol (Figure 1-4). Progesterone is mainly produced by the corpus luteum during the luteal phase of the menstrual cycle and partially produced by the ovary and the adrenal gland; it is converted from pregnenolone inside the mitochondria then released into the bloodstream [25].
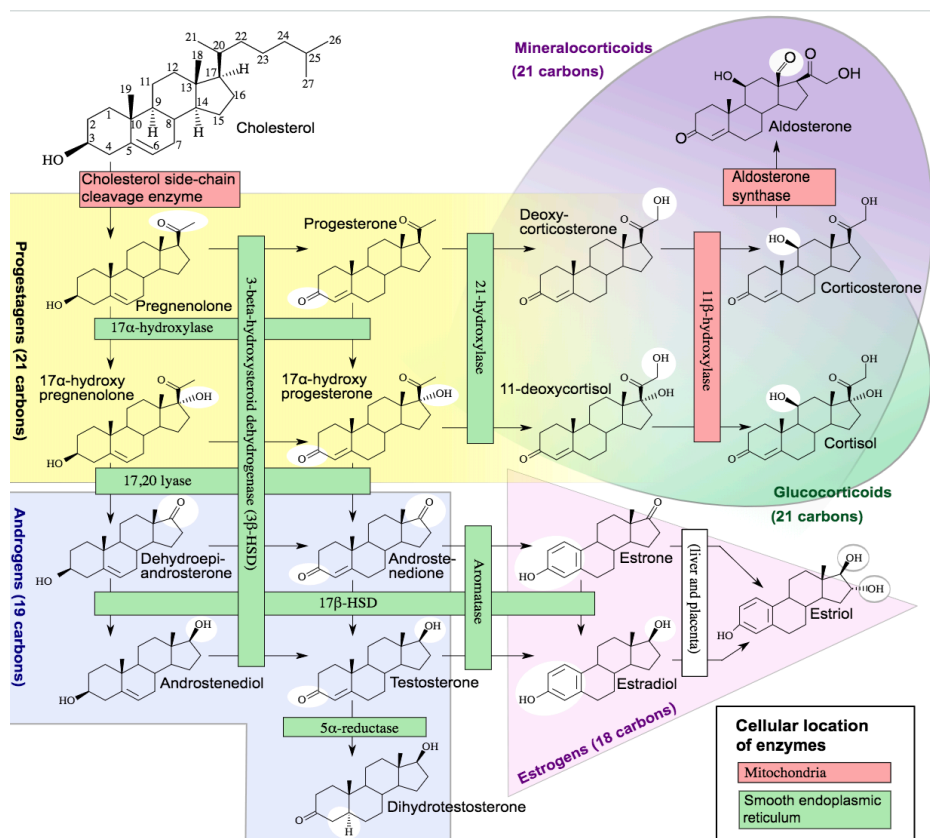


**Figure 1-4 Steroidogenesis overview**

Cholesterol is the precursor of all steroid hormones. Progesterone is converted from pregnenolone in mitochondria. (Adapted from Wikipedia, created by Mikael Häggström)

Progesterone is involved in the menstrual cycle, pregnancy, embryogenesis, as well as the pathogenesis of reproductive tract diseases. Progesterone level fluctuates throughout the menstrual cycle: at the start of the menstrual cycle, progesterone level is low and is maintained throughout the follicular phase [26]; progesterone is the dominant hormone after ovulation during the luteal phase (Figure 1-5). It reaches its peak in the middle of the luteal phase [27]. Progesterone prepares the endometrium for the potential of pregnancy after ovulation. If conception does not occur, the corpus luteum breaks down, lowering the progesterone levels in the body; this change sparks menstruation [25]. If pregnancy occurs, progesterone maintains at a high level to stimulate blood vessel development in the endometrium layer supporting the growth of the embryo and preventing the production of other oocytes.
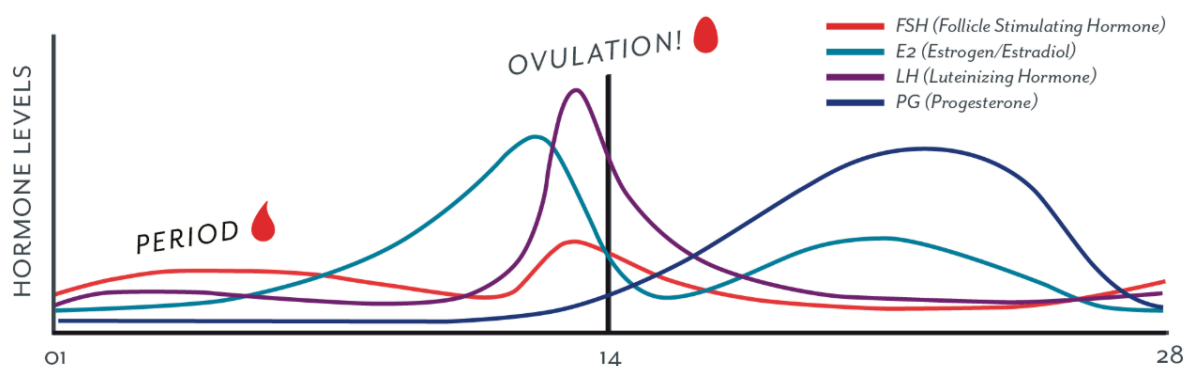


**Figure 1-5 Hormone production during the menstrual cycle**

At the start of the menstrual cycle, progesterone level is low and is maintained throughout the follicular phase; progesterone is the dominant hormone after ovulation during the luteal phase. (Adapted from https://helloclue.com/articles/cycle-a-z/progesterone-101)

## 1.2.2　Progesterone receptor
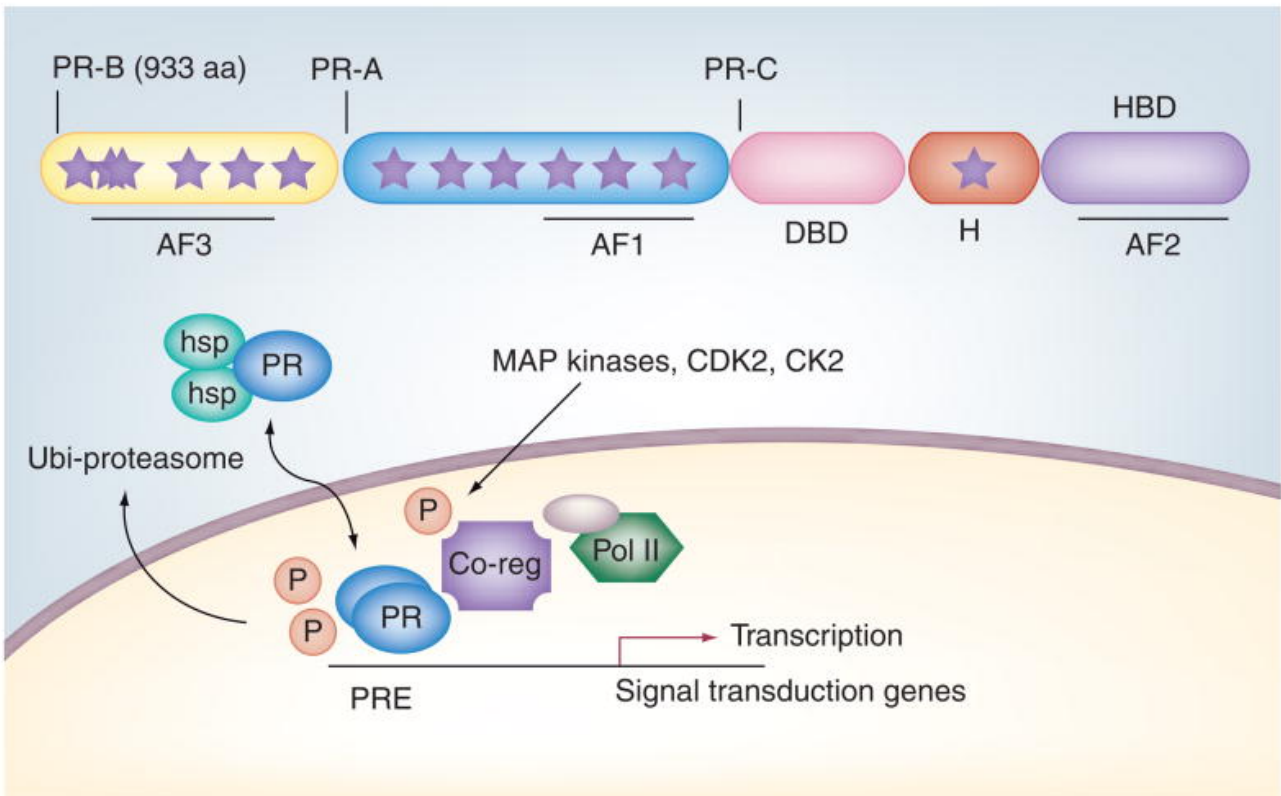


**Figure 1-6 Progesterone receptor isoforms**

PR-B is the full-length isoform of the *PGR* gene, containing the amino-terminal domain (AF3 and AF1), the DNA binding domain (DBD), the hinge (H) domain, and the ligand-binding domain (AF2). PR-A is the N-terminus truncated isoform missing the first 164 amino acids. PR-C is a further truncated version starting at the DBD. (Adapted from ref. [28])

Progesterone functions through its nuclear receptor (PR/PGR). The *PGR* gene contains two major isoforms, PR-A and PR-B. The two isoforms are transcribed from the same gene, but translated via alternate start sites: PR-A lacks 164 amino acids that are present at the N-terminus of PR-B [29] (Figure 1-6). The expression of PR-A and PR-B is governed by isoform-specific regulatory elements.  However, the expressions of PR-A and PR-B are both estrogen-dependent.  For example, PR-A and PR-B  are regulated

by two ERE (estrogen response element)-containing promoters: −711 to +31 (PR-A) and +464 to +1105 (PR-B) of the transcription start site [30]. Estrogen can also regulate the expression of PR via long-range transcriptional mechanisms.  EREs have been found ranging from 311 Kb upstream to 4 Kb downstream of the *PGR* transcription start site [31]. The *PGR* gene can also produce a less common isoform, PR-C, which is truncated still further downstream by the use of an additional AUG codon within the DNA-binding domain. Limited research demonstrated that up-regulated PR-C levels can serve as an inhibitor of PR-B action in laboring myometrium, but more detailed function of PR-C grants further investigation. [32]

Upon progesterone stimulation, PR-A and PR-B homo- or heterodimerize and translocate to the nucleus to conduct transcriptional regulation of its downstream targets [29]. PR-A and PR-B share some of their targets but also have isoform-specific activities under both ligand-dependent and ligand-independent conditions. Due to this fact, PR isoforms are functionally distinct and tissue-specific patterns for PR-A and PR-B ratio contributes to the diversity of progesterone effects [28].  Under normal conditions, PR-A and PR-B are commonly expressed at a similar levels; dysregulated PR isoform ratios can be used as an indicator of disease development (e.g. breast cancer) [33].

Unliganded progesterone receptors are cytoplasmic and bind to heat shock proteins. Upon progesterone or progestin binding, liganded receptors are released from the heat shock proteins, dimerize and translocate to the nuclei. Dimerized progesterone receptors bind to DNA regions that contain a specific palindromic sequence called progesterone responsive element (PRE). At those sites, PR recruits other transcription factors or the transcription complex to facilitate transcription activities. PR can also conduct transcriptional regulation in a PRE-independent manner via tethering to other transcription factors, such as signal transducers and activators of transcription (STATs), specificity protein 1 (SP1) and activator protein 1 (AP1) [28].  Ligand binding changes the conformation of the progesterone receptor and promotes its interaction with other coactivators. More than 300 coregulators in multiple signal networks have been described to interact with PR to form a functional PR activation complex and conduct downstream transcriptional regulation. The co-regulators of PR are highly diverse in function, which includes  p160 co-activators (e.g.

SRC-1/NCoA-1, SRC-2/GRIP1/TIF2, SRC-3/AIB1/ACTR/TRAM1/ p/CIP), histone acetylases (including p300/CBP), DNA helicases (including components of the SWI/SNF complex), ubiquitin ligases (such as E6-AP), methylases (e.g. CARM1) and the steroid receptor-specific RNA activator SRA [34]. The combination of the PR activation complex depends on the abundance of the co-regulators in the tissue type, which indicates a tissue-specific regulation patterns [35, 36].

### 1.2.3   Progesterone receptor in disease progression and cancer.

Dysregulation of progesterone and its receptor PR have been associated with the development of several diseases in the female reproductive tract. Unlike estrogen, the role of progesterone is more complicated and very tissue specific. Progesterone/PR signals have been found to regulate proliferation and differentiation through paracrine action under different circumstances. For instance, in the endometrial tissue, both normal endometrium and endometrial cancer, progesterone inhibits the proliferation of endometrial epithelial cells via PR expressed in stromal cells [2]. PR can induce expression of the basic helix-loop-helix transcription factor Hand2 in the uterine stroma suppresses the production of several fibroblast growth factors (FGFs) that act as paracrine mediators of mitogenic effects of estrogen on the epithelium [37].

While being a proliferation inhibitor in endometrium tissue, the role of progesterone/PR signaling in the mammary gland is less clear. Preclinical and clinical research demonstrated both pro-proliferation and anti-proliferation roles of progesterone/PR signaling in normal adult breast and breast cancer cells. Similar to in the endometrium tissue, the proliferation stimulating effect of progesterone is mostly carried out by paracrine action in the mammary gland. In normal breast epithelial cells, progesterone can stimulate the proliferation of PR negative cells as a result of paracrine actions, whereas PR-expressing cells do not proliferate; in breast cancer cells, with the increased expression of ER-alpha and PR, progesterone acts directly on PR-expressing cells and promotes proliferation [38].  Researchers showed that progesterone can drive mammary secretory differentiation via RANKL-mediated induction of ELF5 in the luminal

progenitor cells, which contribute to the growth of PR-negative mammary epithelial cells [39, 40].  On the other hand, the progesterone receptor can also serve as a tumor suppressor in breast cancer cells by modulating the action of estrogen receptors [41]. PR has also been shown to inhibit the proliferation of human breast cancer cells through the induction of MAPK phosphatase 1 [42, 43].

LM and breast cancer share many similarities with respect to responsiveness to steroid hormones [2, 44]. Under *in vitro* conditions, progesterone also showed both promoting and inhibiting effects on the proliferation of LM cells [45-47].  Due to the focus on progesterone itself, most *in vivo* and clinical studies emphasized the importance of progesterone during the growth and development of LM. For instance, the labeling index for proliferation markers such as Ki67 and proliferating cell nuclear antigen in LM peaks at the luteal phase, when progesterone is dominant [2, 48]. In parallel, using a xenograft mouse model, researchers also demonstrated that presence of progesterone is essential for the maintenance and growth of LM [49].  Further investigation on PR signal is needed to reveal the underlying mechanisms as to why progesterone can also result in anti-proliferation effects in LM cells.

## 1.3     Role of stem cells in uterine leiomyoma

Somatic stem cells, also known as adult stem cells, are a subtype of undifferentiated cells found among differentiated cells in tissues. These stem cells can conduct self-renewal as well as differentiate to some or all of the major specialized cell types within its specific tissue. The primary roles of somatic stem cells are to maintain and repair the tissue they reside in. Somatic stem cells are not only critical in the generation of complex multicellular organisms, but also important for the development of diseases especially tumors [50].

### 1.3.1     Tumor stem cells

A tumor is not a gathering of homogeneous malignant cells; rather, a tumor is a complex ecosystem containing tumor cells following a hierarchical differentiation order, as well as various infiltrating endothelial, immune, stromal, and other cell types that can influence the function of the tumor as a whole [51]. These extraneous cell types can influence tumor cells directly and can create metabolic changes such as a hypoxic environment and nutrient fluctuations, which contribute to heterogeneity in the function of malignant cells [51]. Studies from the past decades have demonstrated that there is a small population of tumor-initiating cells with stem cell characteristics within variety of cancer types, including leukemia, brain tumors, and breast cancers [52-54]. Similar to the embryonic stem cells, these tumor-initiating cells can conduct self-renewal, are capable to develop into multiple lineages, and have the potential to proliferate extensively. Thus, eradication of the stem-cell compartment of a tumor can be essential to achieve stable long-lasting remission and even a cure for the tumor.

The origin of cancer stem cells remains unclear, but studies in the past have described several hypotheses, including genomic instability, microenvironment influence, cell fusion, etc [55]. Genomic instability is one of the most studied and well-known hypotheses. Genome instability and alterations are known as the fundamental basis of cancer initiation. Mutations taken place in the normal stem cell niche transform the normal stem cells into cancer stem cells; with a high ability to proliferate and differentiate, the cancer stem cell population passes on the mutations among all of its descendants (Figure 1-7) [50, 56]. Studies in prostate cancer also showed that, under treatment stimulation, fully differentiated cancer cells can also undergo de-differentiation and regain stem cell-like characteristics [57].
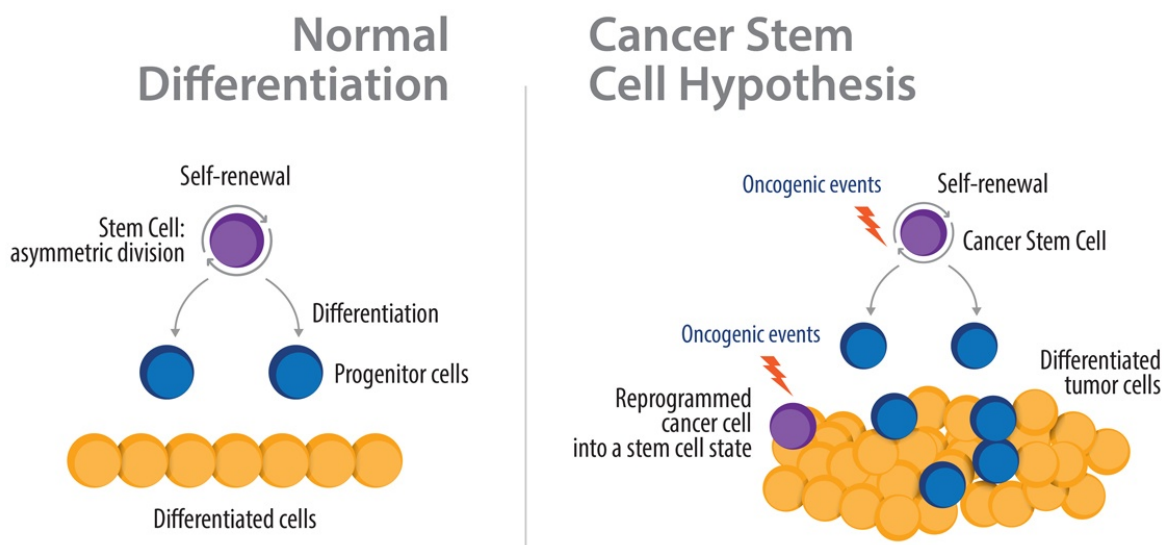
**Figure 1-7 Formation of cancer stem cells**

Cancer stem cells can emerge through two routes: 1) The oncogenic transformation of normal stem cells, most likely due to mutations or microenvironmental stimulations; 2) The dedifferentiation of mature tumor cells driven by the activation of morphogenic pathways during cancer progression. (Adapted from https://blog.crownbio.com/cancer-stem-cell-models)

Due to the importance of cancer stem cells during tumorigenesis, cancer researchers have been focusing on identifying and targeting this critical population to discover potential treatment avenues. To identify cancer stem cells, flow cytometry analysis is commonly used to isolate possible cancer stem cells based on cell surface marker staining [58].   The potential isolated stem cell population will undergo further examinations to validate its stem cell properties. In order to experimentally verify the "stemness" of a given cancer cell subset, researchers rely on xenotransplantation assays in immunocompromised mice using human tumor-derived cancer subpopulations. Using this experimental approach, tumorigenic populations may be capable of transferring human disease into immunodeficient murine hosts and recapitulated the phenotype and morphology of the original patient tumors. Throughout years of research, people characterized stem cell populations in diverse human malignancies, including breast cancer, colon cancer,

central nervous system tumors, ovarian cancer, etc (Figure 1-8) [58].

| Tumor type | Cell surface marker(s) | Reference |
|---|---|---|
| Breast | $CD44^+CD24^-$/low Lineage$^-$ESA$^+$ | Al-Hajj et al.(17) |
| CNS | $CD133^+$ | Singh et al.(18) |
| Colon | $CD133^+$ | O'Brien et al.(20) |
| | $CD133^+$ | Ricci-Vitiani et al.(21) |
| | $ESA^{high}CD44^+$ Lineage$^-$ ($CD166^+$) | Dalerba et al.(19) |
| Ewing's | $CD133^+$ | Suva et al.(28) |
| Head and neck | $CD44^+$Lineage$^-$ | Prince et al.(22) |
| Melanoma | $ABCB5^+$ | Schatton et al.(26) |
| Liver | $CD90^+CD45^-$ ($CD44^+$) | Yang et al.(27) |
| Ovarian | $CD44^+CD117^+$ | Zhang et al.(23) |
| Pancreas | $CD44^+CD24^+ESA^+$ | Li et al.(25) |
| | $CD133^+$ | Hermann et al.(24) |

**Figure 1-8. Surface markers identified for human tumors.**
List of cell surface markers that have been used to isolate tumor stem cells from human samples from past research. Stem cell characteristics of the cells with the molecular profiles were also tested using a mouse xenograft model. (Adapted from ref. [58]).

Many signaling pathways that contribute to the survival, proliferation, self-renewal, and differentiation process of normal stem cells are abnormally activated or repressed in cancer stem cells, including the WNT/beta-catenin signaling, Notch signaling, NF-κB pathway, and TGF/SMAD signaling pathway. For instance, the WNT/beta-catenin pathway is critical in multiple aspects of the tumorigenesis process. This pathway not only transforms dormant cancer stem cells to active stem cells to promote cell cycle

progression, but it also plays a very important role in dedifferentiation of the cancer stem cells [59]. In colon cancer, WNT/beta-catenin activation can decrease the expression level of important differentiation factors, such as PMP22 [60]. The NF-κB pathway is another pathway that showed great importance during cancer stem cell research. It has an essential function in regulating inflammation, self-renewal, maintenance, and metastasis of cancer stem cells. The inflammatory mediator prostaglandin E2 (PGE2) contributes to tumor progression by activating NF-κB in colorectal cancer stem cells [61].  In addition, the NF-κB pathway can also suppress transcription factors which are inhibitors of self-renewal and metastasis [59].

Due to the heterogeneity of cancer composition, the effectiveness of cancer therapy can also vary among different cell types within tumors. Cancers can recur after an initial response to conventional therapies (e.g. chemotherapy or radiation therapy); the recurred tumors are likely to be resistant towards the initial treatment or even maybe multiple other therapies. There are several hypotheses toward this recurrence phenomenon and cancer stem cells may be highly involved (Figure 1-9).  Cancer stem cells are prone to have increased resistance to chemo- and radiation therapy. Researchers think the drug resistance of cancer stem cell can be both intrinsic and acquired. In an intrinsic model, the tumor contains a small population of drug-resistant stem cells. Following chemotherapy, differentiated cells are killed, but the stem cells survive, which can repopulate the entire tumor with a heterogeneous composition. During treatment, cancer stem cells can also acquire mutations, which contributes to the development of the resistance. Tumors repopulated from the mutated cancer stem cells are likely to display the same drug-resistant phenotype as the stem cell population.
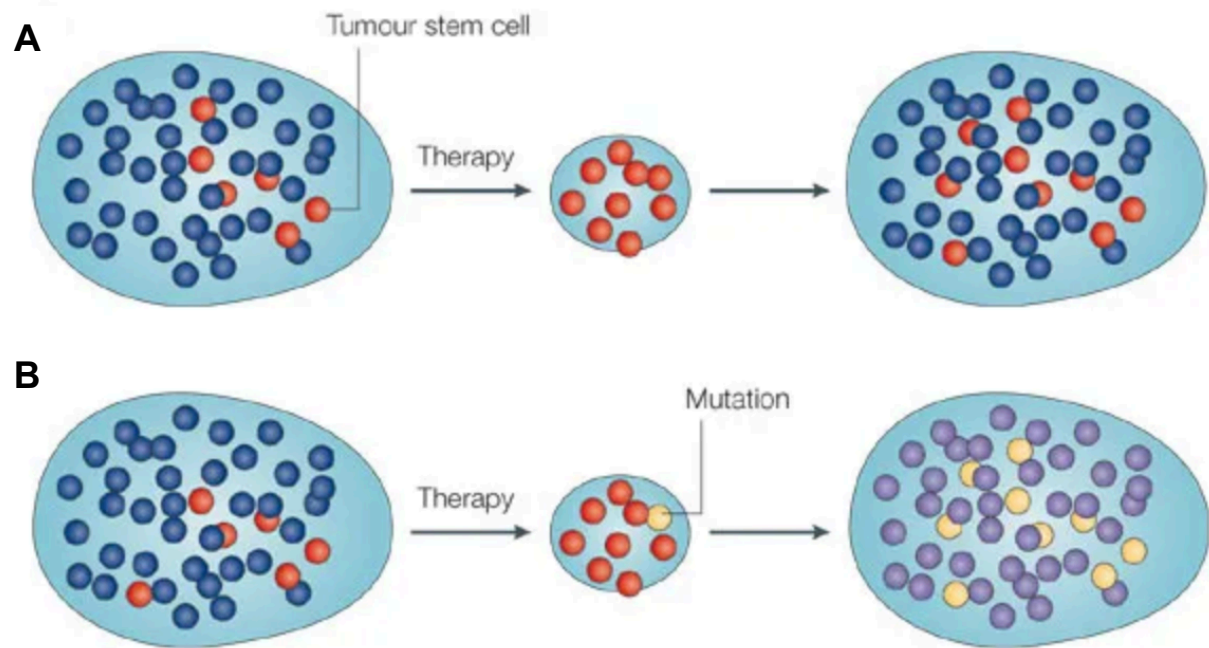
**Figure 1-9 Model of cancer therapy resistance**

**A)** In the cancer-stem-cell model, drug resistance can be mediated by the intrinsic qualify of stem cells. In this model, tumors contain a small population of tumor stem cells (red) and their differentiated offspring (blue). Following chemotherapy, the differentiated cells are killed, but the stem cells survive. These cells repopulate the tumor, resulting in a heterogeneous tumor composed of stem cells and committed but variably differentiated offspring. **B)** In the 'acquired resistance' stem-cell model, the tumor stem cells (red) survive the therapy, whereas the committed but variably differentiated cells are killed. Mutation(s) in the surviving tumor stem cells (yellow) and their descendants (purple) can arise (by mechanisms such as point mutations, gene activation, or gene amplification), conferring a drug-resistant phenotype. (Adapted from ref. [62])

To eliminate or reduce the stemness of the cancer stem cell population, novel stem cell therapies are developed to specifically target dysregulated characteristics in the cancer stem cell populations, such as epigenetic modification. Epigenetic therapies have shown promising clinical outcome when targeting cancer stem cells. In ovarian cancer, epigenetic aberrations, especially DNA methylation, silence tumor-suppressor and differentiation-associated genes that regulate the survival of ovarian cancer stem-like cells

[63]. In one study, researchers demonstrated that DNA-hypomethylating reagent can reset ovarian cancer stem-like cells towards differentiation phenotype and re-sensitize the ovarian cancer stem-like cells to platinum treatment [63].

### 1.3.2 Leiomyoma stem cells

Previous studies have shown that a small stem-cell like population is responsible for the initiation of LM [64, 65]. LM stem cells (LSC) were first identified as a side population based on Hoechst-33342 dye extrusion, a universal feature of somatic stem cells [50, 65, 66]. Cancer/tumor stem cells can extrude dyes such as Hoechst-33342; thus Hoechst-33342 dye was commonly used to identify the stem cell population when the presence of surface marker is unknown [67]. LM side population comprised approximately 1% of all LM cells. And the side population can regenerate a significantly larger tumor compared with the main population using a mouse xenograft model, validating the stem cell characteristic of the side population (Figure 1-10) [65].
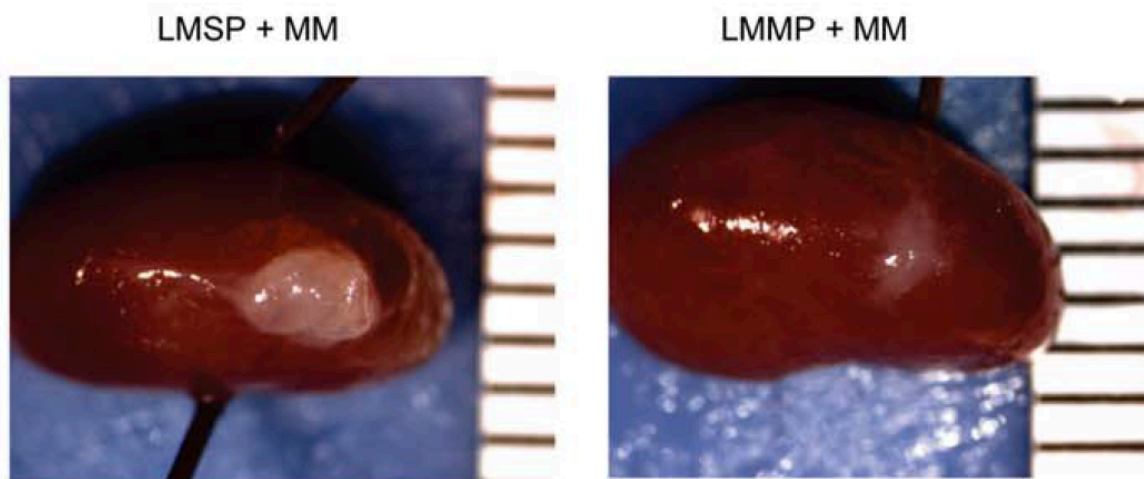


**Figure 1-10 Xenograft tumor of LM side population vs the main population.**

Macroscopic visualization of the transplanted site 12 weeks after xenotransplantation of LM side population (LMSP) and main population (LMMP). The LM side population regenerates a significantly larger tumor compared with the main population via the use of a mouse xenograft model (Adapted from ref. [65]).

To avoid the toxicity of Hoechst staining, the LSC isolation protocol was refined using fluorescence-activated cell sorting (FACS or flow cytometry) to sort LM cells based on the expression of the cell surface markers [64]. Using a microarray to compare the transcriptome of the LM side population and main population, the Bulun lab identified that the side population contained a significantly higher expression of CD34 and CD49b markers. CD34 and CD49b fractionate LM cells into three populations with distinct gene expression profiles and possible differentiation order: CD34+/CD49b+ (5.9%$\pm$ 1.2%, LM stem cell, LSC), CD34+/CD49b- (7.1%$\pm$ 1.3%, LM intermediate cells, LIC), and CD34-/CD49b- (87.1%$\pm$ 2.1%, LM differentiated cells, LDC). The CD34+/CD49b+ population is most likely to be the stem cell enriched population (Figure 1-11 A). The xenograft model was used to examine the tumor initiation capacity of each population. The LSC population is capable of regenerating a significantly larger tumor compared with the LDC population (Figure 1-11 B). While the majority of cells in LM tumors are LDCs, LSCs are indispensable for proliferation, regeneration, and robust tumor growth [68, 69]. The LSC population also contains the highest expression of the stem cell markers and the lowest expression of steroid hormone receptors (e.g.: ESR1 and PGR) and smooth muscle markers (e.g.: $\alpha$-SMA), which may explain the observed frequent recurrence of LM after hormone-based therapies (Figure 1-11 C).

Due to the importance of the LM stem cell population, a microarray was used to profile the global transcriptome of each leiomyoma population identified with the markers CD34 and CD49b. Each LM cell population contains a distinct transcriptome with thousands of differentially expressed genes. The transcriptome analysis also supports the previously hypothesized differentiation order from LSC to LIC then to LDC. The LIC population, as the first stop after LSC differentiation, serves as a feeder progenitor population to the stem cell population and is critical in maintaining the development of LM. It expresses high levels of PR, cytokines (e.g. RANKL), growth factors (e.g. IGF2), and extracellular matrix (ECM) proteins, and plays a critical role in transducing progesterone-mediated paracrine mitotic signals to PR-deficient LSC to stimulate their self-renewal and proliferation [68, 69]. Although thousands of genes were identified to be differentially expressed among LM populations, the underlying mechanism of how LSC is regulated remained unclear.
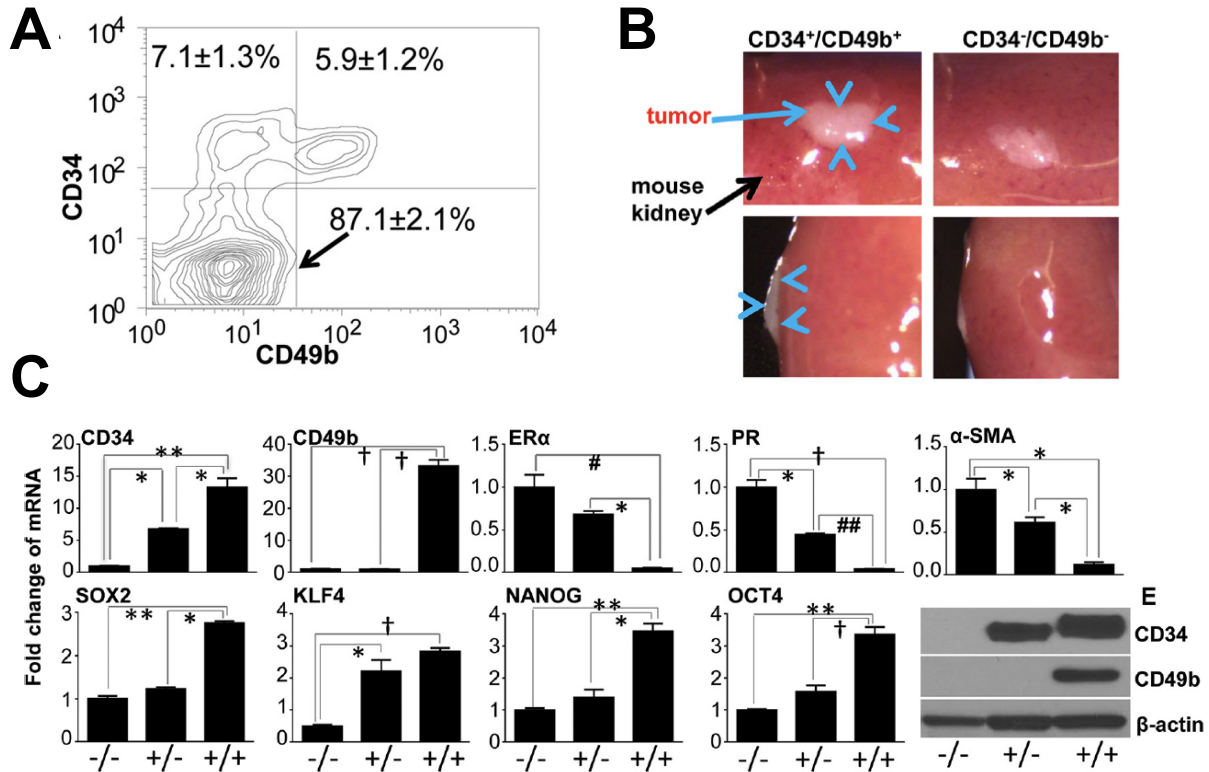
**Figure 1-11 Identifying LM stem cells using CD34 and CD49b.**

**A**) Differential expression of CD34 and CD49b fractionates the total leiomyoma cells into three subpopulations: CD34-/CD49b-, CD34+/CD49b-, and CD34+/CD49b+ cells. **B**) Macroscopic visualization of the regenerated tumor (blue arrows) 8 weeks after engrafting using CD34+/CD49b+ cells and CD34-/CD49b- cells. **C**) Quantitative real-time RT-PCR was performed to determine the mRNA levels of cell surface markers (CD34 and CD49b), stem cell factors (KLF4, NANOG, SOX2, and OCT4), steroid receptors (ERα/ESR1 and PR), and smooth muscle cell marker α-SMA in the three FACS-sorted subpopulations of leiomyoma cells. Right bottom panel shows representative western blots of differential protein expression of CD34 and CD49b between the three leiomyoma cell populations. (Adapted from ref. [64])

## 1.4    The roles of DNA methylation

In the mammalian genome, DNA methylation is the most common and most studied epigenetic mark. DNA methylation is a methylation modification that taken place on the fifth carbon of cytosines. In eukaryotes, DNA methylation is mostly found in the context of symmetrical CpG dinucleotides [70]. DNA methylation is essential for mammalian development and has notable roles in gene silencing, protection against spurious repetitive element activity, genomic stability during mitosis and parent-of-origin imprinting [71].

### 1.4.1    Regulation of DNA methylation

There are three phases of DNA methylation: establishment (*de novo* DNA methylation), maintenance and demethylation. The methylation of the DNA cytosine is catalyzed by DNA methyltransferases (DNMT), which transfer a methyl group from the donor S-adenosyl methionine (SAMe) to the 5' position of the pyrimidine ring. There are 5 members in the DNA methyltransferase protein family: DNMT1, DNMT2, DNMT3a, DNMT3b, and DNMT3L. Among those, DNMT3a and DNMT3b are the *de novo* DNA methyltransferases, which contain a highly conserved DNMT domain in the c-terminus and two chromatin reading domains, ATRX-DNMT3-DNMT3L (ADD) and PWWP [72]. DNA methylation at promoter regions is mostly associated with gene repression. ADD binds to the K4 residue of H3 to conduct its enzymatic activities, which can be blocked by H3K4me3 histone modification enriched at active promoters (Figure 1-12 A) [73]. In contrast, DNA methylation taken place in the gene body is usually hypothesized to be correlated with active transcription. The PWWP domain can interact with the active H3K36me3 histone mark in the gene body to secure DNMT recruitment to the DNA (Figure 1-12 B) [72]. The maintenance of DNA methylation is most conducted by DNMT1. This enzyme copies pre-existing methylation patterns onto the new DNA strand during the S-phase of the cell cycle during DNA replication. DNMT1 interacts with E3 ubiquitin-protein ligase UHRF1 to specifically binds hemimethylated CpG dinucleotides at replication forks to maintain the methylation pattern in the parental cells (Figure 1-12 C) [72].
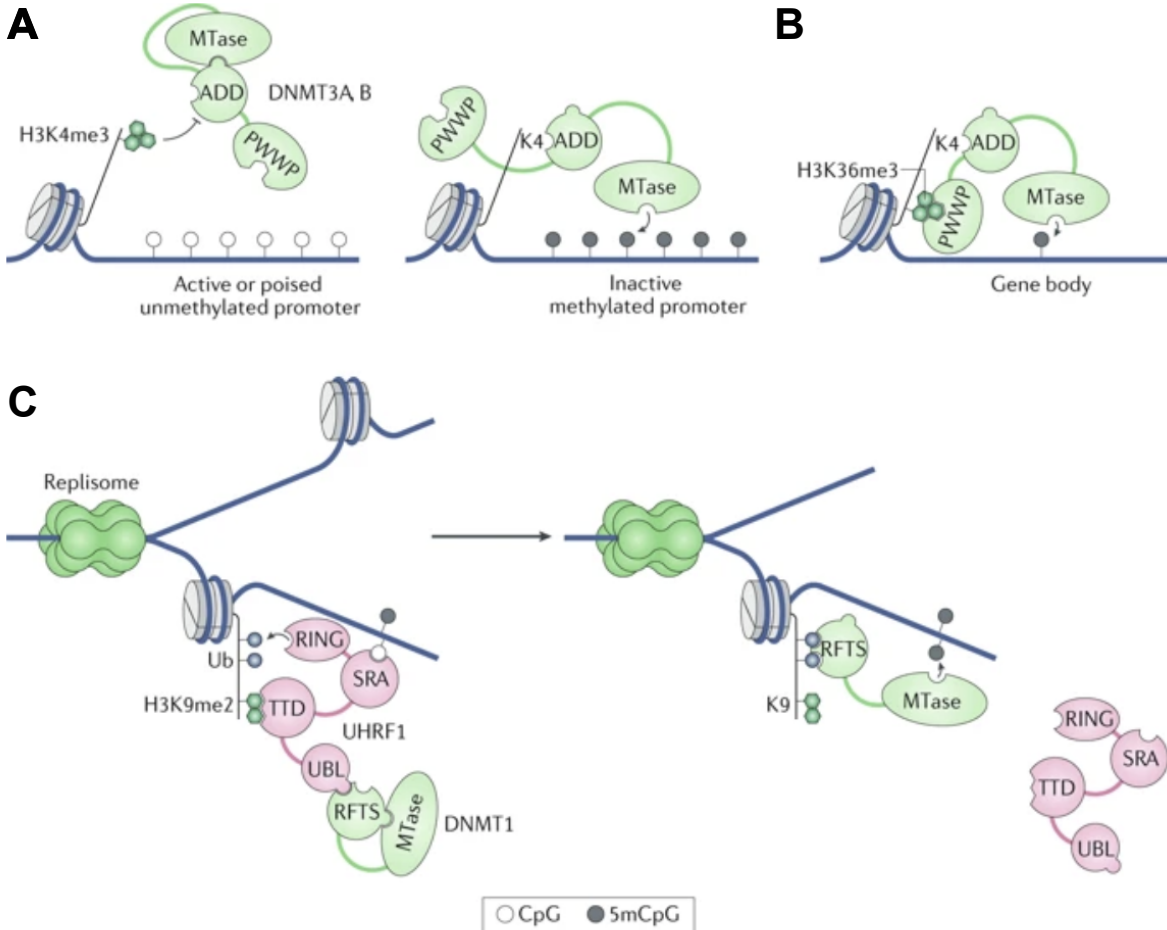
**Figure 1-12 DNA methylation mechanisms.**

**A**) DNA methylation at promoters. H3K4me3, which marks active promoters, prevents DNMT3A and DNMT3B binding to chromatin. **B**) DNA methylation at gene bodies. In gene bodies, H3K36me3 serves as a recruitment module for the DNMT3 PWWP domain. **C**) Maintenance of DNA methylation. E3 ubiquitin-protein ligase UHRF1 is recruited to replicating DNA which binds hemimethylated CpG sites. UHRF1 recruits DNMT1 through an interaction between its ubiquitin-like (UBL) domain and the DNMT1 RFTS. (Adapted from ref. [72])

Genome-wide DNA demethylation is conducted by TET (Ten Eleven Translocation) family enzymes. Mammalian cells express three TET genes: TET1, TET2, and TET3, which share significant sequence homology at their C-terminal catalytic domains [74]. TETs can catalyze the conversion of modified genomic bases 5-methylcytosine to 5-hydroxymethylcytosine, which play a key role in activating the repressed genes. The three TETs share similar catalytic activity having 70% sequence homology around the catalytic domain, but the distinct domain in each TET also suggests its unique binding to DNA sequences and interacting proteins (Figure 1-13). Although TET1 and TET3 share an N-terminal CXXC zinc finger domain which facilitates their binding to DNA, they each have specific roles: TET1 is associated with distal regulatory elements as well as genome-wide diffuse 5hmC deposition and TET3 is involved in cytosine oxidation at a wide range of genomic elements which takes place during the development of the embryo [75-77]. Without the CXXC domain, the mechanism underlying TET2 binding to the genome is less clear. Previous research has suggested TET2's interaction with the genome is mediated through transcription factors that have broad binding activities throughout the genome, such as EBF1 or WT1 [78, 79].
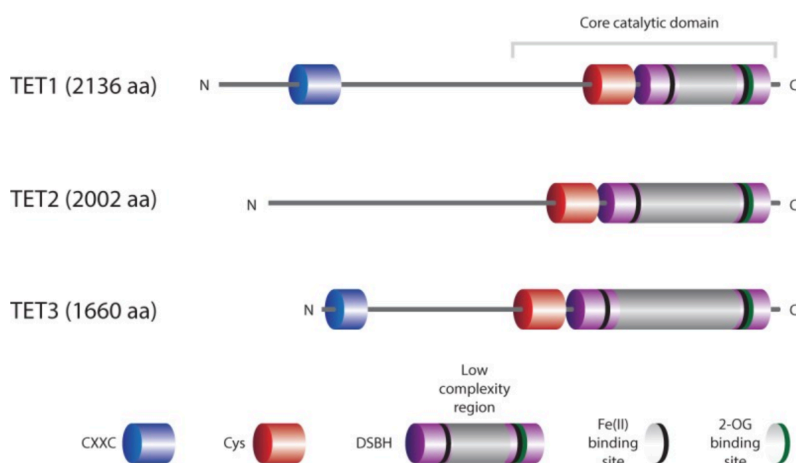


**Figure 1-13 Structure of TETs**

TETs share a similar catalytic group containing DSBH domain a cysteine-rich (Cys) domain, and binding sites for the Fe(II) and 2-OG cofactors. TET1 and TET3 have an N-terminal CXXC domain that can bind directly to DNA and facilitate recruitment to genomic target sites, while TET2 binding is mainly facilitated by other transcription factors. (Adapted from ref. [77]).

### 1.4.2  DNA methylation changes during stem cell differentiation

DNA methylation plays an extremely important role during the differentiation process of stem cells, especially during embryonic development. Global methylation patterns change drastically following fertilization (Figure 1-14). Post-fertilization, the embryo loses gamete-specific DNA methylation patterns inherited from the oocyte and sperm to acquire pluripotency. TET3 is activated in the fertilized zygote, resulting in hydroxymethylation and active DNA demethylation. The global methylation level reaches its low point in the blastocyst stage [72]. After implantation, DNA methyltransferase DNMT3a and DNMT3b are activated towards the epiblast stage. The re-methylation of the epiblast genome is rapid. Even though the epiblast still has a certain degree of pluripotency, the DNA methylation pattern in the epiblast can be inherited throughout life and be maintained as epigenetic memory in early embryogenesis. Elevated DNA methylation, as one of the most important factors, rapidly reduces pluripotency and drives embryonic development toward specific lineages [71]. Besides in germ cells, changes in genome-wide methylation levels after the epiblast stage is relatively small. Nevertheless, there are widely-spread discrepancies in methylation patterns among various tissues, which indicate *de novo* methylation and demethylation play critical roles during differentiation in specific tissues [72].

Different from embryonic stem cell development, changes of DNA methylation during adult stem cell differentiation is not commonly reflected on the global DNA methylation level, but rather specific changes surrounding genes related to differentiation and stemness preservation. The hematopoietic system remains one of the most carefully dissected lineages regarding changes in DNA methylation. This system has provided some of the most comprehensive information on how DNA methylation functions within adult lineages, from stem cells to terminally-differentiated cells, and are representative of other developmental transitions [71]. Within the hematopoietic system, DNMT1 is expressed in all blood lineages, while DNMT3A and DNMT3B are predominantly found within hematopoietic stem cells (HSCs) and lymphoid lineages [80]. While the loss of DNMT3s in HSCs results in defects in the self-renewal process, HSCs with DNMT1 defects suffer from self-renewal defects and dramatic dysregulation of myeloid versus lymphoid compartments [81]. During the differentiation from HSC, DNA methylation levels decrease at certain lineage-specific regulatory

elements and increase at regulatory elements for other lineages [82]. These findings indicate that HSCs are highly sensitive to DNA methylation defects and preservation of DNA methylation is critical for maintaining stemness and proper lineage differentiation.
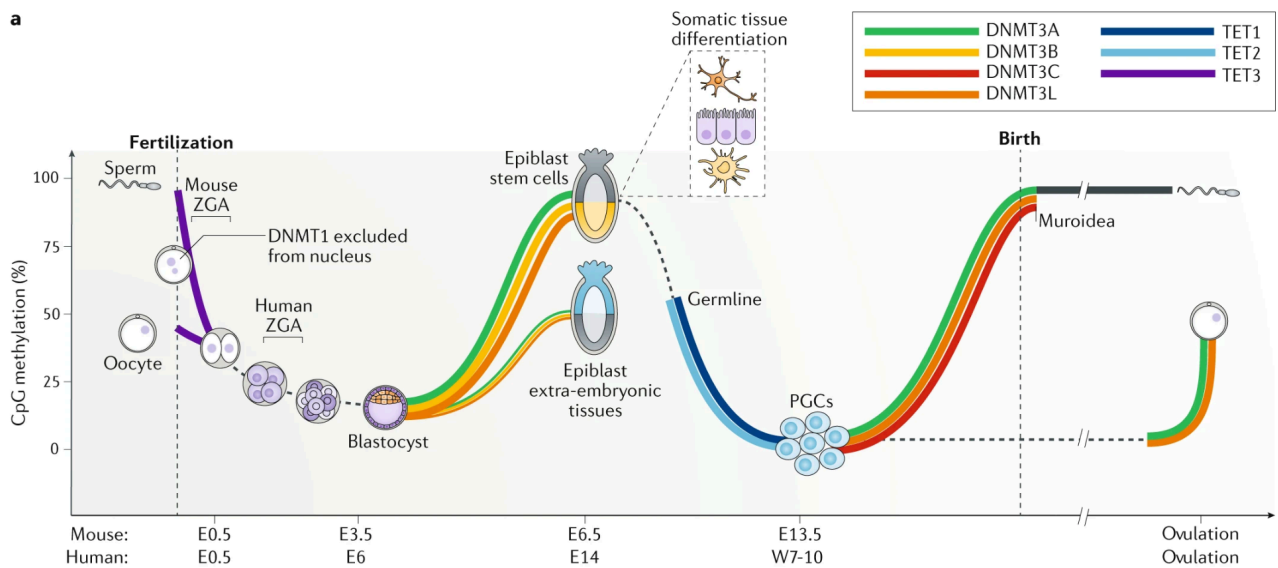


**Figure 1-14 Methylation changes during embryonic development.**
Upon fertilization, the global DNA methylation level decreases via TET3 and reaches the lowest level at the blastocyst stage to acquire pluripotency. With the increased expression of DNMTs, DNA methylation increases till the epiblast stage. This increase reduces pluripotency and drives embryonic development toward specific lineages. (Adapted from ref. [72])

### 1.4.3   The role of DNA methylation during tumor development

DNA methylation is essential for development and proper cell functioning; abnormalities in DNA methylation can lead to various diseases, including cancer. Through genome-wide methylation analyses, the tumor methylome was shown to display a distinct pattern when compared to the adjacent normal tissues. Both hypomethylation and hypermethylation have been seen during cancer development. This phenomenon causes genomic instability, oncogene activation, and, more commonly, silencing of tumor suppressors (Figure 1-15) [83].

## Frequent Target DNA Sequences for Cancer-Associated Hypermethylation or Hypomethylation

| DNA hypermeth. | DNA hypomethylation | | | |
|---|---|---|---|---|
| | High-copy repeats | Moderate-copy repeats | | Unique seqs. |
| CpG islands | heterochromatin repeats | interspersed repeats | e. g., latent viruses or retroviruses and sequences with sperm-specific hypomethylation | various genes, including testes-specific genes |
| e.g., TSG promoters | e.g., satellite DNA | e.g., LINE-1 | | |
| | Could affect chromatin structure and genome stability | Might affect genome stability | | Could increase expression of cancer-promoting genes |
| Often → repression of TSGs | Might have trans effects on transcription in other parts of the genome | | Could lead to transcription activation | |

**Figure 1-15 Dysregulated DNA methylation in cancer**

A summary of the most frequent types of sequences affected by cancer-specific DNA hypermethylation or hypomethylation. TSG, tumor suppressor gene. (Adapted from ref. [84])

In normal cells, pericentromeric heterochromatin is hypermethylated. Satellite sequences (repetitive genomic sequences such as LINE, SINE, IAP, and Alu elements) are silenced, thereby ensuring genomic integrity and stability. In a variety of tumors, this protection mechanism is disrupted by hypomethylation surrounding these inactivated regions, which leads to undesired mitotic recombination and mutagenesis [84]. Loss of DNA methylation may also activate latent viral sequences in the genome, which may also contribute to the development of tumors. For example, the genome of genital human papillomaviruses (HPVs) is suppressed in the host due to hypermethylation. Hypomethylation can induce the activation of HPV and stimulate the progression of cervical cancer [85]. Compared to the genome-wide phenomenon, hypomethylation at specific genes that contribute to cancer progression is less common.

Hypermethylation plays a more critical role in regulating cancer-related gene expression. During tumorigenesis, CpG islands around tumor suppressor promoters are often found to be hypermethylated and repressed [84]. This type of dysregulation was firstly identified in retinoblastoma which contains hypermethylation at the promoter of the retinoblastoma tumor suppressor gene RB1 [86]. Since then, studies focusing on this phenomenon identified a large number of tumor suppressor genes to be hypermethylated in tumor tissues. Genes hypermethylated in cancer are involved in a variety of critical pathways. For instance, hypermethylation of CDKN2A and CDKN2B, two cell cycle regulators, are found in multiple cancers, including colon cancer and breast cancer [87, 88]. It is now commonly believed that DNA methylation-associated silencing plays a crucial role in tumorigenesis and is a hallmark of all types of human cancer.

Abnormal methylation in cancer progression is linked with dysregulation or mutation of the DNA methylation machineries, including DNMTs and TETs [89]. Somatic mutations in DNMT3A and TET2 are linked to enhanced proliferation of immature myeloid cells and the development of adult hematological malignancies [90, 91]. R882 mutation in DNMT3A has a dominant-negative effect on the methyltransferase activity resulting in genome-wide hypomethylation, which stimulates the onset of acute myeloid leukemia [92]. Similarly, TET2 mutations promote hypermethylation in myeloid malignancies, which are mostly found at enhancer regions [93]. Moreover, restoration of TET2 function can reverse cancer progression in the hemopoietic system, further supporting a causative role for the enzyme in the etiology of the disease [94].

### 1.4.4 Dysregulated DNA methylation in LM.

Consistent with other tumor models, dysregulated methylation has been observed in LM. In a study from the Bulun lab in 2012, genome-wide DNA methylation and transcription activity were profiled in LM and matched adjacent normal myometrial tissue from 18 patients [95]. Compared to the adjacent normal myometrial cells, there were 55 genes with differential promoter methylation and mRNA expression in LM. Eighty percent of the identified genes showed an inverse relationship between DNA methylation status and mRNA expression in LM and the majority of genes (62%) displayed hypermethylation associated with gene

silencing. Among those, the promoter region of several known tumor suppressors, such as KLF11, DLEC1, and KRT19, were hypermethylated with significantly lower expression in LM (Figure 1-16). Other research groups identified similar results, with significant genome methylation difference between LM and myometrium. Some of the differentially-methylated genes are also involved in cancer processes, including transformation-related gene IRS1 and collagen-related genes such as COL4A1, COL4A2, and COL6A3 [96].

To explain the methylome difference between LM and myometrium, the Bulun group followed up with another study focusing on the DNA methylcytosine dioxygenase TETs [97]. Navarro et al demonstrated a genome-wide increase of 5-hmC levels in LM tissue compared to normal myometrial tissue. The increase in 5-hmC levels was associated with the upregulation of theTET1 or TET3 level in LM. When knocking down TET1 or TET3 or treating with a TET enzyme inhibitor, the proliferation of LM cells was significantly reduced. This study demonstrated the importance of the DNA methylation system during the proliferation of LM, but the underlying mechanism of how DNA methylation regulates LM development and progression needs further investigation.

| Gene symbol | Gene name | DNA Methylation[1] | P-value | mRNA level[2] | P-value |
|---|---|---|---|---|---|
| KRT19 | Keratin 19 | 12.81 | 4.92E-10 | −2.5166 | 5.68E-05 |
| NUAK1 | NUAK family, SNF1-like kinase, 1 | 7.327 | 9.98E-07 | −1.5532 | 6.70E-06 |
| KLF11 | Kruppel-like factor 11 | 7.1816 | 1.40E-07 | −1.6765 | 1.32E-05 |
| DLEC1 | deleted in lung and esophageal cancer 1 | 6.3828 | 1.01E-07 | −1.51 | 1.44E-03 |
| CORIN | Corin, serine peptidase | 5.9463 | 2.85E-09 | −1.5615 | 2.17E-05 |
| EFEMP1 | EGF-containing fibulin-like extracellular matrix protein 1 isoform b | 5.6828 | 4.02E-09 | −3.32 | 9.22E-06 |
| MBP | Myelin basic protein | 5.5791 | 1.83E-06 | −1.7654 | 1.50E-04 |
| TMEM173 | Transmembrane protein 173 | 5.1523 | 7.89E-07 | −1.6351 | 2.56E-04 |
| TNFSF10 | Tumor necrosis factor (ligand) | 4.9483 | 1.51E-08 | −1.9423 | 1.16E-07 |
| BST2 | Bone marrow stromal cell antigen 2 | 4.5596 | 9.82E-11 | −2.1626 | 1.44E-06 |
| C1orf115 | Chromosome 1 open reaing frame 115 | 3.9093 | 9.45E-07 | −1.7191 | 4.16E-08 |
| HOXA5 | Homeobox A5 | 3.7552 | 8.49E-08 | −1.7155 | 7.35E-06 |
| TEK | TEK tyrosine kinase, endothelial | 3.7477 | 9.69E-09 | −1.8461 | 1.37E-06 |
| RBP1 | Retinol binding protein 1, cellular | 3.686 | 1.10E-08 | −2.3092 | 1.14E-04 |
| RASIP1 | Ras interacting protein 1 | 3.5885 | 1.06E-11 | −1.7365 | 9.28E-05 |
| GRAMD3 | GRAM domain containing 3 | 2.8256 | 1.19E-05 | −1.5505 | 1.03E-05 |
| CCDC109B | Coiled-coil domain containing 109B | 2.645 | 1.60E-05 | −1.5931 | 4.92E-05 |
| APOLD1 | Apolipoprotein L domain containing 1 | 2.6009 | 1.45E-12 | −1.9577 | 6.72E-05 |
| CALCRL | Calcitonin receptor-like | 2.5571 | 2.18E-06 | −1.856 | 3.63E-07 |
| SERPINF1 | Serpin peptidase inhibitor, clade F | 2.4166 | 3.88E-05 | −1.8676 | 9.71E-04 |
| TM4SF1 | Transmembrane 4 L six family | 2.2707 | 7.66E-08 | −1.9102 | 3.64E-07 |
| CD34 | CD34 molecule | 2.2474 | 1.41E-09 | −1.585 | 1.51E-05 |
| CFB | Complement factor B | 2.2336 | 3.08E-04 | −1.9627 | 4.53E-06 |
| SRGN | Serglycin | 2.2307 | 2.10E-05 | −1.8865 | 1.33E-06 |
| LYVE1 | Lymphatic vessel endothelial hyaluronan receptor 1 | 2.2275 | 9.19E-07 | −2.4155 | 1.90E-05 |
| LCN6 | Lipocalin 6 | 2.1991 | 2.38E-10 | −1.5541 | 2.34E-06 |
| PCOLCE | Procollagen C-endopeptidase enhancer | 2.1615 | 2.83E-05 | −2.2239 | 1.87E-05 |
| DARC | Duffy blood group, chemokine receptor | 2.1597 | 1.44E-09 | −3.5596 | 4.90E-06 |
| CLDN5 | Claudin 5 | 2.1438 | 1.44E-09 | −1.6468 | 7.82E-05 |
| S1PR1 | Sphingosine-1-phosphate receptor 1 | 2.1421 | 1.44E-09 | −1.592 | 2.52E-06 |
| CCL21 | Chemokine (C-C motif) ligand 21 | 2.1272 | 2.65E-07 | −1.9191 | 5.97E-05 |
| HTATIP2 | HIV-1 Tat interactive protein 2, 30 kDa | 2.1085 | 1.19E-04 | −1.6651 | 1.88E-06 |
| SOX18 | SRY (sex determining region Y)-box 18 | 2.0969 | 9.81E-08 | −1.699 | 5.79E-05 |
| CREG1 | Cellular repressor of E1A-stimulated genes 1 | 2.0799 | 4.55E-05 | −1.6091 | 3.35E-05 |
| PECAM1 | Platelet/endothelial cell adhesion molecule | 2.0659 | 3.79E-08 | −1.6298 | 9.27E-06 |
| CRIM1 | Cysteine rich transmembrane BMP regulator 1 | 2.0092 | 1.84E-08 | −1.6596 | 1.51E-05 |

**Figure 1-16 Table of hypermethylated genes in LM compared to MM.**

Summary of hypermethylated and transcriptionally downregulated genes in uterine leiomyoma compared with adjacent normal myometrium (Adapted from ref. [95]).

## 1.5    Introduction of the *RANKL* gene

### 1.5.1    Identification of RANK/RANKL pathway

The gene encoding the Receptor Activator of Nuclear factor Kappa-B Ligand (RANKL) protein, *TNFSF11*, is located on the human chromosome 13q14. RANKL has two major variants, a type II transmembrane protein and a secreted protein; they are related to several different functions: an apoptosis regulatory gene, a ligand for RANK, and a binding partner of osteoprotegerin (OPG)  [98].

RANKL and its receptor RANK, both from the TNF superfamily, were first identified as key regulators of osteoclast development and bone metabolism [99] (Figure 1-17). Bones are constantly undergoing remodeling through the synthesis of bone matrices by osteoblasts and resorption of bone by osteoclasts. RANKL, RANK, and osteoprotegerin (OPG) are highly involved in the homeostasis of bone development. OPG is the decoy receptor for RANKL and inhibits osteoclast development. Upon stimulation by the bone-resorbing factors, such as PTHrP, IL-1, IL-11, IL-17, and TNF-α, RANKL was produced and released from osteoblasts to the niche, which can bind to RANK present at the surface of osteoclast [99]. This paracrine action can drive osteoclast development from hematopoietic progenitor cells as well as activate mature osteoclast [98]. OPG, also present in the niche, negatively regulates RANKL binding to RANK, inhibiting bone turnover by osteoclasts.

OPG expression can be regulated by estrogen, providing possible connections between sex hormones and osteoporosis in postmenopausal women. After menopause, there is a drastic decrease in estrogen levels in the blood, leading to a reduction of OPG production and an overly-active RANKL signal. As a result, osteoclast numbers increase and bone turnover enhances, which causes osteoporosis, a significant health problem in the elderly population [98].
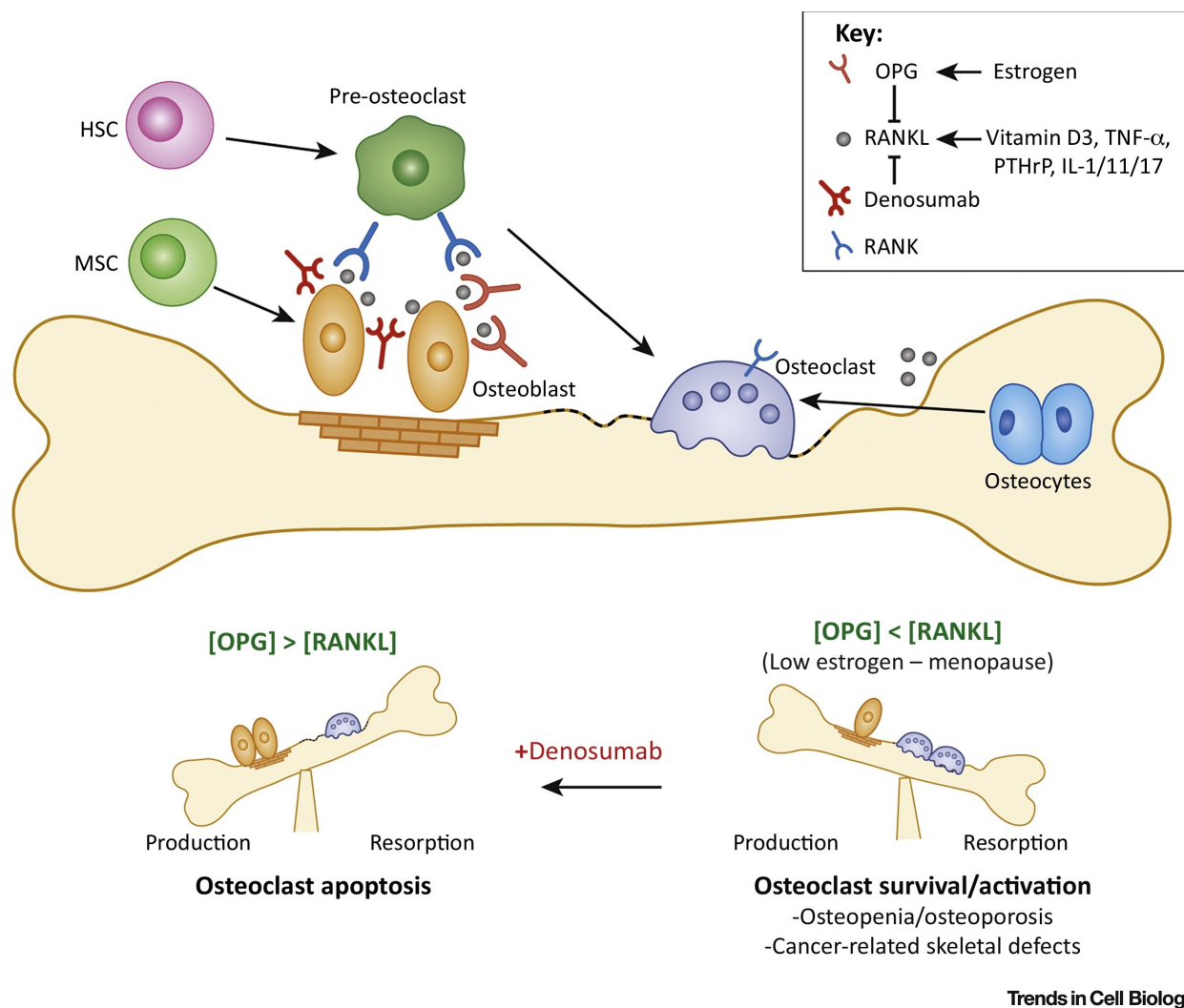
**Figure 1-17 RANKL/RANK/OPG pathway during bone remodeling.**

Upon stimulation by the bone-resorbing factors, such as PTHrP, IL-1, IL-11, IL-17, and TNF-α, RANKL was produced and released from osteoblasts to the niche, which can bind to RANK present at the surface of osteoclast. This paracrine action can drive osteoclast development from hematopoietic progenitor cells as well as activate mature osteoclast. OPG can be regulated by estrogen, IL-4 and TGF-β, and et al and negatively regulates RANKL binding to RANK inhibiting bone turnover by osteoclast. The balance between OPG and RANKL is the key to the homeostasis during bone remodeling. (Adapted from ref. [99]).

RANKL is expressed in a variety of tissues: it is highly expressed in lymph nodes, the thymus, and lungs, but it is also expressed at a lower levels in the spleen, bone marrow, peripheral blood, leukocytes, heart, placenta, skeletal muscle, stomach and the thyroid [98]. Due to this widely spread expression, RANKL is regulated by different factors. In osteoblasts, RANKL is known to be regulated by parathyroid hormone (PTH) and 1,25(OH)2D3 [100]. The CREB binding site was identified upstream of RANKL's transcription start site. Upon PTH or 1,25(OH)2D3 stimulation, CREB is recruited to the RANKL promoter to activate its expression. WNT signals serve as a negative regulator of bone resorption by suppressing RANKL expression and upregulating OPG levels [101]. RANKL expression is also governed by DNA methylation in bone tissue. Two CpG-rich regions were identified, one upstream located at -14,415 bp from the TSS of the isoform I and one downstream located -260 bp from isoform I TSS. The level of DNA methylation at these two regions is inversely correlated with RANKL mRNA levels [102].

### 1.5.2   Paracrine action of RANK/RANK pathway in mammary gland

The RANKL/RANK pathway was recently reported to play a critical role in the development of the mammary gland and breast cancer carcinogenesis. Several studies support the notion that RANKL expression is regulated by progesterone signaling in the mammary gland. During murine pregnancy, changes in mammary biology and structure are associated with RANKL expression. RANKL protein is mainly observed in the regions of lobular branching; its level is positively linked with the progesterone level in serum [103]. Administration of PR agonist medroxyprogesterone acetate (MPA) resulted in a 3000-fold increase of RANKL in mammary epithelial cells in mouse model [104]. A similar phenomenon is observed in humans, with the highest RANKL level in the luteal phase of the menstrual cycle when progesterone levels peak [105].
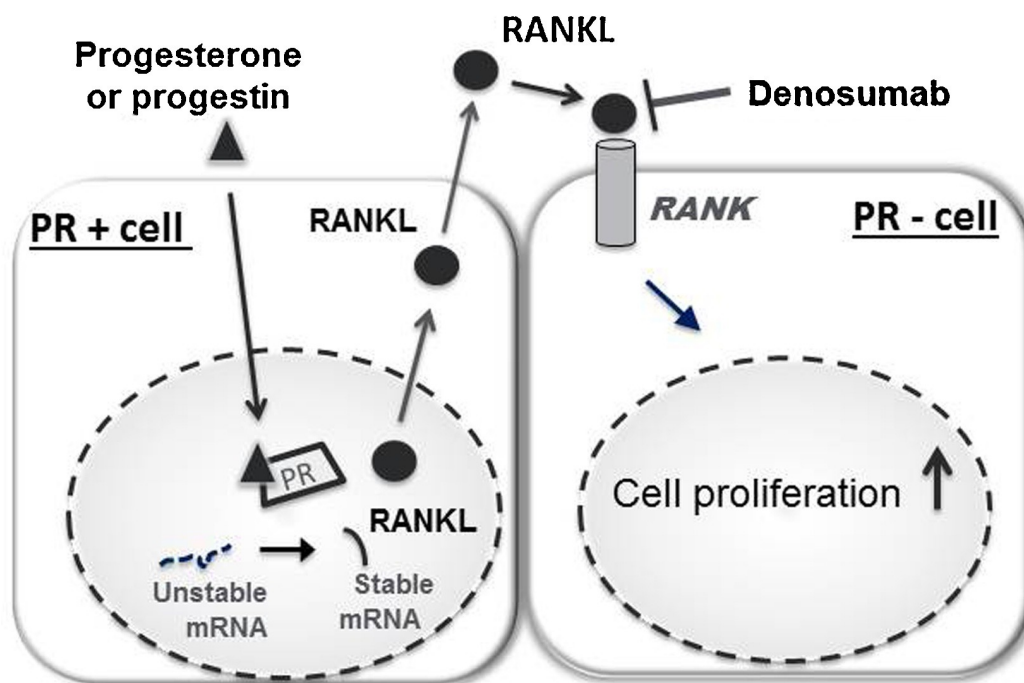
**Figure 1-18 Mechanisms of signal transduction of RANKL/RANK and PR in PR+ and PR- breast luminal cells.**

Progesterone stimulates RANKL expression in the PR positive cell. RANKL binds to its receptor RANK, which is expressed on the surface of the PR-negative cells. Through RANKL/RANK signaling and paracrine signal transduction, progesterone is able to stimulate proliferation of the PR-negative cells. (Adapted from ref. [106])

During investigations of the RANKL/RANK pathway in the mammary gland, researchers found a similar phenomenon as the bone remodeling process: the RANKL/RANK pathway is important as a paracrine mediator of progesterone-induced proliferation in the adult mouse mammary gland. There are two major epithelial cell types in mammary tissue: one is luminal epithelial responsible for the production of milk and the other is basal myoepithelial cells having a contractile function during lactation (Figure 1-18) [107]. Among the epithelial cells, only 30-50% are positive for estrogen and progesterone receptors [108]. The RANKL/RANK pathway has an indispensable role in mediating the crosstalk between the PR+ and PR- cells upon steroid hormone stimulation, which leads to the proliferation of epithelial cells. RANKL is mostly

expressed in PR+ cells, while its receptor RANK is expressed in the PR- cells. Upon progesterone stimulation, RANKL is released and binding of RANKL to RANK activates NF-κB, PI3K/AKT, and terminal MAPK kinases P38, ERK, and JNK, and regulates mammary epithelial cell proliferation and survival. In the murine mammary gland, RANKL mediates signaling from PR+ luminal cells to PR- stem cells, leading to the proliferation of mammary stem cells in the normal mammary gland and breast cancer [106, 109]. By deleting RANKL expression or inhibiting RANKL function, researchers observed a marked decrease of epithelial cell proliferation, hormone-induced tumorigenesis, and breast cancer incident rate [104, 110]. These indicate that the RANKL/RANK system contributes greatly to the incidence and onset of progesterone-driven breast cancer.

### 1.5.3   Roles of RANK/RANKL pathway in leiomyoma.

Similar roles of the RANKL/RANK pathway were also discovered in LM compared with the mammary gland. The Bulun lab found a specific pattern of RANKL/RANK expression in LM cell populations. PR-deficient LSCs (CD34+/CD49b+ cells) express the highest RANK levels, whereas the levels of RANKL peak in PR-rich differentiated populations (LIC: CD34+/CD49b- cells) (Figure 1-19 A) [68]. RANKL mRNA is significantly increased by progestin R5020, specifically in the LIC population. Denosumab is an FDA-approved human neutralizing monoclonal antibody for RANKL. It has shown promising results when treating metastatic bone disease or osteoporosis, which are known to be caused by a high RANKL level. Ikhena et al demonstrated that denosumab is able to block progesterone-dependent proliferation signals and LM growth under both *in vitro* and *in vivo* conditions (Figure 1-19 B). An interesting picture is emerging, whereby activation of the RANKL/RANK pathway by estrogen and progesterone action and growth via stem cell renewal ultimately give rise to clonal expansion of LM (Figure 1-19 C). The detailed mechanism of RANKL-dependent proliferation and the regulation of RANKL expression in LM are evaluated in this dissertation.
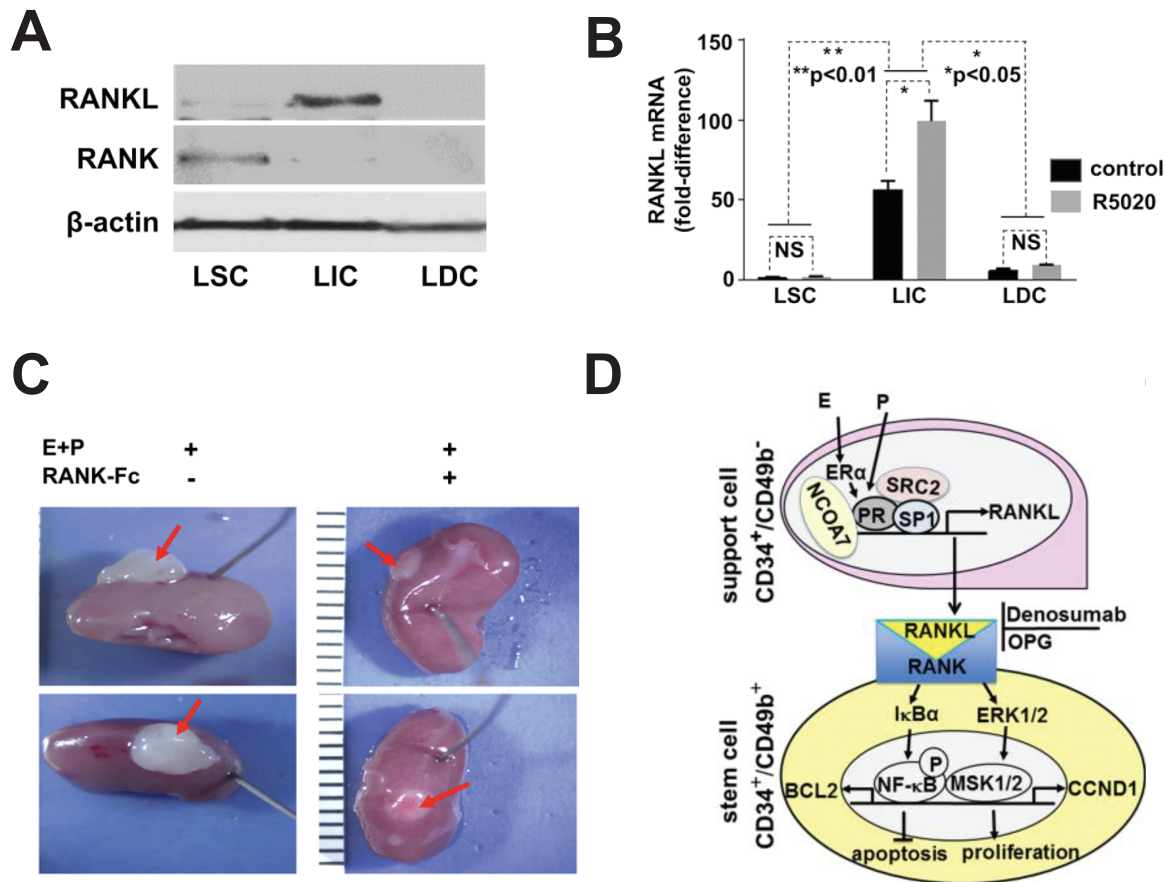
**Figure 1-19 RANKL/RANK pathway in LM**

**A**) Western blot analysis shows the highest RANKL and RANK protein levels in LICs and LSCs, respectively. **B**) R5020 significantly induced RANKL expression in LICs. LSCs, LICs, and LDCs were sorted from leiomyoma explants treated with R5020 for 24 hours. RANKL mRNA levels were then quantified in each cell population. Data indicate representative experiments performed in triplicate from three independent tissue samples. **C**) Cell pellets consisting of total leiomyoma cells were grafted underneath kidney capsules of ovariectomized NSG mouse hosts and mice were treated with vehicle or RANK-Fc (10 mg/kg subcutaneously twice a week for 4 weeks). Significantly smaller tumors were excised from mice treated with RANK-Fc compared with mice treated with vehicle. **D**) Proposed interactions between leiomyoma stem cells and intermediately differentiated support cells via RANKL/RANK for estrogen and progesterone stimulation of cell survival and proliferation. Denosumab (a monoclonal anti-RANKL antibody) and OPG (decoy receptor of RANKL) are anticipated to disrupt this paracrine pathway. (Adapted from ref. [68])

CHAPTER 2      Mechanisms of dysregulated *RANKL* Gene Expression in

Uterine Leiomyoma: Involvement of progesterone action, epigenetic

modification and *MED12* gene mutations

## 2.1 Introduction

Uterine leiomyomas (LM, a.k.a., fibroids) represent the most common gynecological tumors in women. By age 50, about 80% of women will develop at least one LM, and 15-30% will develop severe symptoms, including excessive uterine bleeding, recurrent pregnancy loss, and pelvic pain; these symptoms may mimic or mask malignant tumors [3]. Approximately 200,000 hysterectomies and 30,000 myomectomies are performed to treat LM in the US annually, costing up to $34.4 billion [111]. Considering the heavy socioeconomic burden and associated long-term health problems, there is an urgent need to develop non-surgical treatments for LM.

Progesterone (P4) and its receptor, PR, play essential roles in LM development; however, the cellular and molecular mechanisms responsible remain unclear [49]. Anti-progestins have proven useful in the medical management of LM; however, long-term usage is limited by their side effects and rapid tumor recurrence after treatment cessation [2, 49]. Thus, understanding the molecular mechanisms underlying P4/PR action in LM pathogenesis is critical for the development of targeted and effective therapies for this disease.

Receptor activator of nuclear factor κB ligand (RANKL), a P4/PR target gene, recently became a key actor in oncology [112]. The molecular axis of RANKL and its receptor RANK is involved in all stages of tumorigenesis [112]. In mouse mammary gland, progestins substantially stimulate RANKL expression in PR-positive cells, leading to paracrine activation of PR-negative mammary stem cells and progression of progestin-driven mammary cancer [104, 113, 114]. Later studies in human have demonstrated the positive correlation between serum P4 levels and RANKL mRNA and protein levels in normal and malignant breast tissues [109, 115]. We recently demonstrated that blocking the RANKL/RANK pathway inhibits steroid hormone-mediated LM development in a xenograft mouse model, indicating that this pathway plays a critical role in LM pathogenesis [68]. However, the mechanisms responsible for the hormone responsiveness of RANKL gene expression and the role of RANKL in LM formation require further elucidation.

Epigenetic alterations, such as DNA methylation, play crucial roles in regulating gene expression and disease progression [116]. DNA methylation interferes with the interactions between DNA and specific transcription factors (TFs) and chromatin proteins [117]. Abnormal DNA methylation and expression of its regulatory enzymes have been reported in LM [97, 118]. However, the role of DNA methylation in P4/PR-mediated transcriptional regulation in LM has not been explored.

Mutation of mediator complex subunit 12 (MED12) is the most prevalent mutation (mut-MED12) in LM, occurring in over 70% of all LMs [16]. LM-associated mut-MED12 may alter interactions with proteins involved in transcriptional co-activator pathways [119]. For instance, mut-MED12 disrupts the MED12-Cyclin C binding interface, leading to a loss of mediator-associated CDK activity [22]. Introduction of the most frequent mut-MED12 subtype, G44D, in mice causes LM development in the presence of estradiol (E2) and P4. So far, the hormone-dependent mechanism whereby mut-MED12 causes LM formation remains unclear.

Here, we used RANKL as a representative P4/PR target gene to study the genetic and epigenetic mechanisms underlying dysregulated RANKL expression in LM, which will also shed mechanistic light on our understanding of RANKL gene regulation and PR action in breast cancer and other steroid-responsive tumors [112].

## 2.2    Results

### 2.2.1    RANKL is highly overexpressed in LM versus adjacent normal MM

We examined *in vivo* RANKL mRNA levels in 86 matched MM and LM tissues (n=43 patients). RANKL mRNA was significantly higher in LM compared with matched MM tissues (Figure 2-1 A). Immunoblot analysis showed that RANKL protein level was also markedly upregulated in LM (Figure 2-1 B). Immunohistochemistry revealed more intense and abundant levels of RANKL protein distributed in the cytoplasm and intercellular matrix in LM tissues (Figure 2-1 C).

### 2.2.2    RANKL stimulates LM stem cell proliferation

There are three distinct cell populations in LM based on CD34 and CD49b expression: CD34$^+$/CD49b$^+$ (LSC, LM stem/progenitor cells), CD34$^+$/CD49b$^-$ (LIC, intermediate cells), and CD34$^-$/CD49b$^-$ (LDC, differentiated cells).[64] Based on RANKL's established function in breast cancer, we hypothesized that RANKL plays a critical role in LM pathogenesis by stimulating LSC proliferation. Hence, we treated LM tissue explants with synthetic human RANKL or vehicle for 24h and analyzed the percentage of LM cell populations. RANKL treatment markedly expanded the LSC population compared to control (8.84±1.23% vs 4.87±0.33%) but decreased the percentage of LDC population (64.97±1.88% vs 69.96±0.53%); whereas the LIC population was not significantly affected (Figures 2-2 A and B).

To evaluate the mechanism underlying the percentage change of LM cell populations after RANKL treatment, we conducted proliferation (CCK8 assay) and apoptosis (Annexin V staining) assays. RANKL specifically increased LSC proliferation (Figure 2-2 C) without affecting cell survival (Figure 2-2 D). Furthermore, RANKL significantly increased expression of proliferation gene (Cyclin D1) and stem cell factors (KLF4 and NANOG) particularly in LSC, but did not affect the expression levels of differentiation markers (ESR1, PGR, and ACTA2) (Figure 2-2 E). These findings indicate that RANKL promotes LM growth through inducing proliferation selectively in the stem cell population.
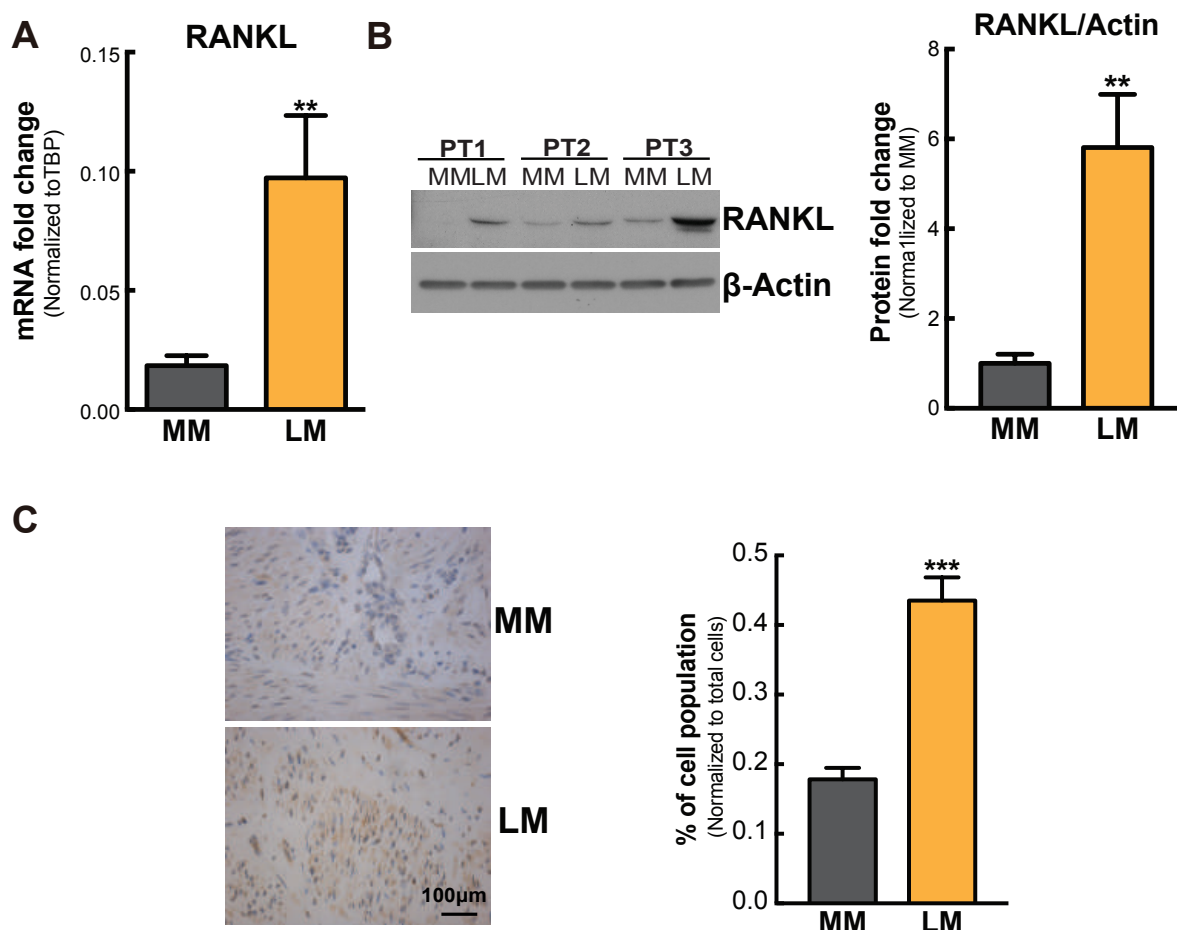
**Figure 2-1. RANKL is highly expressed in LM**

**A**) Histogram showing mRNA levels of RANKL in LM and MM tissues (means ± SEM, n=43, **P<0.01, paired t-test). **B**) Left panel, representative immunoblot showing RANKL protein levels in LM and MM tissues; Right panel, ImageJ quantification of immunoblots (means ± SEM, n=6, **P<0.01, paired t-test). PT: patient. **C**) LM and MM tissue sections were stained with anti-RANKL antibody. Left panel, representative images of immunohistochemical staining of RANKL in MM and LM tissues; Right panel, percentage of RANKL-positive cells among total cells counted (means ± SEM, n=10, ***P<0.001, paired t-test).
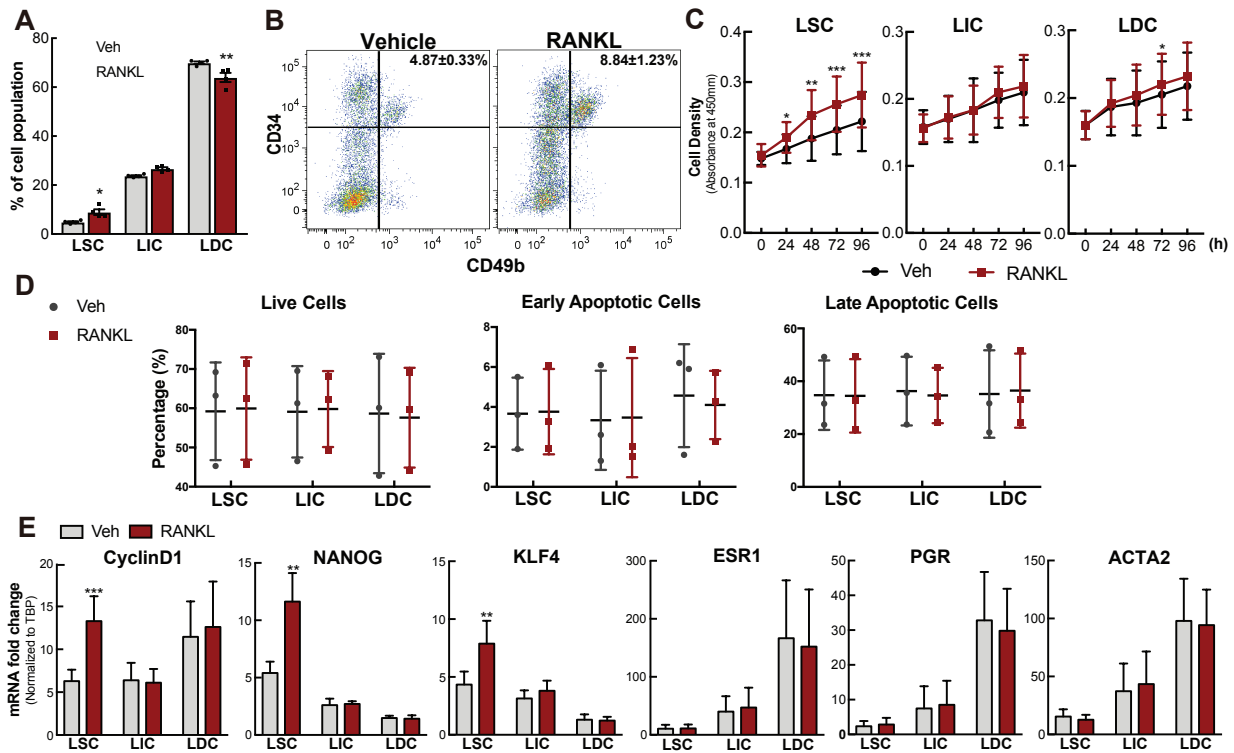
**Figure 2-2. RANKL stimulates LSC proliferation.**

**A**) Percentage of LM cell populations isolated from tissue explants after 24h incubation with vehicle (0.1% BSA in PBS) or RANKL (100 ng/ml) (means ± SEM, n=4, **$P<0.01$ and *$P<0.05$, paired t-test). LM cells were stained with anti-CD45 (depleted hematopoietic cells), anti-CD34, anti-CD49b, and PI (depleted dead cells) and analyzed by the LSR Fortessa system. **B**) Representative FACS scattergram for panel D. **C**) Line graph showing the proliferation rate of each LM cell population treated with vehicle (0.1% BSA in PBS) or RANKL (100 ng/ml) for different time indicated on the graph. (means ± SEM, n=3, ***$P<0.001$, **$P<0.01$, and *$P<0.05$, paired t-test). **D**) Scatter graph showing apoptosis of each LM cell population after 24h incubation with vehicle (0.1% BSA in PBS) or RANKL (100 ng/ml) (means ± SEM, n=3, paired t-test). LM cells were stained with anti-Annexin V and PI (50μg/ml) and analyzed by the LSR Fortessa system. **E**) Regulation of proliferation, stem cell, and differentiation related gene expression by RANKL treatment in each LM cell population determined by real-time PCR (means ± SEM, n=5, **$P<0.01$ and ***$P<0.001$, paired t-test).

### 2.2.3   RANKL expression is differentially regulated by P4/PR signaling in LM versus MM

Our finding that RANKL plays a critical role in regulating LSC function and LM tumorigenesis prompted us to further explore the mechanism underlying dysregulated RANKL gene expression in LM.[68] RANKL is a known P4/PR target gene in mammary glands [120, 121]; therefore, we postulated that the RANKL gene responds to P4/PR signaling differentially between LM and MM tissues, leading to higher RANKL expressions in LM. To test this, we quantified RANKL mRNA levels in MM and LM tissue explants treated with the P4 agonist R5020 for 48h. R5020 increased RANKL expression in both LM and MM explants, but the fold change was significantly higher in LM, suggesting RANKL gene transcription activity is more sensitive to R5020 in LM cells than in MM cells (Figure 2-3 A, 13.10±1.02 LM vs 7.75±0.63 MM). The R5020-mediated induction of RANKL expression in both tissues was completely blocked by co-treatment with RU486 (P4 antagonist) in explants (Figure 2-3 A) or PR knockdowns (Figure 2-3 B) in cells, suggesting that R5020 regulates RANKL expression through PR.

We examined our previously published PR ChIP-sequencing (ChIP-Seq) data in LM cells and found multiple PR-binding sites (PRBS) around RANKL gene locus.[44]  A region 87kb upstream (-87,360/-86,731bp) of RANKL transcription start site (TSS) has been identified to display the strongest PR-binding activity among all binding sites (Figure 2-4 A) and is enriched for DNA binding motifs of SP-1 and progesterone response element (PRE) half sites (PROMO V3.0.2), which are important for PR-chromatin interaction.[122] We designed four sets of primers covering the entire distal PRBS and performed ChIP-qPCR to examine the enrichment of PR-binding using chromatin isolated from fresh-frozen LM and MM tissues. We detected higher PR-binding activity in LM than in MM on four sets of primers spanning Chr13:43,060,931-43,061,560 (Figures 2-4 B and C). Regression analysis demonstrated that the enhanced PR-binding activity in this region was positively correlated with RANKL mRNA level ($R^2$=0.8252), suggesting that the distal PRBS is an enhancer region that regulates RANKL transcription (Figures 2-4 D).

Using PR ChIP-Seq, we also detected PR binding activity in the RANKL promoter region (-2,200/-1,000bp). This region has strong gene transcription regulatory activity and is enriched with DNA binding motifs of AP-

1 and PRE half sites, which are important for PR recruitment to DNA.[123-125] To elucidate whether the -87kb distal PRBS was associated with the proximal promoter, we first determined whether PR was recruited to the promoter region in LM using ChIP-qPCR. We found that PR enrichment in the RANKL promoter (-1,256/-1,118bp) was not only higher in LM versus MM tissue, but also positively correlated ($R^2$=0.7407) with PR-binding at the distal PRBS (Figures 2-4 E and F).

Next, we performed chromatin conformation capture (3C)-qPCR assay to further confirm the chromatin interaction between the distal PRBS and proximal promoter region. As shown in Figure 2-5 A, we evaluated 4 fragments (En1, En2, En3, and En4) around PRBS, and found the interaction strength was significantly stronger in LM versus MM with the highest interaction observed between promoter and En2, which is the region overlapped with the distal PRBS. No interaction was detected between the RANKL promoter and En4 region, which is 3.8 kb away from the distal PRBS. Sanger sequencing of the amplicon of promoter and En2 confirmed the expected ligation product from loop formation (Figure 2-5 B). This suggests that the distal PRBS regulates RANKL gene expression through interaction with the proximal promoter. PR was not recruited to the RPL30 housekeeping gene promoter, and IgG did not contain enriched chromatin in the RANKL or the RPL30 gene, indicating that PR enrichment in the RANKL regulatory region is specific (Figure 2-5 C).
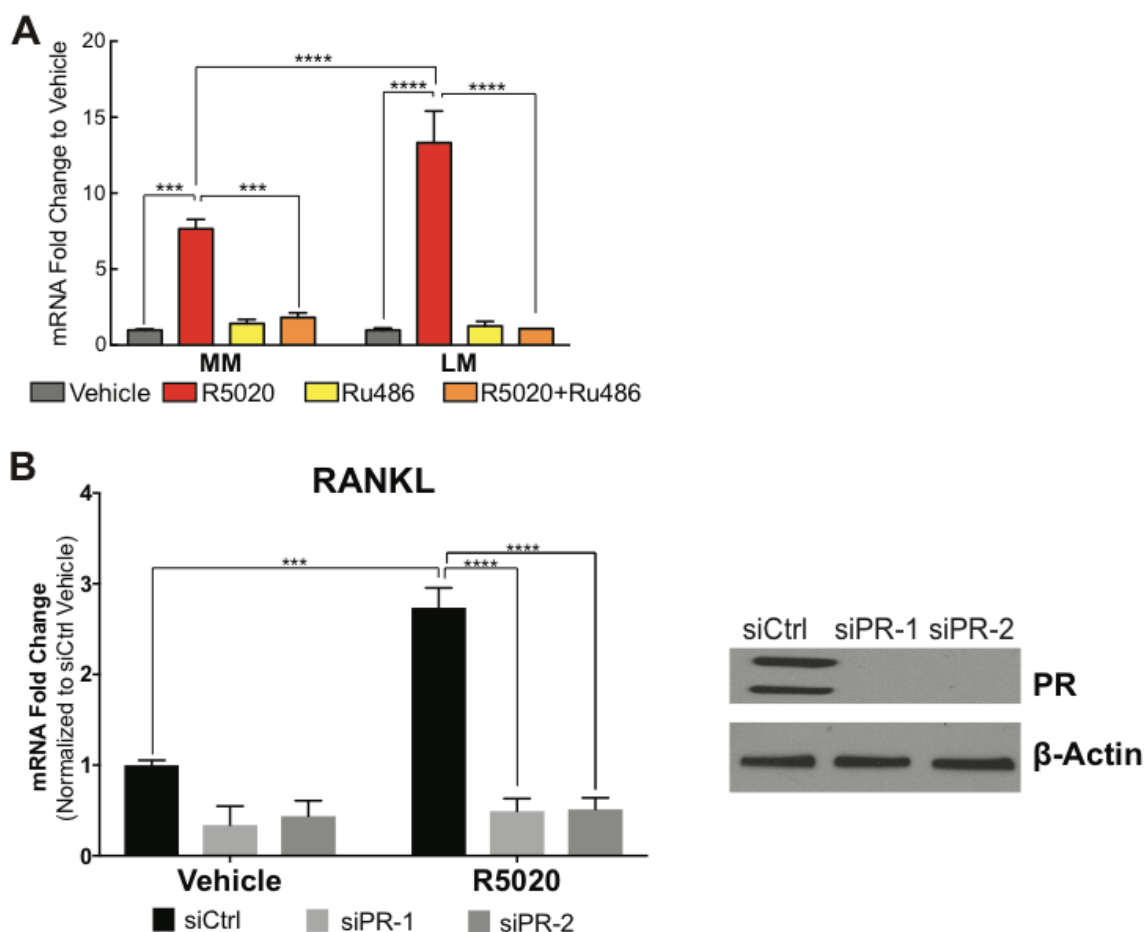
**Figure 2-3. RANKL mRNA level is regulated by progestin.**

**A**) Relative RANKL mRNA levels normalized to vehicle-treated cells were determined by real-time PCR in MM and LM tissue explants treated with R5020 (10-7 M) or vehicle (ethanol) in the presence or absence of RU486 (10-6 M) for 48h (means ± SEM, n=6, ***P<0.001 and ****P<0.0001, paired two-way ANOVA). **B**) Left panel: Bar graph showing qPCR results of RANKL mRNA levels in LM cells transfected with 2 different PR siRNAs (siPR-1 and siPR-2)  or scrambled control siRNA (siCtrl)  for 96 h followed by treatment with or without R5020 for 24h. (means ± SEM, n=4, ***P<0.001 and ****P<0.0001, paired two-way ANOVA). Right panel: Representative immunoblot showing PR level with or without PR knockdown.
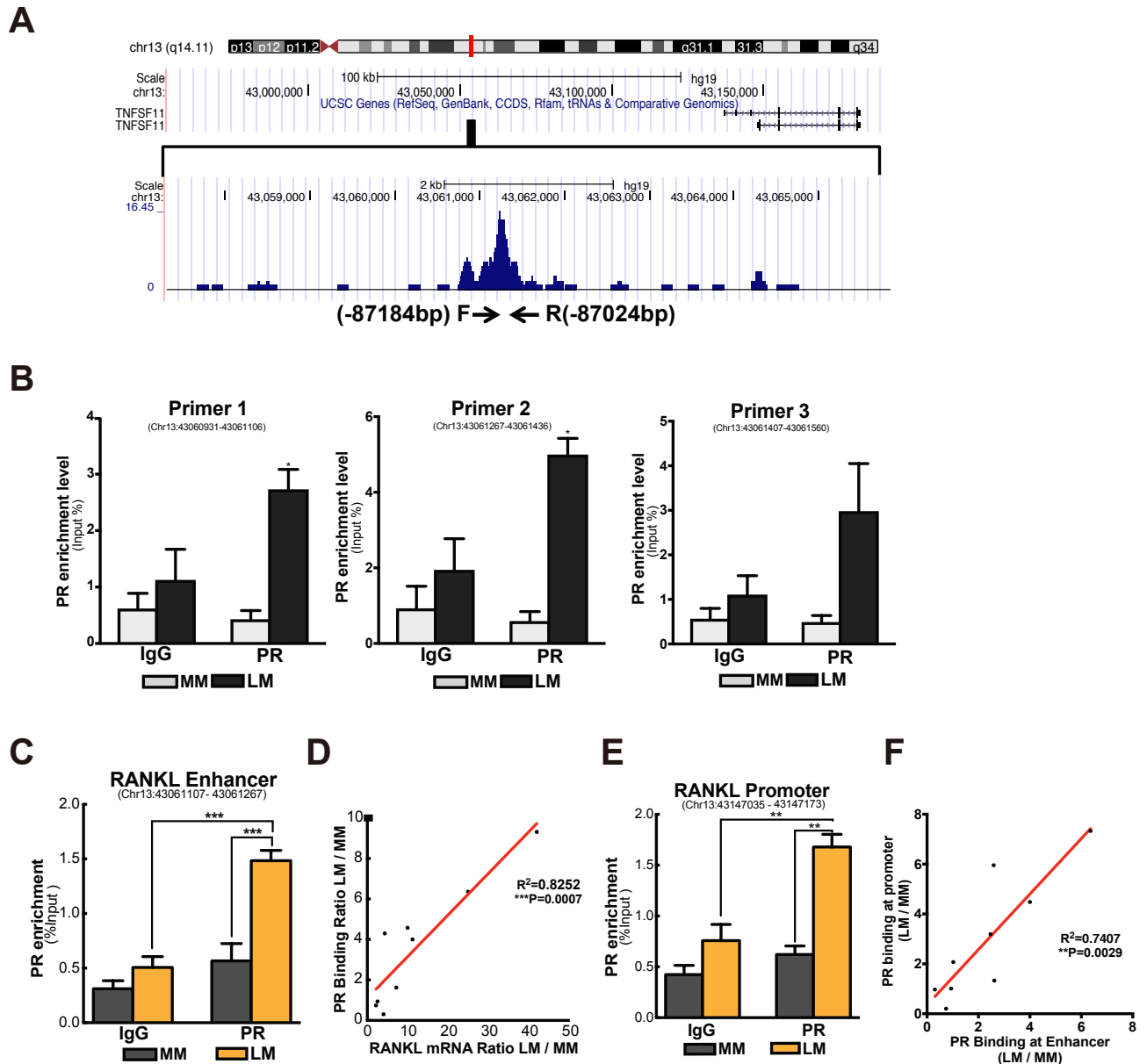
**Figure 2-4. RANKL expression level is positively correlated with PR enrichment at PRBS.**

**A)** UCSC Genome Browser ChIP-Seq Track showing PR-binding peak located 87kb upstream of the RANKL transcription start site in LM. PROMO (V3.0.2) revealed enriched SP-1 motif and half-PRE sites in the region. **B**) Bar graph showing the validation results of PR ChIP-Seq. Primers were designed to target 3 DNA fragments around the RANKL distal PR-binding sites: -87,360bp/-87,185bp (chr13:43,060,931-43,061,106), -87,024bp/-86,855bp (chr13:43,061,267-43,061,436) , and -86,884bp/-86,731bp (chr13:43,061,407-43,061,560) upstream of the RANKL transcription start site. (means ± SEM, *P<0.05, paired two-way ANOVA, n=3). **C**) ChIP-qPCR showing PR enrichment at the RANKL distal PR-binding site

(PRBS, -87,184bp/-87,024bp, Chr13:43,061,107- 43,061,267) (means ± SEM, n=10, \***P<0.001, paired two-way ANOVA). **D)** Pearson correlation showing positive correlation between PR enrichment at the distal PRBS and RANKL mRNA levels. Data represents PR enrichment and RANKL mRNA levels in LM normalized to those in matched MM from nine subjects ($R^2$=0.8252, \***P<0.001, Pearson correlation). **E**) ChIP-qPCR showing PR enrichment at the RANKL proximal promoter (-1,256/-1,118bp, Chr13:43,147,035- 43,147,173) (means ± SEM, n=10, \*\*P<0.01, paired two-way ANOVA). **F**) Scatter plot showing PR enrichment at the RANKL distal PRBS positively correlates with PR enrichment at the RANKL promoter (n=9, $R^2$=0.7407, \*\**P*<0.01, Pearson correlation).
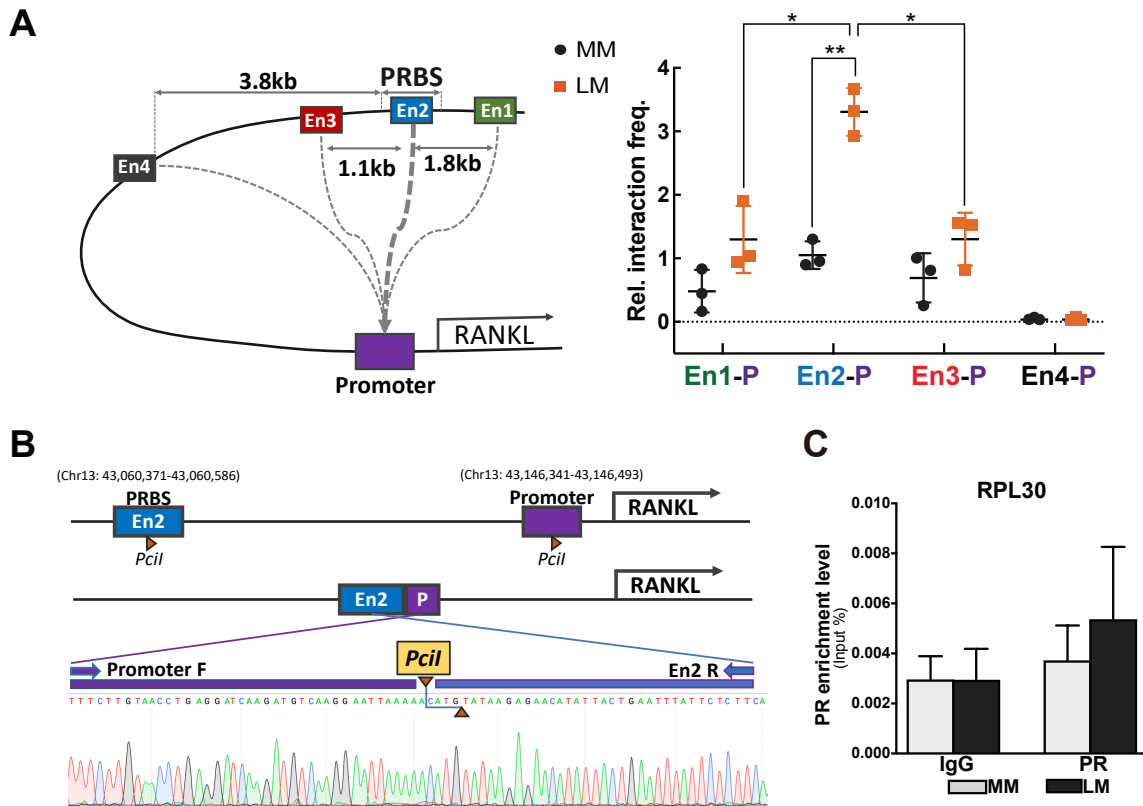
**Figure 2-5. PRBS interacts with RANKL promoter.**

A) Left panel: illustration of examined chromatin interactions of the RANKL promoter and distal PRBS. The promoter region was designed as the "bait" and four "interrogated fragments" were showed around the PRBS with En2 overlapped with PRBS. All investigated regions were located within 150bp of *PciI* sites. Right panel: 3C-qPCR (dot plot) showing RANKL promoter and distal enhancer interactions in MM and LM primary cells (means ± SEM, n=3, *$P<0.05$, **$P<0.01$, paired two-way ANOVA). Results are presented as relative interaction frequencies normalized by GAPDH as an internal control. P: promoter; En1, En2, En3, and En4 represent 4 DNA fragments evaluated. B) Schematic diagram of RANKL promoter and En2 regions following *PciI* excision and ligation. Sequence chromatogram of excised band with regions of interest denoted. D) Bar graph showing IgG and PR enrichment at the RPL30 gene. (means ± SEM, n=3)

**A**



**B**
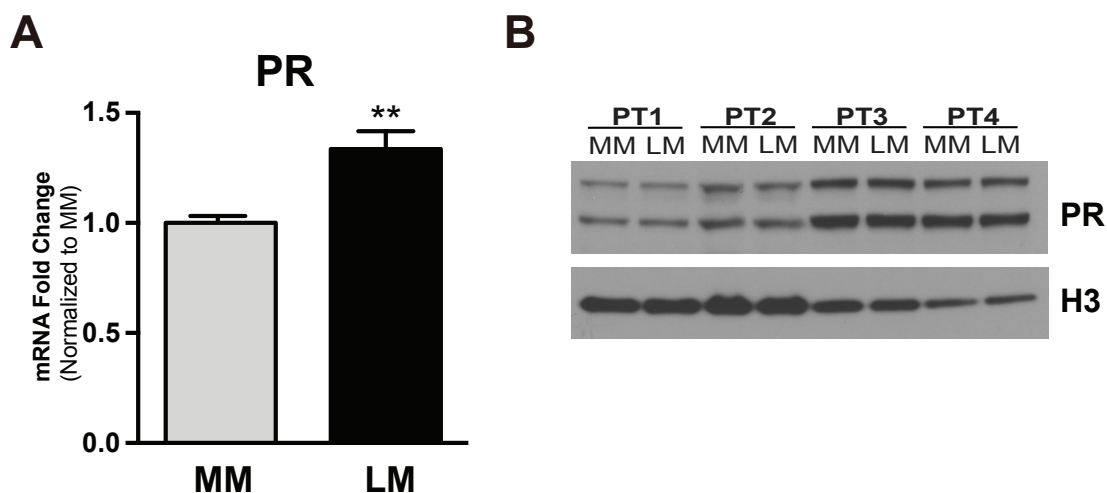


**Figure 2-6. PR protein level between MM and LM.**

**A**) Relative PR mRNA level in MM and LM tissue (means ± SEM, n=13, **P<0.01, paired Student t-test).

**B**) Immunoblot showing PR protein levels in chromatin isolated from LM and MM tissues from 4 patients. H3 was used as loading control of chromatin. PT: patient.
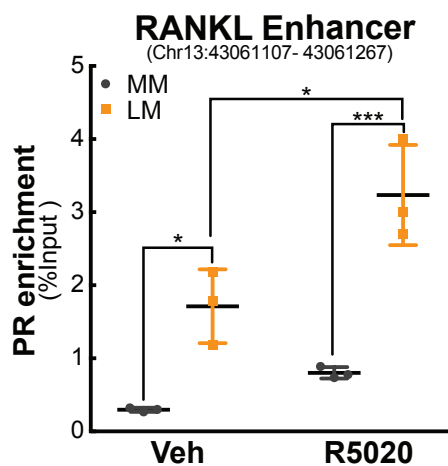


**Figure 2-7. PR recruitment to PRBS with or without P4 stimulation.**

PR ChIP-qPCR showing PR recruitment to the distal PRBS in primary MM and LM cells exposed to vehicle or R5020 for 1h (means ± SEM, n=3, *P<0.05 and ***P<0.001, paired two-way ANOVA).

To examine whether the enhanced PR-binding around the RANKL gene in LM was due to differential PR expression levels between the two tissues, we measured PR mRNA and protein levels in the tissues used in the ChIP-qPCR. The PR mRNA level in LM was only marginally higher (1.3-fold) than in MM tissue (Figure 2-6 A), with the comparable PR protein level between MM and LM tissue chromatin (Figure 2-6 B). These findings suggest that the specific epigenetic context of the distal enhancer region, not the PR expression level, leads to increased recruitment of PR to the RANKL gene in LM tissue.

To determine whether PR-binding activity at the RANKL distal PRBS is ligand-dependent, primary LM and MM cells were treated with R5020 for 1h followed by ChIP-qPCR. Consistent with our findings in LM and MM tissues, basal PR-binding affinity at the enhancer region was higher in LM versus MM cells (Figure 2-7). R5020 increased PR-binding to the distal enhancer in both LM and MM cells; however, R5020-mediated PR enrichment was still higher in LM cells.

### 2.2.4 Active histone marks are highly enriched in distal enhancer and proximal promoter regions of the RANKL gene

Post-translational modification of histone tails plays an important role in regulating gene transcription. High levels of H3K4me3 are commonly found in the promoters of actively transcribed genes, whereas H3K27me3 enrichment is associated with inactive promoters.[126] ChIP-qPCR of histone modifications revealed that H3K4me3 enrichment at the RANKL promoter was higher in LM than in MM, whereas H3K27me3 levels were higher in MM (Figures 2-8 A and B). In the distal PRBS, H3K27Ac (active enhancer mark) was more highly enriched in LM versus MM (Figure 2-8 C), confirming the enhancer characteristic of this region. These histone modification patterns not only support our observation of higher RANKL transcription in LM, but also indicate a more accessible chromatin structure adjacent to the RANKL gene in LM.
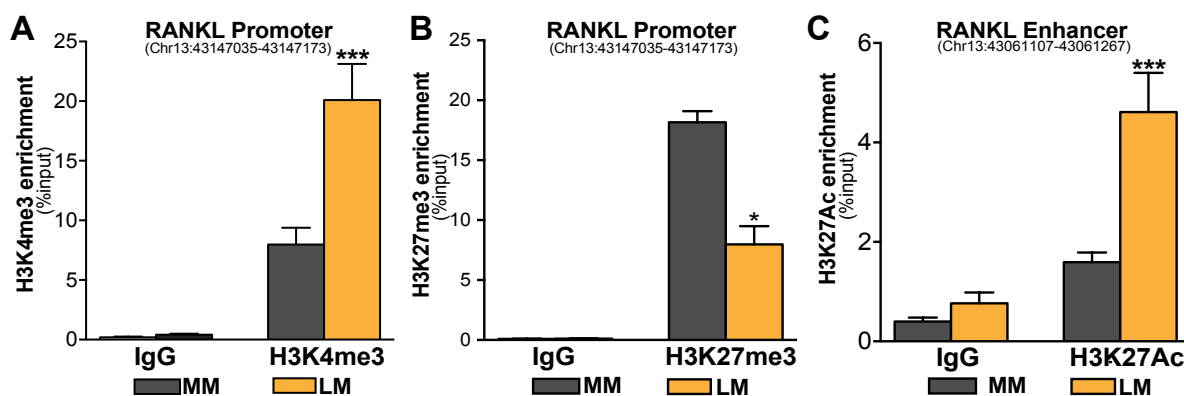
**Figure 2-8. Differential histone modification patterns at the RANKL proximal promoter and distal PRBS in MM and LM tissues.**

ChIP-qPCR assay showing histone marks H3K4me3 **(A)** and H3K27me3 **(B)** enrichment at the RANKL proximal promoter, and H3K27Ac **(C)** enrichment at distal PRBS in MM and LM tissues (n=5, means ± SEM, *$P$<0.05 and ***$P$<0.001, paired two-way ANOVA).

### 2.2.5    DNA methylation mediates RANKL gene expression

Studies have shown that DNA methylation at the CpG island adjacent to the RANKL promoter mediates RANKL expression in bone and mammary gland tissues.[127, 128] To evaluate whether RANKL transcription is regulated by DNA methylation in MM and LM, we treated primary LM and MM cells with the DNA methyltransferase inhibitor 5-aza-2′-deoxycytidine (5'-aza) for 4 days. We found that RANKL mRNA levels in MM cells were more dramatically increased by 5'-aza compared with LM (Figure 2-9 A). The mRNA levels of RANK and osteoprotegerin (OPG, a decoy RANKL receptor) were not affected by 5'-aza, suggesting that DNA methylation specifically regulates RANKL expression in MM and LM cells (Figures 2-9 B and C). In addition, we demonstrated that 5'-aza consistently induced RANKL expression in MM cells in the presence of the protein synthesis inhibitor cycloheximide, suggesting that 5'-aza regulates RANKL expression by directly modulating its DNA methylation (Figure 2-9 D).
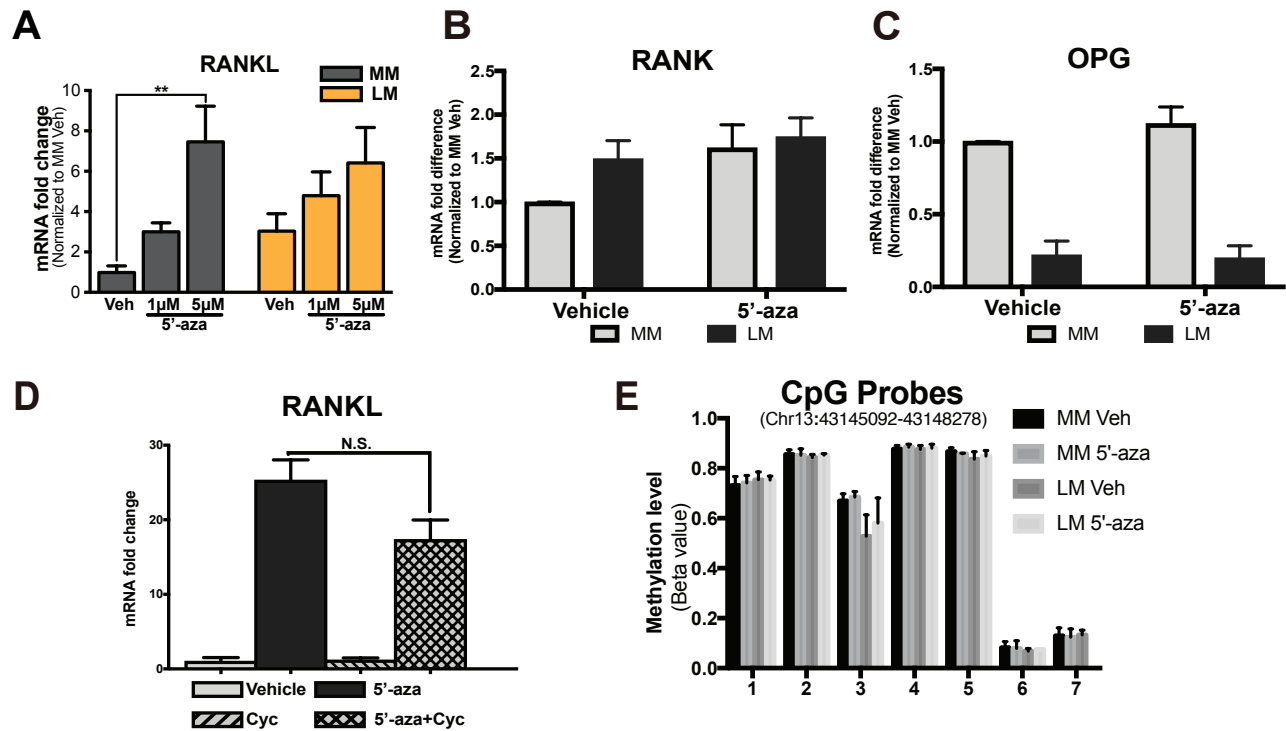
**Figure 2-9. RANKL expressed is inhibited by DNA methylation.**

**A)** Histogram showing RANKL mRNA levels in primary MM and LM cells treated with or without 5'-aza for four days (means ± SEM, n=5, **$P<0.01$, paired one-way ANOVA). Bar graphs showing mRNA levels of RANK (**B**) and OPG (**C**) genes in primary MM and LM cells treated with or without 5'-aza for four days. (means ± SEM, n=3). **D**) Bar graph showing RANKL mRNA levels treated with 5'-aza (5μM, 4 days) in the presence or absence of cycloheximide (100 μg/ml, 24h). N.S. = not significant. (means ± SEM, n=3). **E**) Bar graph showing methylation ratio (Beta value of HumanMethylation450K Array) of the RANKL promoter methylation level in MM and LM primary cells treated with or without 5'-aza (5μM) for four days. (means ± SEM, n=3, paired two-way ANOVA). Each number on the X-axis indicated a CpG probe on HumanMethylation450K Array near RANKL promoter region. **F**) Bar graph showing MethylCap-qPCR result of RDRE in each FACS-sorted LM cell population (means ± SEM, n=4, *$P<0.05$ and **$P<0.01$, paired one-way ANOVA).

To determine whether 5'-aza-mediated RANKL expression through altering the CpG island DNA methylation as previously reported [127, 128], we performed HumanMethylation450K array. No DNA methylation change was detected for all probes around RANKL gene locus (Figure 2-9 E).

### 2.2.6 DNA methylation potentially alters global progesterone responsiveness in LM

To define the mechanism underlying 5'-aza-mediated RANKL expression change, we isolated genomic DNA from LM and MM tissues and performed MethylCap-Seq to profile genome-wide methylation landscapes of MM and LM cells treated with or without 5'-aza. We discovered 21,086 differentially methylated regions (DMRs) between untreated MM and LM cells. Intriguingly, 44.23% of DMRs were found to be located in intergenic regions (Figure 2-10 A). Network enrichment analysis indicated that DMR-associated genes, especially those close to hypomethylated regions in LM, were highly enriched for genes encoding signaling components that control the Wnt/$\beta$-catenin pathway, cytoskeleton regulation, skeletal muscle development, and the estrogen receptor 1 (ESR1) nuclear pathway (Figure 2-10 B), which have been shown to play crucial roles in LM development [129-131].

Motif enrichment analysis of all DMRs between LM and MM showed that PREs (PR and androgen receptor (AR) bind to essentially identical motifs), together with several other critical TF binding elements such as HIF-1b and Smad2, were highly enriched at DMRs (Figure 2-10 C). Since PREs were enriched around DMRs, we investigated the methylation status of genome-wide PRBSs by integrating MethylCap-Seq data with PR ChIP-Seq data in matched MM and LM tissues. We found that DNA methylation levels at PRBSs in LM were lower versus those in MM (Figure 2-10 D), particularly in regions with higher PR-binding activity in LM (Figure 2-10 E).

**A**) Pie chart showing the genomic context of 21,086 differentially methylated regions (DMRs) identified using MethylCap-Seq **B**) Network enrichment analysis (Metacore) of DMRs. **C**) Motif analysis (Homer) of all DMRs between LM and MM. **D**) Genome-wide PR-binding regions in LM are more hypomethylated in LM versus MM. All PR-binding sites identified by PR ChIP-Seq in LM tissue were scanned from 20kb upstream to 20kb downstream of the binding region and the average DNA methylation levels at these regions were determined using the MethylCap-Seq data of untreated MM (green) and LM (blue) cells. **E**) DNA regions with higher PR-binding activity had lower DNA methylation levels in LM compared with MM. Box plot showing normalized DNA methylation levels (log2(RPKM$_{IP}$+1) – log2(RPKM$_{Input}$+1)) at 2662 PR-binding regions with higher PR enrichment in LM versus MM tissue (**$P$<0.01, paired t-test). Line within the box indicates median DNA methylation levels of all regions.

**Figure 2-10. PR targets have higher DNA methylation level in MM compared with LM.**

**2.2.7 The RANKL gene distal enhancer region is hypomethylated and controls RANKL transcription by interfering with PR action**
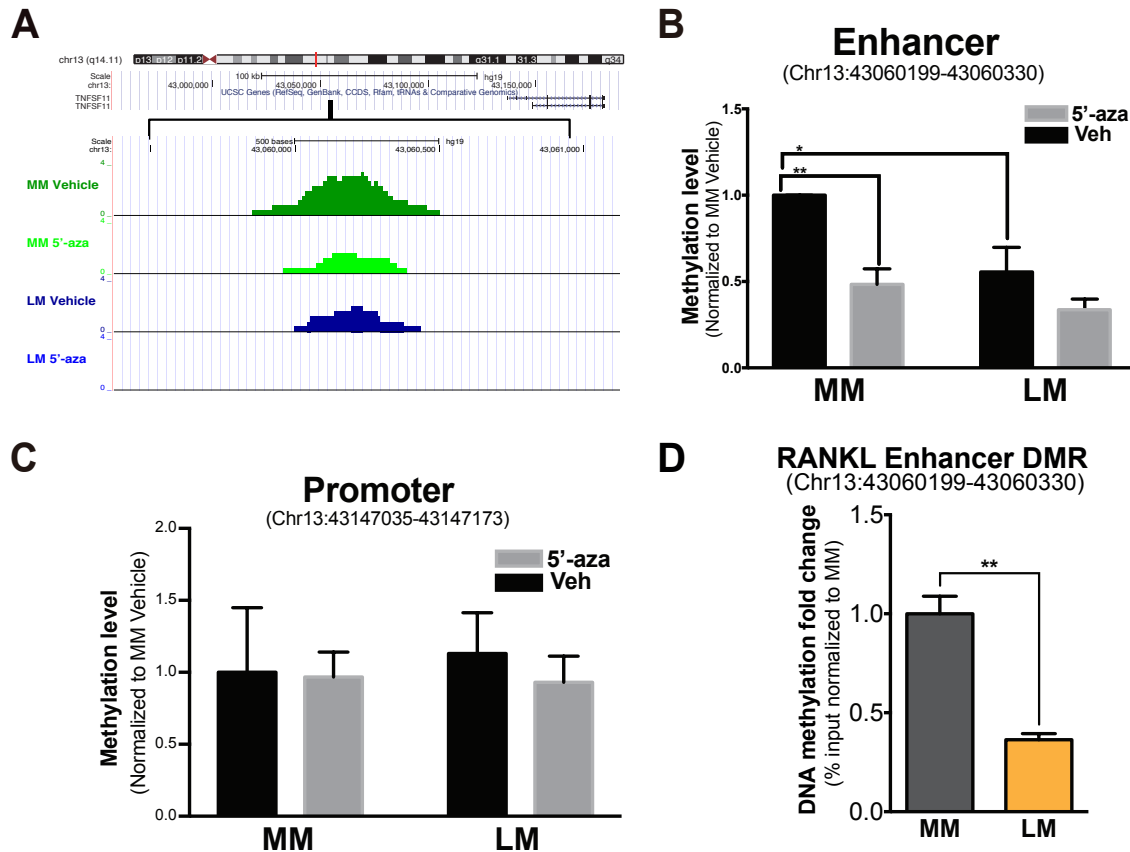


**Figure 2-11. RANKL enhancer is hypermethylated in MM compared with LM.**
**A**) DNA methylation levels of MM and LM cells exposed to 5'-aza (5 µM) or DMSO were evaluated using MethylCap-Seq. Representative UCSC Genome Browser Track view showing a DMR adjacent to the distal PRBS. **B-C**) Bar graph showing MethylCap-qPCR analysis of the DNA methylation levels of RDRE (-88,092bp/-87,961bp, B) and the RANKL promoter (-1,256bp/-1,118bp, C) in MM and LM primary cells treated with or without 5'-aza (5µM) for four days (means ± SEM, n=4, *$P<0.05$ and **$P<0.01$, paired two-way ANOVA). **D**) MethylCap-qPCR confirmed the higher DNA methylation level of DMR (-88,092bp/-87,961bp, Chr13: 43,060,199-43,060,330) adjacent to the distal PRBS of RANKL gene in MM versus LM tissue. Bar plots showing relative DNA methylation levels in LM after normalization to MM using genomic DNA from five different subjects (means ± SEM, n=5, **$P<0.01$, paired t-test).

Our findings suggested that DNA methylation may contribute to the global difference in progesterone responsiveness of LM and MM through interference of PR-binding activity. We specifically examined DMRs in the RANKL gene locus and discovered a hypermethylated region located 500bp upstream of the distal PRBS in MM versus LM cells, which was de-methylated after 5'-aza treatment, but no methylation difference was detected adjacent to the RANKL promoter (Figures 2-11 A-C). Using MethylCap-qPCR, we validated our findings using genomic DNA isolated directly from matched LM and MM tissues of five subjects. DNA methylation levels at this DMR was significantly lower in LM versus MM tissues (Figure 2-11 D).

To test whether DNA methylation affects PR-binding activity at the distal PRBS of the RANKL gene, we treated primary MM and LM cells with 5'-aza (5 µM) or vehicle (DMSO) for 4 days and examined PR recruitment to this region. In vehicle-treated cells, PR-binding activity was significantly higher in LM than in MM cells (Figure 2-12 A). Consistent with the effect of demethylation on RANKL transcription, 5'-aza robustly enhanced PR binding in MM but not in LM cells (Figure 2-12 A), resulting similar binding activity between LM and MM under 5'-aza treatment condition.

Furthermore, we evaluated the effects of 5'-aza and R5020 treatment on chromatin interaction activity between the distal PRBS and the RANKL promoter using 3C-qPCR. In MM, the looping activity between these two regions in vehicle-treated cells was weak, increased slightly with either 5'-aza or R5020 treatment, and peaked when the two treatments were combined (Figure 2-12 B). Consistent with the results shown in Figure 2-5 A, in vehicle-treated cells, the interaction activity was higher in LM versus MM cells and 5'-aza alone did not significantly change the interaction strength in LM (Figure 2-12 B). Consequently, 5'-aza triggered the interaction strength between the distal PRBS and RANKL promoter in MM to the levels observed in LM, which is consistent with 5'-aza-mediated RANKL mRNA expression (Figure 2-9 A) and PR recruitment towards the distal PRBS (Figure 2-12 A).
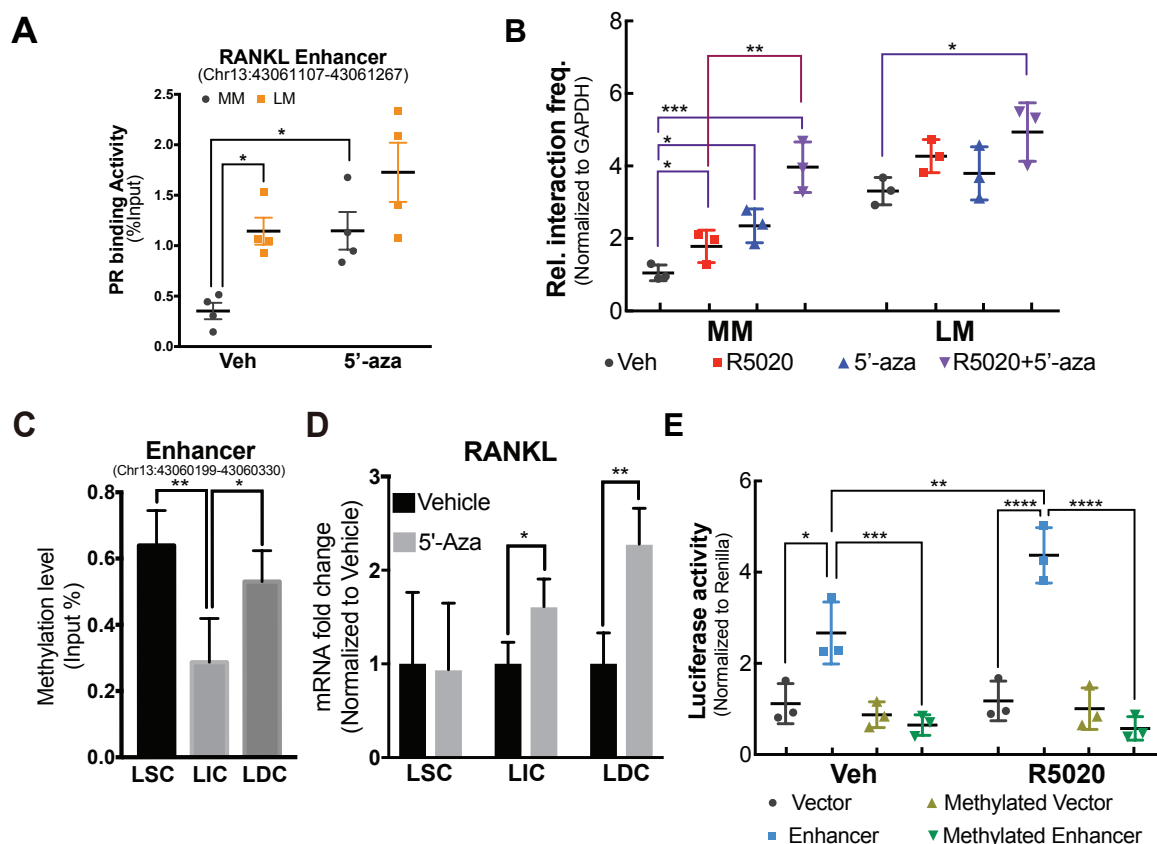
**Figure 2-12 5'-Aza increased PR binding at the RANKL enhancer.**

**A**) Dot plot of ChIP-qPCR of PR association with the RANKL gene distal PRBS in primary MM and LM cells exposed to 5'-aza (5μM, 96h) or vehicle (DMSO) (means ± SEM, *P<0.05, paired two-way ANOVA, n=3). All groups were stimulated by R5020 ($10^{-7}$ M, 1h) prior to harvesting. **B**) Dot plot showing 3C-qPCR results of the RANKL promoter and distal PRBS (En2 fragment shown in Figure 2-5 A) interaction in MM and LM primary cells treated with 5'-aza (5μM, 96h) or vehicle (DMSO) in the presence or absence of R5020 ($10^{-7}$ M, 1h) (means ± SEM, n=3, *P<0.05, **P<0.01, ***P<0.001, paired two-way ANOVA). **C**) Bar graph showing MethylCap-qPCR result of RDRE in each FACS-sorted LM cell population (means ± SEM, n=4, *P<0.05 and **P<0.01, paired one-way ANOVA). **D**) Bar graph showing qPCR result of RANKL expression levels in each FACS-sorted LM cell population treated with or without 5'-aza for 96h. (means ± SEM, n=3, **P<0.01, paired t-test). **E**) DNA fragment spanning distal PRBS and the DMR (-88,347bp/-86,851bp, chr13: 43,059,944 - 43,061,440, 1496bp, designated as RDRE) was cloned into a CpG-free luciferase construct. RDRE-containing or empty vector (control) were treated with or without methylation reagents *in vitro*. LM primary cells were transiently transfected with the treated plasmids, exposed to vehicle (ethanol) or $10^{-7}$ M R5020 for 48h, and harvested for luciferase assay (n=3, means ± SEM, *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, paired two-way ANOVA).

We recently reported that RANKL is preferentially expressed and specifically upregulated by R5020 in LIC [68], we went on investigating whether the same mechanism (DNA methylation difference in the DMR adjacent to the distal PRBS of RANKL gene) was also involved. We examined the DNA methylation levels and the effects of 5'-aza treatment on RANKL expression in each LM cell population. DNA methylation level in the DMR was significantly lower in LIC versus LDC, accompanied with higher RANKL mRNA levels in LIC (Figure 2-12 C). 5'-aza stimulated RANKL expression more robustly in LDC (Figure 2-12 D). We did not detect significant correlation between DNA methylation and RANKL expression in LSC, which is deficient of PR [64].

Together, our findings indicate that the novel DMR upstream of the distal PRBS may mediate RANKL expression. We designated the DNA fragment spanning the novel DMR and the distal PRBS as the RANKL distal regulatory element (RDRE). To determine whether the RDRE has a regulatory function in RANKL gene transcription, we cloned the 1496bp RDRE (chr13: 43,059,944 - 43,061,440) into a CpG-free luciferase construct and performed luciferase reporter assays.[132] As shown in Figure 2-12 E, the vector carrying the RDRE showed significant promoter-enhancing activity in vehicle-treated cells, which was further increased by R5020 treatment. *In vitro* methylation of this fragment completely blocked its promoter-enhancing activity. These results indicate that RDRE has progestin-inducible transcriptional regulatory activity and can be blocked by DNA methylation.

### 2.2.8   MED12 mutation status influences RANKL transcription

Mutations in MED12 alter its interaction with transcriptional co-activators and induce E2+P4-dependent LM formation [119, 133]; this prompted us to hypothesize that MED12 mutations dysregulate its interaction with PR, leading to increased RANKL transcription in LM. Due to RANKL's effect on LSC function, we investigated whether mut-MED12 affects RANKL expression and PR recruitment to the distal PRBS. To minimize the potential variation caused by different mutations of MED12, we focused on the most common mutation type, *i.e.*, G44D mut-MED12 [16].

We observed that RANKL mRNA levels and recruitment of MED12 in the RANKL distal PRBS were significantly higher in LM tissues carrying G44D mut-MED12 than those expressing wild type (WT) MED12 (Figures 2-13 A and B) (MM tissues always express WT-MED12). Regression analysis demonstrated that MED12-binding activity at the enhancer region positively correlated with RANKL expression (Figure 2-13 C, $R^2$=0.6282) and PR-binding affinity in this region (Figure 2-13 D, $R^2$=0.7054) in LM with G44D mut-MED12. Supporting the pro-proliferation role of RANKL in LSC, we detected a marginally larger ($P$=0.0551) LSC population in LM harboring G44D mutation versus wild type (WT) LM (Figures 2-13 E and F). In addition, immunoprecipitation-immunoblot (IP-IB) analysis showed that PR interacted with MED12 and the interaction was stronger in LM bearing G44D mut-MED12 (Figure 2-13 G). These findings suggest that MED12 mutation status influences RANKL transcription and its interaction with PR.
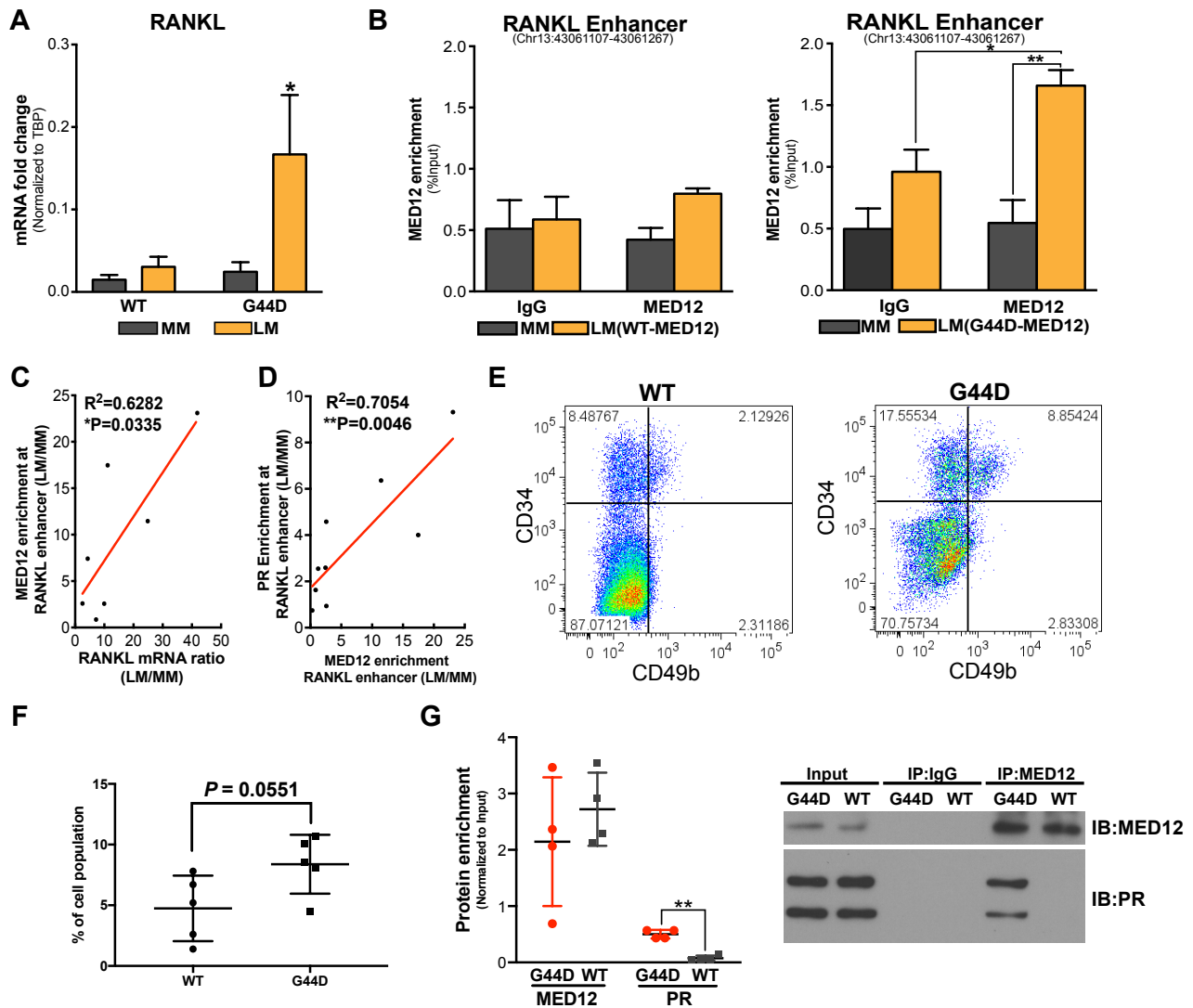
**Figure 2-13. RANKL transcription level is influenced by MED12 mutation status.**

**A**) RANKL mRNA levels in wild-type (WT) (n=12) or G44D MED12 mutation-containing (n=12) LM and their matched MM tissues (mean ± SEM, *$P<0.05$, paired two-way ANOVA). RANKL mRNA levels were normalized by TBP expression. **B**) ChIP-qPCR of MED12 recruitment towards the distal PRBS of RANKL gene using chromatin isolated from fresh-frozen MM and LM tissues. Results were stratified based on MED12 mutation status, including four subjects expressing WT MED12 and nine subjects bearing G44D MED12 mutation (means ± SEM, *$P<0.05$, **$P<0.01$, paired two-way ANOVA). **C** and **D**) Pearson correlation showing that MED12 enrichment at RANKL distal PRBS positively correlates with RANKL expression level (**C**, n=7, $R^2=0.6282$, *$P<0.05$, Pearson correlation) and PR-binding activity in LMs carrying the G44D MED12 mutation (**D**, n=9, $R^2=0.7054$, **$P<0.01$, Pearson correlation). Data represent MED12 enrichment and RANKL mRNA levels in LM normalized to those in matched MM. **E**) Representative FACS

scattergram showing LM cell population distribution. Cells isolated from WT LM and G44D LM tissues were stained with anti-CD45 (depleted hematopoietic cells), anti-CD34, anti-CD49b, and PI (depleted dead cells) and analyzed by the LSRFortessa system. **F**) Scatter graph showing quantification of the percentage of each LM cell population showed in A. **G)** IP-IB assay showing higher physical interaction activity between MED12 and PR in LMs expressing G44D MED12 mutation compared with those expressing WT MED12. Whole-cell protein extracts were immunoprecipitated with either MED12 or IgG followed by immunoblotting for PR. Left panel: ImageJ quantification of immunoblots (means ± SEM, n=3, **$P$<0.01, paired t-test); Right panel: Representative blot.

## 2.3    Conclusion

We uncovered a novel cis-regulatory element (RDRE) downstream of P4/PR signaling that regulates RANKL transcription. We demonstrated that RANKL expression is regulated by converging signals from steroid hormone (P4/PR), genetic factors (MED12 mutation), and epigenetic modifications (DNA methylation and histone modification) and contributes to LSC proliferation and tumor formation. We summarize our proposed model of P4/PR-mediated RANKL gene regulation in Figure 2-14.
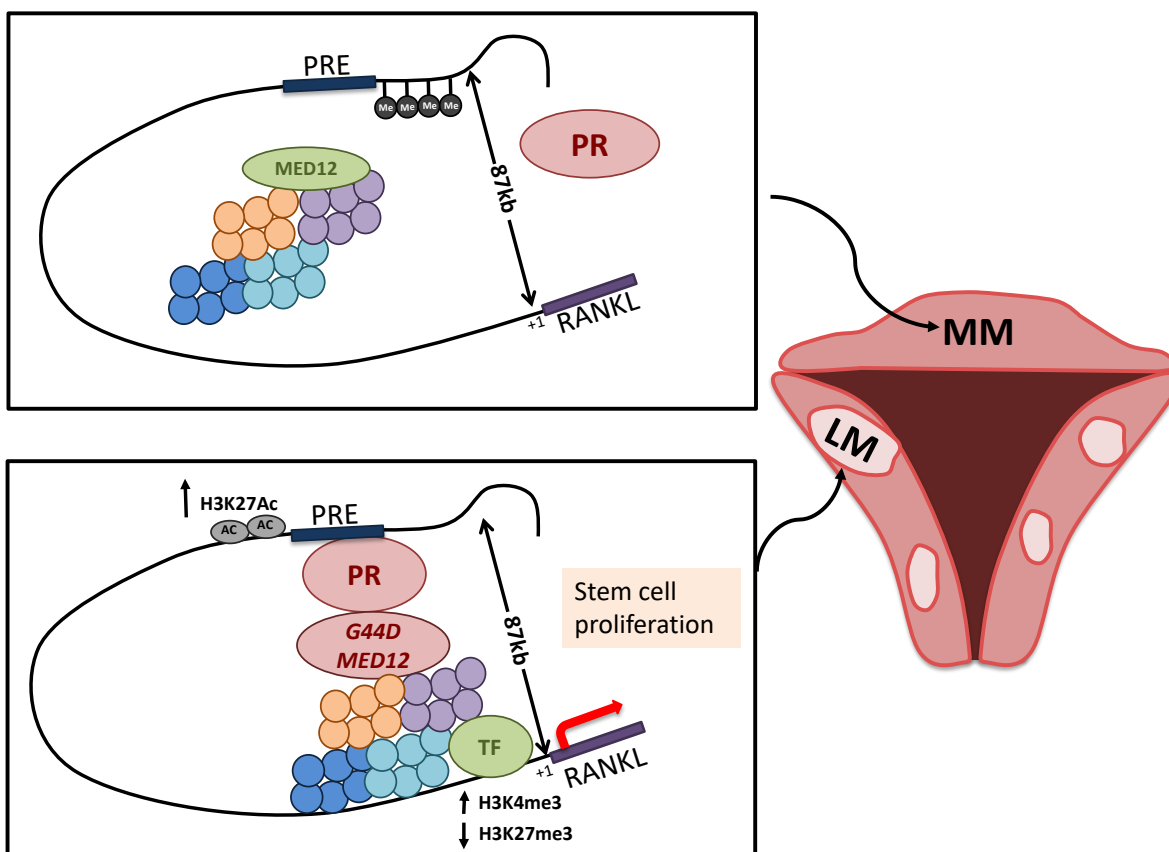


**Figure 2-14. RANKL regulation model in MM and LM.**
Schematic of the proposed RANKL gene regulatory model in MM and LM. In MM, higher DNA methylation at RDRE blocks PR binding, resulting in low RANKL expression levels. In LM, RDRE is hypomethylated, leading to higher PR binding activity. G44D mutated MED12 further stabilizes PR binding at the RDRE.

## 2.4    Materials and Methods

**Tissue collection**

Northwestern University's Institutional Review Board approved the use of human tissue. LM and matched MM tissues were obtained from premenopausal women undergoing either myomectomy or hysterectomy (age 38±9 years, range 28–49 years). Patients receiving hormone treatment six months prior to surgery were excluded. Tissues were dissociated and cells isolated as previously described [68].

**Primary cell and tissue explant culture**

Primary cells and *ex vivo* explant culture of MM and LM tissue were performed as previously described [68, 109]. For 5'-aza treatment experiments, if not indicated in the figures, cells were treated with 5 µM 5'-aza for 96h with medium refreshed every 24h (A3656, Sigma-Aldrich, St Louis, MO, USA). Tissue explants were treated with vehicle or R5020 in the presence or absence of RU486 ($10^{-6}$ M) for 48h. To elucidate the role of RANKL in LSC function, the explants were subjected to vehicle or RANKL (CYT-334, ProSpec Bio, Rehovot, Israel) treatment for 24h and analyzed using flow cytometry.

**Stem cell culture**

Each freshly FACS-sorted population of LM cells was cultured in mesenchymal stem cell growth medium (PT-3238, Lonza, Basel, Swizerland) in low-attachment 96-well plates (07-201-680, Fisher Scientific, Chicago, IL, USA) to maintain stem cell characteristics. After five days, cells were starved for 24h using basal mesenchymal stem cell growth medium followed by treatment with RANKL or vehicle for 24h. Apoptosis assay was performed using Annexin V-FITC Apoptosis Detection Kit (ab14085, Abcam, Cambridge, UK) on FACS-sorted cells treated with vehicle or RANKL for 24h and analyzed using flow cytometry. Proliferation assay was performed using Cell Counting Kit-8 (CK04-05, Dojindo, Rockville, MD, USA).

**PR siRNA knockdown**

LM passage zero cells were transfected with 2 different PR siRNAs (D-003433-03-0010 and D-003433-01-

0010, Dharmacon, Lafayette, CO, USA) or control siRNA (D-001810-10-05) for 96h, and stimulated with vehicle or R5020 for 24h before harvesting.

**Antibodies and primers**

All antibodies and primers used in this study are listed in Appendix Table 1.

**Antibody-based cell sorting**

CD34$^+$/CD49b$^+$, CD34$^+$/CD49b$^-$, and CD34$^-$/CD49b$^-$ LM cells were FACS-sorted as previously described [64].

**Real-time PCR**

0.2 to 1 µg of total RNA was reverse transcribed and quantified using real-time PCR as previously described [68].

**Immunohistochemistry**

Paraffin-embedded LM and MM tissues were sectioned and immunohistochemistry was performed by the Northwestern University Pathology Core Facility to detect RANKL expression as previously described [68]. The score was independently calculated by two individuals who were blinded to the treatment group allocation.

**Chromatin immunoprecipitation assay**

0.2-0.5g of frozen LM and matched MM tissues were used for ChIP using the SimpleChIP Kit (#9005, Cell Signaling Technology, Danvers, MA, USA). Chromatin was isolated and incubated with antibodies against PR, MED12, H3K4me3, H3K27Ac, or H3K27me3. Normal rabbit IgG was used as a negative control.

**MethylCap-qPCR**

Genomic DNA was extracted from LM and MM tissues or cells using DNeasy Blood and Tissue Kit (69504, Qiagen, Hilden, Germany) and fragmented to 300-500bp using Covaris M220 (Covaris, Woburn, MA, USA). Methylated DNA fragments were captured using the MethylCap Kit (C02020010, Diagenode, Denville, NJ, USA).

**3C-qPCR**

3C assay was performed and analyzed following previously published protocol [134]. Briefly, LM and MM primary cells were fixed with 2% formaldehyde. Nuclei were digested using *PciI* (400U) followed by ligation with T4 ligase in diluted condition. Ligated DNA was then de-cross-linked (overnight at 65°C) and purified by classical phenol extraction procedures. All primers were designed within 150bp of *PciI* sites and detailed information was provided in Appendix Table 1. A control template containing all ligation products in equal amount was used to optimize qPCR. All results were normalized using primer set located in GAPDH gene. To test product purity, we sequenced the 3C-PCR product using Sanger sequencing.

**Next-generation sequencing and analysis**

Next-generation sequencing libraries for MethylCap-Seq and PR ChIP-Seq were prepared using the KAPA Hyper Prep Kit (KK8502, KAPA Biosystems, Wilmington, MA, USA) and KAPA Single-Indexed Adapter Kit (KK8710, KAPA Biosystems). The libraries were sequenced at the Northwestern University NUSeq Core Facility using the NextSeq 500 system (Illumina, San Diego, CA, USA) with 20-30 million reads per sample (75bp single-end for PR ChIP-Seq and 75bp paired-end for MethylCap-Seq). Sequences were aligned to the hg19 reference genome using Bowtie2. We performed following analysis using Homer: peak calling (-factor for PR ChIP-Seq and -histone for MethylCap-Seq), differential binding analysis (getDifferentialPeaks) and motif analysis (findMotifsGenome.pl) [135]. Pathway enrichment analysis was performed using Metacore V6.34 (Thomson Reuters). Sequencing tracks were visualized using UCSC Genome Browser. Visualization of DNA methylation at the PR-binding regions was performed using NGSplot [136]; the methylation level at PR-binding regions was quantified using normalized RPKM value. The GEO accession number for the sequencing data reported in this paper is GSE113108.

**Immunoprecipitation (IP)**

0.2-0.5g of frozen LM and MM tissues were lysed in non-denaturing buffer (0.5% NP-40, 50mM Tris-HCl pH 8.0, 250mM NaCl, 5mM EDTA) containing protease inhibitor (11836170001, Sigma-Aldrich) for 2h at 4°C. 150µg protein lysate was used for IP following previously published protocol [137]. 5% of input was used as loading control. Proteins were boiled for 10 min and separated by SDS-PAGE using precast 4-12% Bis-Tris gels (Thermo Fisher Scientific, NP0315BOX). IP with normal rabbit IgG was used as a negative control.

**Immunoblotting**

Whole-cell extracts were prepared and western blot analysis was performed as previously described.[68] ImageJ was used to qualify the intensity of blots.

**Luciferase assay**

The RANKL gene distal cis-regulatory element spanning the novel DMR and distal PRBS (RDRE) was amplified by PCR using primers listed in Appendix Table 1 and cloned into a CpG-free luciferase construct hCpGL-EF1A-Basic [44, 138]. The plasmid was *in vitro* methylated and tested as described previously [138]. LM and MM primary cells were transfected with 2µg reporter plasmid and 0.2µg PRL-TK-Luc using Lipofectamine 3000 Reagent (L3000008, Fisher Scientific). Cell extracts were prepared using passive lysis buffer (E1910, Promega, Madison, WI, USA). Results were normalized to PRL-TK-Luc using a dual luciferase reporter assay system.

**Statistical Analysis**

All of the statistical analysis in this study were performed using GraphPad Prism 7 (GraphPad Inc., La Jolla, CA, USA) with detailed statistic test description in each figure legend. Sample size was determined based on previous publications.[64, 68] No sample was excluded during analysis. Values were considered statistically significant when $P<0.05$. Similarity of variance across compared groups was tested by: F-test

in two groups, Sum of Square in multiple groups. Normal distribution of data was assessed by Shapiro–

Wilk test. Error bar represented SEM for biological replicates of 3 technical replicates within 1 patient. All

experiments are repeated with at least 3 patients, with patient number stated in figure legend.

CHAPTER 3      Targeting DNA methylation and progesterone action

during leiomyoma stem cell differentiation

**3.1    Introduction**

Uterine leiomyoma (LM, a.k.a. fibroid) is the most common tumor in women, affecting approximately 80% of women [2, 3]. It affects African-American women disproportionately at an earlier age and with a higher prevalence [1, 2]. Although benign in origin, LMs can cause significant morbidity, including excessive uterine bleeding, recurrent pregnancy loss, and pelvic pain; these symptoms may also mimic or mask malignant tumors [3]. More than 200,000 surgeries are performed in the US each year to remove LMs, at an estimated cost of $5.9-34.4 billion [111, 139]. Despite their prevalence and impact on women's health, no new medical treatments have been approved in the US for LM since the 1990s [7].

Stem cells are critical for normal tissue and disease development. We identified three molecularly and functionally distinct cell populations in LM tissues with a possible hierarchical differentiation order: LM stem cells (LSC, 5%), LM intermediate cells (LIC, 7%), and terminally differentiated LM cells (LDC, 88%) [69]. Progesterone, progesterone receptor (PR/PGR), and LSC are indispensable for LM growth; however, LSC are deficient for PR, which may explain the observed frequent recurrence of LM to current hormone-based therapies [66, 140]. Progesterone/PR signaling plays dual roles in normal mammary gland development and tumor growth, regulating mammary stem cell differentiation and proliferation through paracrine interaction with adjacent differentiated cells [39, 113]. LM and breast cancer share some similarities with respect to responsiveness to steroid hormones [2, 44]. In LM, we also observed that progesterone, via PR expressed in differentiated LM cells, stimulates paracrine signaling that targets the PR-deficient LSC to induce their proliferation, leading to tumor growth [68, 69, 140]. But whether PR regulates LSC differentiation remains unknown.

Epigenetic alterations, such as DNA methylation, play crucial roles in stem cell regulation and disease progression [116, 141]. Within tumor stem cells, transcription factors required for differentiation are suppressed by DNA methylation, leading to abnormal clonal expansion [142-144]. It remains unclear how DNA methylation affects transcription factor function in LSC, particularly that of nuclear hormone receptors, and how this alters LSC differentiation and tumorigenesis.

In this study, we characterized the global transcriptome, methylome, and histone modifications (H3K4me3 and H3K27Ac) in the three LM cell populations and their relationship with PR function during the LSC differentiation process. We hypothesized that PR plays a critical role in LSC differentiation and the epigenetic patterns of LSC suppress PR activity, leading to the lack of progesterone responsiveness in these cells. Removing this epigenetic suppression would induce direct PR action and drive LSC differentiation, possibly resulting in the depletion of this population.

**3.2    Results**

**3.2.1    The LSC population has a unique DNA methylation landscape**

To evaluate the genome-wide DNA methylome and transcriptome, we performed high-throughput sequencing analyses, including MethylCap-Seq and RNA-Seq, on freshly sorted LM cell populations from patient LM tissues (Figure 3-1 A).  MethylCap-Seq characterized the global DNA methylation landscape of each LM cell population.  LSC showed a unique DNA methylome compared to LIC and LDC (Figure 3-1 B, n = 3 patients). On average, LSC had the highest DNA methylation level, especially at intergenic regions, whereas LIC had the lowest DNA methylation level (Figure 3-1 C). In LSC, we identified 7020 regions with higher DNA methylation and 1842 with lower DNA methylation compared with LIC; and 8364 regions with higher DNA methylation and 1080 with lower DNA methylation compared with LDC (Figure 3-1 D). LIC and LDC showed a more similar DNA methylome with only 288 differentially methylated regions detected, indicating that DNA methylation plays more important role during the initial differentiation of LSC.

We evaluated the genomic distribution of differentially methylated regions between LSC and the non-stem cell populations. Compared with randomly picked genomic regions (0.7% at CpG island and 3.5% at shores, Figure 3-1 E), differentially methylated regions were highly enriched in CpG islands and shores (Figure 3-1 F). This feature was more significant in regions with lower DNA methylation in LSC:  19.9% CpG islands and 35.7% shores were found in comparison with LIC; 29% CpG islands and 47.9% shores were found in comparison with LDC (Figure 3-1F, right panels). These findings were consistent with the DNA methylation pattern changes during intestine stem cell differentiation observed using whole-genome shotgun bisulfite sequencing [145].
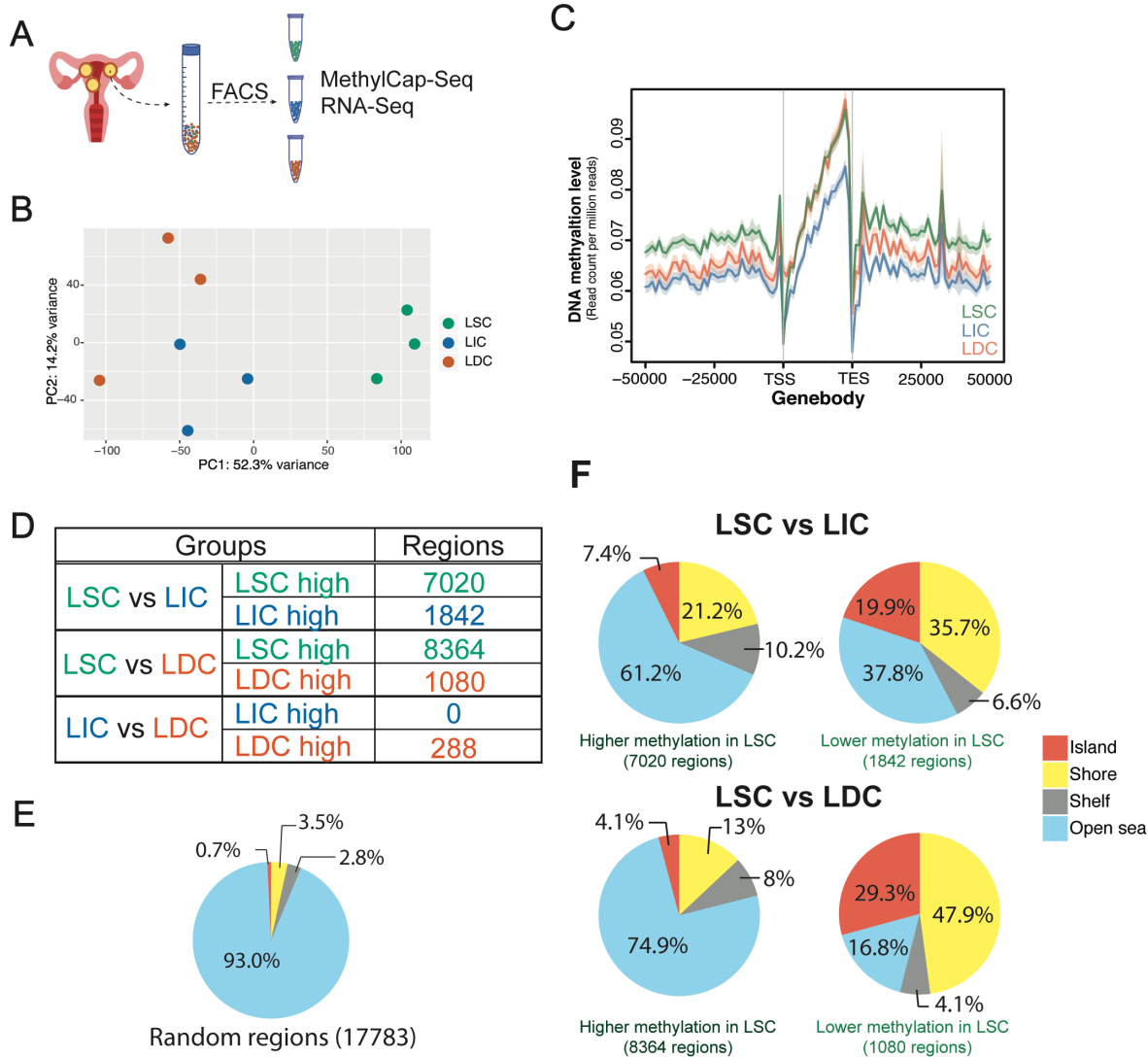
**Figure 3-1. The LSC population has a distinct methylation landscape.**

**A**) Workflow of LM cell population characterization. **B**) Principal component analysis (PCA) plot showing the DNA methylation landscape of each LM population (n = 3, top 10,000 methylated regions with the largest variance among the 9 samples). **C**) Average line plot showing the DNA methylation level around the gene bodies in each cell population (TSS: transcription start sites; TES: transcription end sites). **D**) Numbers of differentially methylated regions (DMRs) identified among the three LM populations. **E**) CpG island annotation of randomly picked 17783 genomic regions. (CpG shore: within 2 kb from CpG island; CpG shelf: within 2-4 kb from CpG island; CpG open sea: more than 4 kb away from CpG island). **F**) CpG island annotation of differential methylated regions identified between LSC vs LIC or LDC.

A

**Hypermethylated in LSC**

| Motif | Match | p.value |
|---|---|---|
| | EGR1 | 1e-19 |
| | ARNT | 1e-18 |
| | FOSL2 | 1e-13 |
| | CDX4 | 1e-13 |
| | JUN | 1e-9 |

B

**Hypomethylated in LSC**

| Motif | Match | p.value |
|---|---|---|
| | OTX2 | 1e-12 |
| | ZNF467 | 1e-10 |
| | GFI1B | 1e-10 |
| | RUNX2 | 1e-10 |
| | SMAD4 | 1e-9 |

**Figure 3-2. Motif analysis of differential methylated regions in LSC.**

Motif analysis on differentially methylated region in LSC. **A**) List of top enriched motifs identified in the 3471 overlap hypermethylated regions in LSC comparing to both LIC and LDC. **B**) List of top enriched motifs identified in the 353 overlap hypomethylated regions in LSC comparing to both LIC and LDC.
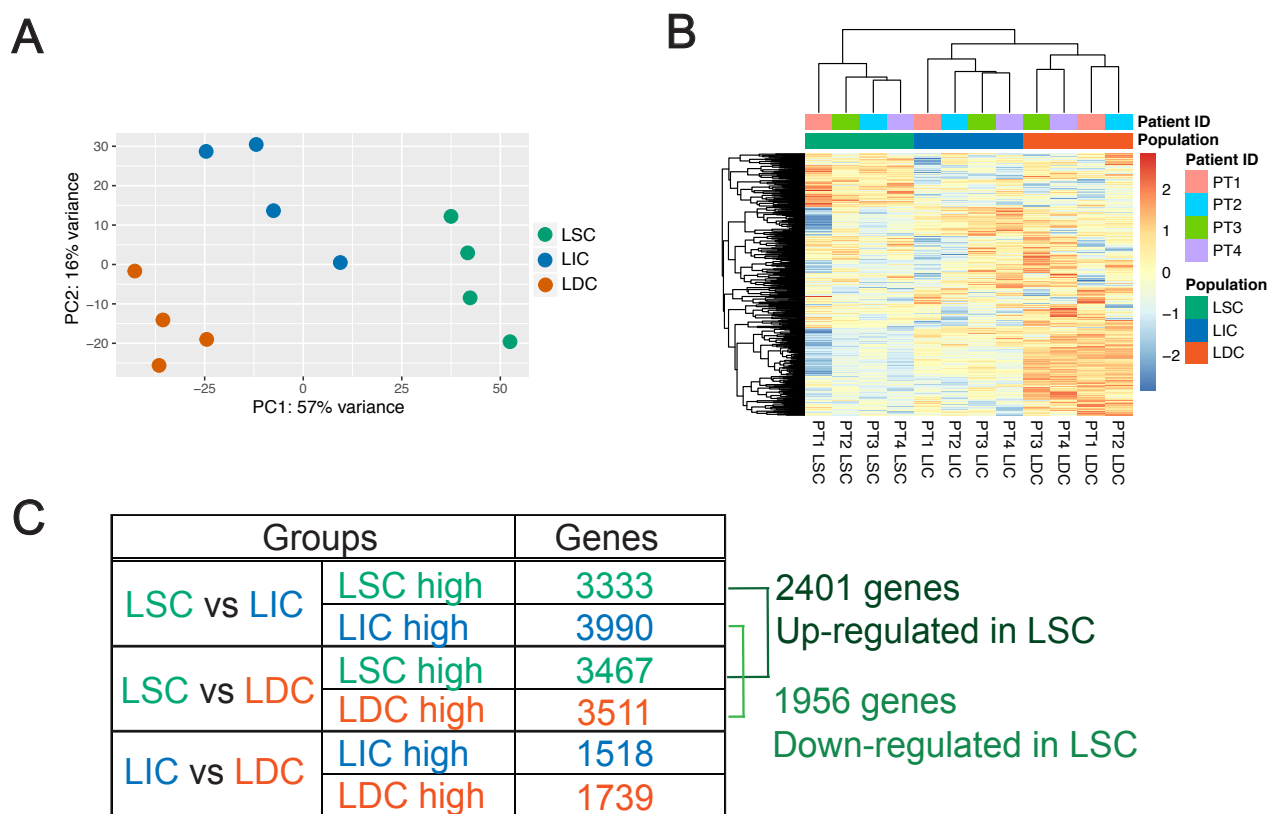
**Figure 3-3. The LSC population has a distinct transcriptome.**

Principal component analysis (PCA) plot (**A**) and hierarchical clustering heatmap (**B**) showing the transcriptional profiles of the three LM cell populations (n = 4, top 1000 genes with the largest variance among the 12 samples). **C**) Numbers of differentially expressed genes (DEGs) identified among the three LM populations using RNA-Seq. 2401 genes were identified higher in LSC compare with either LIC or LDC; 1956 genes were lower in LSC compare with either LIC or LDC.

DNA methylation interferes with the interactions between DNA, specific transcription factors, and chromatin proteins [146]. To understand the consequences of the altered methylation patterns during differentiation, we performed motif analysis on the differentially methylated regions in LSC. We identified EGR1, ARNT, FOSL2, CDX4, and JUN as the highly enriched motifs within hypermethylated regions and OTX2, ZNF467, GFI1B, RUNX2, and SMAD4 as the highly enriched motifs within hypomethylated regions (Figure 3-2). The EGR1 and TGFβ/SMAD signaling pathways have established roles in LM tumorigenesis [147, 148]. FOSL2 and JUN (subunits of the AP1 transcription factor complex) are tethering factors for PR DNA binding and methylation changes around these motifs may alter PR recruitment to DNA, affecting progesterone signaling and LM progression [49, 123].

We then performed RNA-Seq to evaluate the effects of DNA methylation on transcriptome and find possible underlying mechanisms for the DNA methylation changes during LSC differentiation. Principal component and hierarchical clustering analyses of RNA-Seq data demonstrated distinct clustering of the 12 samples (n = 4 patients) into three cell populations and a matching pattern between transcriptome and methylome (Figures 3-3 A). Consistent with previous observations based on microarray data, the heatmap clustering (Figure 3-3 B) indicated a significant transcriptome difference between LSC and the non-stem cell populations [69]. Compared to both LIC and LDC, differential gene expression analysis detected 2401 genes higher and 1956 genes lower in LCS (Figure 3-3 C). Supporting the distinct methylation landscape identified by MethylCap-Seq, genes involved in DNA methylation and demethylation process (GO:0044728) maintained a unique expression pattern in LSCs compared with the other two populations (Figure 3-4 A). qPCR validated that the mRNAs of TET1 and TET3, two DNA demethylases, were expressed at the lowest levels in LSC (Figure 3-4 B), providing a possible molecular explanation for the hypermethylation observed in the LSC population.
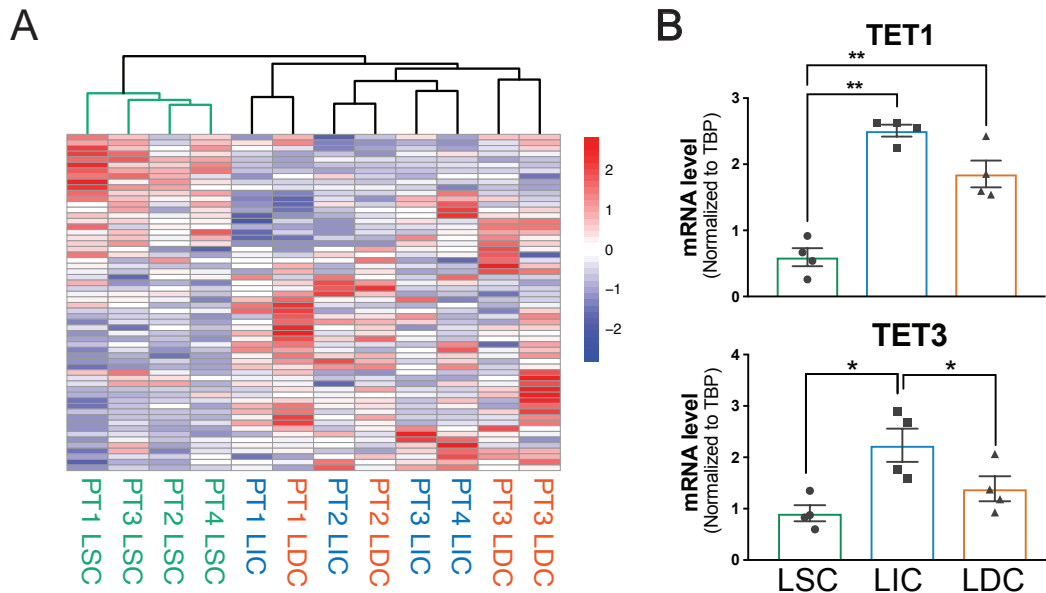
**Figure 3-4. DNA methylation machineries are differentially expressed in LSC.**

**A**) Hierarchical clustering heatmap showing mRNA expression pattern of genes involved in DNA methylation and demethylation process (GSEA GO term M14518) among the 12 RNA-Seq samples from the 3 sorted LM cell types (n = 4). **B**) Bar graphs showing real-time qPCR validation of the differential expression of TET1 (top) and TET3 (bottom) among the three LM cell populations (means ± SEM, n = 4, *P<0.05, **P<0.01, paired one-way ANOVA).

### 3.2.2 DNA methylation controls the expression of key genes during LSC differentiation

To investigate potential influences of DNA methylation on gene expression changes during LSC differentiation, we integrated the RNA-Seq and MethylCap-Seq data. We compared differentially methylated and expressed genes between LSC vs LIC/LDC. 8862 differentially methylated regions, which annotated to 5396 genes, were identified between LSC and LIC. Among these, 3230 regions were found to be related with 1772 genes differentially expressed between LSC and LIC (24.1% of all DEGs, Figure 3-5 A, upper panel). Similarly between LSC and LDC, 3446 out of 9444 differential methylated regions were annotated to 1741 gene differentially expressed between LSC and LDC (24.9% of all DEGs, Figure 3-5 A, lower panel). The majority of these overlapped genes were hypermethylated with lower expression in LSC (blue dots), suggesting an inversed relationship between DNA methylation and gene expression changes (Figure 3-5 B). Pearson correlation analysis further confirmed that DNA methylation loss during differentiation was significantly correlated with expression increase of the nearest gene (LSC_vs LIC: r= -0.38, p.value < 2.2e-16; LSC_vs LIC: r= -0.30, p.value < 2.2e-16). In order to further investigate the possible role for DNA methylation in mediating the transcriptional activates of DNA regulatory elements, we performed ChIP-seq for the active histone marks H3K4me3 and H3K27Ac in sorted three cell populations (n=3 patients). As expected, we also found that, during differentiation from LSC to LIC and LDC, loss of DNA methylation was strongly associated with enrichment of active H3K4me3 and H3K27Ac histone marks (Figures 3-6 A and B).

We next performed enrichment analysis to identify pathways and networks enriched with the genes found to be differentially methylated and expressed during LSC differentiation. We identified several processes that play crucial roles in LM tumorigenesis, including the Wnt/β-catenin pathway, muscle development, BMP/TGFβ signaling, and steroid hormone-related signals (Figure 3-7) [49, 140, 149]. Notably, genes hypermethylated and downregulated in LSC (blue bars) were more enriched in pathways regulating LM-specific functions (e.g. muscle contraction, muscle development, hormone action, and ESR1 signaling), while genes hypomethylated and upregulated (yellow bars) were involved in stem cell-related pathways.

Previous studies have shown that genes governing the differentiation process are repressed by Polycomb in human embryonic stem cells (HESCs) [144]. We found significantly higher DNA methylation and lower mRNA levels at these Polycomb targets in LSC vs. LIC and LDC [144], suggesting that DNA methylation might lock differentiation-associated genes in a repressive state in LSC (Figure 3-8 A).  Furthermore, we found that ESR1 and PR were among the genes both differentially expressed and methylated between LSC and the two non-stem cell populations (LIC and LDC). The *PGR* and *ESR1* genes were hypermethylated at several intronic regions in LSC; and the *ESR1* genes was also hypermethylated at the promoter region in LSC (Figure 3-8 B). Converse to the DNA methylation status, LSC expressed the lowest levels of PR and ESR1 compared to LIC and LDC (Figure 3-8 C). Previous research has demonstrated that PR and ESR1 expression levels were related with DNA methylation status in ovarian endometrioma [150]. Here, to investigate whether PR and ESR1 expression levels are regulated by DNA methylation in LM, we treated the sorted cell populations with vehicle (DMSO) or a DNA methylation inhibitor 5′-Aza (100 nM) for 96 hours. We found that 5′-Aza increased both PR and ESR1 mRNA levels, specifically in the LSC population (Figure 3-8 D).
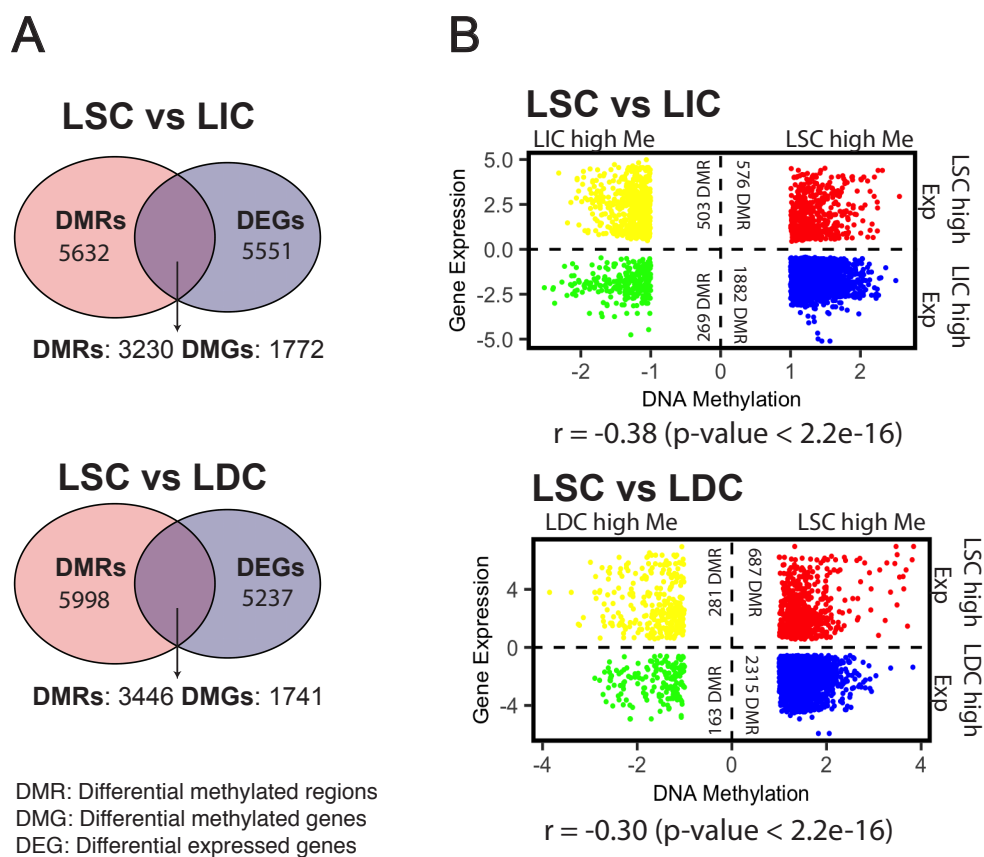
**Figure 3-5. mRNA change is inversely correlated with DNA methylation change.**

**A**) Venn diagram showing the overlapping genes between differentially methylated region (DMR)-associated genes and differentially expressed genes (DEGs) in the LSC vs LIC and LSC vs LDC comparisons. **B**) Dot plot showing gene expression and DNA methylation log2 fold changes in the LSC vs LIC and LSC vs LDC comparisons. Colored dots represent genes with significant changes (FDR adjusted *P*<0.05) in both gene expression and DNA methylation. Correlation coefficient between log2 fold changes of gene expression and DNA methylation were calculated for both pairs.
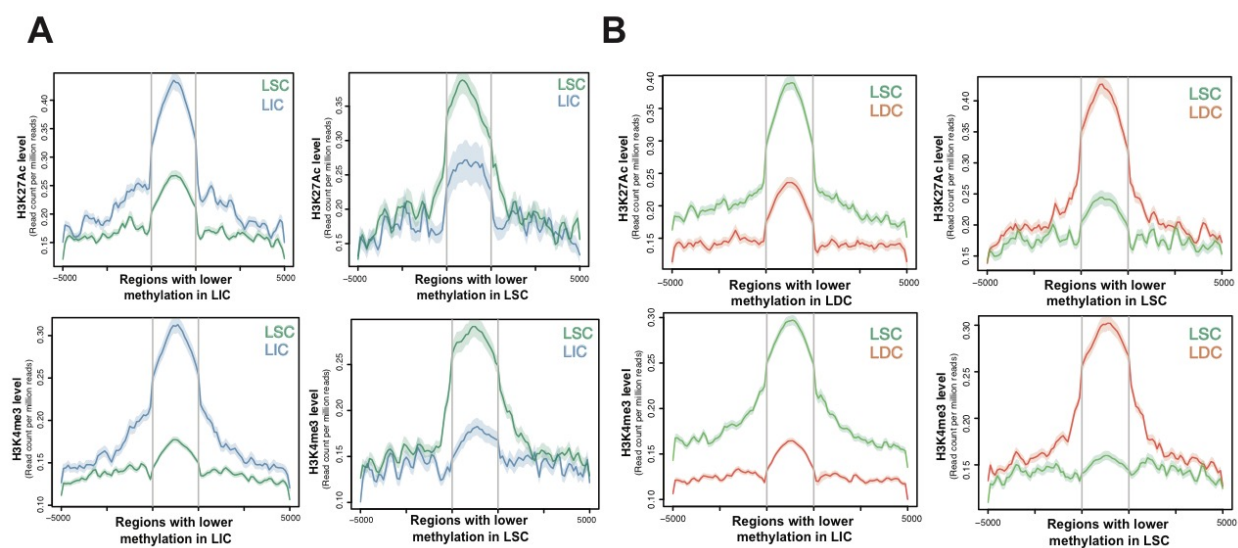
**Figure 3-6. DNA methylation change is inversely correlated with active histone mark enrichment**

(**A**) Average line plots show the H3K27Ac and H3K4me3 level around (±5,000bp) differentially methylated regions between LSC and LIC. (**B**) Average line plots show the H3K27Ac and H3K4me3 level around (±5,000bp) differentially methylated regions between LSC and LDC.
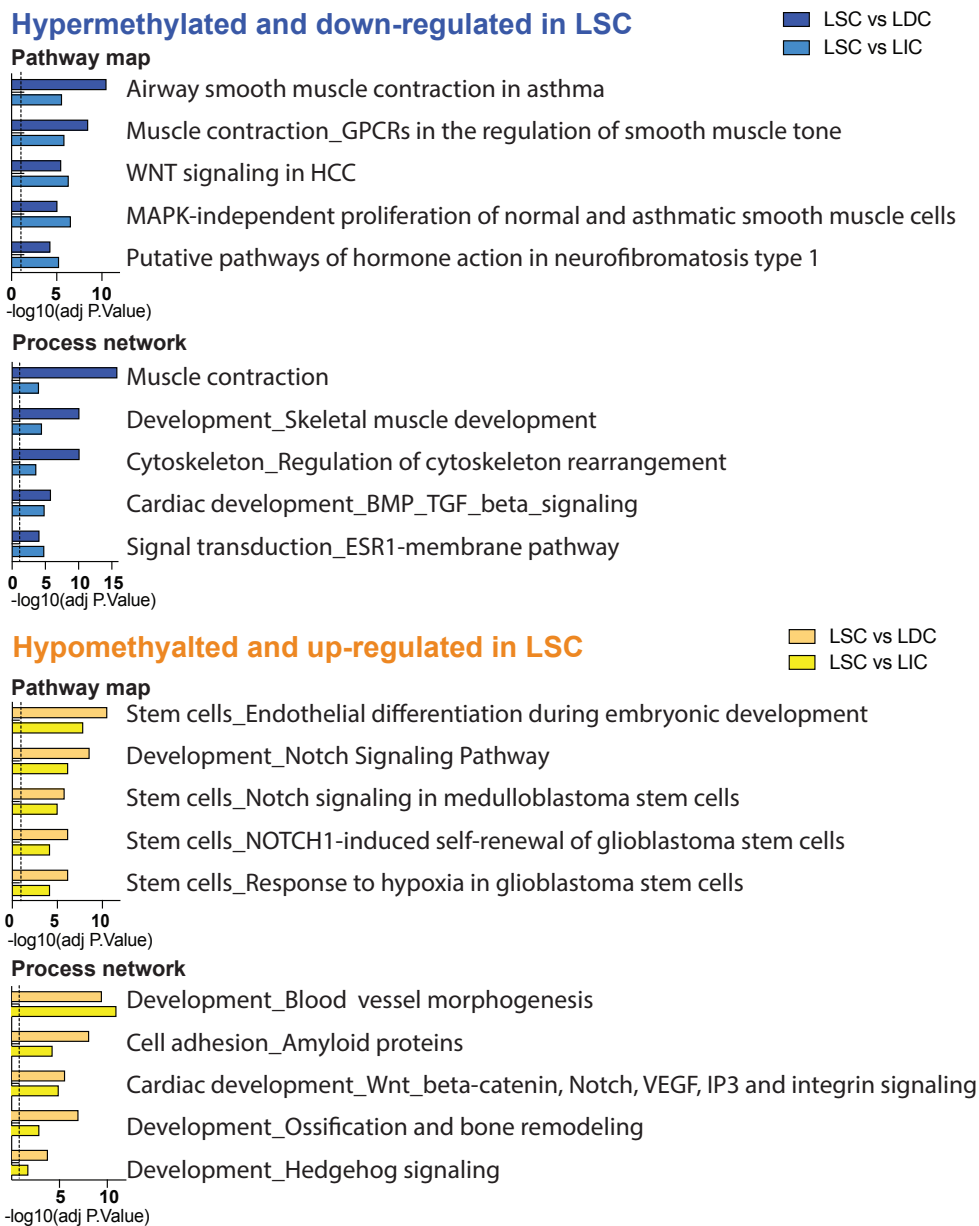
**Figure 3-7 Pathway analysis of genes both differentially expressed and methylated in LSC**

Enrichment analysis (Metacore) showing pathways and networks highly enriched in genes that are hypermethylated/downregulated and hypomethylated/upregulated in the LSC compared with LIC or LDC.
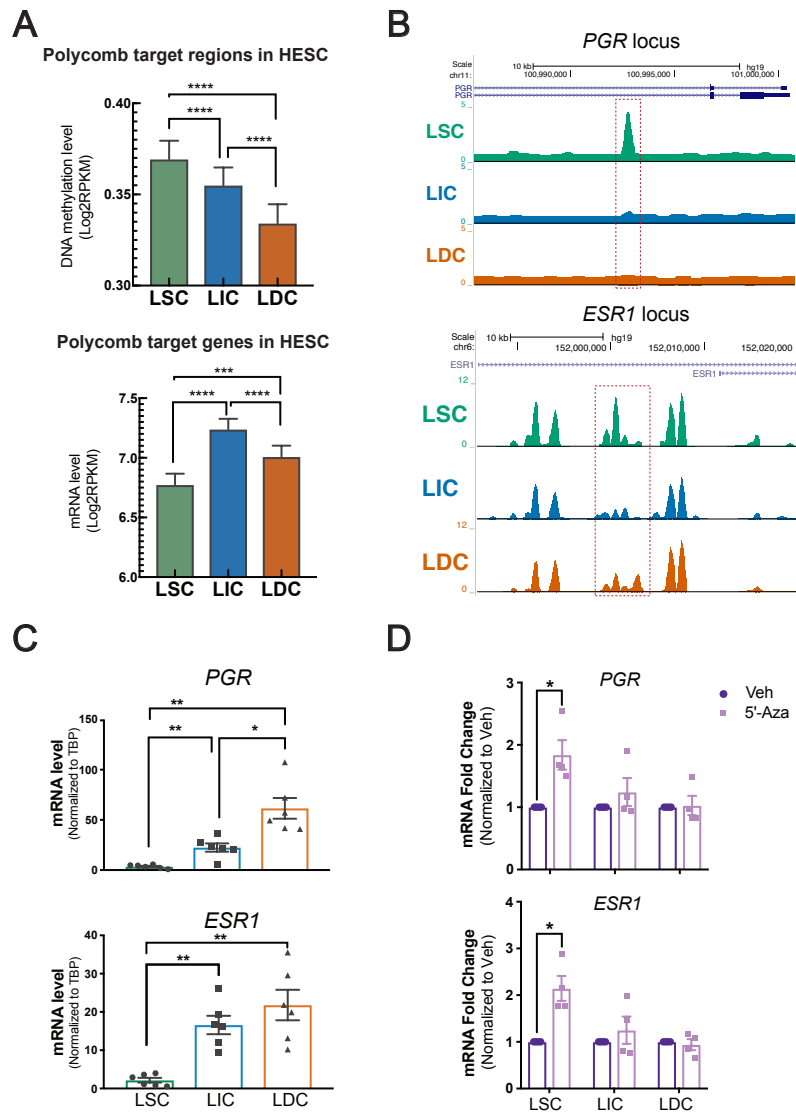
**Figure 3-8. PR and ESR1 are inhibited by DNA methylation in LSC.**
(**A**) Bar graph showing the average DNA methylation (left, region number = 2,808) and mRNA (right, gene number = 1,139) levels in LM cell populations at Polycomb target regions/genes identified from human embryonic stem cells (means ± SEM, n = 3 patients, ***$P<0.001$, ****$P<0.0001$, one-way ANOVA).   **B**) Representative genome browser view showing the methylation levels of the most significant DMRs around *PGR* (Chr11: 100,992,528 – 100,993,028) and *ESR1* (Chr6: 152,000,309 – 152,000,935) gene loci among three cell populations. **C**) Bar graph showing real-time qPCR validation of the mRNA levels of PR and ESR1 in each LM population (means ± SEM, n = 6, **$P<0.01$, *$P<0.05$ paired one-way ANOVA). **D**) Bar graph showing the mRNA fold changes of ESR1 and PR in LM cell populations treated with DNA methylation inhibitor 5′-Aza (96 hours, 100 nM) vs. vehicle (DMSO) (means ± SEM, n = 4, *$P<0.05$, paired t-test).

### 3.2.3 PR target genes are involved in LSC differentiation

Our findings suggest that DNA methylation influences *PGR* expression level; upregulation of PR is a hallmark of LSC differentiation to LIC/LDC. Although PR plays critical roles in LM development, the cellular and molecular mechanisms underlying PR action, especially on LSC differentiation, remain unclear [49]. Here, we identified PR downstream target genes using PR ChIP-Seq in LM tissues and PR knockdown in total LM primary cells followed by RNA-Seq (Figure 3-9 A). Two different PR siRNAs, siPGR-1 and siPGR-2, both significantly downregulated PR mRNA and protein levels (Figure 3-9 B). RNA-Seq showed significant transcriptome shift caused by PR knockdowns identifying 2832 PR target genes regulated by both siPGR-1 and siPGR-2, with 1601 genes downregulated and 1231 genes upregulated (n = 3 patients, Figures 3-9 C and D). Interestingly, focusing on the 4357 genes (2401 plus 1956, Figure 3-3 C) differentially expressed between LSC and LIC/LDC, PCA  and heatmap clustering analysis indicated that, upon PR knockdowns, the transcriptome of total primary LM cells (2 out of 3 patients) shifted toward that of LSC vs LIC/LDC (Figure 3-9 E and F). Gene set enrichment analysis (GSEA) also revealed that PR knockdown inhibited the ECM regulation, a critical signature of differentiated LM cells, and activated genes related to cell cycle and proliferation (Figure 3-9 G). These findings suggest that PR may be a critical transcription factor governing the transcriptional shift during LSC differentiation.
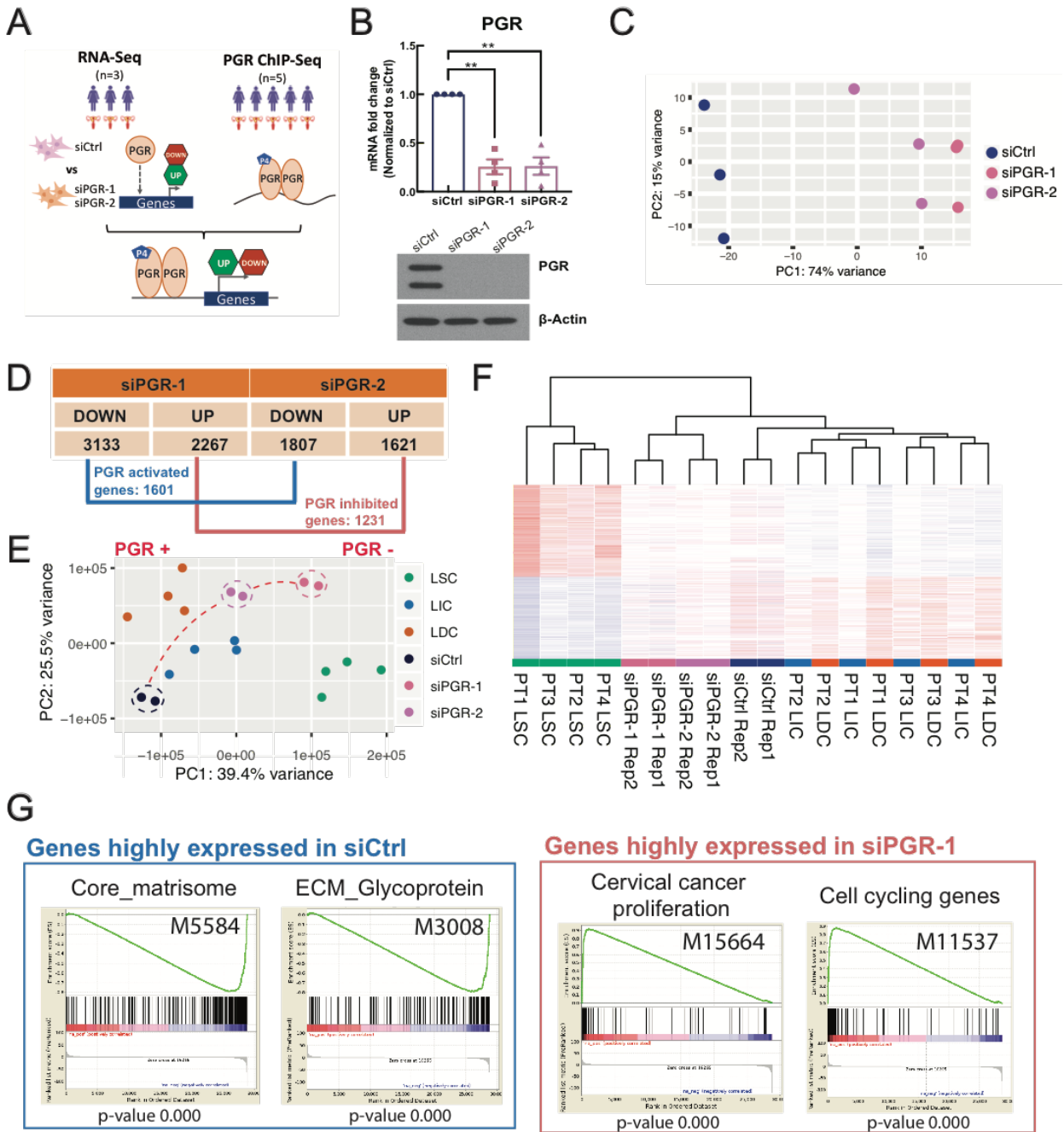
**Figure 3-9. PR is critial in maintaining genes important for LSC differentiation.**

**A**) Schematic showing the workflow of identifying PGR direct target genes using PR ChIP-Seq of patient LM tissues and RNA-seq of siCtrl- vs siPGR-transfected primary LM cells. **B**) Upper panel: Bar graphs showing qPCR results of PR mRNA levels in LM cells transfected with 2 different PR siRNAs (siPGR-1 and siPGR-2) or scrambled control siRNA (siCtrl) for 96 hours (means ± SEM, n=4, **$P<0.01$, paired one-way ANOVA). Lower panel:  Representative immunoblot showing PR levels with or without PR knockdown. **C**)

PCA plot showing transcriptome clustering among siCtrl and siPGRs samples (n = 3). **D**) Table showing up- and down-regulated genes after siPGRs. **E**) PCA plot showing clustering of LM populations (n = 4) and LM total primary cells transfected with siCtrl or siPGRs (n = 2) based on the mRNA levels of the 4357 differentially expressed genes between LSC and LIC/LDC. **F**) Hierarchical clustering heatmap showing mRNA expression pattern of the 4357 genes differentially expressed between LSC and LIC/LDC (n = 4) and LM primary cells transfected with siCtrl or siPGRs (n = 2). **G**) GSEA analysis of siCtrl and siPGR-1 RNA-Seq data in **C** based on C2 (curated gene sets) collection in the GSEA database. Top and bottom panels represent pathways of genes highly expressed in siCtrl and siPGR-1, respectively.

To identify direct PR targets, we then performed PR ChIP-Seq using fresh-frozen LM tissue samples. Using a stringent strategy, we identified PR-binding regions that were only present in all five LM tissues and found 6893 such regions (annotated to 3926 genes; 100kb from TSS). Integrative analysis of PR tissue ChIP-Seq and PR knockdown RNA-Seq data found that PR-binding sites were present near 651 PR-regulated genes, 436 activated and 215 inhibited, suggesting that they may be direct PR target genes (Figure 3-10 A).

Next, we investigated whether PR target genes contribute to the LSC differentiation process from LSC to LIC. PR target genes were compared against the 4357 genes commonly differentially expressed between LSC vs LIC and LDC assuming that these genes harbor the ones important for LSC transition to LIC or LDC. We found 137 direct and 240 indirect PR targets were upregulated while 26 direct and 105 indirect PR targets were downregulated during LSC differentiation (Figure 3-10 B).  Using hypergeometric test, we found a significant enrichment of PR activated gene in the cohort of upregulated genes in LIC/LDC (representation factor 1.7), indicating PR serves an important role as a transcription activator during the differentiation process. The 651 genes directly regulated by PR were highly enriched in pathways important for LM development, including myogenesis, progesterone signaling, progenitor cell differentiation, extracellular matrix remodeling, and interestingly DNA methylation (Figure 3-10 C). These data suggest that progesterone/PR signaling pathway activation might play an important role in LSC differentiation.

We followed up with the enriched pathway related with DNA methylation (Figure 3-10 C). We discovered two key DNA demethylases, TET1 and TET2, were affected by PR activity: RNA-Seq identified their downregulation by PR depletion; and PR tissue ChIP-Seq revealed significant PR binding around TET1 and TET2 promoter regions. TET1 and TET2 regulation by PR and the binding of PR at their loci were validated by qPCR and ChIP-qPCR, respectively (Figure 3-11 A and B). Together with the data that TET1 and TET3 were expressed at the lowest level in LSC (Figure 3-4 B), our findings suggested that increased PR expression during LSC differentiation might contribute to the methylation differences between the LSC and LIC/LDC via transcriptional regulation of DNA demethylases.

A



1165 | 436 | 3275 | 215 | 1016

- PR activated genes
- PR inhibited genes
- PR binding genes

B

**UP regulated genes in PR-rich LIC/LDC population**
1956

**PR activated genes**
1601

**Indirect** 1165

**Direct** 436

**Indirect** 240

**Direct** 137

Representation factor: 1.7 (p < 6.040e-30)

**DOWN regulated genes in PR-rich LIC/LDC population**
2401

**PR inhibited genes**
1231

**Indirect** 1016

**Direct** 215

**Indirect** 105

**Direct** 26

Representation factor: 0.8 (p < 6.477e-04)

C



Transcription_HIF-1 targets
Main growth factor signaling cascades in multiple myeloma cells
Role of HDAC and CaMK in control of skeletal myogenesis
Putative pathways of hormone action in neurofibromatosis type 1
MAPK-independent proliferation of normal and asthmatic smooth muscle cells
PR action in breast cancer: stimulation of cell growth and proliferation
Development_Regulation of lung epithelial progenitor cell differentiation
Role of DNA methylation in progression of multiple myeloma
Role of cell-cell and ECM-cell interactions in oligodendrocyte differentiation and myelination
Reproduction_Gonadotropin-releasing hormone (GnRH) signaling

0 2 4 6 8
-log10(adj P.Value)

**Figure 3-10. PR target genes are differentially expressed in LSC.**

**A**) Venn diagram showing overlapping genes between siPR-regulated genes identified in **Figure 3-9 D** and PR-binding genes identified by PR ChIP-Seq (n = 5). **B**) Identifying overlapping genes between PR-targeted genes and genes differentially expressed in LSC comparing with either LIC or LDC. Top: overlapping genes between PR-activated genes and upregulated genes during LSC differentiation. Bottom: overlapping genes between PR-inhibited genes and downregulated genes during LSC differentiation (n = 3). Hypergeometric p-value was calculated to evaluate the significance of the overlap (total gene number = 17611). **C**) Pathway and network enrichment analysis (Metacore) of the 651 PR direct target genes.
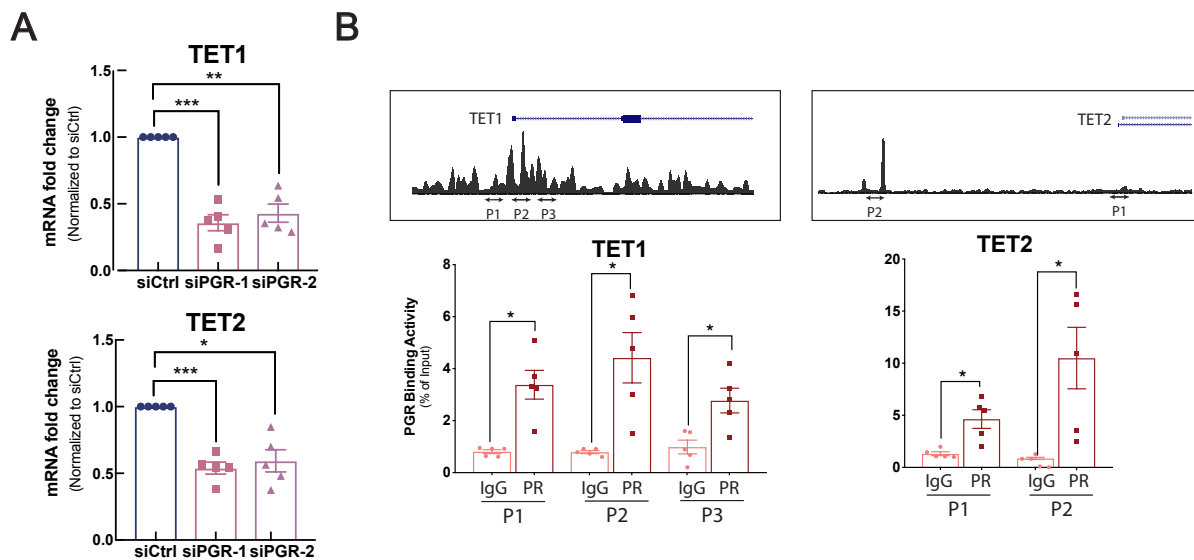
**Figure 3-11. PR regulates TETs expression levels.**

**A**) Bar graphs showing mRNA levels of TET1 and TET2 in siCtrl- and siPGR-transfected primary LM cells (means ± SEM, n = 5, *$P<0.05$, **$P<0.01$, ***$P<0.001$, paired one-way ANOVA). **B**) PR binding around *TET1* and *TET2* gene loci in primary LM cells. Upper: Representative genome browser track view of PR ChIP-seq in LM tissue; Lower: ChIP-qPCR quantifying PR binding around *TET1* and *TET2* in primary LM cells (means ± SEM, n = 5, *$P<0.05$, paired t-test).

**PR target genes are hypermethylated and suppressed in LSC**
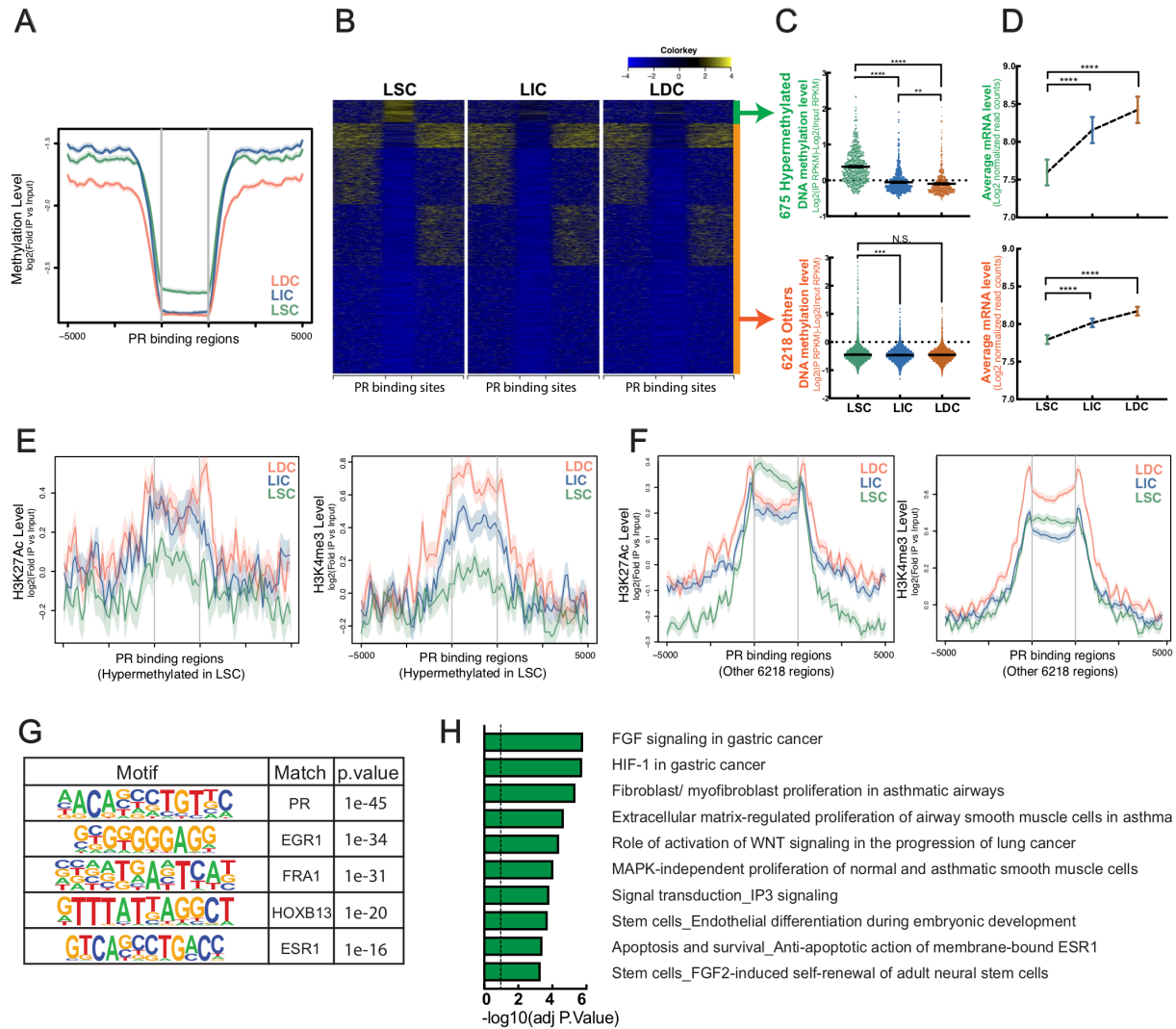


**Figure 3-12. PR target genes are inhibited by hypermethylation in LSC.**

**A**) Average line plots showing DNA methylation levels of 6893 PR-targeted regions in the three LM cell populations. **B**) Heatmaps showing the DNA methylation level around PR target regions (n = 5, 6893 PR peaks ± 5000bp) in each LM population. Yellow indicates hypermethylation and blue indicates hypomethylation. **C**) Dot plots showing quantification of DNA methylation levels (log2[RPKM$_{DNA\ methylation}$] − log2 [RPKM$_{Input}$]) of PR-targeted regions in each LM population. Top: PR-targeted regions with hypermethylation in LSC; Bottom: PR-targeted regions with hypomethylation in LSC. (means ± SEM, ****$P$<0.0001, ***$P$<0.001, **$P$<0.01, paired one-way ANOVA) **D**) Average plots showing the mRNA levels (RPKM) of PR-targeted genes in different LM populations. Top: genes associated with hypermethylated PR-targeted regions in LSC; Bottom: genes associated with hypomethylated PR-targeted regions in LSC.

(means ± SEM, ****$P$<0.0001, paired one-way ANOVA). **E**) Average line plot showing histone modification levels (Left: H3K27Ac level; Right: H3K4me3 level) of hypermethylated PR-targeted regions in LSC. **F**) Average line plots showing histone modification levels (Left: H3K27Ac level; Right: H3K4me3 level) of hypomethylated PR-targeted regions in the three LM cell populations. **G**) Motif analysis (Homer) of PR-targeted regions related with hypermethylation in LSC. **H**) Pathway and network enrichment analysis (Metacore) of genes adjacent to hypermethylated PR-targeted regions.

To evaluate the role of DNA methylation in PR signaling during LSC differentiation, we integrated the MethylCap-Seq data with PR tissue ChIP-Seq data. Among the three populations, LSC had the highest overall DNA methylation level in regions at or adjacent (<1 kb) to the 6893 PR-targeted sites discovered via PR tissue ChIP-Seq (Figure 3-12 A). Using K-means clustering, we pinpointed a cohort of PR-targeted regions (675, 9.8% of 6893) hypermethylated in LSC compared with LIC and LDC (Figures 3-12 B and C top panel). We observed low DNA methylation levels with small difference among the three populations at other PR-targeted regions (Figure 3-12 C, bottom panel). We further evaluated mRNA levels of the genes associated with PR-targeted regions in each LM population. Compared to LIC and LDC, LSC expressed the lowest levels of the genes adjacent to PR-targeted regions (Figure 3-12 D), and the expression difference of the genes associated with the 675 hypermethylated PR-targeted regions was particularly dramatic possibly due to the hypermethylation-related suppressions in LSC. Additionally, we found the lowest enrichment of the active histone marks (H3K27ac, H3K4me3) associated with these hypermethylated PR-targeted regions in LSC compared to LIC and LDC, confirming the inhibitory status of the neighboring genes (Figure 3-12 E and F).

To evaluate what other transcription factors, besides PR, might be interfered by the hypermethylation, we conducted a motif analysis to investigate the consensus sequences enriched in these hypermethylated PR-targeted regions. Motif enrichment showed that progesterone response elements (PREs) and several other hormone-related transcription factor binding elements, such as FRA1/AP-1 and ESR1, were highly enriched in the hypermethylated PR-targeted regions in LSC (Figure 3-12 G). This suggested that the disruption of

hormone signaling in LSC was accomplished by both PR deficiency and hypermethylation of hormone response elements. Importantly, pathways important for LM stem function and disease development, including FGF and WNT signaling in cancer, stem cell self-renew and differentiation, and ECM-reglated proliferation, were enriched in those genes associated with the hypermethylated PR-targeted regions (Figure 3-12 H) [140, 151].

### 3.2.4   Progesterone signaling is interrupted at key PR-targeted genes by hypermethylation

To investigate the mechanistic link between DNA methylation and progesterone signaling during LSC differentiation, we focused on the regulation of several key genes with established roles important for stem cell differentiation or LM pathogenesis, including TIMP3, ROR2, GREB1, MYH11, and WNT5A. TIMP3, tissue inhibitor of metalloproteinase 3, regulates ECM at the tumor–stromal interface [152], and is critical in driving differentiation and decreasing tumorigenesis in the mammary gland and muscle tissue [153, 154]. ROR2, a receptor tyrosine kinase-like orphan receptor, regulates osteoblast differentiation from mesenchymal stem cells [155]. GREB1, MYH11, and WNT5A play roles in hormone signaling, muscle development, and stem cell regulation [156-158].

MethylCap-Seq data showed that, compared to the other two cell populations, LSC had the highest DNA methylation levels around the PR-binding sites of *TIMP3*, *GREB1*, *MYH11*, *ROR2*, and *WNT5A* gene loci (Figure 3-13 A and B). Opposite to the DNA methylation status, mRNA levels of TIMP3, ROR2, GREB1, MYH11, and WNT5A were the lowest in LSC (Figure 3-13 C). To assess the effect of DNA methylation on the transcriptional activities of these genes, we treated individual cell populations with 5′-Aza (100 nM) for 96 hours. We found that 5′-Aza treatment significantly increased the mRNA levels of these genes in LSC. We observed a trend of increased expression of these gene in LIC and LDC, which did not reach statistical significance (Figure 3-14 A). Although LM harbors a small population of LSC, 5′-Aza (100 nM, 6 days) treatment of total primary LM cells also significantly stimulated the expression of these genes Figure 3-14 B). Furthermore, 5′-Aza also upregulated PR expression level and enhanced PR recruitment to its target regions in primary LM cells (Figures 3-14 C and D), contributing to the mRNA increase of PR target genes.

We then evaluated whether PR is necessary for 5′-Aza-mediated upregulation of the target genes. We performed PR knockdowns in total primary LM cells treated with vehicle or 5′-Aza (100nM, 6 days). PR depletion not only significantly decreased mRNA levels of the five candidate target genes in the absence of 5′-Aza, but also completely blocked 5′-Aza-stimulated up-regulation of GREB1, ROR2, and WNT5A, indicating PR is critical for the regulation of these target genes (Figure 3-14 E).

Interestingly, these five genes were also targeted by TET1 around the hypermethylated regions in LSC adjacent to the PR-targeted regions (Figure 3-15 A). PR knockdown in primary LM cells decreased TET1 recruitment at these regions, presumably through PR-mediated TET1 gene regulation. We further examined the DNA methylation level at differential methylated regions between LSC and LIC/LDC with or without PR knockdown. 72h of PR knockdown resulted DNA methylation level in total LM primary cell to slightly increase at regions with higher DNA methylation in LSC (Figure 3-15 B). Together, these findings suggest that during early LSC differentiation, DNA methylation decreases and PR level increases, leading to the enhanced recruitment of PR to its target genes and the increase of TET1 level, thus furthering the loss of the global DNA methylation. This positive feedback loop increased expression of genes which might be important for LSC differentiation.
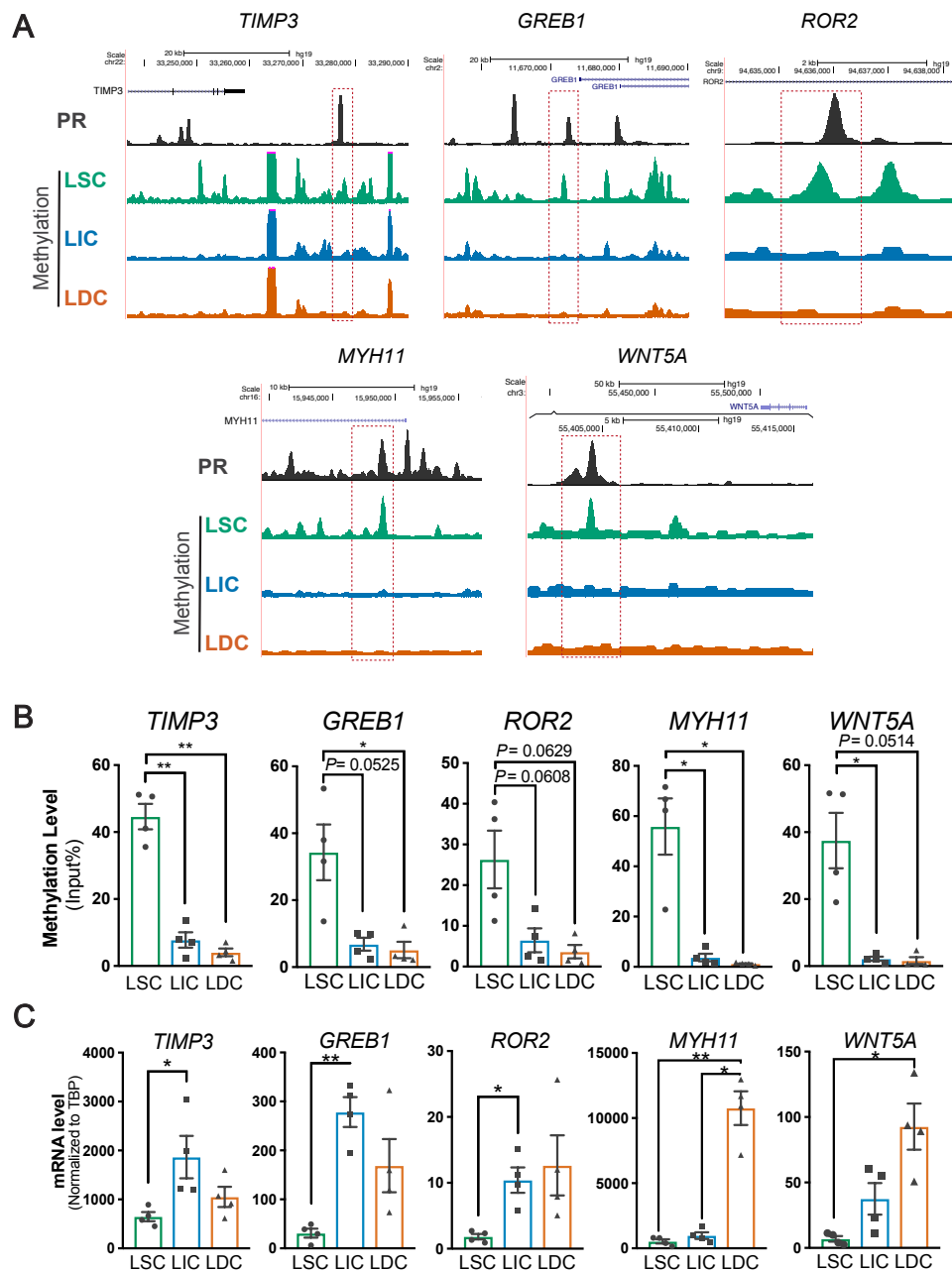
**Figure 3-13. PR target genes are hypermethylated in LSC.**

**A**) Representative genome browser view showing PR-binding activities and DNA methylation levels at PR target genes. (Gray represents PR ChIP-Seq; Green, blue, and orange represent LSC, LIC, and LDC MethylCap-Seq, respectively). **A**) Bar graphs showing MethylCap-qPCR quantification of DNA methylation levels in each LM cell population at DMRs around hypermethylated PR-targeted regions (means ± SEM, n = 4, *P<0.05, **P<0.01, paired one-way ANOVA). **B**) Bar graph showing mRNA levels of PR-targeted genes in each LM population (means ± SEM, n=4, *P<0.05, **P<0.01, paired one-way ANOVA).
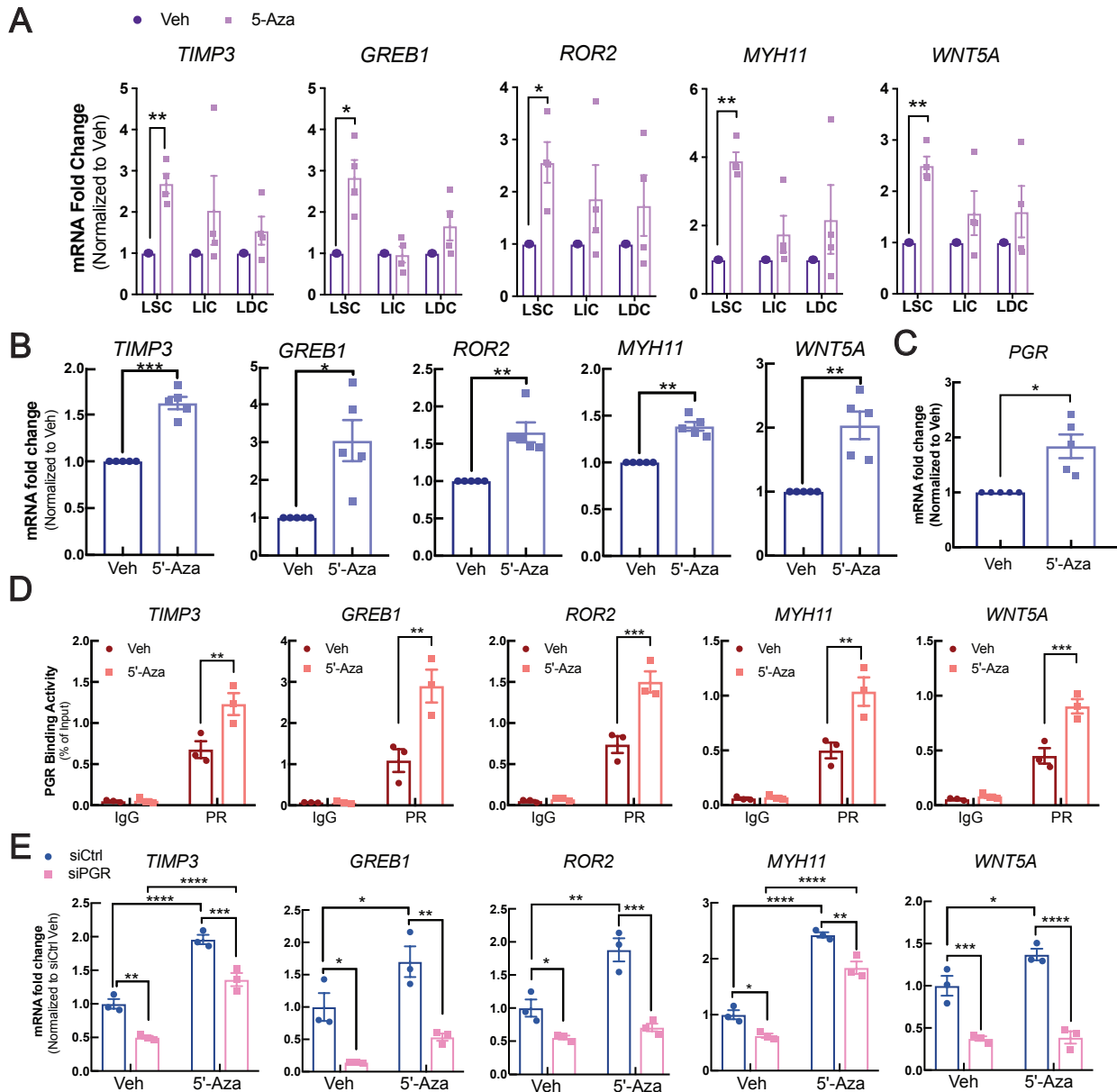
**Figure 3-14. PR target genes can be upregulated by DNA methylation inhibitor 5'-Aza**

**A**) Bar graph showing the mRNA fold changes of PR target genes in LM cell populations treated with DNA methylation inhibitor 5′-Aza (96 hours, 100 nM) vs. vehicle (DMSO) (means ± SEM, n = 4, *$P<0.05$, **$P<0.01$, paired t-test). **B**) Bar graphs showing mRNA levels of PR-targeted genes in vehicle or 5′-Aza treated primary LM cells (means ± SEM, n = 5, *$P<0.05$, **$P<0.01$, ***$P<0.001$, paired t-test). **C**) Bar graph showing qPCR results of PR mRNA levels in LM cells treated with or without 5′-Aza (100 nM) for 6 days (means ± SEM, n = 5, *$P<0.05$, paired t-test). **D**) ChIP-qPCR showing PR recruitment to PR binding sites adjacent to the DMRs around the PR-targeted genes in primary LM cells exposed to vehicle (DMSO) or 5′-

Aza (6 days, 100nM) (means ± SEM, n = 3, **$P<0.01$, ***$P<0.001$, two-way ANOVA). Samples were all treated with R5020 ($10^{-7}$ M) for 1 hour before harvesting for ChIP. **E**) Bar graphs showing mRNA levels of PR-targeted genes in siCtrl- and siPGR1-transfected (50nM, 72h) primary LM cells with or without 5′-Aza treatment (6 days, 100nM) (means ± SEM, n = 3, *$P<0.05$, **$P<0.01$, ***$P<0.001$, ****$P<0.01$, two-way ANOVA).
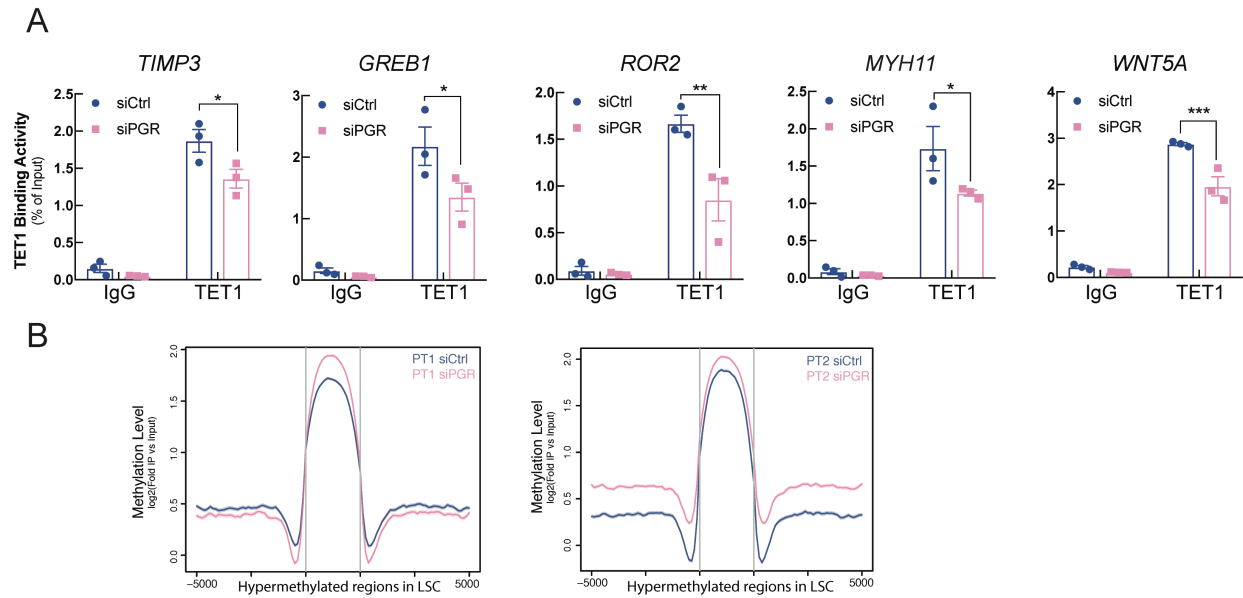
**Figure 3-15. PR contributes to the DNA methlation change during LSC differentiation.**

**A**) ChIP-qPCR showing TET1 recruitment to the DMRs around the PR-targeted genes in primary LM cells transfected with siCtrl or siPGR-1 (50nM, 72h) (means ± SEM, n = 3, *$P<0.05$, **$P<0.01$, ***$P<0.001$, two-way ANOVA). **B**) Average line plot showing DNA methylation level at hypermethylated regions in LSC compared with LIC/LDC in LM total primary cells with or without PR knockdown.

### 3.2.5  5′-Aza reduces LSC stemness and sensitizes LSC to antiprogestin treatment
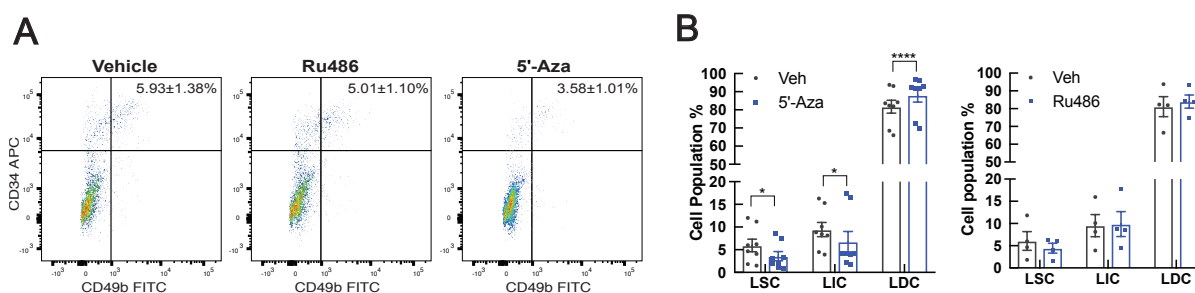


**Figure 3-16. 5'-Aza influences the proportion of each LM population.**

**A**) Representative flow cytometry scattergrams showing the LM cell populations isolated from LM tissue explants after 48-hour incubation with vehicle (DMSO), 5′-Aza (100 nM), or RU486 (1 $\mu$M).  **B**) Bar plots quantifying the percentage of each LM cell population in LM explants treated with vehicle or 5′-Aza (100 nM, n = 8) or Ru486 (1 $\mu$M, n=4) (means ± SEM, *$P$<0.05, ****$P$<0.0001, paired two-way ANOVA).

We demonstrated that the crosstalk between DNA methylation and PR action contributes to the expression changes of critical genes during LSC differentiation. We further investigated the function of 5′-Aza in LM tumor progression and compared its effect with that of RU486, a progesterone antagonist shown to inhibit LM growth [49]. To test the ability of 5′-Aza to regulate LSC function, we treated LM tissue explants with vehicle (DMSO), RU486 (1 $\mu$M), or 5′-Aza (100 nM) for 48 hours and analyzed the proportions of each LM cell population. 5′-Aza treatment decreased the LSC and LIC populations and increased the LDC population compared to vehicle, whereas RU486 did not significantly change the LM cell composition (Figure 3-16 A and B). We also tested the effect of RU486 or 5′-Aza on primary LM cells' clonogenic activity, an indicator of tumor stem cells [159]. Cells were treated with vehicle (DMSO), RU486 (1 µM), or 5′-Aza (25 nM, 50 nM, or 100 nM) for six days, and 500 viable cells from each treatment group were plated in each well of a twelve-well plate and cultured for 21 days without further treatment. We found that pretreatment with 5′-Aza had markedly decreased colony formation in primary LM cells even at a very low dose (25 nM), whereas RU486

did not have a significant effect (Figure 3-17). In addition, we tested the tumor initiation ability of primary LM cells (1x10$^6$ viable cells), pretreated with vehicle, 5′-Aza, or RU486 for six days. We found that primary LM cells pretreated with 5′-Aza (36.30±3.57% of vehicle size,) regenerated significantly smaller tumors compared to RU486-pretreated (76.31±1.86% of vehicle size) or vehicle-pretreated primary LM cells (Figure 3-18 A).
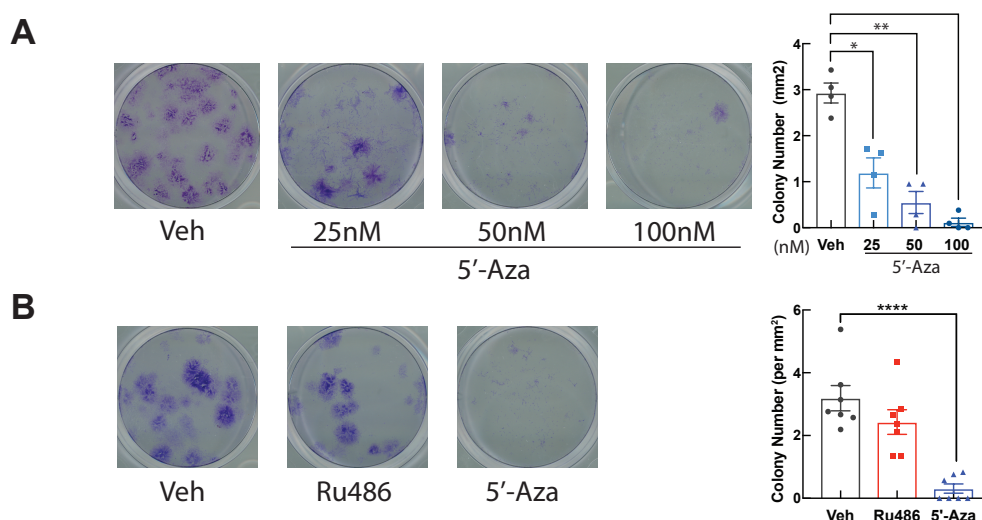


**Figure 3-17. 5'-Aza inhibited the colony formation ability of primary LM cells.**

**A**) Colony formation assay of LM primary cells pretreated with vehicle (DMSO) or 5′-Aza (25 nM, 50 nM, 100 nM) for 6 days. After treatment, 500 viable cells were seeded in each well of 12-well plates and cultured for 21 days without further treatment for colony formation assay. Left panel: representative pictures of the colony formation assay; right panel: quantification of clone numbers per mm$^2$ of culture area under each treatment. (means ± SEM, n = 4, *$P<0.05$, **$P<0.01$, ***$P<0.001$, paired one-way ANOVA) **B**) Colony formation assay of primary LM cells pretreated with vehicle (DMSO), 5′-Aza (100 nM), or RU486 (1 μM) for 6 days. Left panel: representative pictures of the colony formation assay; right panel: quantification of clone numbers per mm$^2$ of culture area under each treatment. (means ± SEM, n = 7, ****$P<0.0001$, paired one-way ANOVA).

We found that 5′-Aza modulates PR expression and alters PR downstream target gene expression. As a progesterone antagonist, RU486 competes with progesterone for receptor binding and alters receptor conformation [160]. Therefore, we then tested whether 5′-Aza alters the inhibitory effect of RU486 on the growth of existing LM tumors. Primary LM cell xenografts ($1 \times 10^6$ viable cells) were established for three weeks, then the mice were administered with vehicle (PBS) or 5′-Aza twice per week followed by treatment with or without RU486 (Figure 3-18 B). 5′-Aza alone partially decreased LM growth, but co-treatment with RU486 completely blocked LM regeneration (Figure 3-18 B). The combined 5′-Aza and RU486 treatment significantly decreased the number of Cyclin D1-positive cells (proliferation marker) (Figure 3-18 B). Together, these findings suggest that inhibition of DNA methylation decreases LSC stemness and drives their differentiation, thereby stimulating LM cell responsiveness to antiprogestin therapies with decreased cell proliferation.

We further explored the molecular mechanisms underlying 5′-Aza–mediated LSC differentiation and LM growth inhibition. We performed RNA-Seq on vehicle (DMSO) or 5′-Aza (100 nM, 6 days) treated primary LM cells. The global transcriptome changed drastically with 5′-Aza treatment, with clear clustering based on treatment status (n = 3 patients, Figure 3-19 A). We identified 3046 genes upregulated and 5120 genes downregulated in the 5′-Aza group compared to the vehicle group. Pathway enrichment analysis revealed that genes downregulated by 5′-Aza were related to cell cycle regulation, while genes upregulated by 5′-Aza were enriched in immune response activation, stem cell differentiation and ECM remodeling pathways (Figure 3-19 A). GSEA showed similar results, with 5′-Aza treatment significantly downregulating proliferation-related while upregulating differentiation- and endocrine-related genes. (Figure 3-19 B).

**A**

Vehicle

Ru486 ($10^{-6}$M)

5'-Aza ($10^{-7}$M)

Cultured for 6 days

$10^6$ cells → 4 weeks → Harvest tumors

Vehicle    Ru486    5'-Aza

2mm    2mm    2mm

Fold Change (Normalized with Veh)

Veh    Ru486    5'-Aza

*

****

**B**

Tumor initiation    Priming    Treatment

Vehicle

3 weeks → Vehicle    1 week    Placebo +Vehicle
5'-Aza ($10^{-7}$M)    2 IPs    Ru486 pellet + Vehicle    2 weeks    Harvest
5'-Aza ($10^{-7}$M)         Placebo + 5'-Aza    4 IPs
              Ru486 pellet + 5'-Aza

Vehicle    Ru486    5'-Aza    Ru486+5'-Aza

2mm    2mm    2mm    2mm

Tumor Volumn Fold Change (Normalized to Veh)

Ru486    −    −    +    +
5'-Aza    −    +    −    +

****  ****  ****  **  ***

**C**

Vehicle    Ru486

50µm    50µm

5'-Aza    Ru486+5'-Aza

50µm    50µm

% of CyclinD1 positive cells

Ru486    −    −    +    +
5'-Aza    −    +    −    +

****  ****  ****  **  *

**Figure 3-18. 5'-Aza inhibits *in vivo* tumor growth.**

**A**) Top: workflow of the *in vivo* xenograft mouse experiments using primary LM cells pretreated with vehicle (DMSO), RU486 (1 µM), or 5′-Aza (100 nM) for 6 days. Bottom left: representative pictures of regenerated LM tumors. Bottom right: quantification of the tumor volumes (means ± SEM, patient number = 3, vehicle group kidney graft = 17, RU486 group kidney graft = 6, 5′-Aza group kidney graft = 17, *$P<0.05$, ****$P<0.0001$, one-way ANOVA). **B**) Top: workflow of the *in vivo* xenograft mouse experiments using primary LM cells. Bottom left: Representative pictures of regenerated tumors under each treatment. Bottom right: quantification of the tumor volumes (means ± SEM, patient number = 3, vehicle group kidney graft = 22, RU486 group kidney graft = 21, 5′-Aza group kidney graft = 18, RU486+5′-Aza group kidney graft = 20, **$P<0.01$, ***$P<0.001$, ****$P<0.0001$, two-way ANOVA). **C**) Left panel, representative images of immunohistochemical staining of Cyclin D1; Right panel, percentage of Cyclin D1-positive cells among total cells counted in a 40X field (means ± SEM, n = 10 slides, *$P<0.05$, **$P<0.01$, ****$P<0.0001$, two-way ANOVA).

**A**



**B**
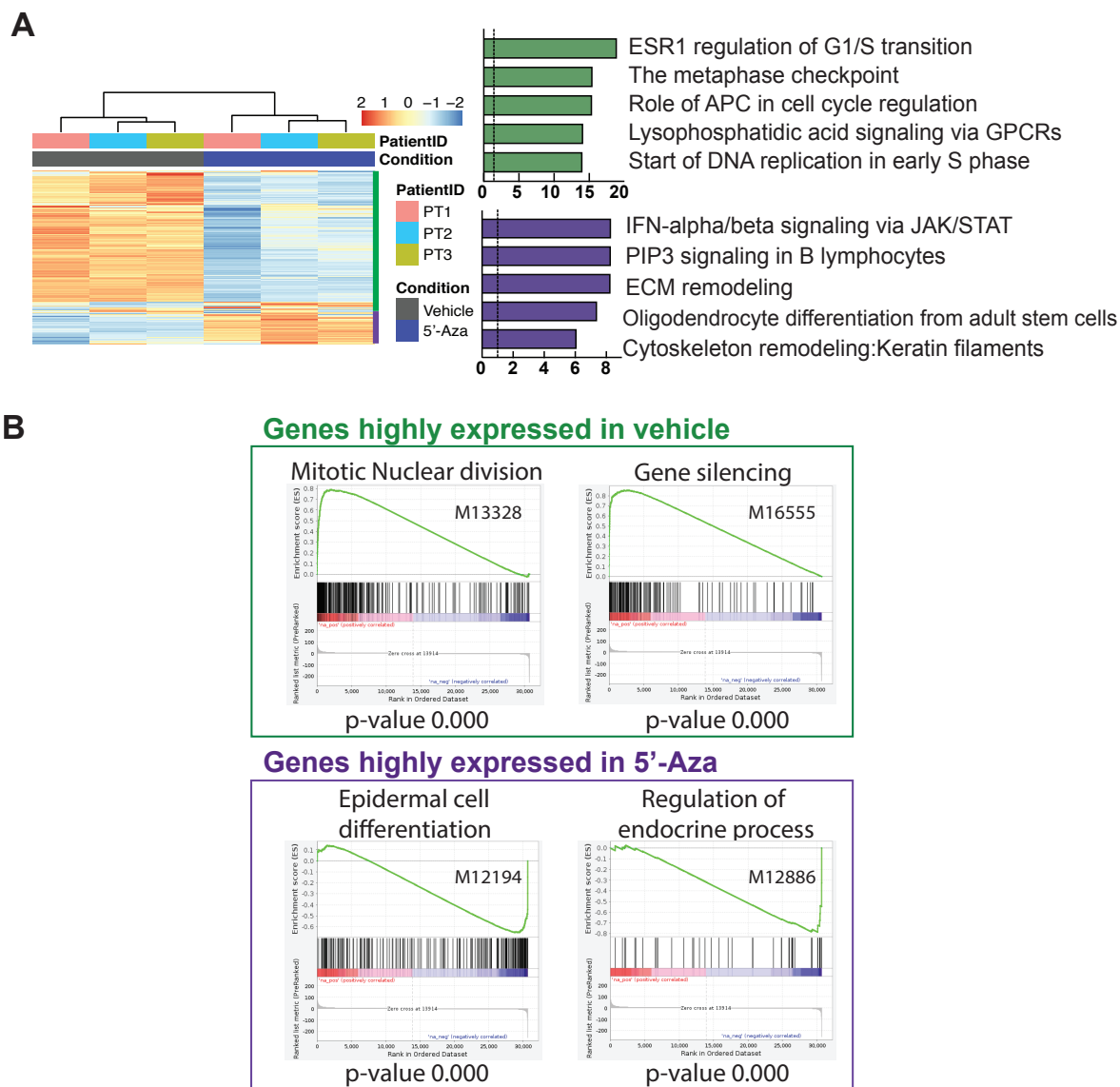


**Figure 3-19. 5'-Aza shifts the genome-wide transciptome in primary LM cells.**

**A**) Left: hierarchical clustering heatmap of RNA-Seq showing the transcriptome profiles of primary LM cells treated with vehicle (DMSO) or 5′-Aza (100 nM) for 6 days. Top and bottom right panels represent enrichment analysis (Metacore) of genes downregulated and upregulated by 5′-Aza, respectively. **B**) GSEA analysis of RNA-Seq data in G based on C5 (Gene ontology (GO) terms) collection in the GSEA database. Top and bottom panels represent GO terms of genes downregulated and upregulated by 5′-Aza, respectively.

### 3.3 Conclusion
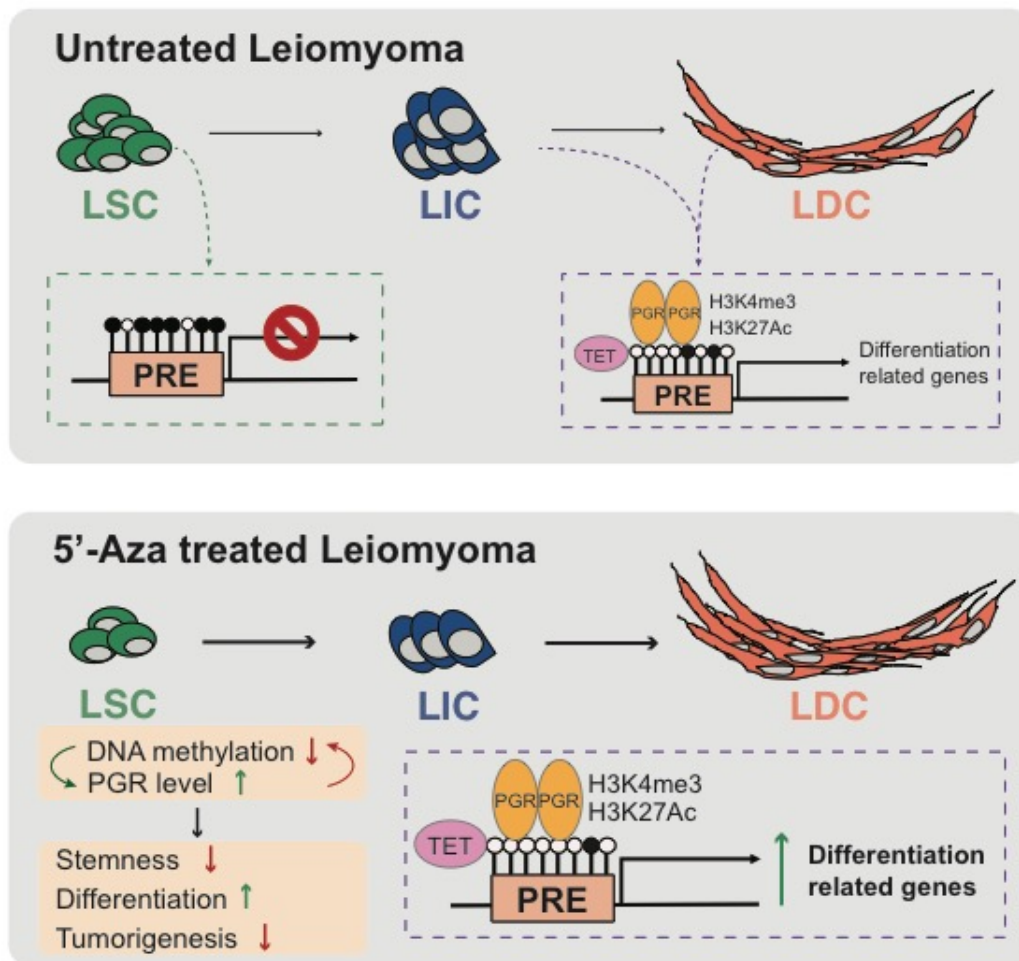


**Figure 3-20. Schematic of the proposed role of DNA methylation during LSC differentiation.**

We demonstrated that LSC harbored a unique transcriptome and DNA methylome. Globally, hypermethylation in LSC suppresses key genes important for stem cell differentiation, including ESR1 and PR. PR target genes are highly enriched in stem cell differentiation-related processes and differentially expressed in LSC vs. LIC and LDC. In LSC, some of the PR target genes are inhibited by hypermethylation at PR regulatory regions, contributing to stemness preservation. The DNA methylation inhibitor 5′-Aza upregulated PR transcriptional activities at its target genes, decreased the stem cell population by stimulating LSC differentiation, and reduced LM regeneration (Figure 3-20).

## 3.4    Materials and Methods

**Tissue collection**

Northwestern University's Institutional Review Board approved the use of human tissue. LM tissues were obtained from premenopausal women undergoing either myomectomy or hysterectomy (age 38±9 years, range 28–49 years). We have obtained informed consent from all participants. Patients receiving hormone treatment six months prior to surgery were excluded. Tissues were dissociated and cells isolated as previously described [64].

**Antibody and primer information**

All antibodies and primers used in this manuscript are listed in Appendix Table 2.

**Antibody-based cell sorting**

CD34$^+$/CD49b$^+$, CD34$^+$/CD49b$^-$, and CD34$^-$/CD49b$^-$ LM cells were FACS-sorted as previously described [64, 68].

**RNA isolation, real-time qPCR, and RNA-Seq**

Total RNA was isolated using the Qiagen Allprep RNA/DNA mini kit (Qiagen, 80204) or Qiagen RNeasy RNA micro kit (Qiagen, 74004). cDNA was synthesized using qScript cDNA SuperMix (Total RNA >100 ng) (VWR International, 95048-100) or SuperScript VILO Master Mix (Total RNA <100 ng) (Thermo Fisher Scientific, 11754050). mRNA levels were quantified using real-time qPCR normalized to TBP as previously described [161]. Total RNA quality was examined using Bioanalyzer: RNA Pico assay. RNA-Seq libraries were prepared using the NEBNext Ultra II RNA Library Prep with Sample Purification Beads (NEB, #7775).

**MethylCap-Seq and MethylCap-qPCR**

Genomic DNA was extracted from LM cells using DNeasy Blood and Tissue Kit (Qiagen, 69504) or Qiagen Allprep RNA/DNA mini kit (Qiagen, 80204), and fragmented to 300-500bp using Covaris M220 (Covaris). 500 ng of fragmented genomic DNA was incubated with H6-GST-MBD fusion proteins that can bind

methylated cytosines. Previous study has shown that a single fully methylated CpG is sufficient for the interaction between the H6-GST-MBD fusion protein and methylated DNA [162]. The protein-DNA complex is then precipitated with antibody-conjugated beads that are specific to the protein tag. Deep sequencing provides greater genome coverage, representing the majority of MBD-bound methylated DNA. Real-time qPCR were also performed to validate the sequencing result. qPCR primer sequences are listed in Appendix Table 2.

**Chromatin immunoprecipitation assay**

0.2-0.5g of frozen LM tissues were used for ChIP-Seq using the SimpleChIP Kit (Cell Signaling Technology, 9005). $5 \times 10^6$ P0 LM cells were used for ChIP-qPCR using the SimpleChIP Kit. For PR ChIP, cells were treated with R5020 for 1 hour before harvesting. $10\mu g$ of chromatin was isolated and incubated with $3\mu g$ antibodies (PR or TET1). Normal rabbit IgG was used as a negative control.

**Native ChIP-Seq for histone modification**

We followed the protocol published by Peter van Galen et al. for native ChIP-Seq for histone modifications [163]. Briefly, we FACS-sorted each population of LM cells ($3 \times 10^5$ cells for each population), which were then resuspended in 20µl PBS with 2x proteinase inhibitor cocktail (PIC, Thermo Fisher, 1862209) and 20µl 2x Lysis Buffer containing 100u of Micrococcal Nuclease (NEB #M0247S). Cells were lysed on ice for 20 min and chromatin was digested for 10 min at 37°C. Chromatin from each cell type was labeled with T7 adapters using the Fast-Link DNA Ligation Kit (Lucigen, LK6201H). Then, the labeled chromatin from three cell populations was pooled and the sample volume was brought up to 200µl/IP with Dilution Buffer containing PIC. Samples were then split and incubated with 1.5µl H3K4me3 antibody (Millipore, 07-473), 2µl H3K27Ac antibody (Active Motif, 39133), or 3µl H3 antibody (Active Motif, 39763, as a control) overnight at 4°C on a tube rotator. 25µl of protein G Dynabeads (Thermo Fisher, 10004D) were washed two times with Dilution Buffer/PIC, added to each IP, and incubated for an additional 3 hours at 4°C on a rotator. Samples were magnetized and the unbound portion was discarded. Beads were washed twice in 200µl of ice-cold RIPA buffer, once in 200µl ice-cold RIPA/High Salt, once in 200µl ice-cold LiCl buffer, and once in

100μl TE. Washed beads were resuspended in 100μl Elution Buffer and then incubated with 0.5μl of RNase Cocktail (Invitrogen, AM2286) and 0.5μl of protease K (10mg/ml) at 37°C for 30 min and 63°C for 1 hour. DNA fragments were purified using 1X SPRI beads (Beckman Coulter, A63881). DNA fragments were then *in vitro* transcribed to RNA using HiScribe T7 Quick High Yield RNA Synthesis Kit (NEB, E2050S) and reverse transcribed to cDNA using SuperScript III kit (Thermo Fisher, 18080-400) to add sequencing adapters. cDNA sequencing library of each IP was amplified using PFU Ultra II HS 2× Master mix (Agilent, 600850). All barcode information was provided in the original publication [163].

**Next-generation sequencing**

Next-generation sequencing libraries for MethylCap-Seq and PR ChIP-Seq were prepared using the KAPA Hyper Prep Kit (KK8502, KAPA Biosystems) and KAPA Single-Indexed Adapter Kit (KK8710, KAPA Biosystems). The libraries were sequenced at the Northwestern University NUSeq Core Facility using the NextSeq 500 system (Illumina, San Diego, CA, USA) with 20-40 million reads per sample. Sequencing methods were 75bp or 50bp single-end for RNA-Seq, 75bp single-end for PR ChIP-Seq, and 75bp paired-end for MethylCap-Seq and histone modification ChIP-Seq.

**Bioinformatic analysis**

Sequences were aligned to the hg19 reference genome using TopHat for RNA-Seq, Bowtie2 for PR ChIP-Seq and MethylCap-Seq, and Burrows-Wheeler Aligner (BWA) for histone modification ChIP-Seq [164-166]. Differential gene expression from RNA-Seq was detected using DESeq2 following the cutoff: FDR-adjusted $P$<0.05. We performed the following MethylCap-Seq and ChIP-Seq analyses using Homer: peak callings (-factor for PR ChIP-Seq and -histone for MethylCap-Seq), differential methylation analyses (getDifferentialPeaksReplicates.pl, FDR-adjusted $P$<0.05, fold change>2), overlapping region identifications (mergePeaks -d 100), motif analyses (findMotifsGenome.pl), and peak annotations (annotatePeaks.pl) [135]. Pathway enrichment analysis was performed using Metacore V6.34 (Thomson Reuters) and GSEA V4.0.1 [167, 168]. Bam files from replicates were merged by samtools before NGSplot and UCSC Genome Browser visualizations [169]. Sequencing tracks were visualized using the UCSC

Genome Browser. Visualization of DNA methylation and histone modification average levels at specified regions (PR-binding regions and gene body regions) were performed using NGSplot [136]. Hypermethylated PR-targeted regions in LSC were identified using the K-mean clustering option (-GO km, -KNC 5) included in NGSplot. Methylation levels at PR-binding regions and Polycomb target regions were quantified on bam files using normalized RPKM value ($log2(RPKM_{IP}+1) - log2(RPKM_{Input}+1)$).

**Primary cell and tissue explant culture**

Primary LM cell and *ex vivo* explant culture of LM tissue were performed as previously described [68]. For 5′-Aza treatment, if not indicated otherwise in the figure legends, cells were treated with 100 nM 5′-Aza for 6 days with medium refreshed every day (Sigma-Aldrich, A3656). To elucidate the role of 5′-Aza and RU486 in LSC function, the explants were subjected to vehicle, 5′-Aza (100 nM), or RU486 treatment ($10^{-6}$ M) for 48 hours and analyzed using flow cytometry based on their expression of CD34 and CD49b, as described previously [68, 161].

**Stem cell culture**

Each freshly FACS-sorted population of LM cells was cultured in mesenchymal stem cell growth medium (Lonza, PT-3238) in low-attachment 96-well plates (Fisher Scientific, 07-201-680) to maintain stem cell characteristics. Cells were recovered in basal mesenchymal stem cell growth medium for 3 days and followed by treatment with 5′-Aza (100nM) or vehicle for 96 hours.

**PR siRNA knockdown**

LM passage zero cells were transfected with two different PR siRNAs (Dharmacon, D-003433-03-0010 and D-003433-01-0010) or control scrambled siRNA (D-001810-10-05) for 72 hours, and then stimulated with R5020 ($10^{-7}$M) for 24 hours before harvesting.

**Colony formation assay**

Primary LM cells freshly isolated from LM tissue were treated with vehicle (DMSO) or 5′-Aza (100 nM) for 6 days. After treatment, 500 viable cells were seeded in each well of 12-well plates and cultured for 21 days. Plates were then washed with 2 mL PBS, fixed with 3 mL 10% formalin (Sigma) for 15 minutes, and stained with crystal violet for 5 minutes (0.025% w/v; Sigma). The number of colonies was counted in each well, excluding small colonies (<50 cells) [170].

**Animal studies**

Northwestern University's Animal Care and Use Committee approved all procedures involving animals. Primary LM cells freshly isolated from LM tissue were suspended in rat-tail collagen (type 1) solution (BD Bioscience) at $10^6$ cells/10 μL and cultured for 48 hours in floating cultures. Then, the cell pellets were grafted underneath the kidney capsules of ovariectomized nonobese diabetic-scid IL2Rgnull mouse hosts (NSG, Jackson Laboratory) supplemented with subcutaneous implantation of 50 mg progesterone plus 50 μg estradiol 60-day slow-release pellets (Innovative Research of America Inc.), as a combination of estradiol and progesterone is indispensable for significant LM growth and regeneration in the xenograft mouse model.[49, 64] To test whether 5′-Aza affects the tumor initiation capacity of primary LM cells, cells were treated with vehicle (DMSO), 5′-Aza (100 nM), or RU486(1μM) for 6 days before xenograft. $10^6$ live cells were used for each xenograft. Mice were euthanized 4 weeks post-surgery. To test how 5′-Aza may affect LM growth, 3 weeks after surgery, mice were treated with vehicle or 5′-Aza (0.25 mg/kg, two i.p. injections/week) for 1 week. Then, the mice were subcutaneously implanted with placebo or a RU486 slow-release pellet (12.5 mg/pellet, 30-day release) and the tumors were allowed to grow for another 2 weeks. Mice were euthanized 6 weeks post-surgery. Images of regenerated tumors on the kidney surface were taken from the x-, y-, and z-axes using a dissecting microscope connected to a computer with Leica Application Suite, version 3.8 software (Leica Microsystems Inc.). Tumors were measured by two individuals who were blinded to the treatment group allocation; the average measurement of the two individuals were used for data plotting. Tumor volume was quantified using the following formula: volume ($mm^3$) = 0.52 (derived from $\pi/6$) * length * width * height ($mm^3$) [49].

**Immunohistochemistry**

Paraffin-embedded mouse kidneys xenografted with the primary LM cells were sectioned and immunohistochemistry (IHC) was performed by the Northwestern University Histology and Phenotyping Laboratory to detect the proliferation marker Cyclin D1. Briefly, after deparaffinization, endogenous peroxidase was blocked with 3% hydrogen peroxide in double-distilled water. Endogenous biotin was blocked with avidin, and nonspecific proteins were blocked with 5% normal donkey serum. The tissue sections were then incubated with a primary antibody against Cyclin D1 (Appendix Table 2), followed by incubation with the biotinylated secondary antibody and 3'-diaminobenzidine. Images were captured with a Leica microscope (Leica Microsystems Inc.). Cells stained positive for Cyclin D1 were counted in 40X fields by two individuals who were blinded to the treatment group allocation; the average counts of the two individuals were used for data plotting.

**Immunoblotting analysis**

Protein was extracted from the leiomyoma primary cells using radioimmunoprecipitation assay buffer, followed by quantification using the bicinchoninic acid protein assay reagent (ThermoFisher Scientific, 23225) per the manufacturer's protocols. Then, the protein was diluted with reducing 4X LDS sample buffer (ThermoFisher Scientific, NP0007), electrophoresed on a 4% to 12% Novex Bis-Tris polyacrylamide precast gel (ThermoFisher Scientific, NP0321BOX), and transferred onto polyvinylidene difluoride membrane. Incubation with primary antibodies (PR or $\beta$-actin) (Appendix Table 2) was performed at 4°C in 5% nonfat milk overnight. The membranes were then washed and incubated with the appropriate horseradish peroxidase-conjugated secondary antibodies for 1 hour at room temperature. Detection was performed using Luminata Crescendo horseradish peroxidase substrate (Millipore, WBLUR0100).

**Statistical analysis**

All statistical analyses in this study were performed using GraphPad Prism 8 (GraphPad Inc.) and R (3.6.0); detailed statistical test descriptions are reported in each figure legend. No sample was excluded during the analysis. Values were considered statistically significant when $P<0.05$. Similarity of variance across

compared groups was tested by F-test in two groups and Sum of Square in multiple groups. Normal distribution of data was assessed by the Shapiro–Wilk test. All experiments were repeated with samples from at least three patients, with the patient number (n) noted in the figure legends. Data points in the bar plots represent biological replicates from different patient or mouse and error bars represent SEM.

CHAPTER 4      Discussion and Conclusion

### 4.1.1    Discussion - Mechanisms of dysregulated RANKL Gene Expression in Uterine Leiomyoma: Involvement of progesterone action, epigenetic modification and MED12 gene mutations

The RANKL/RANK pathway is known for its crucial role in paracrine signaling between stem and differentiated cells in breast cancer, where RANKL, released from PR-positive cells upon P4 stimulation, binds to its receptor on PR-negative progenitor cells and stimulates tumorigenesis through expanding the progenitor cell population [171]. RANKL is mainly expressed in LIC, whereas RANK is predominantly present in LSC, suggesting a similar paracrine action of RANKL/RANK in LM tumorigenesis [68]. Here, we found that RANKL treatment induces Cyclin D1 expression and cell proliferation specifically in LSC, leading to LSC expansion. The tumorigenic function of the RANKL/RANK signaling pathway was further supported by our findings that blocking this pathway inhibits *in vivo* steroid hormone (E2 and P4)-dependent LM growth in a xenograft mouse model [68].

Previous studies have demonstrated that RANKL transcription is directly regulated by DNA methylation at its promoter in bone and mammary gland tissues.[102, 128] However, the same region was not differentially methylated between LM and MM, suggesting tissue specificity of methylation-mediated RANKL expression [95]. Using genome-wide MethylCap-Seq, we discovered a DMR adjacent to the distal PRBS identified previously via PR ChIP-Seq in LM cells [44]. *In vitro* de-methylation of this region increased R5020-dependent recruitment of PR to the adjacent PRBS, whereas methylation repressed fused reporter gene luciferase activity, suggesting that the DMR functionally regulates PR access to its responsive element as well as its transcriptional activity. Methyl-CpG has been proposed to interfere with TF binding directly and indirectly [172]. The direct model posits that the specific TF sees a methyl-CpG as a mutation in its binding site and thus is unable to bind. The indirect model proposes that methylated DNA is bound by a nuclear protein(s), such as MeCP-1, which interrupts the particular TF's interaction with its target gene. Our finding that the DMR resides outside but adjacent to the PRBS is compatible with the indirect model of transcriptional interference, although the detailed mechanism needs further investigation [172].

Furthermore, gene transcription is also controlled by extensive crosstalk and interactions between DNA methylation and histone modifications. For example, in embryonic stem cells, H3K4 methylation (including mono-, di-, and tri-methylation) can prevent the binding of DNA methylation machinery DNMT3L, resulting in lower methylation levels at nearby CpG sites.[173] Consistent with this notion, higher levels of active histone marks (H3K27Ac and H3K4me3) and lower levels of a repressive mark (H3K27me3) were detected in the RANKL distal enhancer region with lower DNA methylation in LM versus MM.

The human mediator complex was first identified through an intracellular ligand-dependent association with thyroid hormone receptor [174]. Since then, many nuclear receptors have been shown to interact with the mediator subunits [175]. For example, ESR1's interaction with MED1 is required for ESR1-mediated gene transcription, E2/ESR1-dependent mammary progenitor/stem cell function, and breast cancer cell growth [176, 177]. To our knowledge, our study is the first to report the interaction of MED12 with PR. Importantly, we found that MED12-binding activity at the RANKL distal PRBS not only positively correlated with PR-binding affinity in this region but also positively correlated with RANKL expression. Furthermore, we demonstrated that mut-MED12 increases its interaction with PR, leading to higher RANKL expression in LM carrying mut-MED12. These observations suggest that PR-MED12 interaction in LM is functionally active in regulating P4/PR-mediated gene expression and mut-MED12 may affect chromatin interaction with PR, leading to preferential transcription of genes and pathways that favor fibroid tumorigenesis.

In support of our findings, previous researches have demonstrated that LM-associated MED12 mutations could alter its interaction with proteins involved in transcriptional coactivator pathways, such as the loss of mediator-associated CDK activity [22, 119]. MED12 has also been shown to interact directly with P300, a PR coactivator and histone acetyltransferase, to maintain an active state for the enhancer [178]. Detailed studies are underway to determine whether and how mut-MED12 alters genome-wide PR binding in LM cells to potentially discover the mechanisms of LM tumorigenesis.

In summary, our findings suggest that DNA methylation, histone modification, and MED12 mutation status together constitute a complex regulatory network that influences P4/PR-mediated RANKL transcription. Using RANKL as a representative PR target gene, our study provides novel and important insights into the regulation of PR action and provides rationale for further elucidation of P4/PR-mediated transcriptional regulation in LM. Dissecting the relationship among TF binding, histone modification, and DNA methylation will improve our understanding of the mechanisms underlying LM development. Considering the heterogeneous genetic background of LM pathogenesis, our studies also represent a key step towards a better understanding of mechanisms underlying the pathogenesis of specific LM subtypes and the necessity for personalized therapeutic strategies.

### 4.1.2 Discussion- Targeting DNA methylation and progesterone action during leiomyoma stem cell differentiation

DNA methylation is a common mechanism for cell programming in both normal tissue growth and tumor development [179]. Integrated analyses of genome-wide DNA methylation maps and expression profiles have revealed a hierarchy of cellular differentiation in hematopoietic tissue and the breast [82, 180]. Transcriptomic analysis of LM populations also strongly suggested a hierarchical differentiation order: LSC→LIC→LDC, while MethylCap-seq analysis showed a similar pattern of alterations in DNA methylation, in particular a transition from LSC to LIC or LDC. We found, however, very little difference between LIC and LDC with respect to DNA methylation, suggesting that DNA methylation change is a critical epigenetic modification during the initial differentiation of LSC. Furthermore, we uncovered that regions lost or gain methylation during LSC differentiation were overrepresented in CpG islands and shores, consistent with the DNA methylation pattern changes during intestine stem cell differentiation observed using whole-genome shotgun bisulfite sequencing [141]. The functional significance of this regional preference in stem cell differentiation warrants further investigation. Importantly, hypermethylated regions in LSC were functionally associated with transcriptional suppression of genes enriched in cell differentiation pathways, consistent with other tumor stem cell systems [142, 143].

In the complex biology of LM, progesterone and PGR play dual roles [39, 113]. On the one hand, progesterone/PGR in LIC/LDC induces paracrine signaling that acts on PGR-deficient LSC to stimulate self-renewal and proliferation [68, 69, 140]. Our current findings are suggestive of another intriguing role of PGR: differentiation. PGR target genes are enriched in pathways critical for LSC differentiation and LM development. Expression of PGR and some of its downstream targets were inhibited by hypermethylation in the LSC population. Furthermore, depletion of endogenous PGR shifted the transcriptome of total primary LM cells towards that of LSC. Our data indicates that PGR signaling influences LSC differentiation not only through regulating differentiation-associated genes, but also through affecting DNA methylation process. At a single gene level, Verde et al recently reported that PGR maintains ESR1 expression by preserving a low DNA methylation at the ESR1 promoter in breast cancer cells [181]. We found that PGR regulated the

expression of TET1 and TET2 via binding to their promoter regions and transient PGR depletion upregulated DNA methylation levels at genomic regions hypermethylated in LSC. Given that TETs are lowly expressed in LSC, these findings suggested that PGR activation may contribute to the global DNA methylation loss during LSC differentiation. It is therefore reasonable to suggest that targeting the feedback interaction between DNA methylation and PGR signaling may be a potential strategy to accelerate LSC differentiation and exhaustion, leading to tumor eradication.

Consistent with previous reports showing that DNA hypomethylating agents can remove the epigenetic "brakes" and allow stem cells to undergo differentiation [63], we found that the DNA methylation inhibitor 5′-Aza not only upregulated PGR and its downstream target gene expression, but also depleted the LSC population and its colony formation ability via inducing their differentiation [39, 68]. The 5′-Aza-induced LSC differentiation was also supported by the fact that 5′-Aza significantly increased the expression of markers for fully differentiated smooth muscle cells, such as MYH11 [182]. Similar to breast cancer stem cells [183], deficiency of PGR and ESR1 in LSC may account for the poor sensitivity of LM tumors to hormonal treatments and high frequency of LM tumor recurrence. 5′-Aza seemed to sensitize LSC to RU486 treatment through the upregulation of PGR gene expression. As a demethylation reagent, high-dose 5′-Aza is toxic [184]; however, primary LM cells were sensitive to a very low dose 5′-Aza treatment (25 nM). Intrauterine delivery of low-dose 5′-Aza may be a potential strategy to prevent LM growth and potentiate hormone treatments, with less systemic toxicity.

The strength of our study is that we took the challenge to perform genome-wide studies in each freshly sorted population of LM cells to explore the mechanisms underlying LSC differentiation in hope of targeting this population for potential therapeutic options. The findings derived from the patient samples with minimal levels of experimental manipulation better represents the *in vivo* condition with a higher translational possibility. However, the technical challenge of long-term *in vitro* culture and propagation of LSC prohibited us from performing detailed downstream mechanistic studies specifically on stem cells. Further work needs

to be done to optimize and establish a LSC culture system, which will help us better understand the functional properties of these cells.

In summary, we found that hypermethylation maintains LSC stemness through suppression of PGR and its downstream target genes. This finding not only sheds mechanistic light on the crosstalk between epigenetic and hormonal regulation during LSC differentiation but also suggests a treatment strategy in which LSC can be sensitized to antiprogestin treatments. Our findings are unique because they collectively dissect the complex interaction between progesterone action and DNA methylation during tumor stem cell differentiation and provide new insights into decoding the mechanisms behind the epigenomic regulation of hormone-dependent tumor growth.

### 4.1.3 Summary

Uterine leiomyomas (LM) are the most common tumors and cause significant morbidity (heavy menstrual bleeding, anemia, significant pain, and infertility) in women. Progesterone (P4) and its receptor (PR) are essential for LM growth but anti-P4 treatments have limited efficacy. Thus, surgery is the only effective treatment option. Understanding how LM develops is essential for identifying novel non-surgical treatments.

The Bulun lab previously defined three molecularly and functionally distinct cell populations in LM: stem (LSC), intermediate (LIC), and differentiated cells (LDC). These different cell populations play unique and interdependent roles in the pathogenesis of LM. This thesis projects focus on exploring the molecular mechanisms underlying the LSC differentiation process and identifying potential regulators of stem cell activity that may serve as novel therapeutic targets.

Utilizing high-throughput sequencing methods, we characterized the cistrome of PR, transcriptome, methylome, and histone modifications (H3K4me3, H3K27Ac, H3K27me3) of the three LM cell populations. RNA-Seq (transcriptome) revealed that thousands of genes were differentially expressed in the three cell populations and suggested a hierarchical differentiation order as they differentiate from LSC to LIC to LDC. MethylCap-seq (methylome) showed that LSC displayed a distinct DNA methylation pattern compared with the other two populations. Via integrative analyses of transcriptome and methylome data, we found significant inverse correlations between DNA methylation and gene expression changes during the differentiation process of LSC. Pathway analysis of differentially methylated and expressed genes uncovered several pathways that play crucial roles in LM tumorigenesis, including the WNT/beta-catenin pathway, PI3K/AKT pathway, ECM receptor regulation, and steroid hormone-related signals.

After diligently screening genes enriched in the above key pathways, we discovered that LIC, which expresses high levels of PR, plays a critical role in transducing P4-mediated paracrine mitotic signals to PR-deficient LSC to stimulate their self-renewal and proliferation. For example, *RANKL*, a critical PR-target gene, was preferentially expressed in LIC while its receptor RANK was mainly expressed in LSC. We further

characterized the function of RANKL in LSC and demonstrated that RANKL specifically stimulated LSC proliferation. Our findings established a completely new mechanistic concept that the LIC population serves as support cells to stimulate LSC tumorigenesis. Through investigating transcriptional regulation of *RANKL*, we also found that PR action in LM can be influenced by DNA methylation and *MED12* mutation status, a somatic mutation that occurs in 70% of LM. This study not only revealed potential targets for drug development, but also set the foundation for pharmacogenomics and precision medicine in LM.

Furthermore, via vigorous bioinformatic analysis of PR and histone modification ChIP-seq and DNA methylation data, we discovered that PR locus and PR-bound regions were hypermethylated in the PR-deficient LSC compared with PR-rich LIC and LDC. Genes associated with these hypermethylated regions had the lowest enrichment of the active histone mark (H3K27ac) and the lowest transcription levels, indicating an inhibitory status of these genes in LSC. Gene Ontology analysis uncovered that these genes were highly enriched in key pathways for LSC function, including smooth muscle development, stem cell differentiation, and growth factor stimulation. Treatment with a DNA methylation inhibitor not only upregulated PR expression but also increased the expression of P4/PR target genes via stimulating recruitment of PR to target gene loci. Functionally, treatment of this DNA methylation inhibitor, but not Ru486 (P4 antagonist), significantly reduced LSC number and their colony formation ability. Using an *in vivo* xenograft mouse model, we demonstrated that the DNA methylation inhibitor modestly decreased LM tumorigenic potential, but co-treatment with Ru486 completely blocked LM regeneration ability. These findings suggest that inhibition of DNA methylation can drive LSC differentiation and stimulate their responsiveness to anti-P4 therapy. This study not only provides a novel insight into the crosstalk between epigenetic regulation and hormone action in LM, but also sheds new mechanistic light on the differentiation process of LSC.

In summary, this thesis study not only dissect the molecular mechanism underlying the regulation of stem cells in hormone-sensitive tumors, but also provided opportunities to develop more tractable, less toxic, and personalized approaches to treat uterine fibroids - a disease that affects the majority of women worldwide.

These studies generated three first-authored papers (CHAPTER 2: published in *Oncogene*; CHAPTER 3: two manuscripts under preparation) and two second-authored papers in *The Journal of Clinical Endocrinology and Metabolism* and *BMC Bioinformatics*.

**References**

1.      Walker, C.L. and E.A. Stewart, *Uterine Fibroids: The Elephant in the Room.* Science, 2005. **308**(5728): p. 1589.

2.      Kim, J.J., T. Kurita, and S.E. Bulun, *Progesterone Action in Endometrial Cancer, Endometriosis, Uterine Fibroids, and Breast Cancer.* Endocrine Reviews, 2013. **34**(1): p. 130-162.

3.      Bulun, S.E., *Uterine Fibroids.* New England Journal of Medicine, 2013. **369**(14): p. 1344-1355.

4.      Day Baird, D., et al., *High cumulative incidence of uterine leiomyoma in black and white women: Ultrasound evidence.* American Journal of Obstetrics and Gynecology, 2003. **188**(1): p. 100-107.

5.      Marshall, L.M., et al., *Variation in the incidence of uterine leiomyoma among premenopausal women by age and race.* Obstetrics and gynecology, 1997. **90**(6): p. 967-973.

6.      Cardozo, E.R., et al., *The estimated annual cost of uterine leiomyomata in the United States.* American journal of obstetrics and gynecology, 2012. **206**(3): p. 211.e1-211.e2119.

7.      Sabry, M. and A. Al-Hendy, *Medical treatment of uterine leiomyoma.* Reprod Sci, 2012. **19**(4): p. 339-53.

8.      Elizabeth A Stewart, S.K.L.-T., *Uterine leiomyomas (fibroids): Epidemiology, clinical features, diagnosis, and natural history.* In: UpToDate, Post TW (Ed), UpToDate, Waltham, MA., (Accessed on Dec 13, 2019.).

9.      Munro, M.G., et al., *The FIGO classification of causes of abnormal uterine bleeding in the reproductive years.* Fertility and sterility, 2011. **95**(7): p. 2204-2208.e22083.

10.     Fusco, A. and M. Fedele, *Roles of HMGA proteins in cancer.* Nature Reviews Cancer, 2007. **7**(12): p. 899-910.

11.     Klemke, M., et al., *Overexpression of HMGA2 in uterine leiomyomas points to its general role for the pathogenesis of the disease.* Genes, Chromosomes and Cancer, 2009. **48**(2): p. 171-178.

12.     Mas, A., et al., *Overexpression of the truncated form of High Mobility Group A proteins (HMGA2) in human myometrial cells induces leiomyoma-like tissue formation.* Molecular Human Reproduction, 2014. **21**(4): p. 330-338.

13. Markowski, D., et al., *HMGA2 and p14Arf: Major Roles in Cellular Senescence of Fibroids and Therapeutic Implications.* Anticancer research, 2011. **31**: p. 753-61.

14. Peng, Y., et al., *Antiproliferative Effects by Let-7 Repression of High-Mobility Group A2 in Uterine Leiomyoma.* Molecular Cancer Research, 2008. **6**(4): p. 663.

15. Hunter, D.S., et al., *Aberrant Expression of HMGA2 in Uterine Leiomyoma Associated with Loss of TSC2 Tumor Suppressor Gene Function.* Cancer Research, 2002. **62**(13): p. 3766.

16. Makinen, N., et al., *MED12, the mediator complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas.* Science, 2011. **334**(6053): p. 252-5.

17. Malik, S. and R.G. Roeder, *The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation.* Nature reviews. Genetics, 2010. **11**(11): p. 761-772.

18. Bertsch, E., et al., *MED12 and HMGA2 mutations: two independent genetic events in uterine leiomyoma and leiomyosarcoma.* Modern Pathology, 2014. **27**(8): p. 1144-1153.

19. Conaway, R.C. and J.W. Conaway, *Function and regulation of the Mediator complex.* Current opinion in genetics & development, 2011. **21**(2): p. 225-230.

20. Akoulitchev, S., S. Chuikov, and D. Reinberg, *TFIIH is negatively regulated by cdk8-containing mediator complexes.* Nature, 2000. **407**(6800): p. 102-106.

21. Donner, A.J., et al., *CDK8 Is a Stimulus-Specific Positive Coregulator of p53 Target Genes.* Molecular Cell, 2007. **27**(1): p. 121-133.

22. Turunen, M., et al., *Uterine leiomyoma-linked MED12 mutations disrupt mediator-associated CDK activity.* Cell Rep, 2014. **7**(3): p. 654-60.

23. Khan, K.N., et al., *Changes in tissue inflammation, angiogenesis and apoptosis in endometriosis, adenomyosis and uterine myoma after GnRH agonist therapy.* Human reproduction (Oxford, England), 2010. **25**(3): p. 642-653.

24. Surrey, E.S., *Gonadotropin-releasing hormone agonist and add-back therapy: what do the data show?* Current opinion in obstetrics & gynecology, 2010. **22**(4): p. 283-288.

25. Taraborrelli, S., *Physiology, production and action of progesterone.* Acta Obstetricia et Gynecologica Scandinavica, 2015. **94**(S161): p. 8-16.

26. Stricker, R., et al., *Establishment of detailed reference values for luteinizing hormone, follicle stimulating hormone, estradiol, and progesterone during different phases of the menstrual cycle on the Abbott ARCHITECT analyzer.* Clinical chemistry and laboratory medicine, 2006. **44**(7): p. 883-887.

27. Vermesh, M. and O.A. Kletzky, *Longitudinal evaluation of the luteal phase and its transition into the follicular phase.* The Journal of clinical endocrinology and metabolism, 1987. **65**(4): p. 653-658.

28. Daniel, A.R., C.R. Hagan, and C.A. Lange, *Progesterone receptor action: defining a role in breast cancer.* Expert Review of Endocrinology & Metabolism, 2011. **6**(3): p. 359-369.

29. Jacobsen, B.M. and K.B. Horwitz, *Progesterone receptors, their isoforms and progesterone regulated transcription.* Molecular and Cellular Endocrinology, 2012. **357**(1): p. 18-29.

30. Kastner, P., et al., *Two distinct estrogen-regulated promoters generate transcripts encoding the two functionally different human progesterone receptor forms A and B.* The EMBO Journal, 1990. **9**(5): p. 1603-1614.

31. Bonéy-Montoya, J., et al., *Long-Range Transcriptional Control of Progesterone Receptor Gene Expression.* Molecular Endocrinology, 2010. **24**(2): p. 346-358.

32. Condon, J.C., et al., *Up-Regulation of the Progesterone Receptor (PR)-C Isoform in Laboring Myometrium by Activation of Nuclear Factor-κB May Contribute to the Onset of Labor through Inhibition of PR Function.* Molecular Endocrinology, 2006. **20**(4): p. 764-775.

33. Hopp, T.A., et al., *Breast Cancer Patients with Progesterone Receptor PR-A-Rich Tumors Have Poorer Disease-Free Survival Rates.* Clinical Cancer Research, 2004. **10**(8): p. 2751.

34. Scarpin, K.M., et al., *Progesterone action in human tissues: regulation by progesterone receptor (PR) isoform expression, nuclear positioning and coregulator expression.* Nuclear receptor signaling, 2009. **7**: p. e009-e009.

35. Han, S.J., et al., *Dynamic cell type specificity of SRC-1 coactivator in modulating uterine progesterone receptor function in mice.* Molecular and cellular biology, 2005. **25**(18): p. 8150-8165.

36. Lonard, D.M., R.B. Lanz, and B.W. O'Malley, *Nuclear Receptor Coregulators and Human Disease.* Endocrine Reviews, 2007. **28**(5): p. 575-587.

37.      Li, Q., et al., *The antiproliferative action of progesterone in uterine epithelium is mediated by Hand2.* Science (New York, N.Y.), 2011. **331**(6019): p. 912-916.

38.      Clarke, R.B., et al., *Dissociation between Steroid Receptor Expression and Cell Proliferation in the Human Breast.* Cancer Research, 1997. **57**(22): p. 4987.

39.      Lee, H.J., et al., *Progesterone drives mammary secretory differentiation via RankL-mediated induction of Elf5 in luminal progenitor cells.* Development, 2013. **140**(7): p. 1397-401.

40.      Hilton, H.N., et al., *The antiproliferative effects of progestins in T47D breast cancer cells are tempered by progestin induction of the ETS transcription factor Elf5.* Molecular endocrinology (Baltimore, Md.), 2010. **24**(7): p. 1380-1392.

41.      Finlay-Schultz, J., et al., *Breast Cancer Suppression by Progesterone Receptors Is Mediated by Their Modulation of Estrogen Receptors and RNA Polymerase III.* Cancer Research, 2017. **77**(18): p. 4934.

42.      Chen, C.-C., D.B. Hardy, and C.R. Mendelson, *Progesterone receptor inhibits proliferation of human breast cancer cells via induction of MAPK phosphatase 1 (MKP-1/DUSP1).* The Journal of biological chemistry, 2011. **286**(50): p. 43091-43102.

43.      Moore, N.L., et al., *Anti-proliferative transcriptional effects of medroxyprogesterone acetate in estrogen receptor positive breast cancer cells are predominantly mediated by the progesterone receptor.* The Journal of Steroid Biochemistry and Molecular Biology, 2020. **199**: p. 105548.

44.      Yin, P., et al., *Genome-Wide Progesterone Receptor Binding: Cell Type-Specific and Shared Mechanisms in T47D Breast Cancer Cells and Primary Leiomyoma Cells.* PLoS ONE, 2012. **7**(1): p. e29021.

45.      Maruo, T., et al., *Sex steroidal regulation of uterine leiomyoma growth and apoptosis.* Human Reproduction Update, 2004. **10**(3): p. 207-220.

46.      Kim, J.J. and E.C. Sefton, *The role of progesterone signaling in the pathogenesis of uterine leiomyoma.* Molecular and cellular endocrinology, 2012. **358**(2): p. 223-231.

47.      Yamada, T., et al., *Progesterone down-regulates insulin-like growth factor-I expression in cultured human uterine leiomyoma cells.* Human Reproduction, 2004. **19**(4): p. 815-821.

48.  Kawaguchi, K., et al., *Mitotic activity in uterine leiomyomas during the menstrual cycle.* American Journal of Obstetrics and Gynecology, 1989. **160**(3): p. 637-641.

49.  Ishikawa, H., et al., *Progesterone Is Essential for Maintenance and Growth of Uterine Leiomyoma.* Endocrinology, 2010. **151**(6): p. 2433-2442.

50.  Jordan, C.T., M.L. Guzman, and M. Noble, *Cancer Stem Cells.* New England Journal of Medicine, 2006. **355**(12): p. 1253-1261.

51.  Kreso, A. and John E. Dick, *Evolution of the Cancer Stem Cell Model.* Cell Stem Cell, 2014. **14**(3): p. 275-291.

52.  Lapidot, T., et al., *A cell initiating human acute myeloid leukaemia after transplantation into SCID mice.* Nature, 1994. **367**(6464): p. 645-648.

53.  Singh, S.K., et al., *Identification of human brain tumour initiating cells.* Nature, 2004. **432**(7015): p. 396-401.

54.  Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proceedings of the National Academy of Sciences, 2003. **100**(7): p. 3983.

55.  Bu, Y.W. and D. Cao, *The origin of cancer stem cells.* Frontiers in bioscience (Scholar edition), 2012. **4**: p. 819-30.

56.  Wang, Y., et al., *Expression of mutant p53 proteins implicates a lineage relationship between neural stem cells and malignant astrocytic glioma in a murine model.* Cancer cell, 2009. **15**(6): p. 514-526.

57.  Nouri, M., et al., *Therapy-induced developmental reprogramming of prostate cancer cells and acquired therapy resistance.* Oncotarget, 2017. **8**(12): p. 18949-18967.

58.  Schatton, T., N.Y. Frank, and M.H. Frank, *Identification and targeting of cancer stem cells.* BioEssays : news and reviews in molecular, cellular and developmental biology, 2009. **31**(10): p. 1038-1049.

59.  Yang, L., et al., *Targeting cancer stem cell pathways for cancer therapy.* Signal Transduction and Targeted Therapy, 2020. **5**(1): p. 8.

60.  Cai, W., et al., *PMP22 Regulates Self-Renewal and Chemoresistance of Gastric Cancer Cells.* Molecular Cancer Therapeutics, 2017. **16**(6): p. 1187.

61.     Wang, D., et al., *Prostaglandin E2 Promotes Colorectal Cancer Stem Cell Expansion and Metastasis in Mice.* Gastroenterology, 2015. **149**(7): p. 1884-1895.e4.

62.     Dean, M., T. Fojo, and S. Bates, *Tumour stem cells and drug resistance.* Nature Reviews Cancer, 2005. **5**(4): p. 275-284.

63.     Nephew, K.P., et al., *Epigenetic Targeting of Ovarian Cancer Stem Cells.* Cancer Research, 2014.

64.     Yin, P., et al., *Human Uterine Leiomyoma Stem/Progenitor Cells Expressing CD34 and CD49b Initiate Tumors In Vivo.* The Journal of Clinical Endocrinology and Metabolism, 2015. **100**(4): p. E601-E606.

65.     Ono, M., et al., *Role of Stem Cells in Human Uterine Leiomyoma Growth.* PLoS ONE, 2012. **7**(5): p. e36935.

66.     Mas, A., et al., *Identification and characterization of the human leiomyoma side population as putative tumor-initiating cells.* Fertility and Sterility, 2012. **98**(3): p. 741-751.e6.

67.     Wu, C. and B.A. Alman, *Side population cells in human cancers.* Cancer Letters, 2008. **268**(1): p. 1-9.

68.     Ikhena, D.E., et al., *RANKL/RANK Pathway and Its Inhibitor RANK-Fc in Uterine Leiomyoma Growth.* The Journal of Clinical Endocrinology & Metabolism, 2018. **103**(5): p. 1842-1849.

69.     Moravek, M.B., et al., *Paracrine Pathways in Uterine Leiomyoma Stem Cells Involve Insulinlike Growth Factor 2 and Insulin Receptor A.* The Journal of Clinical Endocrinology & Metabolism, 2017. **102**(5): p. 1588-1595.

70.     Zemach, A., et al., *Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation.* Science, 2010. **328**(5980): p. 916.

71.     Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development.* Nature Reviews Genetics, 2013. **14**(3): p. 204-220.

72.     Greenberg, M.V.C. and D. Bourc'his, *The diverse roles of DNA methylation in mammalian development and disease.* Nature Reviews Molecular Cell Biology, 2019. **20**(10): p. 590-607.

73.     Guo, X., et al., *Structural insight into autoinhibition and histone H3-induced activation of DNMT3A.* Nature, 2015. **517**(7536): p. 640-644.

74. Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.* Science (New York, N.Y.), 2009. **324**(5929): p. 930-935.

75. Putiri, E.L., et al., *Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells.* Genome biology, 2014. **15**(6): p. R81-R81.

76. Wang, L., et al., *Programming and inheritance of parental DNA methylomes in mammals.* Cell, 2014. **157**(4): p. 979-991.

77. Rasmussen, K.D. and K. Helin, *Role of TET enzymes in DNA methylation, development, and cancer.* Genes & development, 2016. **30**(7): p. 733-750.

78. Guilhamon, P., et al., *Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2.* Nature communications, 2013. **4**: p. 2166-2166.

79. Rampal, R., et al., *DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia.* Cell reports, 2014. **9**(5): p. 1841-1855.

80. Tadokoro, Y., et al., *De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells.* The Journal of experimental medicine, 2007. **204**(4): p. 715-722.

81. Trowbridge, J.J., et al., *DNA Methyltransferase 1 Is Essential for and Uniquely Regulates Hematopoietic Stem and Progenitor Cells.* Cell Stem Cell, 2009. **5**(4): p. 442-449.

82. Bock, C., et al., *DNA Methylation Dynamics during In Vivo Differentiation of Blood and Skin Stem Cells.* Molecular Cell, 2012. **47**(4): p. 633-647.

83. Ehrlich, M., *DNA methylation in cancer: too much, but also too little.* Oncogene, 2002. **21**(35): p. 5400-5413.

84. Kulis, M. and M. Esteller, *2 - DNA Methylation and Cancer*, in *Advances in Genetics*, Z. Herceg and T. Ushijima, Editors. 2010, Academic Press. p. 27-56.

85. Badal, V., et al., *CpG Methylation of Human Papillomavirus Type 16 DNA in Cervical Cancer Cell Lines and in Clinical Specimens: Genomic Hypomethylation Correlates with Carcinogenic Progression.* Journal of Virology, 2003. **77**(11): p. 6227.

86.     Greger, V., et al., *Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma.* Human Genetics, 1989. **83**(2): p. 155-158.

87.     Spitzwieser, M., et al., *Hypermethylation of CDKN2A exon 2 in tumor, tumor-adjacent and tumor-distant tissues from breast cancer patients.* BMC Cancer, 2017. **17**(1): p. 260.

88.     Xing, X., et al., *The prognostic value of CDKN2A hypermethylation in colorectal cancer: a meta-analysis.* British journal of cancer, 2013. **108**(12): p. 2542-2548.

89.     Robertson, K.D., et al., *The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors.* Nucleic acids research, 1999. **27**(11): p. 2291-2298.

90.     Russler-Germain, David A., et al., *The R882H DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its Ability to Form Active Tetramers.* Cancer Cell, 2014. **25**(4): p. 442-454.

91.     Langemeijer, S.M.C., et al., *Acquired mutations in TET2 are common in myelodysplastic syndromes.* Nature Genetics, 2009. **41**(7): p. 838-842.

92.     Spencer, D.H., et al., *CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression.* Cell, 2017. **168**(5): p. 801-816.e13.

93.     Ko, M., et al., *Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2.* Nature, 2010. **468**(7325): p. 839-843.

94.     Cimmino, L., et al., *Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression.* Cell, 2017. **170**(6): p. 1079-1095.e20.

95.     Navarro, A., et al., *Genome-wide DNA methylation indicates silencing of tumor suppressor genes in uterine leiomyoma.* PLoS One, 2012. **7**(3): p. e33284.

96.     Maekawa, R., et al., *Genome-wide DNA methylation analysis reveals a potential mechanism for the pathogenesis and development of uterine leiomyomas.* PloS one, 2013. **8**(6): p. e66632-e66632.

97.    Navarro, A., et al., *5-Hydroxymethylcytosine Promotes Proliferation of Human Uterine Leiomyoma: A Biological Link to a New Epigenetic Modification in Benign Tumors.* The Journal of Clinical Endocrinology & Metabolism, 2014. **99**(11): p. E2437-E2445.

98.    Wada, T., et al., *RANKL–RANK signaling in osteoclastogenesis and bone disease.* Trends in Molecular Medicine, 2006. **12**(1): p. 17-25.

99.    Rao, S., et al., *RANKL and RANK: From Mammalian Physiology to Cancer Treatment.* Trends in Cell Biology, 2018. **28**(3): p. 213-223.

100.   Yasuda, H., et al., *Osteoclast differentiation factor is a ligand for osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical to TRANCE/RANKL.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(7): p. 3597-3602.

101.   Glass, D.A., 2nd, et al., *Canonical Wnt signaling in differentiated osteoblasts controls osteoclast differentiation.* Developmental cell, 2005. **8**(5): p. 751-764.

102.   Delgado-Calle, J., et al., *Role of DNA methylation in the regulation of the RANKL-OPG system in human bone.* Epigenetics, 2012. **7**(1): p. 83-91.

103.   Gonzalez-Suarez, E., et al., *RANK Overexpression in Transgenic Mice with Mouse Mammary Tumor Virus Promoter-Controlled RANK Increases Proliferation and Impairs Alveolar Differentiation in the Mammary Epithelia and Disrupts Lumen Formation in Cultured Epithelial Acini.* Molecular and Cellular Biology, 2007. **27**(4): p. 1442.

104.   Schramek, D., et al., *Osteoclast differentiation factor RANKL controls development of progestin-driven mammary cancer.* Nature, 2010. **468**(7320): p. 98-102.

105.   Fata, J.E., et al., *The Osteoclast Differentiation Factor Osteoprotegerin-Ligand Is Essential for Mammary Gland Development.* Cell, 2000. **103**(1): p. 41-50.

106.   Kiesel, L. and A. Kohl, *Role of the RANK/RANKL pathway in breast cancer.* Maturitas, 2016. **86**: p. 10-16.

107.   Ablett, M.P., J.K. Singh, and R.B. Clarke, *Stem cells in breast tumours: Are they ready for the clinic?* European Journal of Cancer, 2012. **48**(14): p. 2104-2116.

108.    Sflomos, G. and C. Brisken, *A new Achilles Heel in breast cancer?* Oncotarget, 2013. **4**(8): p. 1126-1127.

109.    Tanos, T., et al., *Progesterone/RANKL Is a Major Regulatory Axis in the Human Breast.* Science Translational Medicine, 2013. **5**(182): p. 182ra55.

110.    Gonzalez-Suarez, E., et al., *RANK ligand mediates progestin-induced mammary epithelial proliferation and carcinogenesis.* Nature, 2010. **468**(7320): p. 103-107.

111.    Catherino, W.H., E. Parrott, and J. Segars, *Proceedings from theNational Institute of Child Health and Human Development conference on the Uterine Fibroid Research Update Workshop.* Fertil Steril, 2011. **95**(1): p. 9-12.

112.    Renema, N., et al., *RANK–RANKL signalling in cancer.* Bioscience Reports, 2016. **36**(4): p. e00366.

113.    Joshi, P.A., et al., *Progesterone induces adult mammary stem cell expansion.* Nature, 2010. **465**(7299): p. 803-807.

114.    Asselin-Labat, M.-L., et al., *Control of mammary stem cell function by steroid hormone signalling.* Nature, 2010. **465**: p. 798.

115.    Hu, H., et al., *RANKL expression in normal and malignant breast tissue responds to progesterone and is up-regulated during the luteal phase.* Breast Cancer Research and Treatment, 2014. **146**(3): p. 515-523.

116.    Ohm, J.E., et al., *A stem cell–like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing.* Nature Genetics, 2007. **39**: p. 237.

117.    Razin, A. and H. Cedar, *DNA methylation and gene expression.* Microbiol Rev, 1991. **55**(3): p. 451-8.

118.    Dyson, M.T., et al., *Genome-wide DNA methylation analysis predicts an epigenetic switch for GATA factor expression in endometriosis.* PLoS Genet, 2014. **10**(3): p. e1004158.

119.    Banaganapalli, B., et al., *A Computational Protein Phenotype Prediction Approach to Analyze the Deleterious Mutations of Human MED12 Gene.* Journal of Cellular Biochemistry, 2016. **117**(9): p. 2023-2035.

120. Gonzalez-Suarez, E., et al., *RANK ligand mediates progestin-induced mammary epithelial proliferation and carcinogenesis.* Nature, 2010. **468**(7320): p. 103-7.

121. Kiesel, L. and A. Kohl, *Role of the RANK/RANKL pathway in breast cancer.* Maturitas, 2016. **86**: p. 10-6.

122. Faivre, E.J., et al., *Progesterone Receptor Rapid Signaling Mediates Serine 345 Phosphorylation and Tethering to Specificity Protein 1 Transcription Factors.* Molecular Endocrinology, 2008. **22**(4): p. 823-837.

123. Díaz Flaqué, M.C., et al., *Progesterone receptor assembly of a transcriptional complex along with activator protein 1, signal transducer and activator of transcription 3 and ErbB-2 governs breast cancer growth and predicts response to endocrine therapy.* Breast Cancer Research, 2013. **15**(6): p. R118.

124. Hu, Y., et al., *Transcriptional regulation of human RANK ligand gene expression by E2F1.* Biochemical and Biophysical Research Communications, 2008. **370**(3): p. 440-444.

125. Fan, X., et al., *Regulation of RANKL promoter activity is associated with histone remodeling in murine bone stromal cells.* Journal of Cellular Biochemistry, 2004. **93**(4): p. 807-818.

126. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome.* Cell, 2007. **129**(4): p. 823-37.

127. Nishikawa, K., et al., *DNA methyltransferase 3a regulates osteoclast differentiation by coupling to an S-adenosylmethionine-producing metabolic pathway.* Nat Med, 2015. **21**(3): p. 281-7.

128. Li, Y., et al., *A seven-gene CpG-island methylation panel predicts breast cancer progression.* BMC Cancer, 2015. **15**(1): p. 417.

129. Malik, M., J. Segars, and W.H. Catherino, *Integrin beta1 regulates leiomyoma cytoskeletal integrity and growth.* Matrix Biol, 2012. **31**(7-8): p. 389-97.

130. Ono, M., et al., *Paracrine activation of WNT/beta-catenin pathway in uterine leiomyoma stem cells promotes tumor growth.* Proc Natl Acad Sci U S A, 2013. **110**(42): p. 17053-8.

131. Borahay, M.A., et al., *Signaling Pathways in Leiomyoma: Understanding Pathobiology and Implications for Therapy.* Mol Med, 2015. **21**(1): p. 242-56.

132. Maja Klug, M.R., *Functional Analysis of Promoter CPG- Methylation using a CpG-Free Luciferase Reporter Vector.* Epigenetics, 1:3, 127-130, DOI: 10.4161/epi.1.3.3327, 2006.

133. Mittal, P., et al., *Med12 gain-of-function mutation causes leiomyomas and genomic instability.* J Clin Invest, 2015. **125**(8): p. 3280-4.

134. Hagege, H., et al., *Quantitative analysis of chromosome conformation capture assays (3C-qPCR).* Nat Protoc, 2007. **2**(7): p. 1722-33.

135. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.

136. Shen, L., et al., *ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases.* BMC Genomics, 2014. **15**: p. 284.

137. Bonifacino Juan, S., C. Gershlick David, and C. Dell'Angelica Esteban, *Immunoprecipitation.* Current Protocols in Cell Biology, 2016. **71**(1).

138. Klug, M. and M. Rehli, *Functional Analysis of Promoter CPG-Methylation using a CpG-Free Luciferase Reporter Vector.* Epigenetics, 2014. **1**(3): p. 127-130.

139. Cardozo, E.R., et al., *The estimated annual cost of uterine leiomyomata in the United States.* Am J Obstet Gynecol, 2012. **206**(3): p. 211 e1-9.

140. Ono, M., et al., *Paracrine activation of WNT/β-catenin pathway in uterine leiomyoma stem cells promotes tumor growth.* Proceedings of the National Academy of Sciences, 2013. **110**(42): p. 17053-17058.

141. Sheaffer, K.L., et al., *DNA methylation is required for the control of stem cell differentiation in the small intestine.* Genes Dev, 2014. **28**(6): p. 652-64.

142. Widschwendter, M., et al., *Epigenetic stem cell signature in cancer.* Nature Genetics, 2006. **39**: p. 157.

143. Schuebel, K., W. Chen, and S.B. Baylin, *CIMPle origin for promoter hypermethylation in colorectal cancer?* Nature Genetics, 2006. **38**(7): p. 738-740.

144. Lee, T.I., et al., *Control of Developmental Regulators by Polycomb in Human Embryonic Stem Cells.* Cell, 2006. **125**(2): p. 301-313.

145. Sheaffer, K.L., et al., *DNA methylation is required for the control of stem cell differentiation in the small intestine.* Genes & development, 2014. **28**(6): p. 652-664.

146. Tate, P.H. and A.P. Bird, *Effects of DNA methylation on DNA-binding proteins and gene expression.* Current Opinion in Genetics & Development, 1993. **3**(2): p. 226-231.

147. Chegini, N., et al., *The expression of Smads and transforming growth factor beta receptors in leiomyoma and myometrium and the effect of gonadotropin releasing hormone analogue therapy.* Molecular and Cellular Endocrinology, 2003. **209**(1): p. 9-16.

148. Shozu, M., et al., *Decreased Expression of Early Growth Response-1 and Its Role in Uterine Leiomyoma Growth.* Cancer Research, 2004. **64**(13): p. 4677.

149. Norian, J.M., et al., *Transforming Growth Factor β3 Regulates the Versican Variants in the Extracellular Matrix-Rich Uterine Leiomyomas.* Reproductive Sciences, 2009. **16**(12): p. 1153-1164.

150. Maekawa, R., et al., *Aberrant DNA methylation suppresses expression of estrogen receptor 1 (ESR1) in ovarian endometrioma.* Journal of Ovarian Research, 2019. **12**(1): p. 14.

151. Wolańska, M. and E. Bańkowski, *Fibroblast growth factors (FGF) in human myometrium and uterine leiomyomas in various stages of tumour growth.* Biochimie, 2006. **88**(2): p. 141-146.

152. Kessenbrock, K., V. Plaks, and Z. Werb, *Matrix Metalloproteinases: Regulators of the Tumor Microenvironment.* Cell, 2010. **141**(1): p. 52-67.

153. Liu, H., et al., *TIMP3: a physiological regulator of adult myogenesis.* Journal of Cell Science, 2010. **123**(17): p. 2914.

154. Jackson, H.W., et al., *Expansion of stem cells counteracts age-related mammary regression in compound Timp1/Timp3 null mice.* Nature Cell Biology, 2015. **17**: p. 217.

155. Tarfiei, G., et al., *ROR2 Promoter Methylation Change in Osteoblastic Differentiation of Mesenchymal Stem Cells.* Cell journal, 2011. **13**(1): p. 11-15.

156.     Rae, J.M., et al., *GREB1 is a critical regulator of hormone dependent breast cancer growth.* Breast Cancer Research and Treatment, 2005. **92**(2): p. 141-149.

157.     Zhu, L., et al., *Mutations in myosin heavy chain 11 cause a syndrome associating thoracic aortic aneurysm/aortic dissection and patent ductus arteriosus.* Nature Genetics, 2006. **38**(3): p. 343-349.

158.     Zhou, Y., T.J. Kipps, and S. Zhang, *Wnt5a Signaling in Normal and Cancer Stem Cells.* Stem cells international, 2017. **2017**: p. 5295286-5295286.

159.     Rajendran, V. and M.V. Jain, *In Vitro Tumorigenic Assay: Colony Forming Assay for Cancer Stem Cells*, in *Cancer Stem Cells: Methods and Protocols*, G. Papaccio and V. Desiderio, Editors. 2018, Springer New York: New York, NY. p. 89-95.

160.     Allan, G.F., et al., *Ligand-dependent conformational changes in the progesterone receptor are necessary for events that follow DNA binding.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(24): p. 11750-11754.

161.     Liu, S., et al., *Progesterone receptor integrates the effects of mutated MED12 and altered DNA methylation to stimulate RANKL expression and stem cell proliferation in uterine leiomyoma.* Oncogene, 2018.

162.     Kangaspeska, S., et al., *Transient cyclical methylation of promoter DNA.* Nature, 2008. **452**(7183): p. 112-115.

163.     van Galen, P., et al., *A Multiplexed System for Quantitative Comparisons of Chromatin Landscapes.* Molecular Cell. **61**(1): p. 170-180.

164.     Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biology, 2013. **14**(4): p. R36.

165.     Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nature Methods, 2012. **9**: p. 357.

166.     Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics (Oxford, England), 2009. **25**(14): p. 1754-1760.

167. Subramanian, A., et al., *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.* Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15545.

168. Mootha, V.K., et al., *PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nature Genetics, 2003. **34**(3): p. 267-273.

169. Hoekstra, A.V., et al., *Progestins activate the AKT pathway in leiomyoma cells and promote survival.* The Journal of clinical endocrinology and metabolism, 2009. **94**(5): p. 1768-1774.

170. Au - Rafehi, H., et al., *Clonogenic Assay: Adherent Cells.* JoVE, 2011(49): p. e2573.

171. Nolan, E., et al., *RANK ligand as a potential target for breast cancer prevention in BRCA1-mutation carriers.* Nature Medicine, 2016. **22**: p. 933.

172. Boyes, J. and A. Bird, *DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein.* Cell, 1991. **64**(6): p. 1123-34.

173. Ooi, S.K., et al., *DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.* Nature, 2007. **448**(7154): p. 714-7.

174. Fondell, J.D., H. Ge, and R.G. Roeder, *Ligand induction of a transcriptionally active thyroid hormone receptor coactivator complex.* Proc Natl Acad Sci U S A, 1996. **93**(16): p. 8329-33.

175. Malik, S. and R.G. Roeder, *Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells.* Trends in Biochemical Sciences, 2000. **25**(6): p. 277-283.

176. Acevedo, M.L. and W.L. Kraus, *Mediator and p300/CBP-Steroid Receptor Coactivator Complexes Have Distinct Roles, but Function Synergistically, during Estrogen Receptor-Dependent Transcription with Chromatin Templates.* Molecular and Cellular Biology, 2003. **23**(1): p. 335-348.

177. Chen, W. and R.G. Roeder, *The Mediator subunit MED1/TRAP220 is required for optimal glucocorticoid receptor-mediated transcription activation.* Nucleic Acids Res, 2007. **35**(18): p. 6161-9.

178. Aranda-Orgilles, B., et al., *MED12 Regulates HSC-Specific Enhancers Independently of Mediator Kinase Activity to Control Hematopoiesis.* Cell stem cell, 2016. **19**(6): p. 784-799.

179. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev., 2002. **16(1):6-21.**

180.    Bloushtain-Qimron, N., et al., *Cell type-specific DNA methylation patterns in the human breast.* Proceedings of the National Academy of Sciences, 2008. **105**(37): p. 14076.

181.    Verde, G., et al., *Unliganded Progesterone Receptor Governs Estrogen Receptor Gene Expression by Regulating DNA Methylation in Breast Cancer Cells.* Cancers, 2018. **10**(10): p. 371.

182.    Goikuria, H., et al., *Characterization of Carotid Smooth Muscle Cells during Phenotypic Transition.* Cells, 2018. **7**(3): p. 23.

183.    Liu, S. and M.S. Wicha, *Targeting breast cancer stem cells.* J Clin Oncol, 2010. **28**(25): p. 4006-12.

184.    Christman, J.K., *5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy.* Oncogene, 2002. **21**(35): p. 5483-5495.

# Appendix

**Appendix Table 1**

**Antibodies**

| Antibody | Company | Cat. Number |
|----------|---------|-------------|
| CD45 | BD Biosciences | 564105 |
| CD34 | BD Biosciences | 555824 |
| CD49b | BD Biosciences | 555498 |
| RANKL | Proteintech | 23408-1-AP |
| Beta-actin | Proteintech | HRP-60008 |
| PR | Santa Cruz | sc-7208X |
| MED12 | Bethyl Laboratories Inc | A300-774A |
| H3K4me3 | Millipore | 07-473 |
| H3K27Ac | Active Motif | 39133 |
| H3K27me3 | Millipore | 07-449 |
| Rabbit IgG | Cell signaling Tech | 9003S |

**SYBR Primers**

| Name | Forward | Reverse |
|------|---------|---------|
| RANKL_enhancer_PR_binding | GTCATAATATTATATTGCTGGCTTGGC | ATTCCCCTGGCCTATTAGTCATGAAA |
| RANKL_promoter_PR_binding | CTGATATCCATGGAAGACTGGTTCCA | CAAGTGTACAAGAGGATGTGTGTA |
| RANKL_enhancer_PR_binding_Primer1 | GCTACACTCAGTGAGTCATTCCTGAA | AGCAGATCACATCTTGACTTGAAAAAGC |
| RANKL_enhancer_PR_binding_Primer2 | GTACTAAACAAACAGTTCAGCTCGAAC | TGAATGTACACTGTGAAATTACCACTTC |
| RANKL_enhancer_PR_binding_Primer3 | GAAGTGGTAATTTCACAGTGTACATTCA | ACATATTTGCAAGTTTACCTCATAGTCA |
| MethylCap-qPCR | TGGCACATACTTGTAATCCCAGCTGGT | ATGGAGTCTCGCTCTGTTGCCCA |
| Luciferase_assay | AACCCTGCAGGCATTAGAAAATACAGAAATCA | TGGACTAGTCCTATTACCCGAGTAAACTGAAA |
| 3C_RANKL_Promoter_F | GGTCCCATATATGAAGCTCCTG | TCTTCCATTCTTTTTAATGGGTTT |
| 3C_RANKL_Enhancer_E1 | CCATCCTGTGCATTGTAGGA | TGAAGGCAATTTTTCCATGC |
| 3C_RANKL_Enhancer_E2 | GGAATCAGGAAAAACTTTCCAA | TTGTCCATTGAAGAGAATAAATTCAG |
| 3C_RANKL_Enhancer_E3 | TGGACAAGCAGAGAACATGC | TGAAGATTGTTTTGGCTATTTGG |
| 3C_RANKL_Enhancer_E4 | GCTTGGGAAGGTTCTGATTG | ACTGGATGCACCGAAATGTT |

**TaqMan Expression Primers**

| Gene name | Company | Assay  ID |
|---|---|---|
| RANKL | Life Technologies | Hs00243522 |
| RANK | Life Technologies | Hs00921372_m1 |
| OPG | Life Technologies | Hs00900358_m1 |
| CyclinD1 | Life Technologies | Hs00765553_m1 |
| NANOG | Integrated DNA Technologies | Hs.PT.58.21480849 |
| KLF4 | Integrated DNA Technologies | Hs.PT.58.45542593 |
| ESR1 | Life Technologies | Hs00174860_m1 |
| PR | Life Technologies | Hs01556702_m1 |
| ACTA2 | Life Technologies | Hs00426835_g1 |

**Appendix Table 2**

**Antibodies**

| Antibody | Company | Cat. Number | Usage |
|---|---|---|---|
| CD45 | BD Biosciences | 564105 | FACS |
| CD34 | BD Biosciences | 555824 | FACS |
| CD49b | BD Biosciences | 555498 | FACS |
| PR | Santa Cruz | sc-7208X | ChIP-qPCR/-Seq |
| H3K4me3 | Millipore | 07-473 | ChIP-Seq |
| H3K27Ac | Active Motif | 39133 | ChIP-Seq |
| H3 | Active Motif | 39763 | ChIP-Seq |
| Rabbit IgG | Cell signaling Tech | 9003S | ChIP-qPCR |
| TET1 | Active Motif | 61741 | ChIP-qPCR |
| PR | Dako Products (Agilent) | M3569 | Western Blot |
| beta-actin | Proteintech | HRP-60008 | Western Blot |
| Cyclin D1 | Thermo Fisher Scientific | MA5-14512 | IHC |

**ChIP-qPCR primers**

| Name | Forward | Reverse |
|---|---|---|
| WNT5A_PR_binding | TCTTTGGAGCAGTCCCTTTG | GTGCCAACGTGCGTAGTTTA |
| TIMP3_PR_binding | GATGAGAGATGGGCCTCAGA | AAATGTGCTCTCCCATCACC |
| ROR2_PR_binding | CCACCAATGGGAACAAAAAG | TCCAGTGGAGGTGGTGTGTA |
| GREB1_PR_binding | CACTTTGAGCAAAAGCCACA | GACCCAGTTGCCACACTTTT |
| MYH11_PR_binding | TCTCGGAAAAGACCCAGCTA | GCACTGTTTGTCCCCTGATT |
| TET1_PR_binding_P1 | AGGTCCAGGGCCAAATAACT | AGAAGGTGCCAGGTCAGAGA |
| TET1_PR_binding_P2 | CAAGCGTACCCCCTAAACAA | TTCCCTTCTTCCCTGATCCT |
| TET1_PR_binding_P3 | CTGGGCATTTCTGATCCACT | ATTTGGGAGAGGGACGAGTT |
| TET2_PR_binding_P1 | TTCTCTTATGCCGCGAAACT | AGCTTCCCTCTTCCCTCTTG |
| TET2_PR_binding_P2 | TGTGTGGGAGTGGATTTTGA | GCCAAGGCATTAAATCCTGA |

**MethylCap-qPCR primers**

| Name | Forward | Reverse |
|---|---|---|
| WNT5A_methylation | TCTTTGGAGCAGTCCCTTTG | GTGCCAACGTGCGTAGTTTA |
| TIMP3_methylation | GTGAAACGCCTGATGGTCTT | GGCTTCTGAGGAAACCCTCT |
| ROR2_methylation | CCACCAATGGGAACAAAAAG | TCCAGTGGAGGTGGTGTGTA |
| GREB1_methylation | ATCTTGGCTCACTGCAACCT | AGACGGGTGGATCACTTCAG |
| MYH11_methylation | TCTCGGAAAAGACCCAGCTA | GCACTGTTTGTCCCCTGATT |

**mRNA expression primers**

| Gene name | Company | Assay ID |
|-----------|---------|----------|
| TET1 | Integrated DNA Technologies | Hs.PT.58.27802060 |
| TET2 | Integrated DNA Technologies | Hs.PT.58.21240322 |
| TET3 | Integrated DNA Technologies | Hs.PT.58.4763348 |
| ESR1 | Life Technologies | Hs00174860_m1 |
| PR | Life Technologies | Hs01556702_m1 |
| TBP | Life Technologies | Hs00427620_m1 |
| TIMP3 | Integrated DNA Technologies | Hs.PT.58.1756331 |
| GREB1 | Integrated DNA Technologies | Hs.PT.58.26216464 |
| ROR2 | Integrated DNA Technologies | Hs.PT.58.22908006 |
| MYH11 | Integrated DNA Technologies | Hs.PT.58.2909933 |
| WNT5A | Integrated DNA Technologies | Hs.PT.58.22221435 |